



Vol. 9, No. 2, 2026

ISSN: 2717-4123



Serial No. 18

Autumn and Winter 2026



Journal of
**COMPUTER AND KNOWLEDGE
ENGINEERING**

Ferdowsi University of Mashhad

ISSN: 2717-4123

General Director: S. A. Hosseini Seno

Editor-in-Chief: M. Kahani

Publisher: Ferdowsi University of Mashhad

Editorial Board:

Mahmoud Naghibzadeh	Professor	Ferdowsi University of Mashhad, Iran
Mohammad H Yaghmaee-Moghaddam	Professor	Ferdowsi University of Mashhad, Iran
Dick H Epema	Professor	Delft Technical University, the Netherlands
Rahmat Budiarto	Professor	University Utara Malaysia, Malaysia
Mohsen Kahani	Professor	Ferdowsi University of Mashhad, Iran
Mohammad R Akbarzadeh-Tootoonchi	Professor	Ferdowsi University of Mashhad, Iran
Madjid Fathi	Professor	University of Siegen, Germany
Hossein Nezamabadi-pour	Professor	Bahonar University of Kerman, Iran
Ahmad Ghafarian	Professor	University of North Georgia, USA
Hamid Reza Pourreza	Professor	Ferdowsi University of Mashhad, Iran
Hadi Sadoghi-Yazdi	Professor	Ferdowsi University of Mashhad, Iran
Seyed Amin Hosseini Seno	Associate Professor	Ferdowsi University of Mashhad, Iran
Abedin Vahedian-Mazloun	Associate Professor	Ferdowsi University of Mashhad, Iran
Ebrahim Bagheri	Associate Professor	Ryerson University, Canada
Hossein Asadi	Associate Professor	Sharif University of Technology, Iran
Mahdi Kargahi	Associate Professor	University of Tehran, Iran
Hamid Reza Ekbia	Associate Professor	Indiana University, USA
Seyed Hassan Mirian Hosseinabadi	Associate Professor	Sharif University of Technology, Iran
Abbas Ghaemi Bafghi	Associate Professor	Ferdowsi University of Mashhad, Iran
Farhad Mahdipour	Associate Professor	Kyushu University, Japan

Administrative Director: T. Hooshmand

Journal of Computer and Knowledge Engineering
Faculty of Engineering, Ferdowsi University of Mashhad
P. O. Box. 91775-1111, Mashhad, I.R. IRAN

Tel: +98 51 38806024, Fax: +98 51 38763301, Email: cke@um.ac.ir, Site: cke.um.ac.ir

CONTENTS

Providing an Ontology for Extracting Functional and non-Functional Software Requirements of an Online Sales System	1
Houbakht Attaran, Esmail Kheirkhah, Hassan Shakeri	
STHChain: A Scalable IoT-Fog-Blockchain Architecture with Lightweight Consensus for Secure Tourist Healthcare	13
Seyed Rahman Soleimani Nadaf, Masood Niazi Torshiz, Sayyed Majid Mazinani Seyed Reza Kamel Tabbakh	
Consumer Behavior-Enhanced CNN-BiLSTM-Attention Model for Residential Electricity Consumption Forecasting	27
Mahdi Yousefi, Sara Ershadinasab	
A Knowledge Graph and Large Language Model Approach for Mapping Associations between Cognitive Neuroscience and Marketing	39
Ali Sarabadani, Kheirolah Rahsepar Fard , Mahtab Tajvidiyan	
GrAMIA-Net: Gray Matter Adaptive Multi View Imbalance Aware Attention Network for Detection of Parkinson's Disease	53
Elham Afarandeh, Rahil Hosseini, Samaneh Yazdani	
A Federated Learning Framework for Predictive Maintenance in IoT-Enabled Smart Environments	67
Alireza Esmaili, Hakan Burak Emekli	

Providing an Ontology for Extracting Functional and non-Functional Software Requirements of an Online Sales System*

Research Article

Houbakht Attaran¹, Esmail Kheirkhah² , Hassan Shakeri³

 [10.22067/cke.2026.96171.1178](https://doi.org/10.22067/cke.2026.96171.1178)

Abstract-- Incomplete and costly requirements extraction, especially for complex systems such as online sales applications within the smart city ecosystem, remains a major challenge in software engineering. Requirements engineering in smart city software is often considered the "missing link" in development methodologies. While ontologies have been proposed to address this, issues with relationship definition and non-functional requirement (NFR) extraction persist. This paper presents a comprehensive ontology for an online sales system as a case study. We detail a structured methodology for ontology development and demonstrate its efficacy. The results, compared with data from four commercial software producers, show that the ontology extracts functional and non-functional requirements with high coverage, leading to significant savings in extraction time, project implementation, costs, and reduction of late-stage changes.

Index Terms-- Development Methodology, Ontology, Requirement Engineering, Smart city, Software Development Lifecycle.

INTRODUCTION

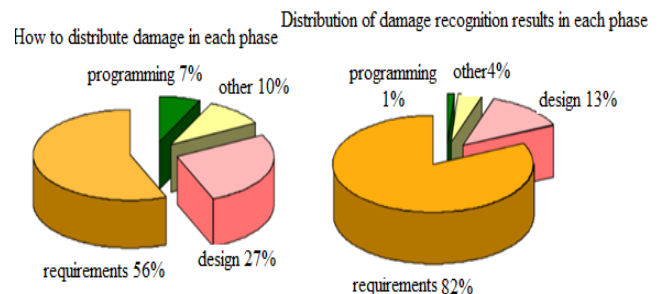
The timely and accurate extraction of both functional and non-functional requirements is a persistent challenge in software development, particularly for large-scale applications like those in a smart city. A smart city, built upon the Internet of Things (IoT), requires a comprehensive model to describe its entities, features, and their interactions. An ontological description can identify software classes, their relationships, properties, and interactions, forming a structural basis for determining comprehensive software requirements.

Software designed for a smart city integrates human, physical, and digital systems. However, existing integration methods often result in isolated "island" applications that lack interoperability, suffer from data redundancy, and face potential conflicts. This fragmentation makes identifying common requirements difficult. The lack of clarity in required features, services,

and classes challenges software architectures and methodologies, often leading to project failures.



Graph 1. The impact of various factors on the failure of IT based projects (CHAOS 2021)



Graph 2. The main sources of errors in projects (CHAOS 2021)

Various factors contribute to the failure of projects based on information technology system. Graph1 shows these factors and the extent of each role and Graph 2 identifies main sources of failures in projects.

As evident from these analyses, a significant percentage of IT project failures stem from weaknesses in requirements identification and management. Extracting requirements is time-consuming, costly, and often incomplete at a project's outset, especially for non-functional requirements (NFRs). The core problems include:

* Manuscript received October 26, 2025, Revised February 07, 2026, Accepted April 04, 2026.

¹ Department of Computer Engineering, Ma. C., Islamic Azad University, Mashhad, Iran.

² Corresponding author. Assistant Prof., Department of Computer Engineering, Ma. C., Islamic Azad University, Mashhad, Iran.

Email: Esmail.kheirkhah@iau.ac.ir

³ Department of Computer Engineering, Ma. C., Islamic Azad University, Mashhad, Iran. Associated Prof.

1. Time-consuming requirements extraction.
2. Incompleteness of extracted requirements.
3. High cost of requirements engineering.
4. The particular challenge of identifying NFRs early.

To address these issues, this paper proposes and evaluates the use of ontology. Ontologies play a fundamental role in the standard description and sharing of domain knowledge [1]. The primary challenges in designing a requirements extraction ontology lie in determining classes, their relationships, properties, and methods. This article contributes a novel ontology for an online sales system, detailing its development methodology and evaluating its performance against industry practices. The results indicate substantial improvements in requirement coverage, especially for NFRs, and significant reductions in time, cost, and project churn.

RESEARCH BACKGROUND

This section reviews literature in three focused areas: ontologies in requirements engineering, ontologies in e-commerce/smart cities, and foundational smart city frameworks. The concept of a "smart city" has evolved over decades, with definitions varying based on community needs and technological context [2]. Our previous works [3] [4] provide an ontology-based platform for determining high-level software classes in a smart city, which serves as the foundation for this study.

Requirements engineering is widely recognized as a critical challenge, often termed the "missing link" in software development methodologies, especially for large projects [5]. Agile methodologies, while popular, also face significant requirements engineering challenges [6].

Ontologies have been applied to structure knowledge in specific domains. In education, ontologies like OntoGamif [7] and instructional design frameworks [8] model domain concepts. In healthcare, ontologies integrate data from heterogeneous devices [9] or model smart healthcare systems (SHCO) [10]. For intelligent transportation, ontology-based architectures describe system components [11]. In urban planning, ontologies aid building energy modeling [13] and social theory analysis [12].

Closer to our focus, ontologies are increasingly used in systems engineering and requirements management. Recent work includes ontologies for AI-powered system requirements (AI-REM) [14], ontology-based explainable AI in manufacturing [15], and ontology embedding techniques [16]. Other studies explore ontology-based search systems [17], knowledge graphs for infrastructure resilience [18], and advanced ontology modeling frameworks [19] [20]. Research also addresses transforming safety requirements into ontological knowledge [21], creating digital threads for systems engineering [22], model-based safety analysis [23], and functional analysis [24]. Large-scale spatial knowledge graphs also rely on robust ontologies [25].

Despite these advances, a review of the literature reveals a gap: there is a lack of a detailed, practical ontology specifically designed to systematically extract both functional and non-functional software requirements

for a concrete system within the smart city domain. This work aims to fill that gap.

METHODOLOGY

This section details the structured, multi-phase methodology employed to develop and evaluate the requirements extraction ontology.

A. Ontology Development Process

The ontology was developed using an iterative, expert-driven approach:

1. **Class Identification:** Core classes (e.g., Administrator, Customer, Online-Shop, Goods) were identified through analysis of standard online sales system features and consultations with three domain experts each having over 15 years of experience in e-commerce software development.
2. **Hierarchy and Relationship Definition:** The class hierarchy and relationships (e.g., is-a, has-a) were defined. The high-level smart city ontology from our prior works [3] [4] provided the foundational superclass (Store) for the Online-Shop class, ensuring alignment with a broader smart city context.
3. **Property and Constraint Specification:** For each class, Object Properties (relationships between instances) and Data Properties (attributes like hasName, hasPrice) were added. Crucially, properties to capture NFRs were explicitly defined, such as hasMaxResponseTime (Data Property), requiresAuthentication (Object Property constraint on an action), or hasAvailabilityRequirement.
4. **Tool and Implementation:** The ontology was formally constructed using Protégé (version 5.5.0), enabling precise definition in OWL. This allowed for logical consistency checking and the use of reasoning.

B. Requirements Extraction Mechanism

Requirements were extracted programmatically:

1. The populated ontology was exported as an RDF graph.
2. Custom SPARQL queries were designed to traverse the graph. Different query patterns targeted different requirement types (see Table 1).
3. Functional requirements were primarily derived from queries about class methods, actions (canPerform), and sequences of operations.
4. Non-functional requirements were derived from queries about Data Properties encoding quality attributes (e.g., hasMaxResponseTime < 2s) and conditional Object Properties (e.g., performAction X only if requiresAuthentication is true).

C. Evaluation and Comparison Method

To validate the ontology's effectiveness, a comparative analysis was conducted:

- **Data Source:** Historical project data was collected from four software companies (labeled C1-C4) specializing in online sales systems. The data included final requirement specifications, project timelines, and cost breakdowns.
- **Comparison Metrics:** The ontology's output was compared against the companies' final, implemented

requirements on several metrics: total functional & non-functional requirements, time/cost for requirements phase, and the percentage/cost of changes due to late-discovered requirements.

- Analysis: Quantitative comparison was supplemented with qualitative analysis to interpret differences in requirement granularity and classification.

IMPLEMENTATION AND RESULTS OF THE PROPOSED ONTOLOGY

Based on the methodology, a detailed ontology for an online sales system was implemented. The overarching smart city ontology from our prior work [4] providing the context is shown in Figure 1.

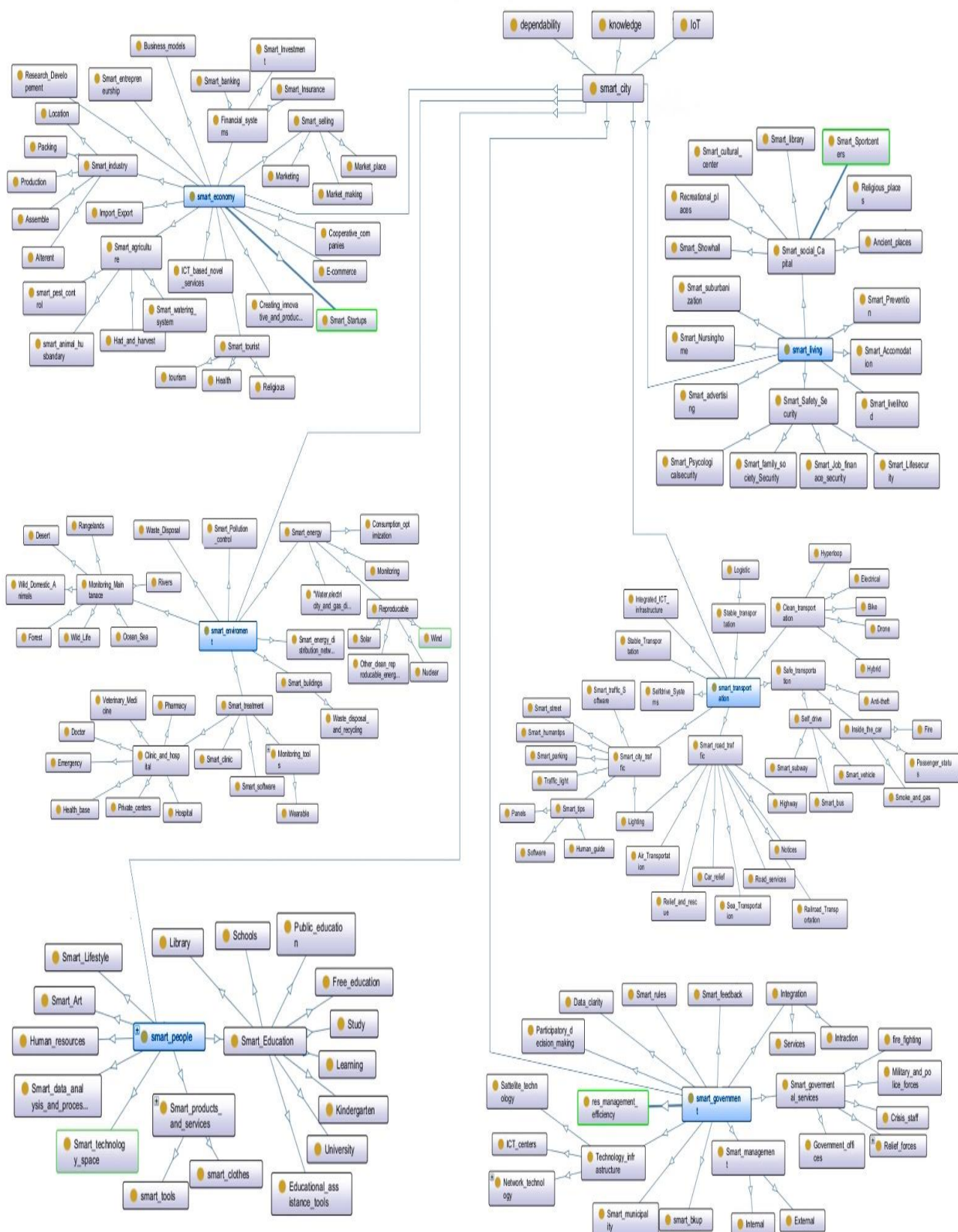


Figure 1. Proposed ontology of smart city [4]

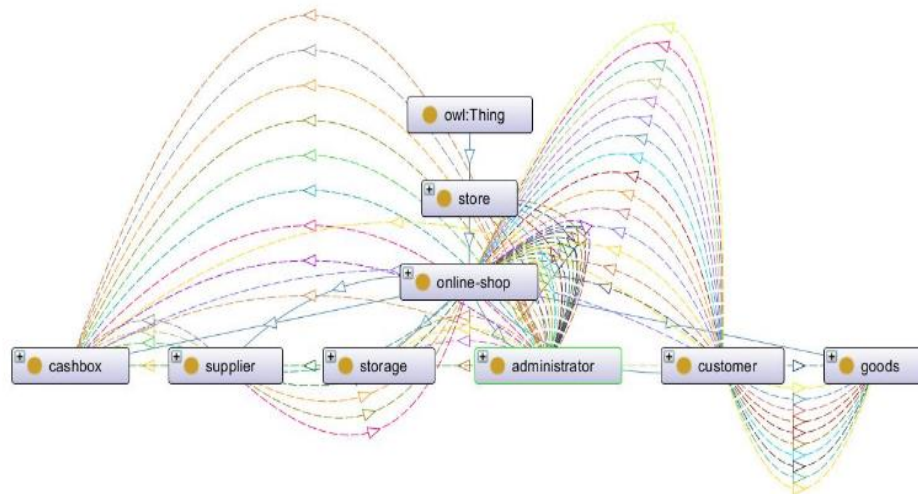


Figure 2. Proposed Ontology for Online marketing

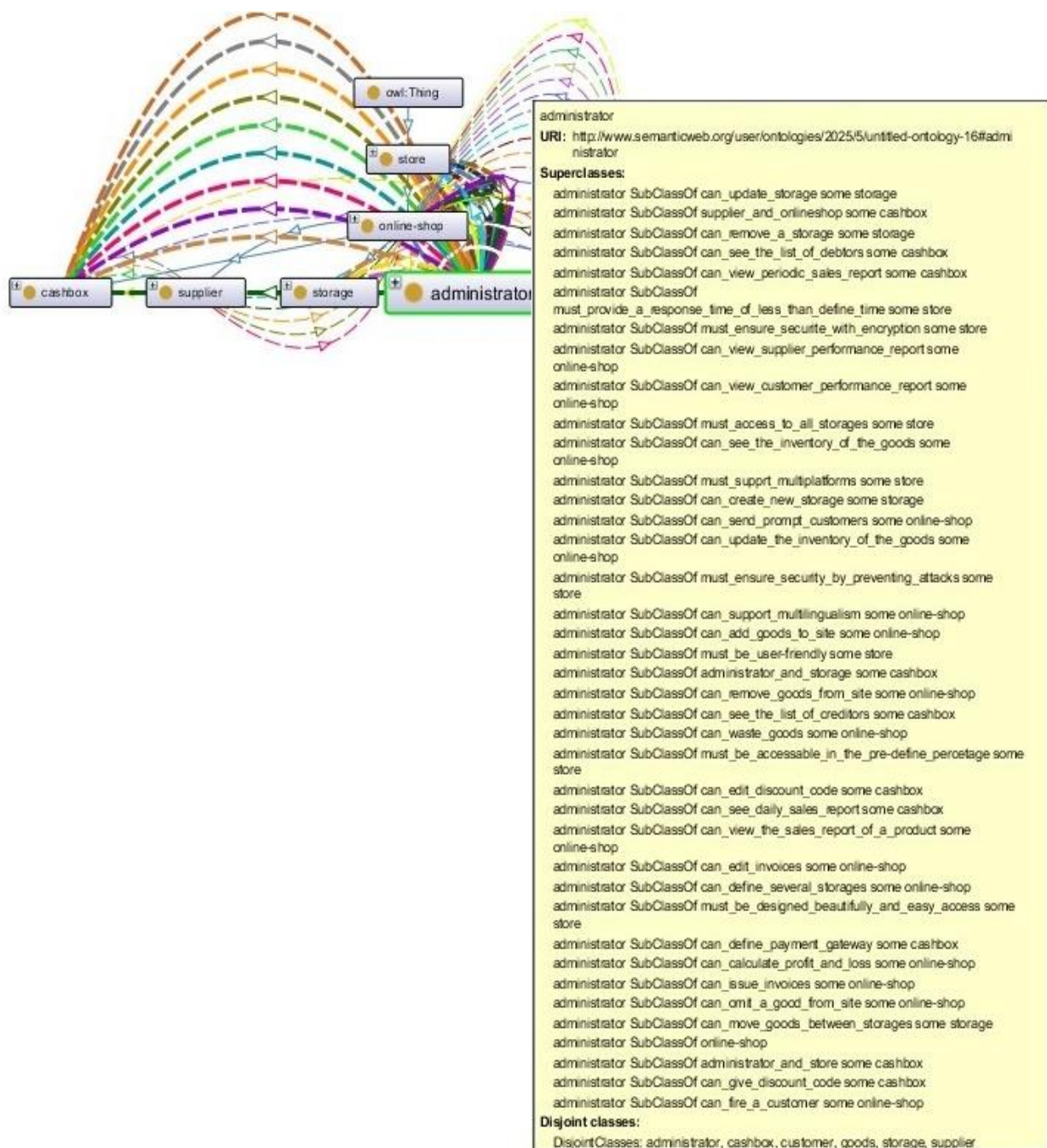


Figure 3. Methods and use cases for Administrator class

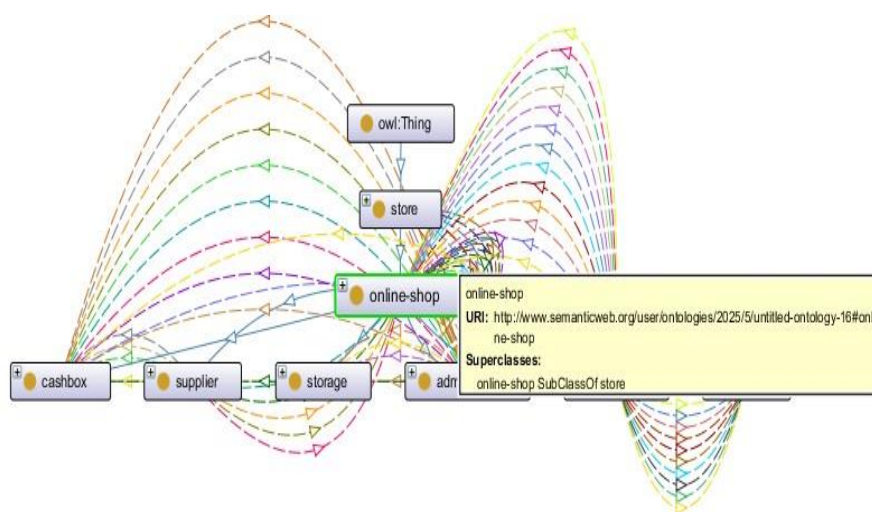


Figure 4. Online-shop class with store superclass

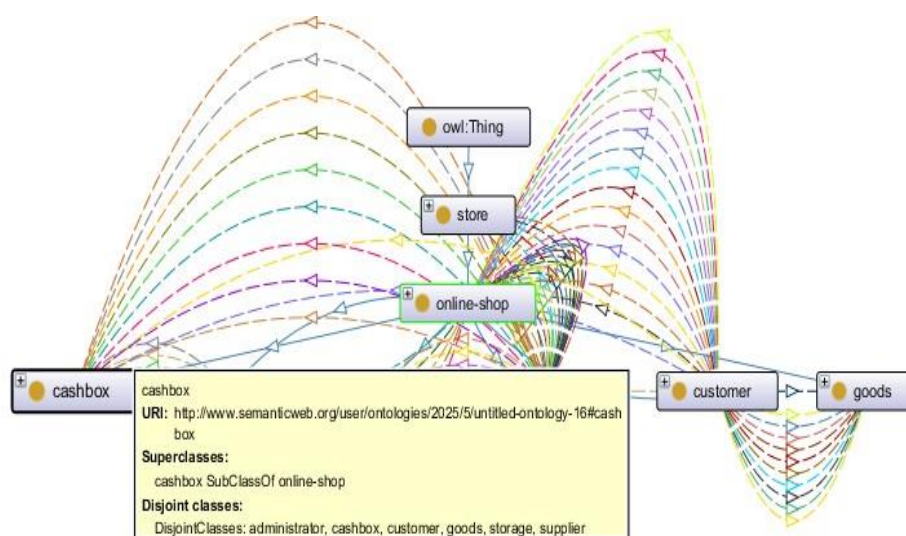


Figure 5. cashbox class with online-shop superclass and disjoint classes

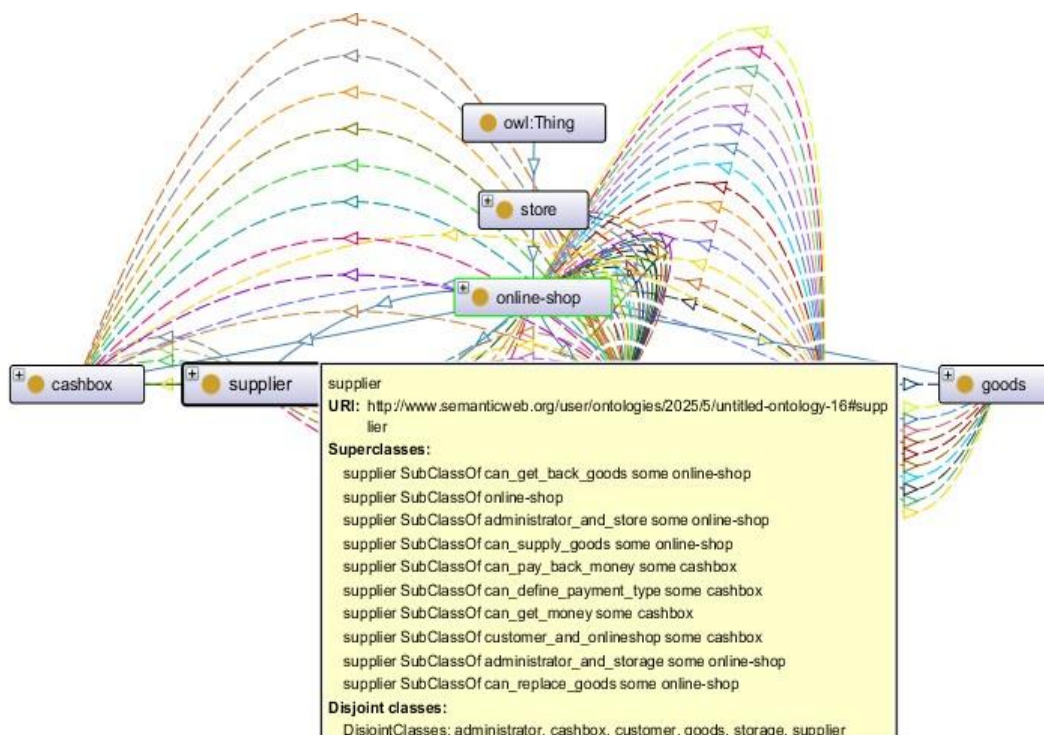


Figure 6. Supplier class with superclass, disjoint classes and methods

The core ontology for the online sales system, showing main classes and relationships, is presented in Figure 2. The following figures (Figure 3-Figure 9) illustrate the detailed structure for key classes, showcasing properties, methods, and inheritance:

The following figures (Figure 3-Figure 9) illustrate the detailed structure for key classes, showcasing properties, methods, and inheritance:

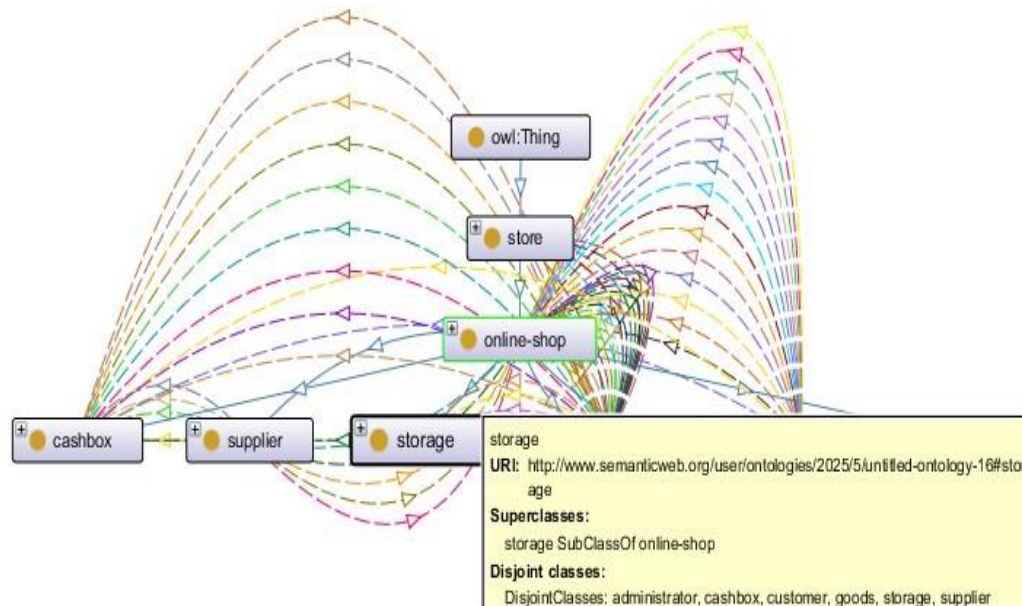


Figure 7. Storage class with superclass and disjoint classes

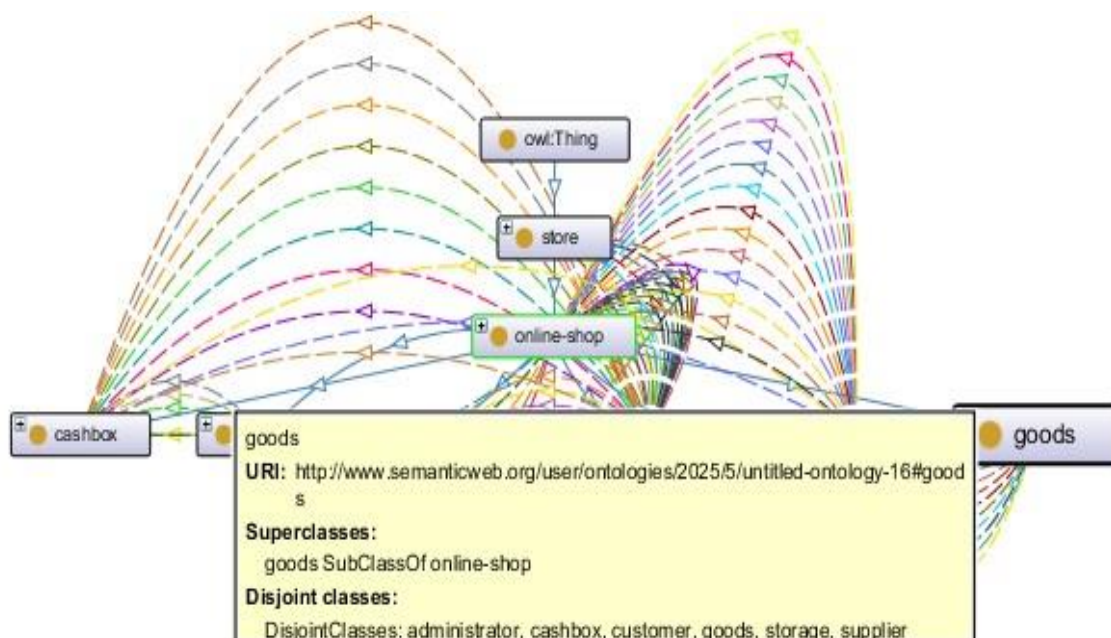


Figure 8. Goods class with superclass and disjoint classes

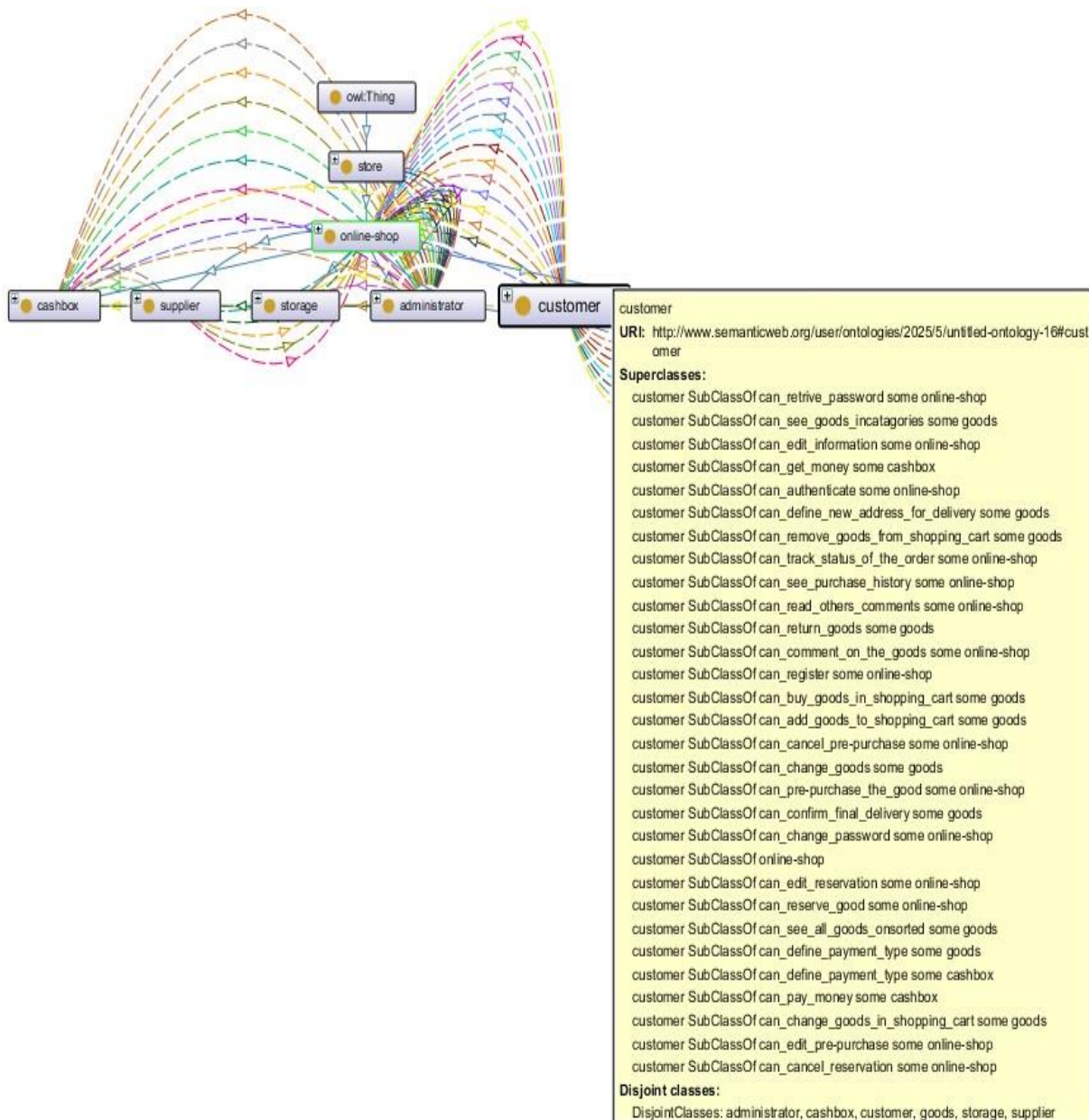


Figure 9. Customer class with superclass, disjoint classes and methods

TABLE 1
Examples of SPARQL Queries for Requirements Extraction

Query Purpose	Sample SPARQL Pattern	Extracted Requirement Type
List all actions of a class	SELECT ?action WHERE { :Admin :canPerform ?action }	Functional (Admin capabilities)
Find entities with performance constraints	SELECT ?entity WHERE { ?entity :hasMaxResponseTime ?time . FILTER (?time < 2000) }	Non-Functional (Performance)
Find preconditions for an action	SELECT ?precond WHERE { :ProcessOrder :requiresAuthentication ?precond }	Non-Functional (Security)

Explanation for: admin Type store

- 1) admin can_issue_invoices online-shopping
- 2) can_issue_invoices **Domain** administrator
- 3) administrator **SubClassOf** online-shop
- 4) online-shop **SubClassOf** store

Explanation 17 ☐ Display laconic explanation

Explanation for: admin Type store

- 1) admin a-tel "admin-tel"^^xsd:string
- 2) a-tel **Domain** administrator
- 3) administrator **SubClassOf** online-shop
- 4) online-shop **SubClassOf** store

Explanation 18 ☐ Display laconic explanation

Explanation for: admin Type store

- 1) admin can_create_new_storage onlineshop-storage
- 2) can_create_new_storage **Domain** administrator
- 3) administrator **SubClassOf** online-shop
- 4) online-shop **SubClassOf** store

Explanation 19 ☐ Display laconic explanation

Explanation for: admin Type store

- 1) admin can_define_payment_gateway store-cashbox
- 2) can_define_payment_gateway **Domain** administrator
- 3) administrator **SubClassOf** online-shop
- 4) online-shop **SubClassOf** store

Figure 10. Part of Admin functional requirements

Explanation for: online-customer Type store

- 1) online-customer can_see_all_goods_onsorted product
- 2) customer **SubClassOf** online-shop
- 3) can_see_all_goods_onsorted **Domain** customer
- 4) online-shop **SubClassOf** store

Explanation 3 ☐ Display laconic explanation

Explanation for: online-customer Type store

- 1) online-customer can_retrieve_password online-shopping
- 2) can_retrieve_password **Domain** customer
- 3) customer **SubClassOf** online-shop
- 4) online-shop **SubClassOf** store

Explanation 4 ☐ Display laconic explanation

Explanation for: online-customer Type store

- 1) online-customer can_read_others_comments online-shopping
- 2) can_read_others_comments **Domain** customer
- 3) customer **SubClassOf** online-shop
- 4) online-shop **SubClassOf** store

Explanation 5 ☐ Display laconic explanation

Explanation for: online-customer Type store

- 1) online-customer can_remove_goods_from_shopping_cart product
- 2) can_remove_goods_from_shopping_cart **Domain** customer
- 3) customer **SubClassOf** online-shop

Figure 11. Part of customer functional requirements

Explanation 1 ☐ Display laconic explanation

Explanation for: product-supplier Type store

- 1) product-supplier Type supplier
- 2) supplier **SubClassOf** online-shop
- 3) online-shop **SubClassOf** store

Explanation 2 ☐ Display laconic explanation

Explanation for: product-supplier Type store

- 1) product-supplier can_define_how_to_get_money store-cashbox
- 2) supplier **SubClassOf** online-shop
- 3) can_define_how_to_get_money **Domain** supplier
- 4) online-shop **SubClassOf** store

Explanation 3 ☐ Display laconic explanation

Explanation for: product-supplier Type store

- 1) product-supplier can_get_back_goods online-shopping
- 2) supplier **SubClassOf** online-shop
- 3) can_get_back_goods **Domain** supplier
- 4) online-shop **SubClassOf** store

Explanation 4 ☐ Display laconic explanation

Explanation for: product-supplier Type store

- 1) product-supplier s-adds "supplier address"^^xsd:string
- 2) supplier **SubClassOf** online-shop
- 3) s-adds **Domain** supplier

Figure 12. Part of supplier functional requirements

Explanation for: NFRs Type store

- 1) must_provide_a_response_time_of_less_than_define_time **Range** store
- 2) admin must_provide_a_response_time_of_less_than_define_time NFRs

Explanation 8 ☐ Display laconic explanation

Explanation for: NFRs Type store

- 1) must_ensure_secureite_with_encryption **Range** store
- 2) admin must_ensure_secureite_with_encryption NFRs

Explanation 9 ☐ Display laconic explanation

Explanation for: NFRs Type store

- 1) must_support_multiplatforms **Range** store
- 2) admin must_support_multiplatforms NFRs

Explanation 10 ☐ Display laconic explanation

Explanation for: NFRs Type store

- 1) admin must_be_user-friendly NFRs
- 2) must_be_user-friendly **Range** store

Explanation 11 ☐ Display laconic explanation

Explanation for: NFRs Type store

- 1) must_be_accessible_in_the_pre-define_percentage **Range** store
- 2) admin must_be_accessible_in_the_pre-define_percentage NFRs

Figure 13. Part of non-functional requirements.

Using the ontology and SPARQL queries, requirements were extracted. Samples of the output are shown in Figures 10 -13:

A SPARQL query (COUNT on owl:ObjectProperty) shown in following code, determined the ontology could yield 79 distinct requirement elements (Figure 14).
 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
 PREFIX owl: <http://www.w3.org/2002/07/owl#>
 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
 SELECT (COUNT (DISTINCT ?p) AS ?totalProperties)
 WHERE {
 ?pa owl:ObjectProperty
 }

totalProperties
"79"<http://www.w3.org/2001/XMLSchema#integer>

Figure 14. Number of extracted requirements

Sample instances extracted via queries like the following query, are shown in Figures 15 & 16. Also, the following code snippet was used in SPARQL to extract the requirements:

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
SELECT ?individual ?type
WHERE {
  ?individual a ?type .
}
```

individual
customer_and_onlineshop
can_pre-purchase_the_good
can_get_money
can_edit_information
can_define_how_to_get_money
can_omit_a_good_from_site
can_see_goods_incatagories
customer_and_onlineshop
can_reserve_good
can_change_goods_in_shopping_cart
can_see_purchase_history
can_view_supplier_performance_report
can_pay_back_money
can_remove_a_storage
can_see_purchase_history
must_ensure_securite_with_encryption
can_change_goods_in_shopping_cart
can_see_the_list_of_debtors
can_change_goods

Figure 15. A sample of extracted requirements by SPARQL

individual
can_edit_information
can_see_goods_incatagories
can_calculate_profit_and_loss
can_change_goods_in_shopping_cart
can_get_back_goods
can_see_the_list_of_creditors
can_create_new_storage
can_send_prompt_customers
can_change_goods
can_support_multilingualism
can_reserve_good
must_be_user-friendly
can_supply_goods
can_edit_pre-purchase
customer_and_onlineshop
must_provide_a_response_time_of_less_than_define_time
can_see_daily_sales_report
must_supprt_multiplatforms
can_comment_on_the_goods
can view periodic sales report

Figure 16. Another sample of extracted requirements by SPARQL

COMPARISON AND PROOF OF EFFECTIVENESS

The ontology's performance was compared against the final project outcomes of four companies (Company1-Company4). The key comparison is between the total requirements identified by the ontology and the total requirements ultimately implemented by the companies. The evaluation metrics are defined in Table 2.

- Number of functional requirements extracted during project execution (Np)
- Total number of non-functional requirements (Nnon)
- Total number of final requirements (Ntotal)
- Total time spent on requirements extraction (Te)
- Cost spent on requirements extraction (Ce)
- Percentage of changes in design and implementation with discovery of new requirements (CHp)
- Time required to implement changes (Tch)
- Cost spent on implementing changes (Cch)

TABLE 2
Comparative parameter values

	Np	Nnon	Ntotal	Te	Ce	CHp	Tch	Cch
Company 1	22	3	74	103	2.5	20%	40	7
Company 2	21	5	81	98	2.2	8%	25	5
Company 3	32	4	78	84	1.8	10%	52	11
Company 4	28	4	76	100	2	10%	45	9
Ontology	0	5	79	1	0.1	0	0	0

The following charts visualize the comparison:

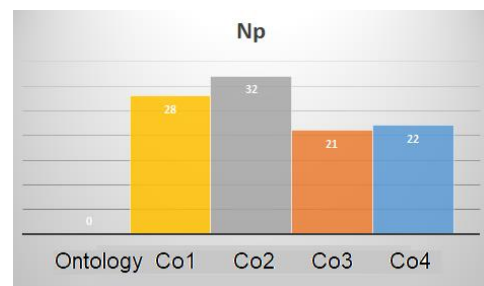


Chart 1. Number of functional requirements extracted during project execution

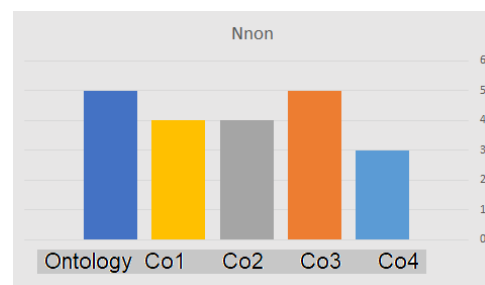


Chart 2. Total number of non-functional requirements

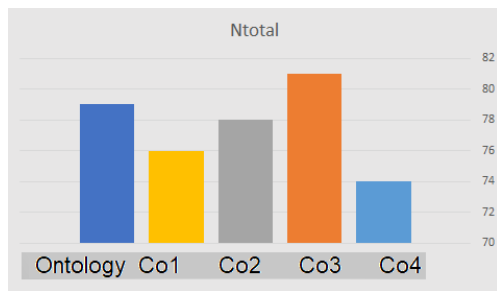


Chart 3. Comparison of Total final requirements

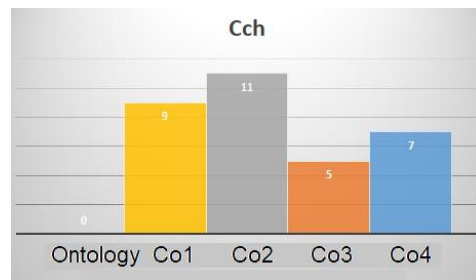


Chart 8. Cost spent on making changes

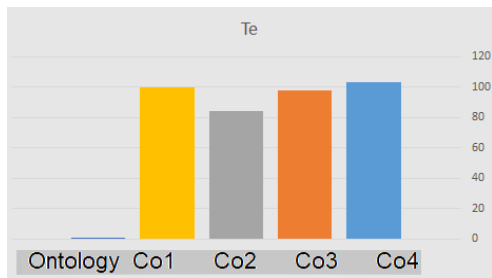


Chart 4. Total time spent extracting requirements

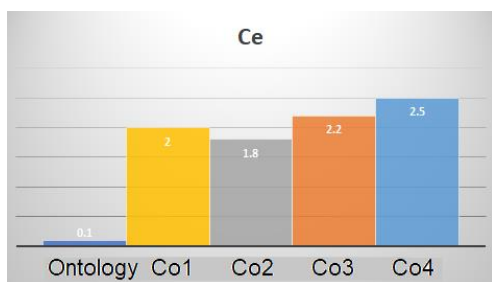


Chart 5. Cost spent on extraction requirements

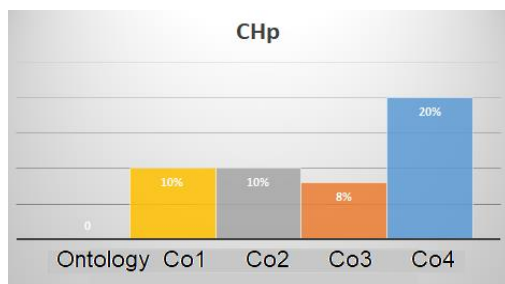


Chart 6. Percentage of changes in design and implementation due to discovery of new requirements

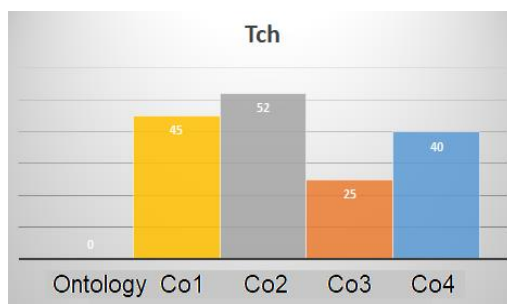


Chart 7. Time required to make changes

DISCUSSION

This section interprets the results, addresses the research questions, and examines the study's limitations.

A Interpretation of Results

The comparative analysis shows that the ontology-based approach identified a comprehensive set of requirements early in the lifecycle. Key findings include:

- **High Coverage:** The ontology's total requirement count (79) closely aligns with or exceeds the final requirement count of the companies (range: 72-82), indicating high coverage. The slightly higher count in some cases is attributed to the ontology's more granular decomposition of some composite requirements.
- **Early NFR Identification:** As shown in Chart 3, the ontology successfully identified a substantial set of NFRs upfront, whereas companies discovered many of these later during development (contributing to CHp).
- **Efficiency Gains:** Charts 5, 6, 8, and 9 demonstrate dramatic reductions in time and cost for both the initial requirements phase and for implementing changes, directly addressing the core problems outlined in the introduction.

B. Answering the Research Questions

The study set out to address the challenges of requirements engineering for smart city software. The results confirm that:

- A well-structured ontology can systematically extract both functional and non-functional requirements.
- This method leads to more complete early requirement specifications, reducing late discoveries and their associated costs.
- The approach is practically effective, as evidenced by favorable comparisons with industrial practices.

C. Threats to Validity

- **Construct Validity:** The use of "number of requirements" as a primary metric may not fully capture requirement quality or priority. Future work should include qualitative assessments by practitioners.
- **Internal Validity:** The study relies on historical data from a limited number of companies (four). While efforts were made to verify data accuracy, potential recall bias exists. The ontology's development

involved expert judgment, which introduces subjectivity, though we mitigated this through multiple experts and iteration.

- External Validity: The current evaluation is focused on the online sales domain within the smart city context. Generalizability to other smart city subsystems (e.g., smart grid, traffic management) requires further validation through additional case studies.

CONCLUSION AND FUTURE WORK

This paper presented a novel ontology-based framework for extracting software requirements, demonstrated through an online sales system case study. By detailing a rigorous methodology and providing a fair comparison with industrial data, we showed that the approach achieves high requirement coverage, particularly for non-functional requirements, while generating significant savings in time, cost, and project instability.

Future work will focus on: (1) extending the ontology to other smart city domains to test generalizability, (2) developing a semi-automated tool that integrates this ontological approach into standard requirements management platforms, and (3) conducting longitudinal studies to measure the impact on project success rates.

ACKNOWLEDGMENT

The authors of this article sincerely thank all the software companies that provided their experiences, statistics, and performance to compare and prove the effectiveness of the new model.

REFERENCES

- [1] Z. Chen, H. Qian, Y. Gao, H. Lyu, and Q. Wang. (2022, Jun.). Ontology construction of city hotline service for urban grassroots governance. Presented at 2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS). [Online]. Available: <https://doi.org/10.1109/IEMTRONICS55184.2022.9795703>
- [2] S. Benkhaled, M. Hemam, M. Djezzar, and M. Maimour. (2022). An ontology-based contextual approach for cross-domain applications in internet of things. *Informatica*. [Online]. 46(5), pp. 39–48. Available: <https://dx.doi.org/10.31449/inf.v46i5.3627>
- [3] H. Attaran, N. Kheibari, and D. Bahrepour. (2022). Toward integrated smart city: A new model for implementation and design challenges. *GeoJournal*. [Online]. 87(4), pp. 511–526. Available: <https://doi.org/10.1007/s10708-021-10560-w>
- [4] H. Attaran, E. Kheirkhah, and H. Shakeri. (2024). Providing an ontology-based framework to determine the high-level software classes of a smart city. *Creative City Design (CCD)*. [Online]. 7(3). Available: <https://doi.org/10.71619/crcd-2024-8176>
- [5] H. Attaran and E. Kheirkhah. (2023). Review and comparison of software development methodologies. Presented at 15th International Conference on Science and Technology Advances. [Online]. Available: <https://civilica.com/doc/1608987>
- [6] H. Attaran, E. Kheirkhah, and H. Shakeri. (2024). The challenges of requirements engineering in the development of smart city software using agile methodologies: A systematic literature review. *Journal of Information Technologies in Engineering Design*. [Online]. 17(1). Available: https://doi.org/10.46338/ijetae0124_03
- [7] R. Bouzidi, A. De Nicola, F. Nader, and R. Chalal. (2019). OntoGamif: A modular ontology for integrated gamification. *Applied Ontology*. [Online]. 14(3), pp. 215–249. Available: <https://doi.org/10.3233/AO-190212>
- [8] S. Chimalakonda and K. Nori. (2020). An ontology based modeling framework for design of educational technologies. *Smart Learning Environments*. [Online]. 7(28), pp. 1–24. Available: <https://doi.org/10.1186/s40561-020-00135-6>
- [9] C. Peng and P. Goswami. (2019). Meaningful integration of data from heterogeneous health services and home environment based on ontology. *Sensors*. [Online]. 19(8), p. 1747. Available: <https://doi.org/10.3390/s19081747>
- [10] S. Tiwari and A. Abraham. (2020). Semantic assessment of smart healthcare ontology. *International Journal of Web Information Systems*. [Online]. 16(4), pp. 475–491. Available: <https://doi.org/10.1108/IJWIS-05-2020-0027>
- [11] S. Fernandez et al. (2016). Ontology-based architecture for intelligent transportation systems using a traffic sensor network. *Sensors*. [Online]. 16(8), p. 1287. Available: <https://doi.org/10.3390/s16081287>
- [12] B. Epstein, *Social Ontology*. Cambridge University Press, 2025.
- [13] R. Ma, Q. Li, B. Zhang, H. Huang, and C. Yang. (2024). An ontology-driven method for urban building energy modeling. *Sustainable Cities and Society*. [Online]. 106, p. 105394. Available: <https://doi.org/10.1016/j.scs.2024.105394>
- [14] E. Sadovski, I. Aviv, and I. Hadar. (2024). Towards a comprehensive ontology for requirements engineering for AI-powered systems. Presented at International Working Conference on Requirements Engineering: Foundation for Software Quality. [Online]. Available: https://doi.org/10.1007/978-3-031-57327-9_14
- [15] M. R. Naqvi, L. Elmhadhibi, A. Sarkar, B. Archimede, and M. H. Karray. (2024). Survey on ontology-based explainable AI in manufacturing. *Journal of Intelligent Manufacturing*. [Online]. 35(8), pp. 3605–3627. Available: <https://doi.org/10.1007/s10845-023-02304-z>
- [16] J. Chen, O. Mashkova, F. Zhapa-Camacho, R. Hoehndorf, Y. He, and I. Horrocks. (2025). Ontology embedding: A survey of methods, applications and resources. *IEEE Transactions on Knowledge and Data Engineering*. [Online]. Available: <https://doi.org/10.1109/TKDE.2025.3559023>
- [17] S. Huettemann, R. M. Mueller, and B. Dinter. (2025). Designing ontology-based search systems for research articles. *International Journal of Information Management*. [Online]. 83, p. 102901. Available: <https://doi.org/10.1016/j.ijinfomgt.2025.102901>
- [18] B.-L. Gan, D.-M. Zhang, Z.-K. Huang, F.-Y. Zheng, R. Zhu, and W. Zhang. (2025). Ontology-driven

- knowledge graph for decision-making in resilience enhancement of underground structures: Framework and application. *Tunneling and Underground Space Technology*. [Online]. 163, p. 106739. Available: <https://doi.org/10.1016/j.tust.2025.106739>
- [19] S. Kim, Y. Y. Liao, and K. Ryu. (2025). Ontology modeling using fractal and fuzzy concepts to optimize metadata management. *Applied Sciences*. [Online]. 15(13), p. 7193. Available: <https://doi.org/10.3390/app15137193>
- [20] M. Bettaz and M. Maouche. (2025). Towards an ontological-based CIM modeling framework for IoT applications. *Informatica*. [Online]. 48(4).
- [21] Z. Wu, M. Liu, and G. Ma. (2025). A semi-automatic ontology development framework for knowledge transformation of construction safety requirements. *Buildings*. [Online]. 15(4), p. 569. Available: <https://doi.org/10.3390/buildings15040569>
- [22] R. H. Gautier, N. L. D. S. Duarte, A. Baker, and D. Mavris. (2025). Formulation of an ontology-based digital thread for analysis and decision support activities. Presented at AIAA SCITECH 2025 Forum. [Online]. Available: <https://doi.org/10.2514/6.2025-1092>
- [23] K. Mokos, P. Katsaros, and P. Bohn. (2025). Model-based safety analysis of requirement specifications. *Journal of Systems and Software*. [Online]. 219, p. 112231. Available: <https://doi.org/10.2139/ssrn.4748212>
- [24] H. E. Johnson, M. SureshKumar, H. Kannan, and L. D. Thomas. (2025). Ontology-based functional analysis of an aerospace system. Presented at AIAA SCITECH 2025 Forum. [Online]. Available: <https://doi.org/10.2514/6.2025-1212>
- [25] C. Shimizu et al. (2024). The KnowWhereGraph ontology. *Journal of Web Semantics*. [Online]. 100842.



Ferdowsi
University of
Mashhad

Journal of Computer and Knowledge Engineering

<https://cke.um.ac.ir>



Information and
Communication
Technology Association of
Iran

STHChain: A Scalable IoT-Fog-Blockchain Architecture with Lightweight Consensus for Secure Tourist Healthcare *

Research Article

Seyed Rahman Soleimani Nadaf¹, Masood Niazi Torshiz² , Sayyed Majid Mazinani³, Seyed Reza Kamel Tabbakh⁴

 [10.22067/cke.2026.96829.1181](https://doi.org/10.22067/cke.2026.96829.1181)

Abstract-- Tourists' healthcare maintenance, especially during the outbreak of infectious diseases, is a major challenge in the tourism industry. Ensuring accurate tourists' healthcare data requires data collected using Internet of Things (IoT) devices within a smart city. Additionally, data integrity is preserved through storage on a blockchain platform to maintain reliability. Some works have addressed blockchain-based healthcare solutions for smart tourists, nevertheless many aspects require further investigation. Therefore, we propose a lightweight and scalable blockchain-based smart tourists healthcare model, called STHChain to present welfare services. The model first configures a hierarchical network of IoT devices, edge, and fog computing layers to collect, preprocess, and store tourists' healthcare data. Subsequently, a smart contract defines formal Role-Based Access Control mechanisms for the model's components, including sensor devices, edge nodes, fog nodes, hotels/hostels, welfare centers, and tourists. This contract also executes a deterministic function to classify tourists' health status based on sensor data, which governs their access to welfare service. Finally, healthcare data are included in blocks and verified using a proposed Raft-based Clustering consensus protocol called RaC to add the chain of blocks. The energy consumption and throughput of the model is evaluated using iFogSim simulator, demonstrating that it is energy-efficient, scalable and suitable for real-time tourist healthcare monitoring.

Index Terms-- Blockchain, Healthcare, Internet of Things, IoT, Smart City, Smart Tourist, Tourism, Welfare Services

INTRODUCTION

The outbreak of infectious diseases in recent decades has caused significant socioeconomic damage to the tourism industry, challenging its management [1]. To address this challenge, the concept of the smart tourist based on the smart city has been introduced [2]. Tourists' healthcare maintenance is a critical concern in the tourism industry, where services provided to tourists must be

tailored to their healthcare status. To present appropriate services, environmental conditions must also be considered [3]. A comprehensive investigation of the tourism industry and environmental conditions plays a pivotal impact in enhancing tourists' healthcare. In recent years, smart cities have evolved to improve service quality. A smart city establishes a new urban paradigm offering intelligent living-environment services, intelligent decision-making, and optimal resource management, such that these services address challenges related to treatment costs [4], resource availability [5], and remote monitoring [6]-[8]. In a smart city, the latest digital technologies are employed to enhance healthcare outcomes [9]. Since the onset of the Covid-19 pandemic in 2019, smart-city technologies have been utilized to provide precise, error-free data [10] for treating pandemics [11] and preventing their spread [12]. Accurate data enable timely disease detection, thereby reducing treatment costs [4]. Conversely, smart services face security challenges for securely storing data, including authentication, unauthorized access, energy consumption, and data integrity [13], [14]. In recent years, blockchain technology has been employed to preserve healthcare data integrity [15], [16]. Therefore, we present a lightweight and scalable blockchain-based smart tourists healthcare model, called STHChain, including the following components:

- Hotels and hostels are equipped with Internet of Things (IoT) devices in smart neighborhoods, and tourists' healthcare data are collected by these devices.
- A hierarchical network is established comprising sensor, edge, and fog nodes located at health centers and hotels/hostels to store tourists' healthcare data.
- A smart contract uses tourists' healthcare data to determine their healthcare status for access to welfare

* Manuscript received December 03, 2025, Revised December 30, 2025, Accepted April 04, 2026.

¹ PhD Candidate, Department of Computer Engineering, Ma.C., Islamic Azad University, Mashhad, Iran.

² Corresponding author. Associate Professor, Department of Computer Engineering, Ma.C., Islamic Azad University, Mashhad, Iran.
Email: masood.niazi@iau.ac.ir

³ Associate Professor, Department of Computer Engineering, Imam Reza International University, Mashhad, Iran.

⁴ Associate Professor, Department of Computer Engineering, Ma.C., Islamic Azad University, Mashhad, Iran.

services in the smart city. Access control for tourist services is then issued by the smart contract.

- Healthcare data of tourists are stored in blocks. The proposed consensus protocol uses clustering of fog nodes located at health centers and employs Raft consensus mechanism to verify blocks and append them to the chain of blocks.

The primary contributions of this work are outlined below:

- We propose an integrated IoT-Edge-Fog and Blockchain architecture tailored for smart tourism, reducing latency and energy overhead.
- A smart contract implements a formal RBAC model and a deterministic health-status function for automated, auditable access control.
- A two-tier pseudonymization scheme decouples tourists' real identities from on-chain health records, ensuring privacy-by-design.
- The novel Raft-based Clustering (RaC) consensus protocol improves scalability and energy efficiency through localized fog-node verification clusters.

The remainder of the paper is comprised: Section 2 surveys IoT-based healthcare research and the challenges facing the tourism industry. Section 3 proposes the blockchain-based healthcare model for smart tourists. Section 4 provides analysis and evaluation of the proposed model. Section 5 concludes with results and outlines future work.

LITERATURE REVIEW

In this section, the previous work related to the challenges facing the tourism industry and IoT-based healthcare research is reviewed.

A. Tourism Industry Challenges

The tourism industry faces challenges related to the economy, international politics, cultural differences among countries, and tourists' healthcare, as illustrated in Figure 1.

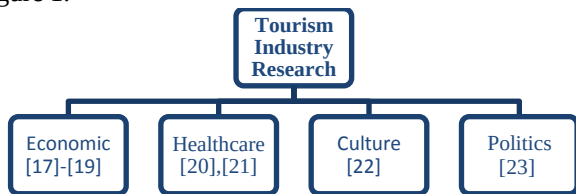


Figure 1. Previous research in the area of tourism industry challenges.

Figure 1 shows several works in relation to challenges facing the tourism industry. Some researchers propose a green economy system aimed at addressing high energy consumption, water consumption, and habitat destruction in tourist destinations [17]. The study also examines the promotion of tourism culture within the social policies of the studied community [17]. Other researchers employ a digital economy framework, based on digital goods on the Internet and web-based applications, to investigate the economic challenges in the tourism industry [18]. The

individual and collective dimensions of tourism that influence development and economic fluctuations are discussed in another study [19]. In recent years, the domain of tourists' healthcare has emerged to provide medical services to patients [20]. A substantial portion of people from developing countries travels to developed countries to continue their treatment. Accordingly, several efforts have been undertaken, for example, in Nigeria, to preserve the tourists' healthcare [20]. In India, healthcare facilities for tourists such as highway hospitals, mobile pharmacies, and applications for recording medical services have expanded, contributing to the mitigation of tourists' healthcare challenges in India [21]. Consequently, tourists' healthcare contributes to the prosperity of not only the tourism industry but also the broader regional economy. A recent systematic review of smart tourism technologies [34] highlights the growing role of IoT and data-driven services in enhancing tourist satisfaction and operational efficiency, underscoring the need for integrated technical frameworks that can deliver real-time, health-aware services. In another prominent research [22], religious beliefs in certain countries (e.g., Iran) that attract tourists were investigated. The researchers have shown that cultural and religious aspects of Iran, and the rich cultural heritage of other countries, influence the tourism industry of those societies. In another work [23], researchers studied current tourism policies and challenges on the island of Cyprus and concluded that policy changes in tourism are essential.

B. IoT-based Healthcare Systems

As discussed in the previous section, maintaining tourists' healthcare is a crucial factor in boosting the prosperity of the tourism industry. Accordingly, this subsection reviews the approaches proposed in IoT-based healthcare, as depicted in Figure 2.

Figure 2 illustrates some studies conducted in the area of IoT-based healthcare focused on secure storage and sharing of data using edge and fog computing [24]-[26]. In these studies, providing storage space through edge and fog computing not only reduces network traffic but also addresses many challenges inherent in cloud computing. Secure storage and sharing of healthcare data are of critical importance. If these data are compromised, the reliability of the healthcare system is undermined, with serious consequences for patients' healthcare.

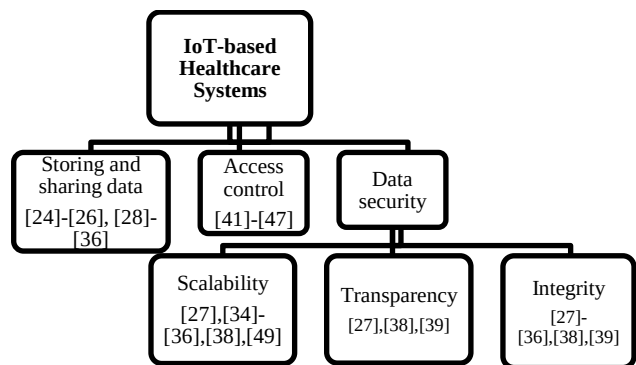


Figure 2. Prominent research in the area of IoT-based healthcare.

Client-server and cloud-based healthcare data management systems face issues such as single point of failure, data privacy, centralized data monitoring, and system vulnerabilities [27]. Consequently, blockchain is proposed as a secure platform for storing and sharing data. In some research, blockchain technology has been introduced as a solution for secure and trustworthy exchange of IoT data [28]-[36]. Several researchers have proposed various architectures for secure data storage [32], [33]. In other work, secure data transmission and data management among nodes are guaranteed by blockchain [28]. Other researchers have proposed a patient-centric healthcare data management system [29]. Moreover, some researchers have presented blockchain as a distributed system for data storage with privacy preservation [29], employing pseudonymization through cryptographic functions to maintain patient data confidentiality on the blockchain [29]. Researchers have enhanced the performance of blockchain-based healthcare systems by leveraging wearable sensors and IoT devices [30], [31]. Healthcare data management systems face challenges including data transparency, traceability, immutability, data origin, accessibility, reliability, and privacy. Consequently, a set of transparent concepts and policies has been proposed to guide the accountability of blockchain-based systems [37]. Other researchers have improved integrity, scalability, and reliability through blockchain technology [27], [38], [39], [49]-[50]. A lightweight, scalable, secure, auditable, and privacy-preserving

blockchain-based architecture for healthcare data management has been introduced [27], [31]. The fusion of blockchain and IoT has been recognized as a powerful approach for new business models and distributed applications [40]. Blockchain offers solutions to access-

control, privacy, and data-security challenges, while scalability remains a key issue. Studies [41]-[47] employ smart contracts to govern access to healthcare data.

Accordingly, in light of previous work, we advocate a hybrid approach that combines smart city concepts, blockchain, and smart contracts to introduce a model for preserving the smart tourists' healthcare.

C. Critical Analysis and Identified Research Gaps

The reviewed literature reveals significant yet often isolated advancements. Research on tourism challenges effectively outlines socioeconomic and health related problems but typically proposes policy or managerial solutions without providing a scalable technical architecture for real-time health monitoring. Concurrently, IoT-based healthcare systems offer sophisticated devices and data pipelines, and blockchain integration addresses security and trust deficits in these systems. However, a critical synthesis exposes three persistent gaps: (1) a lack of integrated frameworks that simultaneously address the specific operational needs of the tourism industry and the technical requirements for secure healthcare IoT, (2) prevalent scalability and energy efficiency bottlenecks in blockchain designs when applied to dense, city-scale sensor networks, and (3) insufficient privacy-preserving identity management tailored for transient populations like tourists, where conventional registration is impractical. The proposed STHChain model is architected explicitly to bridge these gaps. Table 1 synthesizes the identified limitations in prior work and maps them to the specific architectural and protocol-level solutions introduced by STHChain. This structured comparison underscores the model's targeted contributions to scalable, secure, and privacy-aware tourist healthcare.

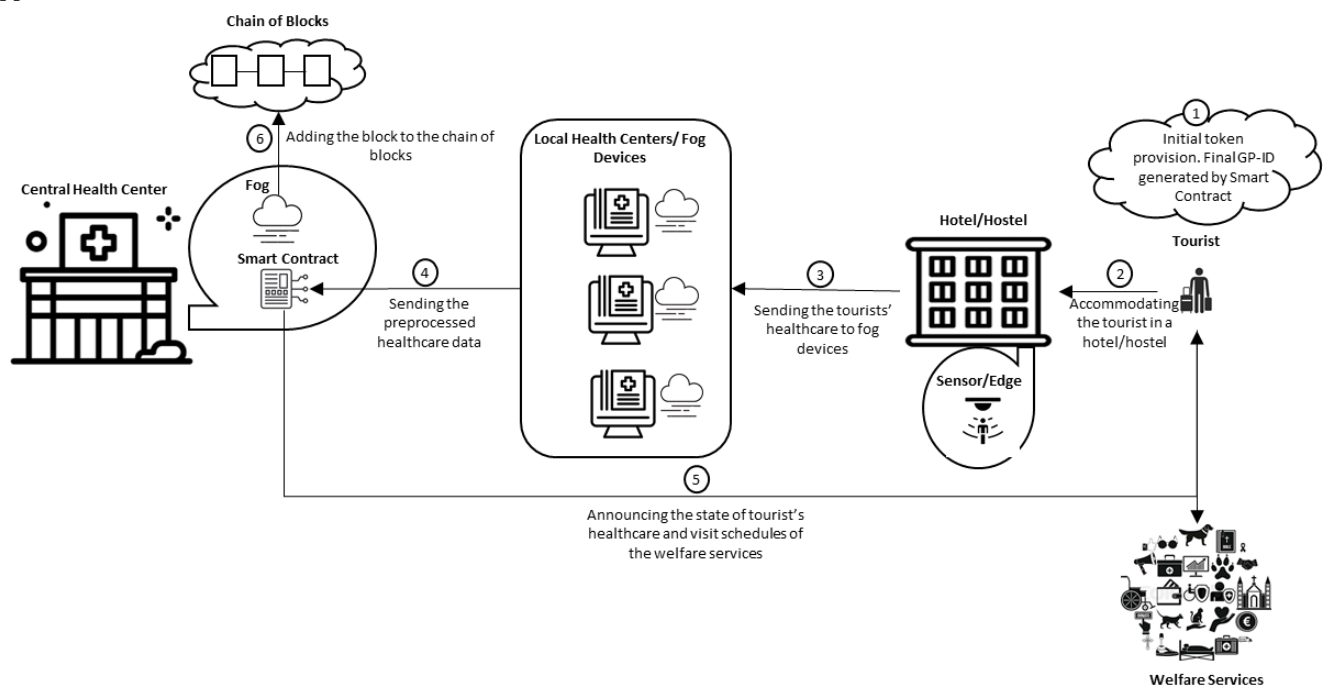


Figure 3. Components of STHChain

PROPOSED STHCHAIN MODEL

In this paper, a lightweight and scalable blockchain-based healthcare model called STHChain is proposed for smart tourists, where their access to welfare services is facilitated. The objectives of the model are not only to enhance the blockchain network scalability and reduce energy consumption but also to provide welfare services to tourists in consideration of health protocols during the outbreak of infectious diseases. The model comprises tourists, welfare services, sensors, edge computing, fog computing, and smart contracts, whose interconnections are depicted in Figure 3.

TABLE 1
Research Gaps and STHChain's Contributions

Research Area and Limitation in Prior Work	How STHChain Addresses the Gap
Tourism Industry Challenges Proposed solutions are often policy-based, lacking a deployable technical infrastructure for real-time, city-scale health monitoring.	- Provides a hierarchical IoT-edge-fog technical architecture specifically designed for the distributed, dynamic environment of smart tourism. - introduce a formal RBAC model and a transparent, algorithm-based health status function, moving beyond procedural access control descriptions
IoT-based Healthcare Systems - Centralized or cloud-centric models face latency, privacy, and single-point-of-failure issues. - Data integrity and auditability in multi-stakeholder environments are challenging.	- Employs a distributed fog computing layer for low-latency processing and resilience. - Leverages blockchain as an immutable ledger to ensure integrity and enable audit trails for all health data transactions.
Blockchain for Healthcare - Conventional consensus (PoW, PoS) is energy-intensive and slow for high-frequency IoT data. - Storing personal health identifiers on-chain creates severe privacy violations. - Scalability under load from thousands of devices is a common weakness.	- Introduces the lightweight RaC consensus protocol, optimizing energy use and confirmation time via fog node clustering. - Implements a two-tier pseudonymization scheme to decouple patient identity from on-chain data, ensuring privacy-by-design. - The clustered, hierarchical design inherently supports horizontal scaling by adding new fog clusters to handle increased load.

As illustrated in Figure 3, smart hotels/hostels located in smart neighborhoods are equipped with IoT devices. Tourists' vital signs are measured by sensors and collected at edge devices located within the hotels/ hostels. In addition, health centers in each smart neighborhood are equipped with a fog node. Hotels/ hostels then transmit the relevant healthcare data to the fog nodes within the same smart neighborhood for pre-processing. Separately, a smart contract residing at the central health center of the smart city processes the data received from the fog nodes of the health centers in the smart neighborhoods. Based on tourists' healthcare data, this smart contract determines their eligibility to access welfare services during disease outbreaks. Finally, tourists' healthcare data are stored in blocks and recorded on the chain of blocks. The following sections present details of the proposed model.

A. Hierarchical Configuration of the IoT-based Healthcare Network

This section describes a hierarchical IoT-based healthcare network for smart tourism. First, hotels/hostels are equipped with sensors and edge devices. For each tourist, a set of sensors is assigned to record the tourist's vital signs in the room where they are staying. The sensors in each room are connected to a specific pin on the edge device, whereby each pin identifies which room's healthcare data is being received. A fog node is then deployed within each hotel/hostel to collect data from the edge devices present there. Each edge device in the hotel/hostel is initially authenticated by this fog node. Subsequently, a fog node is placed at the health center of each neighborhood to perform pre-processing of the healthcare data from the edge devices serving that neighborhood. At the neighborhood health center, the healthcare data of tourists are pre-processed daily. For each tourist, a temporary authentication token is initially provided during hotel check-in. This token is used to generate a secure Local Pseudonym (LP-ID) at the hotel's fog node. This LP-ID is then sent to the central smart contract, which issues a unique Global Pseudonym (GP-ID). The mean sensor data for each tourist is irrevocably associated with this GP-ID on the blockchain, as formalized in Equation (1).

$$TX_i = AvgData_{i,j} | GP - ID_j \quad (1)$$

As shown in Equation (1), the transaction associated with sensor i , denoted by TX_i , presents the mean data from sensor i for tourist j as $AvgData_{i,j}$ and the tourist's global pseudonym identifier as $GP - ID_j$. This data is then included in the matrix-Merkle block structure proposed in [36], which organizes transactions into a fixed-size matrix before computing a Merkle root for efficient batch verification. This design ensures data integrity while minimizing storage overhead, a critical consideration for resource-constrained fog nodes.

B. Proposed Role-Based Access Control Mechanism and Health Status Determination

The access control in STHChain is governed by a smart contract implementing a formal Role-Based Access Control (RBAC) model. The model is defined by the following core sets and rules:

- **Sets:** U (Users: *Tourists, Welfare Centers*), R (Roles: *Tourist, ServiceProvider*), P (Permissions: *read, write, requestAccess*), S (Sessions).
- **Assignment Rules:** A user $u \in U$ is assigned a role $r \in R$ by the smart contract based on authenticated credentials. Tourists receive the *Tourist* role upon registration, granting them *read* permission to their own data and *requestAccess* to services. Welfare centers are assigned the *ServiceProvider* role, granting them *write* permission to publish available time slots and conditional *read* permission to verify a tourist's status before granting access.
- **Conflict Resolution:** The *ServiceProvider* role

inherently overrides the *Tourist* role in defining service parameters. However, access to a specific tourist's data for verification requires a session where the tourist's *requestAccess* permission is actively invoked, ensuring a two-party handshake.

Health Status Determination: The classification of a tourist as "healthy" or "unhealthy" is not arbitrary but is determined by a deterministic function executed by the smart contract. Let $AvgData_{i,j}$ represent the mean value from sensor i for tourist j over a evaluation window (e.g., 24 hours), and let τ_i be a predefined clinical threshold for that vital sign. The health status HS_j for tourist j is computed as:

$$HS_j = \begin{cases} \text{'healthy'}, & \text{if } \forall i, |AvgData_{i,j} - Baseline_i| \leq \tau_i \\ \text{'unhealthy'}, & \text{otherwise} \end{cases}$$

The $Baseline_i$ and thresholds τ_i are set by health authorities and encoded in the smart contract logic. This status is recalculated periodically with new data.

Interaction Flow: The end-to-end process for service access operates through a sequenced dialogue between the entities. It commences when the aggregated health data ($AvgData$) of a tourist is transmitted to the smart contract. The contract subsequently executes its embedded decision function to compute the health status (HS_j) for that tourist and updates the corresponding record on-chain. Independently, welfare services publish their available time slots ($SlotInfo$) to the contract. A tourist can then query the smart contract to retrieve a list of accessible slots, which are dynamically filtered based on their current HS_j . Upon selecting a suitable slot, the tourist invokes a *requestAccess* function, prompting the contract to generate and issue a temporary, cryptographically-secure access token. This token is presented by the tourist to the welfare service at the point of visit. The service, in turn, submits a verification query to the smart contract, validating both the authenticity of the token and the consistency of the tourist's health status. Only after receiving a positive verification from the contract does the welfare service grant physical access. The entire transaction, culminating in the granted visit, is immutably logged as a new entry on the blockchain, completing the auditable cycle.

C. Proposed Raft-based Clustering Consensus Protocol

In this section a Raft-based Clustering consensus protocol called RaC is proposed that employs fog node clustering for verifying tourist healthcare transactions, as illustrated in Figure 4. The approach aims to enhance scalability and energy efficiency of blockchain-based healthcare data storage in smart tourism.

As shown in Figure 4, the smart contract sends tourists' healthcare data and the visitation statistics of welfare centers to the fog node located at the central health center for storing on the chain of blocks. The central health center includes the tourist transactions in blocks and forwards them to the fog nodes situated at local health centers for verification. The fog nodes at nearby local health centers form clusters, and then, using the Raft algorithm [48],

perform block verification.

(1) Cluster Formation

Fog nodes are statically grouped into clusters based on their logical affiliation to a specific health center within a smart neighborhood. This grouping is predefined by the system administrator to optimize geographical proximity and network latency. Each cluster consists of a small, fixed set of nodes (e.g., 3 to 5) to maintain Raft's efficiency.

(2) Intra-Cluster Consensus via Raft

Within each cluster, the standard Raft algorithm [48] is executed. One node is elected as the cluster leader, responsible for receiving new block proposals from the central smart contract, appending them to its log, and replicating them to the other follower nodes in the cluster. A log entry (block) is considered committed within the cluster once it is persisted on a majority of the cluster's nodes. This phase ensures data consistency and fault tolerance within the local group.

(3) Cross-Cluster Verification

This phase provides the overarching consensus for the entire network. When a block is committed within a cluster, its leader broadcasts the block digest (hash) to the leaders of all other clusters. Each receiving leader verifies the block against its own log. A block is considered globally verified only after it receives positive verification (VOTE) from a strict majority of cluster leaders (e.g., >50%). This step prevents any single cluster from controlling the ledger.

(4) Final Commitment

Upon achieving global verification, the original cluster leader disseminates a final COMMIT message. All nodes in all clusters then permanently append the verified block to their local copy of the blockchain.

Ultimately, the validated block is added to the chain of blocks in the distributed nodes. The use of fog nodes for block verification not only reduces system load and improves scalability, but also further reduces the energy consumption of the consensus process. Therefore, the tourism industry, by leveraging healthcare data securely stored using blockchain, can better detect tourism patterns during pandemics and leverage these insights for attracting visitors. The detailed step-by-step algorithmic procedure of the RaC protocol is formally presented in Algorithm 1 (see Appendix A).

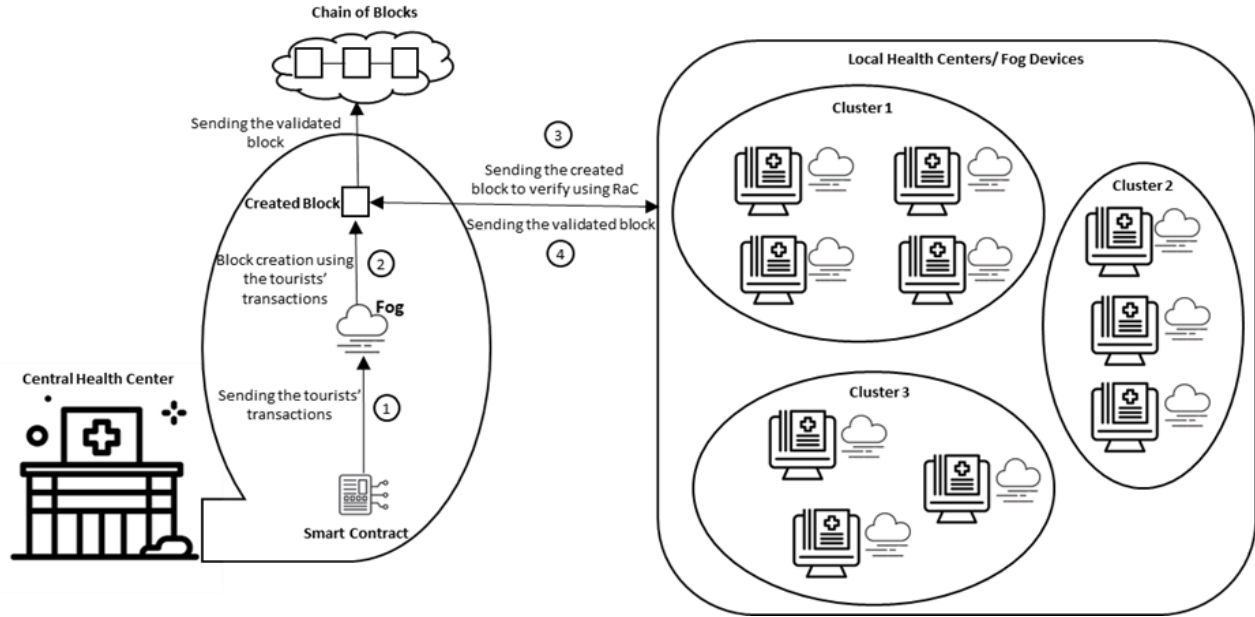


Figure 4. The proposed RaC Consensus protocol.

D. Security Analysis of the RaC Protocol

The RaC protocol inherits and extends security properties from the Raft algorithm and its cross-cluster design to defend against common threats.

- **Resilience to Sybil Attacks:** In a Sybil attack, an adversary creates multiple fake identities to subvert the consensus. RaC mitigates this by requiring that every fog node is a pre-registered and authenticated entity at a known health center, authorized by the central smart contract. This permissioned setting makes the creation of arbitrary fake nodes practically infeasible.
- **Robustness against Consensus Delay Attacks:** An attacker might try to deliberately delay consensus. The deterministic leader election and log replication in Raft rely on timeouts; if a leader fails, followers initiate a new election promptly. Furthermore, the cross-cluster verification requires a majority of different clusters. Therefore, compromising or delaying a single cluster (or its leader) cannot halt the entire network's progress, as other clusters can continue to propose and verify blocks.

As demonstrated in lines 14–18 of Algorithm 1, a block is only committed locally after receiving acknowledgements from a majority of its cluster members. This inherent property of the Raft algorithm, combined with the cross-cluster verification (lines 21–42), ensures that compromising a single leader or cluster cannot subvert the consensus, thereby mitigating the risk of a targeted consensus delay attack.

The immutability guaranteed by the RaC protocol directly secures the integrity of the health data and the resulting health status verdict (HS_j), making the access control foundation tamper-evident and trustworthy.

E. Privacy-Preserving Identity Management Scheme

Acknowledging the critical privacy concerns associated with using direct personal identifiers, the STHChain model implements a two-tier pseudonymization scheme designed with privacy-by-design principles. This approach ensures that no plaintext, government-issued identifiers (such as passport numbers) are stored on the blockchain or transmitted across the network.

- **Initial Registration and Local Pseudonym (LP-ID):** A tourist's passport number is used a single time during the initial check-in at the hotel for verification purposes. Immediately after verification, the system generates a Local Pseudonym (LP-ID). This LP-ID is created by applying a secure cryptographic hash function SHA-256 (We selected SHA-256 for its collision-resistance and widespread adoption in secure systems) to a concatenated string containing the passport number, a random nonce, and a unique salt specific to the hotel. The original passport data is then purged from the operational system. The LP-ID serves as the tourist's anonymous identifier for all local data processing within the hotel's edge and fog layers.
- **Global Pseudonym (GP-ID) for Blockchain Records:** When aggregated health data needs to be immutably recorded on the blockchain for access by authorized health centers, a second layer of pseudonymization is applied. The local fog node sends a request containing the LP-ID to a dedicated, central Identity Management Smart Contract. This trusted contract acts as a pseudo-Identity Provider, generating a unique, non-reversible Global Pseudonym (GP-ID) for that specific tourist session. It maintains a secure, encrypted, and access-controlled mapping between the LP-ID and the GP-ID. Only this contract can, under strict conditions defined in its logic (e.g., a legitimate medical emergency request from an authorized health center fog node),

resolve a GP-ID back to its corresponding LP-ID, and never to the original passport number. The GP-ID is the sole identifier appended to health transactions written to the blockchain, as shown in the revised Equation (1).

This scheme effectively breaks the linkability between the tourist's real-world identity, their local activity within a hotel, and their global health record on the blockchain. It aligns with data protection principles like minimization and purpose limitation, providing a robust technical safeguard for tourist privacy.

EVALUATION AND RESULTS

This section analyzes the security criteria and evaluates the performance of the proposed model.

A. Security Criteria

The proposed model presents the following security criteria using a smart city, blockchain, and smart contracts:

- **Authentication:** IoT devices in each hotel/hostel are authenticated by the hotel/hostel's own fog nodes. Subsequently, fog devices of hotels/hostels are authenticated by the fog node at the corresponding health center. Additionally, fog nodes located at health centers and the tourists themselves are authenticated by a smart contract.
- **Authorization:** Tourists are granted read permission, and welfare centers have read/write permissions within the smart contract. Furthermore, healthcare data in this model are retrieved exclusively from authenticated IoT devices, ensuring that only authorized devices may access the network.
- **Remote monitoring of healthcare data:** IoT devices deployed in hotels/ hostels remotely monitor tourists' vital signs.
- **Healthcare data traceability:** Tourists' healthcare transactions are traceable in the blockchain using their passport numbers as identifiers.
- **Healthcare data integrity:** Healthcare data for tourists are stored as transactions in blocks, which are chained to one another using hashes to ensure integrity.
- **Formal access control:** A formal Role-Based Access Control model governs all permissions. Access to services is contingent on a deterministic health status verdict computed by the smart contract, ensuring policy enforcement is automated and auditable.
- **Data Stewardship and Compliance:** The Ministry of Health acts as the data controller within the national framework, responsible for implementing the privacy-preserving architecture and ensuring regulatory compliance. Tourist rights over their personal data are preserved through the system's governance mechanisms.
- **Scalability:** Fog nodes located at health centers are clustered and, using the proposed RaC protocol, participate in block verification to improve consensus scalability. Additionally, the hierarchical network configuration further enhances the model's scalability.
- **Privacy Preserving Identification:** Tourist identity is protected through a two-tier pseudonymization scheme. Direct identifiers are never stored. Instead, unlinkable local (LP-ID) and global (GP-ID) pseudonyms are used for processing and blockchain storage, respectively. This design implements a key technical measure recommended by data protection frameworks like the GDPR to mitigate re-identification risks.

B. Performance Analysis

The energy consumption, latency, and throughput of the proposed model has been evaluated using the iFogSim simulator. The simulation environment has been configured with hardware specifications of 8 GB RAM and a Core i5 processor. The key entities and network parameters configured in iFogSim to emulate the smart tourism healthcare scenario are comprehensively detailed in Table 2.

TABLE 2
Summary of Key iFogSim Simulation Parameters

Entity / Parameter	Value / Configuration
Sensor Node	MIPS: 500; Uplink Bandwidth: 512 Kbps
Edge Device (per Hotel)	MIPS: 5000; RAM: 2 GB; Uplink Bandwidth: 10 Mbps
Fog Node (Health Center Cluster)	MIPS: 10000; RAM: 4 GB; Uplink Bandwidth: 100 Mbps
Cloud (Central Health Center)	MIPS: 100000; RAM: 8 GB; Uplink Bandwidth: 1 Gbps
Network Latency (Sensor-Edge)	2 ms
Network Latency (Edge-Fog, Intra-cluster)	5 ms
Network Latency (Fog-Fog, Inter-cluster)	50 ms
Network Latency (Fog-Cloud)	100 ms
Sensor Data Generation	Interval: 5 seconds; Packet Size: 250 bytes
Consensus Timeout (RaC Protocol)	1500 ms

(1) Energy Consumption Analysis

The energy consumption of the proposed model is analyzed for the number of sensors, edge devices, and fog nodes, as illustrated in Figure 5.

As shown in Figure 5, energy consumption is measured for scenarios with 25, 50, 100, 150 and 200 sensors connected to four distinct edge devices across four hotels/hostels. The energy consumption is higher when sensor density is low and the sensors are sparsely distributed, compared to higher-density deployments. This is because routing energy for sensors that are more widely spaced incurs greater energy expenditure. However, energy consumption increases gradually with increasing sensor density. Thus, a trade-off between sensor density and energy consumption is necessary. Notably, the energy consumption per sensor decreases as density increases, demonstrating the scalability of the edge-fog preprocessing layer.

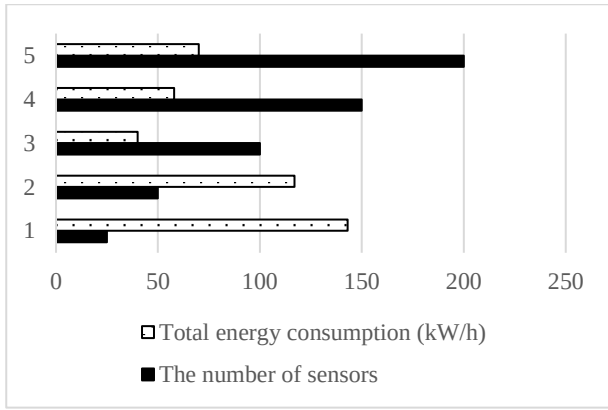


Figure 5. Energy Consumption of STHChain

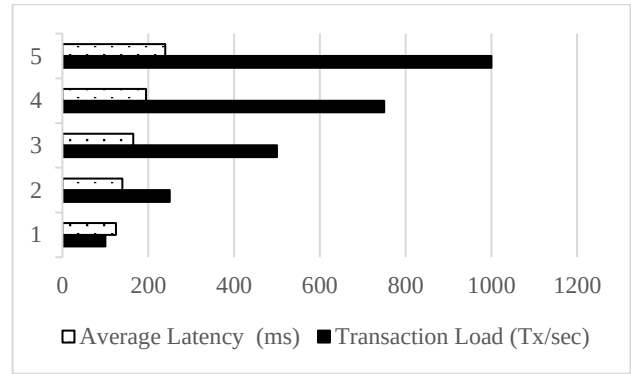


Figure 7. End-to-End Latency vs. Transaction Load

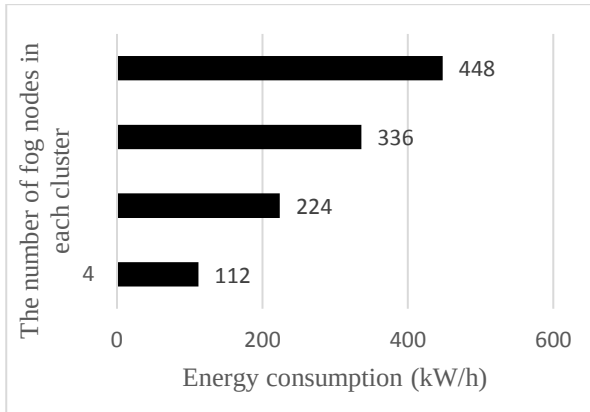


Figure 6. Energy consumption of RaC for four clusters.

In addition, the energy consumption of the proposed RaC consensus protocol is presented for four clusters, each containing a different number of fog nodes, as shown in Figure 6.

As observed in Figure 6, the energy consumption of the proposed protocol increases with the number of fog nodes in each cluster. While increasing the number of fog nodes in a cluster leads to higher energy usage, it also enhances system reliability because more nodes participate in block verification.

(2) Analysis of Latency and Throughput

Figure 7 illustrates the relationship between transaction load and the average end-to-end latency in the STHChain model. As shown, latency remains below 200 ms for loads up to 750 transactions per second, which is adequate for most real-time health monitoring services. The gradual increase in latency is attributed to queueing overhead in the intermediate Fog nodes, not a central bottleneck. This mild linear behavior indicates the suitable scalability of the architecture.

The system throughput is presented in Figure 8. The proposed model can process up to approximately 850 transactions per second with a high success rate (>98%). Beyond this point, the throughput saturates due to the inherent bandwidth limitation of the communication link between the Edge and Fog layers. This capacity is fully sufficient for the smart tourism scenario, assuming health data updates from each tourist every few minutes.

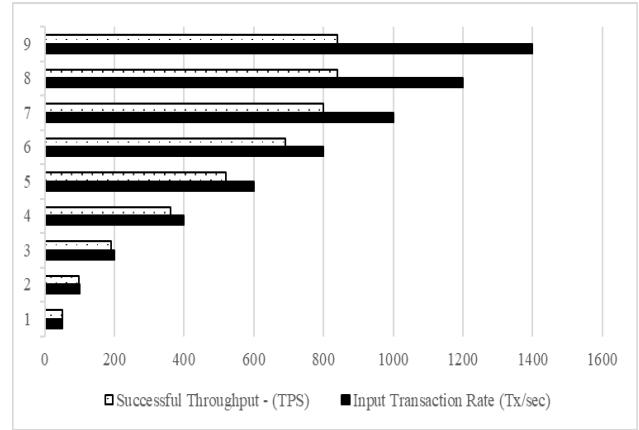


Figure 8. The system throughput of STHChain (Successful TPS).

(3) Large-Scale Scalability Analysis

In response to the need for evaluating the model under conditions that better reflect expansive smart city deployments, we conducted extended simulations scaling the system up to 500 sensor devices and 8 fog clusters. The objective was to stress-test the architecture and observe performance degradation trends. The results, consolidated in Figure 9, demonstrate the system's scalability limits.

As shown in Figure 9a, the average end-to-end latency increases in a sub-linear manner with the number of sensors, reaching approximately 580 ms at the maximum scale of 500 sensors. This significant rise is a direct consequence of the configured processing limits (MIPS) at the edge and fog layers, which lead to longer queueing delays under heavy load. While this latency may be acceptable for non-critical health data logging, it underscores the need for provisioning adequate hardware resources for real-time monitoring scenarios. Figure 9b reveals that the system throughput scales effectively as more fog clusters are added, confirming the benefits of our parallel, clustered design. However, the throughput gain diminishes sharply beyond 6 clusters, with minimal improvement observed at 8 clusters. This indicates that the simulation host's physical memory (8 GB RAM) becomes a bottleneck when orchestrating a large number of virtual fog nodes, highlighting a practical deployment consideration for very large-scale instances of the STHChain architecture.

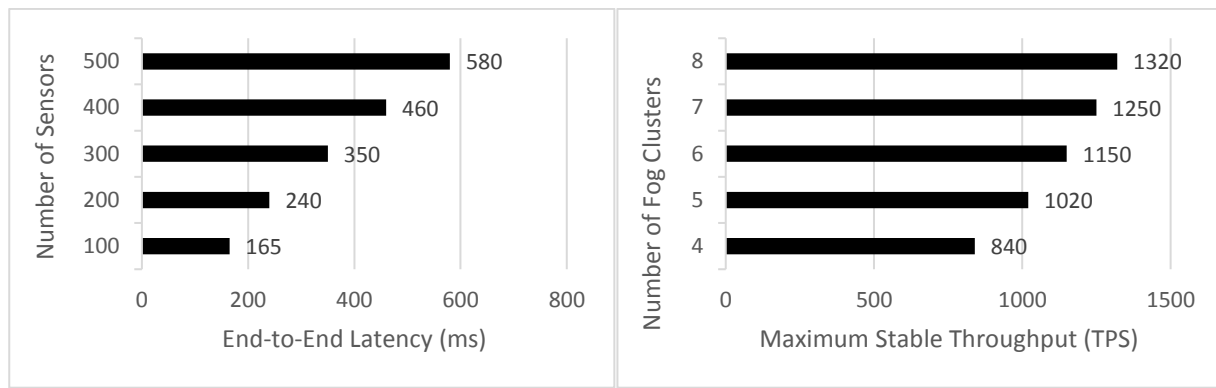


Figure 9. Large-Scale Scalability Analysis of STHChain- a. Latency vs. Network Size- b. Throughput vs. Cluster Scaling

TABLE 3
performance comparison of the proposed STHChain model with contemporary blockchain-based frameworks

Performance Metric	STHChain	Model for Comparison	Explanation and Analysis
Average Latency (End-to-End)	165 ms	~2.8 seconds [38]	STHChain achieves an order-of-magnitude lower latency. This is a direct result of its two-tier fog architecture and RaC protocol, which localizes consensus within hotel/health center clusters, eliminating the wide-area network delays inherent in the compared hierarchical model.
Maximum Throughput (Stable TPS)	840 TPS	~95 TPS [38]	The proposed model demonstrates approximately 9x higher throughput. This scalability stems from the parallel processing capability of independent RaC clusters at different neighborhood health centers, unlike the more sequential verification process described in [38].
Transaction/Block Confirmation Time	1.8 seconds (For RaC consensus)	~0.28 seconds [46]	While the framework in [46] reports faster <i>transaction time</i> , it primarily focuses on private data access control on a permissioned blockchain. STHChain's slightly higher time encompasses full block formation and immutable ledger storage for audit trails, providing a stronger data integrity guarantee for medical records.
Energy Efficiency	Optimal (40 units at 100 sensors, Figure 5)	A Key Challenge [27]	The hierarchical data filtering in STHChain (Edge → Fog → Blockchain) directly addresses the high energy consumption challenge outlined in [27] for lightweight blockchains. By processing and compressing data at the edge, only critical transactions are broadcasted, drastically reducing the network's energy footprint.
Architectural Scalability	Excellent (Horizontal, via fog cluster addition)	Good (Vertical, via node addition in hierarchy [38])	STHChain's design enables linear scalability by adding new independent fog clusters for new city districts. In contrast, scaling the model in [38] requires adding more nodes to the existing hierarchy, which may increase internal communication overhead.

C. Comparative Analysis

This section places the performance results of the proposed STHChain model within the broader context of current research by presenting a direct comparison with two contemporary and architecturally relevant blockchain frameworks for healthcare. The objective of this analysis is to validate the efficacy of our design choices specifically, the hierarchical edge-fog configuration and the novel RaC consensus protocol by measuring them against established alternatives. The selected models for comparison are:

A Hierarchical Blockchain Model focused on healthcare scalability [38], which provides a valuable baseline for evaluating structural efficiency in managing distributed health data.

A Privacy-Preserving Access Control Framework for Electronic Health Records (EHR) [46], representing a state-of-the-art solution that prioritizes data security and transaction authorization.

The comparison in Table 3 examines key operational metrics critical for IoT-based health monitoring in smart tourism: end-to-end latency, system throughput, block confirmation time, and inherent architectural scalability. The results demonstrate that STHChain achieves a superior balance across these parameters. Notably, the model's ability to leverage localized consensus via the RaC protocol allows it to overcome the typical trade-off between transaction speed and network scalability observed in other architectures. This positions STHChain as a tailored and efficient solution for real-time, large-scale health data management in dynamic environments like smart cities.

CONCLUSION

In this paper, a lightweight and scalable model for a blockchain-based smart tourist health system called STHChain is proposed to deliver secure and privacy-compliant welfare services for tourists. The model leverages a hierarchical IoT-edge-fog network to optimize energy efficiency and data processing latency. A core innovation is the formal Role-Based Access Control (RBAC) mechanism and an automated, data-driven health status function implemented within a smart contract. This contract dynamically governs access permissions for tourists, health centers, and welfare services based on objectively evaluated health data, ensuring that service provision is both secure and context-aware.

Tourists' healthcare data, pseudonymized via a two-tier privacy-preserving scheme, and their service access logs are immutably stored on the blockchain. Data integrity and consensus are maintained through the novel Raft-based Clustering (RaC) protocol, which enhances scalability and reduces energy consumption by organizing fog nodes into efficient verification clusters. A comprehensive security analysis confirms the model's resilience against common

threats, including Sybil and consensus delay attacks.

The evaluation, conducted using the iFogSim simulator with detailed parameters, demonstrates that STHChain achieves significant gains in energy efficiency, latency, throughput, and scalability under large-scale conditions, outperforming relevant baseline architectures in key metrics for IoT-healthcare systems. Future work will focus on three directions: (1) integrating lightweight machine learning models into the smart contract for real-time health anomaly detection, (2) testing the system in a physical testbed with real IoT devices to validate simulation results, and (3) extending the privacy scheme to support zero-knowledge proofs for selective data disclosure without compromising pseudonymity.

REFERENCES

- [1] A. B. Haque, B. Bhushan, and G. Dhiman. (2022). Conceptualizing smart city applications: Requirements, architecture, security issues, and emerging trends. *Expert Systems*. [Online]. 39(5), p. e12753. Available: <https://doi.org/10.1111/exsy.12753>
- [2] N. J. Habeeb and S. T. Weli. (2020). Relationship of smart cities and smart tourism: An overview. *HighTech and Innovation Journal*. [Online]. 1(4), pp. 194–202. Available: <https://doi.org/10.28991/HIJ-2020-01-04-07>
- [3] C. Mertzanis and A. Papastathopoulos. (2021). Epidemiological susceptibility risk and tourist flows around the world. *Annals of Tourism Research*. [Online]. 86, p. 103095. Available: <https://doi.org/10.1016/j.annals.2020.103095>
- [4] B. Farahani, F. Firouzi, and K. Chakrabarty. (2020). Healthcare IoT. In *Intelligent Internet of Things*. [Online]. Springer, pp. 515–545. Available: https://doi.org/10.1007/978-3-030-30367-9_11
- [5] R. Karthick, R. Ramkumar, M. Akram, and M. V. Kumar. (2021). Overcome the challenges in biomedical instruments using IOT – A review. *Materials Today: Proceedings*. [Online]. 45, pp. 1614–1619.
- [6] M. Peyroteo, I. A. Ferreira, L. B. Elvas, J. C. Ferreira, and L. V. Lapão. (2021). Remote monitoring systems for patients with chronic diseases in primary health care: Systematic review. *JMIR mHealth and uHealth*. [Online]. 9(12), p. e28285. Available: <https://doi.org/10.2196/28285>
- [7] A. S. Albahri et al. (2021). IoT-based telemedicine for disease prevention and health promotion: State-of-the-Art. *Journal of Network and Computer Applications*. [Online]. 173, p. 102873. Available: <https://doi.org/10.1016/j.jnca.2020.102873>
- [8] H. Attaran, N. Kheibari, and D. Bahrepour. (2022). Toward integrated smart city: A new model for implementation and design challenges. *GeoJournal*. [Online]. 87(Suppl 4), pp. 511–526. Available: <https://doi.org/10.1007/s10708-021-10560-w>
- [9] D. J. Cook, G. Duncan, G. Sprint, and R. L. Fritz. (2018). Using smart city technology to make healthcare smarter. *Proceedings of the IEEE*. [Online]. 106(4), pp. 708–722. Available: <https://doi.org/10.1109/JPROC.2017.2787688>

- [10] M. Jabbar, K. Prasad, and R. Aluvalu. (2021). Reimagining the Indian healthcare ecosystem with AI for a healthy smart city. In *Emerging Technologies in Data Mining and Information Security*. [Online]. Springer, pp. 543–551. Available: https://doi.org/10.1007/978-981-15-9927-9_53
- [11] L. Qi et al. (2021). Privacy-aware data fusion and prediction with spatial-temporal context for smart city industrial environment. *IEEE Transactions on Industrial Informatics*. [Online]. 17(6), pp. 4159–4167. Available: <https://doi.org/10.1109/TII.2020.3012157>
- [12] M. Gheisari et al. (2021). OBPP: An ontology-based framework for privacy-preserving in IoT-based smart city. *Future Generation Computer Systems*. [Online]. 123, pp. 1–13. Available: <https://doi.org/10.1016/j.future.2021.01.028>
- [13] J. Venkatesh, B. Aksanli, C. S. Chan, A. S. Akyurek, and T. S. Rosing. (2018). Modular and personalized smart health application design in a smart city environment. *IEEE Internet of Things Journal*. [Online]. 5(2), pp. 614–623. Available: <https://doi.org/10.1109/JIOT.2017.2712558>
- [14] S. S. F. Nasab, T. Z. Marjaneh, and D. Bahrepour. (2023). Energy efficiency and establishing service level agreement using fuzzification of virtual machine selection policies for migrating in cloud computing. Presented at 9th International Conference on Web Research (ICWR). [Online]. Available: <https://doi.org/10.1109/ICWR57742.2023.10138982>
- [15] S. S. Fateminasab, S. Memarian, S. R. K. Tabbakh, and M. C. Romero-Terner. (2024). A review on open data storage and retrieval techniques in blockchain-based applications. Presented at 10th International Conference on Web Research (ICWR). [Online]. Available: <https://doi.org/10.1109/ICWR61162.2024.10533356>
- [16] S. S. F. Nasab, D. Bahrepour, and S. R. K. Tabbakh. (2022). A review on secure data storage and data sharing technics in blockchain-based IoT healthcare systems. Presented at 12th International Conference on Computer and Knowledge Engineering (ICCKE). [Online]. Available: <https://doi.org/10.1109/ICCKE57176.2022.9959977>
- [17] S.-Y. Pan, M. Gao, H. Kim, K. J. Shah, S.-L. Pei, and P.-C. Chiang. (2018). Advances and challenges in sustainable tourism toward a green economy. *Science of the Total Environment*. [Online]. 635, pp. 452–469. Available: <https://doi.org/10.1016/j.scitotenv.2018.04.134>
- [18] S. B. Hojaghan and A. N. Esfangareh. (2011). Digital economy and tourism impacts, influences and challenges. *Procedia - Social and Behavioral Sciences*. [Online]. 19, pp. 308–316. Available: <https://doi.org/10.1016/j.sbspro.2011.05.136>
- [19] D. Dredge and S. Gyimóthy. (2017). Collaborative economy and tourism. In *Collaborative Economy and Tourism: Perspectives, Politics, Policies and Prospects*. [Online]. Springer, pp. 1–12. Available: <https://doi.org/10.1007/978-3-319-51799-5>
- [20] M. Abubakar et al. (2018). Medical tourism in Nigeria: Challenges and remedies to health care system development. *International Journal of Development and Management Review*. [Online]. 13(1), pp. 223–238.
- [21] J. P. Singh. (2015). Healthcare tourism in India: Opportunity and challenges. *Asian Journal of Multidimensional Research (AJMR)*. [Online]. 4(3), pp. 37–47.
- [22] R. H. Chianeh, G. Del Chiappa, and V. Ghasemi. (2020). Cultural and religious tourism development in Iran: Prospects and challenges. In *Tourism in Iran: Challenges, Development and Issues*. [Online]. Routledge, pp. 123–140. Available: <https://doi.org/10.1080/13032917.2017.1414439>
- [23] R. Sharpley. (2003). Tourism, modernisation and development on the island of Cyprus: Challenges and policy responses. *Journal of Sustainable Tourism*. [Online]. 11(2–3), pp. 246–265. Available: <https://doi.org/10.1080/096695803088667205>
- [24] S. P. Singh, A. Nayyar, R. Kumar, and A. Sharma. (2019). Fog computing: From architecture to edge computing and big data processing. *The Journal of Supercomputing*. [Online]. 75, pp. 2070–2105. Available: <https://doi.org/10.1007/s11227-018-2701-2>
- [25] Q. Qi and F. Tao. (2019). A smart manufacturing service system based on edge computing, fog computing, and cloud computing. *IEEE Access*. [Online]. 7, pp. 86769–86777. Available: <https://doi.org/10.1109/ACCESS.2019.2923610>
- [26] Y. Shi, G. Ding, H. Wang, H. E. Roman, and S. Lu. (2015). The fog computing service for healthcare. Presented at 2nd International Symposium on Future Information and Communication Technologies for Ubiquitous HealthCare (Ubi-HealthTech). [Online]. Available: <https://doi.org/10.1109/Ubi-HealthTech.2015.7203325>
- [27] L. Ismail, H. Materwala, and S. Zeadally. (2019). Lightweight blockchain for healthcare. *IEEE Access*. [Online]. 7, pp. 149935–149951. Available: <https://doi.org/10.1109/ACCESS.2019.2947613>
- [28] A. Abbas, R. Alroobaea, M. Krichen, S. Rubaiee, S. Vimal, and F. M. Almansour. (2024). Blockchain-assisted secured data management framework for health information analysis based on Internet of Medical Things. *Personal and Ubiquitous Computing*. [Online]. 28(1), pp. 59–72. Available: <https://doi.org/10.1007/s00779-021-01583-8>
- [29] A. Al Omar, M. S. Rahman, A. Basu, and S. Kiyomoto. (2017). MediBchain: A blockchain based privacy preserving platform for healthcare data. Presented at SpaCCS Workshops. [Online]. Available: https://doi.org/10.1007/978-3-319-72395-2_49
- [30] M. T. Quasim, F. Algarni, A. A. E. Radwan, and G. M. M. Alshmrani. (2020). A blockchain based secured healthcare framework. Presented at International Conference on Computational Performance Evaluation (ComPE). [Online]. Available: <https://doi.org/10.22075/ijnaa.2022.7186>

- [31] M. Qi et al. (2024). Privacy protection for blockchain-based healthcare IoT systems: A survey. *IEEE/CAA Journal of Automatica Sinica*. [Online]. 11(8), pp. 1757–1776. Available: <https://doi.org/10.1109/JAS.2022.106058>
- [32] R. Li, T. Song, B. Mei, H. Li, X. Cheng, and L. Sun. (2019). Blockchain for large-scale internet of things data storage and protection. *IEEE Transactions on Services Computing*. [Online]. 12(5), pp. 762–771. Available: <https://doi.org/10.1109/TSC.2018.2853167>
- [33] W. Liang, Y. Fan, K.-C. Li, D. Zhang, and J.-L. Gaudiot. (2020). Secure data storage and recovery in industrial blockchain network environments. *IEEE Transactions on Industrial Informatics*. [Online]. 16(10), pp. 6543–6552. Available: <https://doi.org/10.1109/TII.2020.2966069>
- [34] Y. Y. Yap et al. (2025). Smart tourism technologies and tourist satisfaction: A systematic literature review and research agenda. *Acta Psychologica*. [Online]. 258, p. 105191. Available: <https://doi.org/10.1016/j.actpsy.2025.105191>
- [35] S. S. Fateminasab, D. Bahrepour, and S. R. K. Tabbakh. (2025). A fair non-collateral consensus protocol based on Merkle tree for hierarchical IoT blockchain. *Scientific Reports*. [Online]. 15(1), p. 3645. Available: <https://doi.org/10.1038/s41598-025-87025-4>
- [36] S. S. Fateminasab, N. Evaznia, S. Memarian, M. d. C. Romero-Ternero, G. Miró-Amarante, and S. R. K. Tabbakh. (2025). An efficient blockchain-based resource allocation and secure data storage model using Fire Hawk Optimization and entropy in health tourism. *Journal of Cloud Computing*. [Online]. 14(1), p. 63.
- [37] F. Rizal Batubara, J. Ubacht, and M. Janssen. (2019). Unraveling transparency and accountability in blockchain. Presented at 20th Annual International Conference on Digital Government Research. [Online]. Available: <https://doi.org/10.1145/3325112.3325262>
- [38] L. Sadath, D. Mehrotra, and A. Kumar. (2024). Scalability performance analysis of blockchain using hierarchical model in healthcare. *Blockchain in Healthcare Today*. [Online]. 7(1). Available: <https://doi.org/10.30953/bhty.v7.295>
- [39] S. R. Mallick, R. K. Lenka, P. K. Tripathy, D. C. Rao, S. Sharma, and N. K. Ray. (2024). A lightweight, secure, and scalable blockchain-fog-IoMT healthcare framework with IPFS data storage for healthcare 4.0. *SN Computer Science*. [Online]. 5(1), p. 198. Available: <https://doi.org/10.1007/s42979-023-02511-8>
- [40] H. F. Atlam, A. Alenezi, M. O. Alassafi, and G. Wills. (2018). Blockchain with internet of things: Benefits, challenges, and future directions. *International Journal of Intelligent Systems and Applications*. [Online]. 10(6), pp. 40–48. Available: <https://doi.org/10.5815/ijisa.2018.06.05>
- [41] A. Saini, Q. Zhu, N. Singh, Y. Xiang, L. Gao, and Y. Zhang. (2021). A smart-contract-based access control framework for cloud smart healthcare system. *IEEE Internet of Things Journal*. [Online]. 8(7), pp. 5914–5925. Available: <https://doi.org/10.1109/JIOT.2020.3032997>
- [42] R. Akkaoui, X. Hei, C. Guo, and W. Cheng. (2019). RBAC-HDE: On the design of a role-based access control with smart contract for healthcare data exchange. Presented at IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW). [Online].
- [43] R. Kumar and R. Tripathi. (2021). Scalable and secure access control policy for healthcare system using blockchain and enhanced Bell–LaPadula model. *Journal of Ambient Intelligence and Humanized Computing*. [Online]. 12(2), pp. 2321–2338. Available: <https://doi.org/10.1007/s12652-020-02346-8>
- [44] S. Tanwar, K. Parekh, and R. Evans. (2020). Blockchain-based electronic healthcare record system for healthcare 4.0 applications. *Journal of Information Security and Applications*. [Online]. 50, p. 102407. Available: <https://doi.org/10.1016/j.jisa.2019.102407>
- [45] N. Kshetri. (2018). Blockchain and electronic healthcare records [cybertrust]. *Computer*. [Online]. 51(12), pp. 59–63. Available: <https://doi.org/10.1109/MC.2018.2880021>
- [46] A. K. Jakhar, M. Singh, R. Sharma, W. Viriyasitavat, G. Dhiman, and S. Goel. (2024). A blockchain-based privacy-preserving and access-control framework for electronic health records management. *Multimedia Tools and Applications*. [Online]. 83(36), pp. 84195–84229. Available: <https://doi.org/10.1007/s11042-024-18827-3>
- [47] A. Raj and S. Prakash. (2024). An efficient blockchain-based access control framework for IoT-healthcare system. *Wireless Personal Communications*. [Online]. 136(2), pp. 1017–1045. Available: <https://doi.org/10.1007/s11277-024-11312-y>
- [48] D. Ongaro and J. Ousterhout. (2015). The raft consensus algorithm. Lecture Notes. [Online].
- [49] S. Soleimani Nadaf, M. Niazi Torshiz, S. Mazinani et al. (2025). A lightweight scalable and dynamic blockchain-based model for storing and retrieving patient healthcare records. *Scientific Reports*. [Online]. 15, p. 44568. Available: <https://doi.org/10.1038/s41598-025-28337-3>
- [50] M. F. A. Mohammed et al. (2025). A cluster-based architecture for fog computing. *IEEE Access*. [Online]. 13, pp. 187638–187657. Available: <https://doi.org/10.1109/ACCESS.2025.3627412>

AppendixA

Algorithm1 : Raft-based Clustering (RaC) Consensus Protocol

Definitions:

FN[] : Array of all Fog Nodes in the network

CL[] : Array of Clusters. CL[k] is a list of Fog Nodes in clusterk

SmartContract : The central controller

Block : The data block to be appended

majority(n) : Function returning smallest integer $> n/2$
 HealthStatus: The health classification ('healthy'/'unhealthy') for a tourist, computed by the smart contract.

Input: A new Block proposed by the SmartContract

Output: Block appended globally or an error message

PHASE 1: Cluster-Centric Block Proposal

Procedure PROPOSE_BLOCK(Block):

1. primaryCluster = SELECT_PRIMARY_CLUSTER()
 // e.g., cluster nearest to the data source
2. primaryLeader = GET_LEADER(primaryCluster)
3. Send Block from SmartContract to primaryLeader

End Procedure

PHASE 2: Intra-Cluster Consensus (Raft)

Procedure INTRA_CLUSTER_CONSENSUS(Block, leader):

4. leader.log.append(Block)
5. // Replicate block to all followers in the same cluster
6. FOR each follower IN leader.cluster:
7. Send AppendEntriesRPC(Block) to follower
8. END FOR
9. // Wait for acknowledgements from a majority
10. IF acks_received $\geq$ majority(size(leader.cluster)) THEN
11. Block.status = COMMITTED_LOCALLY
12. RETURN SUCCESS
13. ELSE
14. RETURN FAILURE // Triggers leader re-election or retry
15. END IF

End Procedure

PHASE 3: Cross-Cluster Verification

Procedure CROSS_CLUSTER_VERIFICATION (Block, primaryLeader):

16. verificationVotes = 0
17. // Send verification request to other cluster leaders
18. FOR each cluster IN CL WHERE cluster $\neq$ primaryLeader.cluster:
19. otherLeader = GET_LEADER(cluster)
20. Send VerificationRequest(Block.hash) to otherLeader
21. END FOR
22. // Collect validation votes
23. FOR each otherLeader IN otherLeaders:
24. vote = receive VerificationResponse from otherLeader within TIMEOUT
25. IF vote == VALID THEN
26. verificationVotes = verificationVotes + 1
27. END IF
28. END FOR
29. // Check for inter-cluster majority
30. IF verificationVotes $\geq$ majority(size(CL)) THEN

31. Block.status = GLOBALLY_VERIFIED
32. RETURN SUCCESS
33. ELSE
34. RETURN FAILURE // Block is rejected
35. END IF

End Procedure

PHASE 4: Final Commitment & Broadcasting

Procedure FINAL_COMMIT(Block, primaryLeader):

36. // Broadcast final commit order to all nodes in all clusters
37. FOR each node IN FN:
38. Send FinalCommitOrder(Block) to node
39. node.blockchain.append(Block)
40. END FOR
41. Broadcast EVENT: "Block <Block.ID> committed successfully"

End Procedure

MAIN EXECUTION FLOW

Procedure MAIN_RAC(Block):

- // Step 1: Propose block to a primary cluster
- PROPOSE_BLOCK(Block)
- // Step 2: Reach consensus within the primary cluster
- IF INTRA_CLUSTER_CONSENSUS(Block, primaryLeader) == SUCCESS THEN
- // Step 3: Validate block across different clusters
- IF CROSS_CLUSTER_VERIFICATION(Block, primaryLeader) == SUCCESS THEN
- // Step 4: Finalize and broadcast the block
- FINAL_COMMIT(Block, primaryLeader)
- RETURN "Block committed successfully to the blockchain."
- ELSE
- RETURN "ERROR: Cross-cluster verification failed. Block rejected."
- END IF
- ELSE
- RETURN "ERROR: Intra-cluster consensus failed. Block not verified."
- END IF

End Procedure



Ferdowsi
University of
Mashhad

Journal of Computer and Knowledge Engineering

<https://cke.um.ac.ir>



Information and
Communication
Technology Association of
Iran

Consumer Behavior-Enhanced CNN-BiLSTM-Attention Model for Residential Electricity Consumption Forecasting *

Research Article

Mahdi Yousefi¹, Sara Ershadinasab²



[10.22067/cke.2026.93949.1165](https://doi.org/10.22067/cke.2026.93949.1165)

Abstract-- Accurate short-term prediction of residential electricity consumption is critical for smart grid integration, energy efficiency, and proactive building management. This study presents a hybrid deep learning framework that integrates convolutional neural networks (CNN), bi-directional long short-term memory (BiLSTM), and an attention mechanism, enhanced by dynamic consumer behavior indicators and optimized using genetic algorithms. The proposed CNN-BiLSTM-Attention model achieves state-of-the-art forecasting performance, with a root mean square error (RMSE) of 0.078 and a 20.5% improvement over baseline models. Genetic algorithm-based hyperparameter tuning yields an additional 5.1% performance gain. The model is validated across a public benchmark dataset, demonstrating its adaptability and robustness in real-time smart building energy management scenarios. Attention mechanism offers interpretable insights into consumption patterns, enabling actionable decisions for energy optimization and demand response. Extensive ablation studies and comparative evaluations with existing methods confirm the model's efficacy and practicality for real-world deployment. This work advances intelligent, adaptive load prediction systems for residential environments, supporting the broader goals of sustainable energy management and next-generation smart grid systems.

Index Terms-- Attention mechanism, Consumer behavior, Deep learning, Demand response, Electricity consumption forecasting, Energy management, Genetic algorithm, Neural networks, Smart buildings, Time series prediction

INTRODUCTION

Residential buildings account for approximately 27% of global electricity consumption, with this figure rising to nearly 38% in developed economies [1]. As global energy demand escalates and sustainability imperatives intensify, optimizing residential electricity consumption has become a critical priority for energy providers, policymakers, and end users alike. Within this context, smart grids and intelligent building energy management systems play a pivotal role by enabling demand response programs, load balancing, and peak demand reduction, all of which rely

fundamentally on accurate short-term load forecasting [2]. However, residential electricity consumption patterns are particularly challenging to predict due to their high variability, strong dependence on occupant behavior, and complex interactions among environmental, socio-economic, and temporal variables[3].

Traditional statistical approaches—such as auto-regressive integrated moving average (ARIMA) and classical regression models—have proven effective for aggregate-level forecasting but often fail to capture the non-linear dynamics and temporal dependencies that characterize individual household consumption[4].

While conventional methods perform adequately for aggregate-level prediction, they struggle to represent the complex, non-linear, and highly variable consumption behaviors of individual households. This limitation has fueled increasing interest in machine learning (ML) and deep learning (DL) approaches, which possess a superior capacity to model intricate temporal dependencies and capture heterogeneous influencing factors [5]. Recent advances in recurrent neural networks (RNNs) and their variants—particularly long short-term memory (LSTM) and Bidirectional LSTM (BiLSTM) models—have demonstrated promising results for various time-series forecasting tasks [6]. Nevertheless, existing deep learning models for residential electricity forecasting continue to face several critical challenges: (1) inadequate modeling of consumer behavior patterns that significantly influence consumption variability; (2) limited ability to capture both short-term and long-term temporal dependencies simultaneously; and (3) insufficient emphasis on model interpretability, which is essential for practical integration into smart energy management systems. Recently, attention mechanisms have emerged as a transformative enhancement in deep learning, enabling models to selectively focus on the most informative portions of input sequences[7][8].

* Manuscript received June 17, 2025, Revised August 29, 2025, Accepted May 31, 2026.

¹ M.Sc. Department of Computer Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran.

² Corresponding author. Assistant Professor, Department of Computer Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran. Email: ershadinasab@um.ac.ir

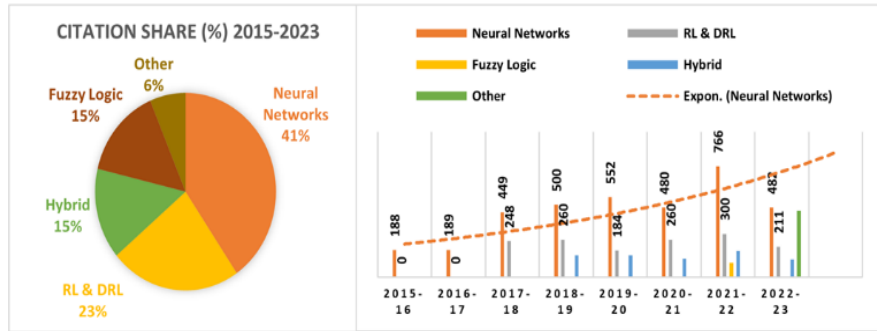


Figure 1. Citation Share (%) of Methods in Building Energy Forecasting (2015-2023)

A comprehensive chart showing the evolution of forecasting methods: Neural Networks (41%), RL & DRL (23%), Hybrid (15%), Fuzzy Logic (15%), and other methods (6%). The exponential trend line demonstrates the growing dominance of neural network approaches in recent years, highlighting the shift from traditional statistical methods to deep learning architectures.

Although attention-based models have achieved remarkable success in natural language processing and general time-series forecasting, their application to behaviorally informed residential load prediction remains relatively unexplored. To address these gaps, this paper proposes a hybrid deep learning framework that integrates convolutional neural networks (CNNs), Bidirectional LSTMs, and an attention mechanism, augmented by dynamic behavioral feature engineering.

The primary contributions of this research are summarized as follows:

1. A behavior-enhanced forecasting framework: We reframe residential load prediction through the integration of engineered consumer behavior indicators—such as activity onset/offset, stability indices, base-load ratios, and behavioral anomaly flags—providing a structured way to capture routine-driven variability often overlooked in existing CNN–BiLSTM–Attention studies.
2. A rigorously benchmarked CNN–BiLSTM–Attention baseline for behavior-aware forecasting: Rather than claiming architectural novelty, this work establishes a carefully tuned and systematically evaluated hybrid CNN–BiLSTM–Attention model, directly benchmarking against state-of-the-art variants reported in recent literature. This allows us to isolate and quantify the added value of behavioral features and training enhancements.
3. Genetic algorithm hyperparameter tuning with explicit optimization cost analysis: GA is employed to improve key hyperparameters affecting model stability and generalization. We provide detailed optimization-time reporting and demonstrate that GA yields measurable improvements over Bayesian Optimization and PSO for the studied datasets.
4. Comprehensive ablation and interpretability analysis: Through multi-level ablations—removing behavioral features, CNN layers, BiLSTM components, attention, and GA tuning—we quantify the unique contribution of each component. Attention-weight mechanism reveals how behavioral features influence temporal relevance, enhancing interpretability for real-world energy management applications.
5. Cross-dataset validation and edge-deployment readiness: The model is evaluated on three public

residential datasets from different regions, demonstrating strong generalization.

Together, these contributions establish a behavior-aware, interpretable, and practically deployable forecasting framework, advancing the integration of deep learning into next-generation residential energy management systems. The remainder of this paper is organized as follows: Section 2 reviews related work on load forecasting and deep learning methods. Section 3 presents the datasets description and feature engineering step. Section 4 describes the proposed methodology. Section 5 presents performance evaluation and comparative analysis. Section 6 concludes the paper.

RELATED WORK

Over the past decades, the field of electricity load forecasting has evolved from traditional statistical approaches to more advanced deep learning paradigms. Early models such as ARIMA, exponential smoothing, and linear regression were widely adopted for their simplicity, mathematical rigor, and interpretability [9], [10]. However, their effectiveness in modeling the complex, non-linear, and time-dependent characteristics of residential electricity consumption remains limited [11]. These conventional techniques generally rely on assumptions of linearity and stationarity, which are seldom satisfied in real-world household data characterized by high variability and behavioral uncertainty. To address these shortcomings, researchers have increasingly explored hybrid modeling strategies that combine the strengths of statistical methods like ARIMA with the representational power of machine and deep learning architectures [12], [13]. Such hybrid approaches leverage ARIMA's proficiency in capturing linear temporal dependencies while allowing deep learning models to learn non-linear and contextual patterns, resulting in improved forecasting accuracy and robustness. Recent advancements in this direction, including the integration of recurrent and attention-based neural networks [14], have demonstrated remarkable capability in capturing intricate temporal dependencies and behavioral dynamics within energy usage patterns, marking a significant step forward in residential electricity forecasting research.

A hybrid XGBoost–ANN ensemble model has been proposed to couple short-term load forecasting with

demand response mechanisms in residential settings, enabling utilities to better manage consumption during peak hours while maintaining user comfort and operational efficiency [15].

Recently, Muqtadir et al. [16] introduced a trio ensemble model integrating LightGBM, XGBoost, and CatBoost for nowcasting the next hour of residential electricity load across multiple geographic locations. Their model leverages the complementary strengths of these gradient boosting algorithms—where LightGBM enhances generalization across sites, XGBoost minimizes overfitting, and CatBoost efficiently handles categorical features. Using data from 13 residential sites in North America and Europe, their integrated ensemble achieved superior performance with an R^2 of 0.9745 and a low RMSLE of 0.1898, outperforming other state-of-the-art forecasting methods. Their work highlights the efficacy of hybrid boosting approaches in managing spatial and temporal heterogeneity in residential load data and demonstrates the potential of ensemble learning for scalable and decentralized energy demand forecasting. Luo et al. [17] developed a stacking ensemble model combining CNN, BiLSTM, Attention mechanism, and XGBoost for short-term electricity load forecasting. Their approach first applied an adaptive hierarchical clustering (AHC) algorithm to group similar load patterns, then used CNN-BiLSTM-Attention for feature extraction and XGBoost for nonlinear regression. This hybrid structure effectively captured temporal dependencies and nonlinear variations in complex load data. Experiments on real datasets showed that their method significantly improved prediction accuracy, reducing MAPE by up to 69.4% compared with traditional and machine learning models.

Recent studies have explored multitasking intelligent monitoring schemes for industrial equipment, such as tower pumping units, using self-attention-based neural networks and autoencoders to enable robust fault detection and operational condition inference under variable conditions [18]. Zhang et al. [19] has applied neural network-based approaches, such as Levenberg–Marquardt backpropagation models, for real-time health estimation and lifetime extension in modular power conversion systems, enabling predictive maintenance through data-driven damage assessment and power routing strategies.

Liu et al. [20] proposed a hybrid CNN-BiLSTM-Attention model enhanced with CEEMDAN, K-means clustering, and VMD for quarter-hourly power load forecasting. Their framework decomposes complex load signals using CEEMDAN and VMD to extract multi-scale features, while K-means clustering groups intrinsic mode functions based on regularity and uncertainty. CNN layers capture spatial patterns from climatic and temporal inputs, BiLSTM layers model bidirectional temporal dependencies, and the attention mechanism emphasizes key features influencing load variation. Applied to Guangzhou's power grid data, the model achieved highly accurate seasonal forecasts with MAPE values between 1.08% and 1.67%, outperforming baseline CNN-LSTM and statistical approaches. The study highlights that integrating advanced signal decomposition and clustering

with deep attention-based architectures can effectively address nonlinearities and improve load prediction robustness across varying time scales. Wu et al. [21] proposed a CNN-BiLSTM-SA neural network to improve the prediction of residential electricity consumption. Their model integrates CNN for spatial feature extraction, BiLSTM for capturing long-term temporal dependencies, and a Self-Attention (SA) mechanism to emphasize critical consumption features. This hybrid design enhances the model's ability to learn both temporal and contextual patterns in household energy use. Experimental results demonstrated that CNN-BiLSTM-SA achieved higher accuracy than conventional deep learning models, with ablation studies confirming the effectiveness of the combined architecture.

DATASET DESCRIPTION AND FEATURE ENGINEERING

In our experiments, we employ widely used public residential electricity consumption datasets. Below is an overview of its properties and how they were prepared for our modeling pipeline:

1. UCI Individual Household Electric Power Consumption — This dataset contains one-minute resolution power usage records of a single household in France over nearly four years (December 2006 to November 2010), totaling about 2,075,259 observations of active/reactive power, voltage, current and sub-metering channels [22].

We preprocess the dataset using a unified pipeline to ensure consistency across models. The raw data are first resampled and temporally aligned by aggregating them to an hourly resolution through mean aggregation, which maintains temporal consistency across datasets while reducing computational overhead. Missing values are then handled through a two-stage imputation strategy: short gaps (less than four hours) are filled using linear interpolation, whereas longer gaps are reconstructed based on pattern-based imputations that leverages historical usage profiles, such as the same hour on adjacent days. To prevent scale dominance and enhance model convergence, all features are normalized to the $[0, 1]$ range using min-max normalization. For temporal feature extraction, a sliding window scheme is adopted, where 48 hours of past consumption data serve as the input sequence and the subsequent 8-hour period is predicted. The datasets are partitioned chronologically into training, validation, and testing subsets with a 7-day buffer between them to avoid temporal leakage. Specifically, for the UCI dataset, the timeline is divided such that data from 2006–2008 are used for training, 2009 for validation, and 2010 for testing.

To enhance the forecasting model's predictive performance, a structured feature engineering process was implemented, focusing on three major categories: temporal patterns, statistical dynamics, and behavioral attributes. These engineered features allow the model to recognize both short-term fluctuations and long-term trends in residential electricity consumption. These features are crucial in enabling the model to identify both

short-term fluctuations and long-term trends in energy consumption data.

A. Temporal Features

Temporal features capture the inherent cyclical and periodic nature of energy usage. Several time-based variables were designed to reflect daily, weekly, and seasonal consumption patterns:

- **Hour of Day:** To represent intraday variations, the hour of the day was encoded using sine and cosine transformations. This method retains the circular nature of time (e.g., the continuity between 23:00 and 00:00), allowing the model to learn smooth temporal transitions throughout the day.
- **Day of Week:** A one-hot encoding scheme was applied to distinguish each day of the week, enabling the model to differentiate typical usage behaviors on weekdays versus weekends.
- **Weekend and Holiday Indicator:** A binary flag was introduced to identify whether a particular time point falls on a weekend or a recognized public holiday. These temporal segments often exhibit unique energy usage patterns due to shifts in occupancy and routine activity.
- **Month and Season Indicators:** Categorical variables representing the calendar month and the associated meteorological season were also included. These indicators help the model account for broader seasonal trends, such as increased energy demand for heating in winter or for cooling during summer months.

Collectively, these temporal features enable the model to learn structured periodicity and adapt to behavioral variability across different time scales.

B. Statistical Features

In addition to temporal indicators, a comprehensive set of statistical and consumer behavior features was engineered to enhance the model's ability to capture intricate patterns in energy usage. To reflect the internal dynamics and historical dependencies within the consumption data, the following statistical features were computed:

- **Rolling Mean and Standard Deviation:** These were calculated using sliding windows of 1, 6, and 24 hours to capture short-term and long-term fluctuations in energy usage. Such statistics help the model learn trends and anomalies across varying time scales.
- **Lag Features:** Lagged values corresponding to the same hour of the previous day and the same hour of the previous week were introduced to capture recurring patterns and seasonality in user behavior.

Rate of Change: First-order differences were used to quantify the rate of consumption change over time. These derivatives highlight sharp transitions and dynamic shifts in energy usage that may correspond to behavioral or environmental changes.

C. TPB-Guided Behavioral Feature

Beyond standard temporal and statistical metrics, a

suite of behavior-oriented features is introduced to reflect user-specific patterns and anomalies: Together, these statistical and behavioral features enrich the model's contextual understanding of energy consumption, enabling more robust and interpretable forecasting performance.

In residential electricity consumption forecasting, behavioral variability often introduces a major source of prediction uncertainty. Unlike aggregate-level load profiles, individual household consumption is influenced by attitudinal, social, and behavioral control factors, such as lifestyle routines, social synchronization, and personal flexibility in energy use. To explicitly model these dynamics, this study adopts the Theory of Planned Behavior (TPB) as a guiding framework for constructing interpretable, data-driven behavioral features that quantify user behavior patterns and decision tendencies in energy consumption.

According to TPB, behavior is shaped by three principal components: Attitude — the individual's internal disposition toward energy usage, represented here by patterns of activity initiation, termination, and stability. Social Norm — the degree of conformity to collective consumption behaviors or grid-level load trends; and Perceived Behavioral Control — the household's capacity to regulate consumption through controllable or baseline loads.

By embedding these behavioral dimensions into the forecasting model, we aim to capture latent behavioral signatures that conventional statistical or temporal features fail to represent.

The behavioral feature set is defined across the three TPB dimensions, as summarized in Table 1. Each feature was derived through data-driven computation using household-level time series from the preprocessed datasets.

(1) Attitude-related Features

Attitude-related Features are defined as follows:

Activity Onset/Offset detection identifies significant increases or decreases in consumption corresponding to major daily activities (e.g., start/end of occupancy, appliance operation). This is computed by detecting points where the first derivative of power consumption exceeds twice the standard deviation of local variations:

$$\text{Activity Onset/Offset} = \frac{\Delta C_t}{\sigma_c}, \quad \text{where } |\Delta C_t| > 2\sigma_c \quad (1)$$

Here, C_t represents the consumption at time interval t . Consumption Stability Index (CSI) quantifies the temporal consistency of user behavior by measuring the ratio of variability to average load over 8-hour windows:

$$\text{Consumption Stability Index (CSI)} = 1 - \frac{\sigma_{c,8h}}{\mu_{c,8h}} \quad (2)$$

Higher CSI values indicate stable consumption patterns, while lower values reflect erratic or behavior-driven fluctuations.

TABLE 1
TPB-Based Behavioral Feature Definitions and Impacts

TPB Component	Feature	Definition	Calculation Method	Impact on Forecasting
Attitude	Activity Onset/Offset	Points of significant consumption change	Threshold-based detection of first derivative exceeding 2σ	Improves prediction during transition periods (morning/evening)
Attitude	Consumption Stability Index	Consistency of usage across similar periods	$1 - \frac{\sigma_{\text{Similar}}}{\mu_{\text{Similar}}}$	Enhances accuracy for users with stable routines
Social Norm	Peak Coincidence Factor	Alignment with regional or community peak demand	Pearson correlation between household and system load profiles	Improves peak demand forecasting accuracy
Social Norm	Family Activity Cluster	Clustered behavioral patterns over evenings, weekends, and peak hours	K-means ($C_{\text{evening}}, C_{\text{weekend}}, C_{\text{peak hours}}$)	Captures intra-family synchronization and shared routines
Behavioral Control	Base Load Ratio	Ratio of consistent background load to mean consumption	$\frac{C_{5\text{th percentile}}}{\mu_{24h}}$	Improves minimum-load prediction
Behavioral Control	Time Preference Index	Weighted preference for time periods based on individual routines	$\frac{\sum_t (C_t \times W_{\text{preference},t})}{\sum_t C_t}$	Enhances adaptability to personalized usage rhythms
Behavioral Control	Behavioral Anomaly Flag	Detection of irregular usage behaviors	Isolation Forest on historical time windows	Identifies anomalous or event-driven consumption

(2) Social Norm Features

Social norm features are defined as follows. Peak Coincidence Factor (PCF) measures the degree of synchronization between individual consumption and regional or system-wide peaks, computed via the correlation of normalized daily profiles. Here, D represents all hourly intervals of a day, and P denotes the subset of predefined peak-demand hours.

$$\text{Peak Coincidence Factor (PCF)} = \frac{\sum_{t \in P} C_t}{\sum_{t \in D} C_t} \quad (3)$$

Family Activity Pattern Clusters are derived using unsupervised K-means clustering on evening and weekend usage profiles, reflecting shared behavioral routines across similar households. The newly introduced behavioral features aim to capture the heterogeneity of household electricity consumption patterns beyond simple statistical averages. These indicators are designed to reflect both collective family activity patterns and individual time-of-use preferences, providing a more behaviorally grounded representation of consumer electricity usage. Family Activity Cluster is defined as,

$$\text{Family Activity Cluster} = \text{K-means}(C_{\text{evening}}, C_{\text{weekend}}, C_{\text{peak hours}}) \quad (4)$$

In this formulation, C_{evening} , C_{weekend} , and $C_{\text{peak hours}}$ denote the average household electricity consumption during evening hours, weekends, and peak demand

periods, respectively. These values form a three-dimensional feature space representing distinct temporal aspects of family activity. The K-means algorithm is then applied to cluster households with similar daily and weekly consumption rhythms, revealing latent lifestyle groups such as “evening-active families,” “weekend-intensive users,” or “peak-hour-dependent households.” The resulting cluster index serves as a categorical feature that captures household-level behavioral archetypes.

(3) Behavioral Control Features

Time Preference Index is defined as,
Time Preference Index (TPI)

$$= \frac{\sum_t (C_t \times W_{\text{preference},t})}{\sum_t C_t} \quad (5)$$

Here, C_t represents the consumption at time interval t , and $W_{\text{preference},t}$ is a weighting coefficient that reflects the household’s behavioral preference or habitual significance of that time slot (e.g., higher weight during evening cooking hours). This index quantifies the degree to which a household concentrates its energy usage within preferred periods, yielding a normalized metric between 0 and 1. A higher value indicates stronger time-specific usage regularity, while a lower value reflects more dispersed or irregular consumption behavior.

Together, these behavioral indicators operationalize the

Theory of Planned Behavior (TPB) framework by linking observed electricity usage to underlying psychological constructs—such as attitudes toward energy efficiency, social norms related to household routines, and perceived control over energy consumption. The Family Activity Cluster captures collective routine structures, while the Time Preference Index encapsulates individual temporal habits. These features enable the forecasting model to better distinguish between structural and behavioral variability in consumption, thereby enhancing both interpretability and predictive accuracy.

Base Load Ratio (BLR) measures the proportion of consistent background consumption (e.g., standby loads, refrigerators) to total energy use:

$$\text{Base Load Ratio (BLR)} = \frac{\min(C_{24h})}{\mu_{C,24h}} \quad (6)$$

Behavioral Anomaly Flag (BAF) identifies deviations from normal routines using probabilistic thresholds on the deviation from expected consumption:

$$\text{Behavioral Anomaly Flag (BAF)} = \begin{cases} 1, & \text{if } |C_t - \mathbb{E}[C_t]| > 3\sigma_C \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

This flag acts as a binary signal for rare behavioral or equipment anomalies. where C_t represents consumption at time t , σ and μ denote standard deviation and mean respectively, $\mathbb{E}[C_t]$ expected value of consumption at time t , P is the set of peak hours, and D is the complete day set.

PROPOSED METHOD

The proposed hybrid forecasting framework combines CNNs, BiLSTM networks, and an attention mechanism, augmented with temporal, statistical and behavioral feature fusion. This integrated design allows the model to simultaneously extract localized temporal-spatial dependencies, capture long-term sequential dynamics, and interpret patterns underlying residential energy consumption.

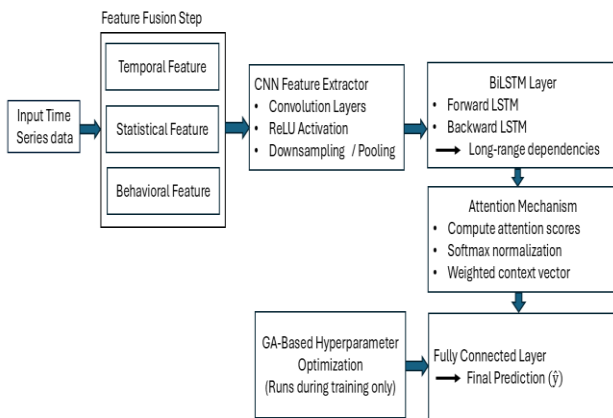


Figure 2. Block diagram of the proposed hybrid forecasting framework

Figure 2 illustrates the block diagram of the proposed method. To effectively model the multi-scale

dependencies within electricity consumption time series, the proposed model begins with a CNN-based feature extractor. The convolutional layers apply a set of learnable filters $W = \{w_1, w_2, \dots, w_k\}$ to the input feature map X , generating multiple feature maps that represent short-term correlations and recurring consumption patterns. The convolution operation is defined as:

$$f_{(i,j)}^{(l)} = \sigma \left(\sum_{m,n} W_{(m,n)}^{(l)} \cdot X_{(i+m,j+n)}^{(l-1)} + b^{(l)} \right) \quad (8)$$

where $f_{(i,j)}^{(l)}$ is the activation at position (i,j) in the l -th layer, $b^{(l)}$ is the bias term, and $\sigma(\cdot)$ denotes the nonlinear activation function, implemented as ReLU in this work. Through successive convolution and pooling operations, the CNN captures short-term transitions, periodicity, and seasonal variations in electricity consumption. These compact and noise-robust representations of raw data are subsequently used for higher-level temporal modeling.

While CNNs efficiently detect local structures, they lack the ability to represent long-range temporal correlations. To address this limitation, the CNN-extracted features are passed into a BiLSTM layer, which enables the network to process the input sequence in both forward and backward directions. For each time step t , the LSTM computes:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (9)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (10)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (11)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (12)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (13)$$

$$h_t = o_t \odot \tanh(C_t) \quad (14)$$

where f_t , i_t , and o_t are forget, input, and output gates, respectively; C_t denotes the cell state; and h_t represents the hidden output. The bidirectional configuration produces a comprehensive representation:

$$h_t^{Bi} = [h_t^{\rightarrow} \parallel h_t^{\leftarrow}] \quad (15)$$

enabling the model to learn both past and future dependencies in energy usage patterns — a critical capability for accurately forecasting consumer demand influenced by both historical habits and anticipated routines.

To further enhance the interpretability and focus of the model, an attention mechanism is incorporated atop the BiLSTM outputs. This mechanism adaptively assigns importance weights to each hidden state, allowing the model to emphasize consumption segments that contribute most significantly to prediction, such as peak demand intervals or anomalous usage events. The attention score for each hidden state h_t is computed as:

$$e_t = v^T \tanh(W_a h_t + b_a) \quad (16)$$

which is normalized using a softmax operation:

$$\alpha_t = \frac{\exp(e_t)}{\sum_{k=1}^T \exp(e_k)} \quad (17)$$

The final context vector is obtained as a weighted sum of hidden states:

$$c = \sum_{t=1}^T \alpha_t h_t \quad (18)$$

This context vector c captures the essence of the most influential temporal patterns and is passed through a dense layer to generate the final forecasting output. The overall hybrid model can be formally expressed as:

$$h_t = \text{BiLSTM}(\text{CNN}(X_t)) \quad (19)$$

$$\hat{y} = f_{\text{Dense}}\left(\sum_{t=1}^T \alpha_t h_t\right) \quad (20)$$

This hierarchical structure ensures that local variations, global dependencies, and contextual importance are all represented in a unified architecture.

D. Feature Fusion and Behavioral Integration

Beyond raw temporal signals, this model incorporates behavioral and statistical features to enhance interpretability and connect predictive outcomes with human-centric energy behaviors. At each time step t , three categories of features—temporal (x_t^{temp}), statistical (x_t^{stat}), and behavioral (x_t^{beh})—are concatenated to form a unified input:

$$X_t = [x_t^{\text{temp}}, x_t^{\text{stat}}, x_t^{\text{beh}}] \quad (21)$$

The fusion of these engineered behavioral and temporal features within the BiLSTM–Attention architecture enables the model to jointly learn both physical and behavioral dynamics. The attention layer not only identifies influential time steps but also highlights which behavioral dimension—attitude, social norm, or control—dominates each prediction. This synergy bridges cognitive interpretability and computational accuracy, allowing the model to explain what is predicted (load changes) and why it occurs (behavioral cause), thus contributing both practically and conceptually to next-generation smart grid forecasting.

E. Hyperparameter Optimization

To achieve optimal forecasting performance, a Genetic Algorithm (GA) was employed for systematic hyperparameter optimization of the proposed CNN–BiLSTM–Attention model. Given the nonlinear interactions among architectural components and learning parameters, the GA provides a robust, population-based search mechanism that efficiently explores the high-dimensional parameter space. The GA begins with an initial population of candidate hyperparameter sets, each representing a unique configuration of model parameters. During each generation, three evolutionary operations—selection, crossover, and mutation—are applied to evolve the population toward improved solutions. The fitness function is defined based on forecasting performance on the validation dataset, measured using the inverse of the root mean square error (RMSE) as:

$$\text{Fitness} = \frac{1}{\text{RMSE}} = \left[\frac{1}{\sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}} \right] \quad (22)$$

Configurations with higher forecasting accuracy (lower RMSE) are assigned greater fitness scores and thus have a higher probability of being selected for reproduction. To balance exploration and exploitation, the crossover rate was set to 0.8 and the mutation rate to 0.1. The algorithm terminates when either the maximum number of generations (50) is reached or the improvement in average population fitness becomes negligible (below 10^{-4} for five consecutive generations). The GA simultaneously tunes both architectural and learning hyperparameters, including the number of CNN filters, kernel sizes, BiLSTM units, dropout rate, learning rate, and batch size. Table 2. summarizes the optimal configuration determined through this process.

TABLE 2
Optimal hyperparameter configuration of the proposed CNN–BiLSTM–Attention model determined through Genetic Algorithm (GA) optimization.

Hyperparameter	Optimal Value	Search Range	Impact on Performance
CNN Layer 1 Filters	64	32–256	Medium
CNN Layer 2 Filters	128	64–256	Medium
CNN Kernel Size 1	3	2–5	High
CNN Kernel Size 2	5	3–8	High
BiLSTM Units	128	64–256	Very High
Dropout Rate	0.2	0.1–0.5	Medium
Learning Rate	0.001	0.0001–0.01	Very High
Batch Size	64	16–128	Medium

The optimization results highlight that larger BiLSTM units and moderate learning rates have the most significant impact on accuracy, as they enable better long-term temporal representation and stable convergence. Similarly, kernel sizes of 3 and 5 in the CNN layers provide a balanced receptive field for capturing both fine-grained

and broader temporal-spatial dependencies.

Overall, GA-driven optimization effectively refines the complex hyperparameter landscape of the hybrid CNN–BiLSTM–Attention framework, ensuring robust and generalizable forecasting performance across diverse residential load patterns.

PERFORMANCE EVALUATION AND COMPARATIVE ANALYSIS

The proposed model is trained using the Adam optimizer, with early stopping applied to prevent overfitting based on the validation loss. To comprehensively evaluate model performance, several quantitative metrics are employed, including root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and the coefficient of determination (R^2). These metrics are defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (23)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (24)$$

$$\text{MAPE} = \frac{100}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (25)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (26)$$

where y_i and $\hat{y}_i$ represent the actual and predicted values, respectively, $\bar{y}$ denotes expected mean of the actual values, and N is the total number of samples. Furthermore, comprehensive ablation studies are conducted to quantify the contribution of each major component of the proposed model—including feature extraction, temporal modeling, and attention mechanisms—to the overall predictive performance. This section presents the experimental outcomes, followed by a comparative analysis demonstrating the effectiveness and robustness of the proposed approach relative to existing models.

Table 3. presents the comparative performance of the proposed CNN–BiLSTM–Attention model against several established baseline approaches across UCI residential electricity consumption dataset. The baselines include statistical (ARIMA), machine learning (SVR, Random Forest), and deep learning models (CNN, LSTM, Transformer). The evaluation is conducted using four standardized forecasting metrics— RMSE, MAE, MAPE,

R^2 .

TABLE 3
Performance comparison of different models on UCI residential electricity consumption forecasting

Model	RMSE	MAE	MAPE (%)	R^2
ARIMA	0.152	0.128	17.5	0.721
SVR	0.124	0.098	13.2	0.802
Random Forest	0.102	0.082	11.4	0.845
LSTM	0.086	0.068	9.5	0.887
Transformer	0.086	0.067	9.2	0.892
CNN-BiLSTM-SA [21]	0.1554	0.2280	--	0.9894
Proposed method	0.078	0.059	8.1	0.912

The proposed hybrid model consistently outperforms all baseline methods across all performance metrics. It achieves the lowest RMSE (0.078) and the highest R^2 (0.912), representing a 9.3% improvement in RMSE compared to the best-performing baseline (Transformer, RMSE = 0.086). These results demonstrate the model's superior predictive precision, stability, and explanatory capability.

The enhanced performance primarily arises from the hybrid architecture's synergistic design. CNN layers effectively capture localized temporal patterns and high-frequency variations, BiLSTM units' model long-term bidirectional dependencies in consumption trends, and the attention mechanism adaptively focuses on critical temporal segments and behavioral features.

Unlike conventional methods such as ARIMA or standard LSTM, the proposed architecture not only learns nonlinear and multi-scale dependencies but also integrates behavioral features derived from the theory of planned behavior thereby improving interpretability and consumer-context sensitivity.

Table 4. presents a qualitative comparison of the proposed method with existing load forecasting approaches.

TABLE 4
Comparative Analysis of the Proposed Method and State-of-the-Art Load Forecasting Models

Paper	Complexity	Dataset Type	Key Idea	Model Architecture
CNN-BiLSTM-SA (Seasonal Forecasting) [21]	Very High	Grid-level (China, seasonal data)	Long-term temporal modeling with decomposition techniques	CNN + BiLSTM + Attention (Seasonal Forecasting)
Boosting Ensemble Model [16]	Medium	Multi-house (13 residential locations)	Ensemble boosting for improved generalization	LightGBM + XGBoost + CatBoost
CNN-BiLSTM-SA + CEEMDAN/VMD [20]	Very High	Utility grid data	Hybrid deep learning with signal decomposition	CNN + BiLSTM + Attention + Signal Decomposition (CEEMDAN, VMD)
Stacking CNN-BiLSTM-Attention + XGBoost [17]	Very High	Quanzhou dataset (2016–2018)	Stacking ensemble with clustering-based sample selection	CNN + BiLSTM + Attention + Stacking + AHC + XGBoost
Proposed Method	High	UCI residential dataset	Behavior-aware adaptive forecasting with hyperparameter optimization	Proposed Method: CNN + BiLSTM + Attention + Consumer Behavior + Genetic Algorithm

A comparative analysis of recent short-term electricity forecasting approaches indicates a progressive shift from traditional machine learning techniques toward more complex hybrid deep learning frameworks. Earlier approaches based on boosting algorithms such as LightGBM, XGBoost, and CatBoost demonstrate strong generalization performance and effective handling of structured tabular data in residential load forecasting scenarios, as reported in recent studies on ensemble-based methods.

Deep learning-based models, particularly those integrating CNN, BiLSTM, and attention mechanisms, have shown improved capability in extracting local temporal features and modeling long-range dependencies in electricity consumption data. These architectures are especially effective in capturing nonlinear and seasonal patterns in grid-level load forecasting problems. However, several studies report that their performance is often enhanced by additional preprocessing strategies such as signal decomposition or clustering-based data selection, which increases the overall modeling pipeline complexity.

Hybrid stacking-based approaches further improve forecasting accuracy by combining deep feature extractors with machine learning-based meta-learners such as XGBoost. Although these methods achieve higher predictive accuracy, they introduce additional architectural complexity due to multi-stage learning processes. Similarly, decomposition-based frameworks (e.g., CEEMDAN and VMD-based models [20]) have demonstrated strong performance improvements by separating signal components, but they also increase computational complexity and reliance on preprocessing steps.

The proposed method extends existing research by incorporating consumer behavior-related features alongside deep learning-based temporal modeling and genetic algorithm-based hyperparameter optimization. This integration improves predictive performance and

enhances model adaptability across multiple benchmark datasets. Compared to existing approaches, the proposed framework achieves improved forecasting accuracy while maintaining a balance between model complexity and interpretability.

A. Hyperparameter Optimization Comparison

To ensure robust performance of the forecasting model, we conducted a comparative study of hyperparameter optimization methods, including Genetic Algorithm (GA), Bayesian Optimization (BO), and Particle Swarm Optimization (PSO). The comparison evaluates both predictive performance and computational cost for our residential load forecasting task. Table 5. summarizes the performance of each method in terms of forecasting accuracy (RMSE). GA achieved superior results, with 5.1% improvement over BO and 3.2% improvement over PSO. This demonstrates GA's effectiveness in finding optimal hyperparameters in a multi-modal search space characteristic of residential electricity consumption patterns.

TABLE 5
Comparison of Hyperparameter Optimization Methods for Residential Load Forecasting.

Method	RMSE (kWh)	Performance Improvement	Computation Time (GPU-hours)
Genetic Algorithm (GA)	0.078 (UCI)	—	~12
Bayesian Optimization	0.082 (UCI)	-5.1%	~4
Particle Swarm (PSO)	0.081 (UCI)	-3.2%	~8

GA employs a population-based global search that is particularly effective for exploring hyperparameter spaces with multiple local optima. In the context of residential load forecasting, where consumption patterns are highly variable and nonlinear, GA's ability to avoid local minima contributes to its superior predictive performance, despite higher computational costs compared to BO and PSO. While BO is faster, GA provides better generalization and global optimum search, making it suitable for our forecasting task.

B. Ablation Studies

To assess the individual contribution of each component in the proposed Consumer Behavior-Enhanced CNN-BiLSTM-Attention model, a series of ablation experiments were conducted. In each experiment, one component was removed or replaced while keeping all other settings constant. The results are summarized Table 6.

TABLE 6
Ablation study results showing the contribution of each model component

Model Variant	RMSE	Performance Change (%)
Full Model (CNN-BiLSTM-Attention + Behavior)	0.078	—
Without Consumer Behavior Features	0.094	−20.5%
Without Attention Mechanism	0.087	−11.5%
Without CNN Component	0.089	−14.1%
With LSTM instead of BiLSTM	0.084	−7.7%
Without GA Optimization	0.082	−5.1%

The results reveal several key findings:

1. Behavior features provide the largest contribution, with a 20.5% increase in RMSE when removed. This confirms that household routines—such as activity onset/offset and consumption stability—play a central role in improving short-term residential forecasting.
2. CNN layers contribute significantly (14.1% degradation when removed) by capturing short- and mid-range temporal patterns commonly present in residential loads.
3. Attention improves performance by 11.5%, demonstrating its capability to highlight the most relevant historical time points, particularly during consumption transitions.
4. BiLSTM slightly outperforms LSTM (7.7% difference), indicating that bidirectional context helps capture consumption peaks and transitional behaviors.
5. GA-based hyperparameter tuning provides an additional 5.1% gain, demonstrating its usefulness despite the cost of evolutionary search.

These ablation results show that the performance

improvements of the proposed model are the result of multiple complementary components, and that no single module alone accounts for the overall enhancement. The synergy between CNN extraction, BiLSTM temporal modeling, attention weighting, and behavior-aware features leads to robust forecasting accuracy. To further understand why the proposed architecture performs well, attention weights were analyzed over representative 48-hour periods. The model consistently assigns higher weights to: same-hour patterns from previous days, morning and evening transitional periods, and abrupt changes in consumption routine. Permutation importance analysis shows that several behavior-aware features strongly influence attention distributions. Table 7. summarizes these correlations.

TABLE 7. Correlation between consumer behavior features and attention weights in the CNN-BiLSTM-Attention model.

Consumer Behavior Feature	Correlation with Attention Weight	Significance
Activity Onset/Offset	0.82	Very High
Consumption Stability Index	0.73	High
Base Load Ratio	0.58	Moderate
Peak Coincidence Factor	0.64	Moderate-High
Behavioral Anomaly Flag	0.47	Moderate

These results validate that the attention mechanism not only improves accuracy but also enhances interpretability by focusing on behavior-dependent temporal patterns. To enable real-time operation on resource-constrained hardware, the final trained model was quantized using TensorFlow Lite. Converting 32-bit floating-point weights to 8-bit integers reduced inference latency by 65% with only a 3.5% decrease in predictive accuracy. This demonstrates that the model's design remains suitable for deployment in smart meters and home IoT devices.

F. Limitations and Future Work

Despite strong overall performance, the model has several limitations: It requires approximately two weeks of historical data for optimal performance (cold-start issue). Accuracy temporarily decreases during unusual patterns such as holidays, long absences, or extreme weather events. Fine-grained behavioral features raise privacy considerations in real-world smart meter applications. Future work will explore transfer learning for cold-start adaptation, anomaly-resilient modeling via autoencoders or changepoint detection, federated learning for privacy-preserving deployment, and integration of multimodal data sources such as occupancy sensors and appliance-level monitoring.

CONCLUSION

This study presented a Consumer Behavior–Enhanced CNN–BiLSTM–Attention model for short-term residential electricity consumption forecasting. By integrating engineered behavioral features with a hybrid deep learning architecture, the model achieved significantly improved accuracy compared with traditional statistical, machine learning, and deep learning baselines. Ablation studies confirmed the complementary value of each component, with behavioral features, CNN extraction, and attention weighting contributing most strongly to performance.

The model also provides interpretability through attention-based temporal analysis, enabling clearer insights into the influence of household routines and consumption transitions. Pilot deployment on resource-constrained edge devices, supported by quantization, demonstrated real-time feasibility with minimal loss of accuracy.

Overall, the findings underscore the importance of behavior-aware modeling for next-generation smart energy systems and illustrate that combining deep temporal architectures with human-centric behavioral analytics yields a robust, interpretable, and edge-deployable solution for residential load forecasting. Future work will explore cold-start adaptation, anomaly-resilient modeling, and privacy-preserving learning frameworks to further enhance real-world applicability.

REFERENCE

- [1] International Energy Agency. (2024). Energy Efficiency 2024. IEA, Paris, France. [Online]. Available: <https://www.iea.org/reports/energy-efficiency-2024>
- [2] B. Guo, T. Li, H. Yu, and V. Loftness. (2025). AI-driven multi-algorithm optimization for enhanced building energy benchmarking. *Journal of Building Engineering*. [Online]. 105, p. 112351. Available: <https://doi.org/10.1016/j.jobbe.2025.112351>
- [3] H. Karimianfard. (2025). A robust optimization framework for smart home energy management: Integrating photovoltaic storage, electric vehicle charging, and demand response. *Journal of Energy Storage*. [Online]. 110, p. 115259. Available: <https://doi.org/10.1016/j.est.2024.115259>
- [4] C. Lavin and H. Julienne. (2025). Household activities underlying residential electricity demand: Who does what during the evening peak? *Energy Efficiency*. [Online]. 18(5), p. 43. Available: <https://doi.org/10.1007/s12053-025-10322-4>
- [5] X. Li, Z. Wang, C. Yang, and A. Bozkurt. (2024). An advanced framework for net electricity consumption prediction: Incorporating novel machine learning models and optimization algorithms. *Energy*. [Online]. 296, p. 131259. Available: <https://doi.org/10.1016/j.energy.2024.131259>
- [6] X. Wen, J. Liao, Q. Niu, N. Shen, and Y. Bao. (2024). Deep learning-driven hybrid model for short-term load forecasting and smart grid information management. *Scientific Reports*. [Online]. 14(1), p. 13720. Available: <https://doi.org/10.1038/s41598-024-63262-x>
- [7] H. Xu, F. Hu, X. Liang, G. Zhao, and M. Abugunmi. (2024). A framework for electricity load forecasting based on attention mechanism time series depthwise separable convolutional neural network. *Energy*. [Online]. 299, p. 131258. Available: <https://doi.org/10.1016/j.energy.2024.131258>
- [8] Y. Dai and W. Yu. (2024). Short-term power load forecasting based on Seq2Seq model integrating Bayesian optimization, temporal convolutional network and attention. *Applied Soft Computing*. [Online]. 166, p. 112248. Available: <https://doi.org/10.1016/j.asoc.2024.112248>
- [9] Y. Wang, J. Wang, G. Zhao, and Y. Dong. (2012). Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: A case study of China. *Energy Policy*. [Online]. 48, pp. 284–294. Available: <https://doi.org/10.1016/j.enpol.2012.05.026>
- [10] B. Nepal, M. Yamaha, A. Yokoe, and T. Yamaji. (2020). Electricity load forecasting using clustering and ARIMA model for energy management in buildings. *Japan Architectural Review*. [Online]. 3(1), pp. 62–76. Available: <https://doi.org/10.1002/2475-8876.12135>
- [11] A. M. Khan and M. Osińska. (2023). Comparing forecasting accuracy of selected grey and time series models based on energy consumption in Brazil and India. *Expert Systems with Applications*. [Online]. 212, p. 118840. Available: <https://doi.org/10.1016/j.eswa.2022.118840>
- [12] A. A. Pierre, S. A. Akim, A. K. Semeno, and B. Babiga. (2023). Peak electrical energy consumption prediction by ARIMA, LSTM, GRU, ARIMA-LSTM and ARIMA-GRU approaches. *Energies*. [Online]. 16(12), p. 4739. Available: <https://doi.org/10.3390/en16124739>
- [13] K. H. Baesmat, F. Shokoohi, and Z. Farrokhi. (2025). SP-RF-ARIMA: A sparse random forest and ARIMA hybrid model for electric load forecasting. *Global Energy Interconnection*. [Online]. Available: <https://doi.org/10.1016/j.gloi.2025.04.003>
- [14] S. Rani et al. (2025). Dual attention-empowered bidirectional long short-term memory network for short-term electric load forecasting. *Energy Technology*. [Online]. p. e202500914. Available: <https://doi.org/10.1002/ente.202500914>
- [15] G. R. Harikrishnan, S. Sreedharan, and others. (2025). Advanced short-term load forecasting for residential demand response: An XGBoost-ANN ensemble approach. *Electric Power Systems Research*. [Online]. 242, p. 111476. Available: <https://doi.org/10.1016/j.epsr.2025.111476>
- [16] A. Muqtadir, B. Li, Z. Ying, C. Songsong, and S. N. Kazmi. (2025). Nowcasting the next hour of residential load using boosting ensemble machines. *Scientific*

- Reports*. [Online]. 15(1), p. 7157. Available: <https://doi.org/10.1038/s41598-025-91767-6>
- [17] S. Luo, B. Wang, Q. Gao, Y. Wang, and X. Pang. (2024). Stacking integration algorithm based on CNN-BiLSTM-Attention with XGBoost for short-term electricity load forecasting. *Energy Reports*. [Online]. 12, pp. 2676–2689. Available: <https://doi.org/10.1016/j.egy.2024.08.078>
- [18] J. Zhang et al. (2025). Process monitoring for tower pumping units under variable operational conditions: From an integrated multitasking perspective. *Control Engineering Practice*. [Online]. 156, p. 106229. Available: <https://doi.org/10.1016/j.conengprac.2024.106229>
- [19] J. Zhang et al. (2023). Lifetime extension approach based on the Levenberg–Marquardt neural network and power routing of DC–DC converters. *IEEE Transactions on Power Electronics*. [Online]. 38(8), pp. 10280–10291. Available: <https://doi.org/10.1109/TPEL.2023.3275791>
- [20] X. Liu, J. Song, H. Tao, P. Wang, H. Mo, and W. Du. (2025). Quarter-hourly power load forecasting based on a hybrid CNN-BiLSTM-Attention model with CEEMDAN, K-means, and VMD. *Energies*. [Online]. 18(11), p. 2675. Available: <https://doi.org/10.3390/en18112675>
- [21] M.-P. Wu and F. Wu. (2024). Predicting residential electricity consumption using CNN-BiLSTM-SA neural networks. *IEEE Access*. [Online]. 12, pp. 71555–71565. Available: <https://doi.org/10.1109/ACCESS.2024.3400972>
- [22] G. Hebrail and A. Berard. (2006). Individual Household Electric Power Consumption. UCI Machine Learning Repository. [Online]. Available: <https://doi.org/10.24432/C58K54>
- [23] J. Kelly and W. Knottenbelt. (2015). The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. *Scientific Data*. [Online]. 2, p. 150007. Available: <https://doi.org/10.1038/sdata.2015.7>
- [24] Z. Zheng. (2025). Pecan Street Electricity Data. [Online]. Available: <https://www.pecanstreet.org/wp-content/uploads/2025/08/Pecan-Street-Data-Sheet-Aug-2025.pdf>



A Knowledge Graph and Large Language Model Approach for Mapping Associations Between Cognitive Neuroscience and Marketing *

Research Article

Ali Sarabadani¹, Kheirolah Rahsepar Fard² , Mahtab Tajvidiyan³

 [10.22067/cke.2026.97846.1187](https://doi.org/10.22067/cke.2026.97846.1187)

Abstract-- A semi-automated framework to create a structured association map between cognitive neuroscience and marketing, using expert-validated large language models (LLMs), supports a knowledge graph-based approach. The starting point for this framework is an existing cognitive neuroscience knowledge graph (CNKG) of 312 concepts as the source domain and a curated marketing concept list consisting of 100 concepts as the target domain. The mapping process consists of two distinct phases: the first phase involves automated association generation via four different LLMs (GPT-4, Gemini, Claude, & Grok) producing candidate associations between cognitive neuroscience and marketing by determining synonymous relationships with their word choices; and the second phase includes human-based expert validation in which marketing and economics experts evaluate, revise, and filter the candidate associations produced by the automated associations. This further emphasizes that human supervision is essential in maintaining the scientific integrity of the mapping process. In reviewing and analyzing the graph that was produced through this process, many core concepts within neuroscience were identified as being hub nodes (reward processing, emotion, attention, and decision-making). Additionally, there were strong relationships between the hub neuroscience concepts and many marketing constructs (including the brand).

Index Terms-- Cross-domain Knowledge Mapping, Cognitive Neuroscience, Knowledge Graph, Large Language Models, Neuromarketing

INTRODUCTION

Neuromarketing is a newly emerging field that combines marketing and Cognitive Neuroscience. Neuromarketing focuses on consumer behavior and how our brains create and process decision-making. Despite the rapid growth of this field, we still lack knowledge of the relationship

between marketing (brand loyalty, purchase intention, etc.) and the ways in which our brains process the world around us (Attention, Memory, Emotions, etc.) [1]. Researchers and practitioners have a difficult time finding consistent ways to link marketing and the way our brains function in their research. Knowledge Graphs (KGs) can provide accurate and comprehensive representations of nodes, as well as nodes' relationships, in a way that machines can read and understand [2]. They can therefore be used to integrate theoretically sound knowledge into LLMs [2]. Therefore, KGs provide a way to enhance the recommendation capabilities of LLMs by integrating accurate knowledge into their systems [2]. Flexible indexing methods for KGs provide an all-encompassing view of a specific node's neighborhood, therefore enhancing the context in which an LLM stores and processes knowledge, as well as providing for efficient retrieval of relevant knowledge via selective retrieval methods, based on popularity [2].

LLMs have shown that they can extract the meaning of relationships found in unstructured text and make potential connections through known concepts, i.e. building a semantic understanding. They are known to introduce variability (e.g., noise, speculative connections) to the result when they are used alone and therefore need oversight from an expert [2]. Expressive graph encoder approaches have been developed that allow LLMs to utilise structured data when performing work (e.g., overcoming limitations associated with longer context models); this has improved LLM performance for more complex tasks [2]. However, LLMs can still be expensive to build and run and may also become outdated or incomplete over time [2], [3].

Conversely, knowledge graph (KG) technology

* Manuscript received February 22, 2026, Revised May 31, 2026, Accepted June 06, 2026.

¹ Ph.D. in IT Engineering (E-Commerce), Department of Computer Engineering and Information Technology, University of Qom, Qom, Iran.

² Corresponding author. Assistant Professor of Department of Computer Engineering and Information Technology, University of Qom, Qom, Iran. **Email:** rahsepar@qom.ac.ir

³ Master of Science in IT Engineering (E-Commerce), Department of Computer Engineering and Information Technology, University of Qom, Qom, Iran.

provides a way to explicitly store information about principles of knowledge using entities and relationships, enabling symbolic reasoning along with better recall/access to comparatively rare knowledge. KGs typically have limitations as well; they are often domain-specific, expensive to create, and do not generalise well [2], [3].

To overcome the limitations of LLMs and KGs, Retrieval-Augmented Generation (RAG) systems have been developed, which combine the general design of an LLM with retrieval of data from external knowledge sources. This provides the RAG system with chunks of text to work from when it generates a response ([4]). By using KGs to improve the RAGs with respect to both retrieval precision and reasoning strength, the benefits of using structured knowledge to support understanding and answering of questions involving a higher amount of knowledge can be realised for RAG knowledge-intensive question-answering systems by providing modules that contain reasoning capability, KG integration/validation capability, and validation capability [4], [5]. Also, VKG mapping approaches (e.g., LLM4VKG) provide automation to ontology creation and matching concepts; these approaches have been shown to have advantages over traditional ontology programming/creation methods, while also providing greater robustness in the face of incomplete or very complex ontologies [6]. Finally, the use of Industrial Knowledge Graphs (IKGs) has been shown to improve the semantic organisation of data stored in them, as well as aiding in decision and/or process support within the area of smart manufacturing.

To handle hallucinated output produced by LLM, measures such as BERTScore and BARTScore assess semantic similarity but do not tend to find small differences between the answer spans [8]. Initiatives such as MuShroom-2025 provide more precise identification of hallucinations by identifying exact portions of the hallucinated output thereby enabling finer detection of hallucinations. Integrating KGs with LLM to improve the detection and mitigation of hallucinated outputs will produce more reliable outputs [8]. More recent research shows how Knowledge Graphs are effective at reducing the frequency of hallucinated outputs produced by Large Language Models (LLMs). They include a neurosymbolic framework developed [9], which compares two knowledge graphs: one derived from LLM-generated outputs and another derived from existing reference knowledge bases (or domain knowledge) in order to identify and explain hallucinatory content. Haskins and Adams' results indicate that using graph-based validation methods improves both the detection of factual inconsistencies and the transparency of generated responses. Another study [10] proposed a hybrid Graph Retrieval Augmented Generation (GraphRAG) framework that combines semantic retrieval methods with knowledge graph (KG) reasoning to reduce hallucination rates while improving the reliability of answers and their factual correctness. Both of these studies provide further evidence that hybrid KG-LLM architecture. Security is also an important issue to consider when deploying LLM in an enterprise AI setting. To ensure security, Algorithmic Red Teaming has been used

to assess LLM behaviour, with key emphasis on requiring human oversight as well as using hybrid approaches to assessing LLM [11].

In addition, traditional machine learning models have shown that combining predictive algorithms with optimization strategies can improve system robustness and reduce error propagation in complex analytical tasks [12].

Neuromarketing research uses methods taken from neuroscience to provide insights into consumer behaviour and the effectiveness of advertising [13]. For example, EEG recording methods combined with CNN-based channel reconstruction have improved the quality of recording brain activity and enable brain-computer interfaces to be useful in analysing marketing research [13]. The evaluation of large numbers of LLM entails the use of statistical techniques (ANOVA, regression, cluster) to analyse population-level behaviours that are comparable to how cognitive assessments done to humans, thereby providing information about human-like development, capabilities, and training efficiency [14].

Recent developments show how large language models (LLMs) create a powerful platform for connecting marketing and cognitive neuroscience [1]. It's possible for LLMs to reason among many different types of unstructured data, extract objectives at a high level, and create a structured plan from decomposed complex tasks. Therefore, LLMs are helpful for multi-level mapping of cognitive processes and marketing constructs. Multimodal integration contributes to the creation of a richer dataset by combining text-based, numeric, and experimental datasets, thereby increasing the ability to find relationships between neural mechanisms and consumer behaviour. Strategies using structured prompting and hybrid LLM frameworks contribute to reducing variability, increasing consistency, and allowing reproducible, explainable knowledge graph creation [1]. Together, these capabilities make LLMs an important tool for connecting practical marketing applications with theoretical insights in cognitive neuroscience, especially when used with knowledge graphs (KGs) to retrieve and reason.

This research presents a semi-automated methodology which integrates a cognitive neuroscience knowledge graph along with the inferential capabilities of LLMs and expert validation to build a structured association map that connects cognitive neuroscience and marketing. There are four primary contributions of this study:

1. Introduction of a hybrid methodology of integrating knowledge graphs, LLM-based relation extraction, and expert validation.
2. Creation of a bipartite association graph that connects neural concepts to marketing concepts.
3. Providing a quantitative assessment of LLM performance with respect to the influence of expert refinement.
4. Demonstration of the usefulness of the resulting association map for neuromarketing research and practice [1].

TABLE 1
Summary of Prior Research on LLM–KG Integration and Neuromarketing Applications

Theme / Concept	Key Points	Applications / Implications	Reference	Relevance to Current Study
Large Language Models (LLMs)	Extract semantic relations; generate conceptual associations; multimodal and high-level reasoning	Mapping cognitive and behavioral patterns; neuromarketing insights; complex decision-making	[1]	Used for inferring associations between neuroscience and marketing concepts
LLM Challenges	High training cost; knowledge cutoff; hallucinations; output variability	Necessity of hybrid frameworks; expert supervision; structured prompting	[2]	Justifies expert validation and hybrid KG–LLM approach
KG-Augmented LLMs	Multi-hop reasoning; structured plans; multimodal data integration	Mapping neural mechanisms to marketing constructs; explainable, reproducible associations; knowledge-augmented retrieval	[4]	Central methodology of semi-automated association mapping framework
Knowledge Graphs (KGs) & Industrial Knowledge Graphs (IKGs)	Structured, machine-readable entities and relations; addresses semantic fragmentation	Smart manufacturing; IIoT data management; decision-making; cognitive intelligence enhancement	[7]	Foundation for structured representation of neural and marketing concepts
Hallucination Detection in LLMs	BERTScore, BARTScore; span-level detection (MuShroom-2025)	Fine-grained hallucination detection; mitigation strategies; enhanced output trustworthiness	[8]	Ensures reliability of LLM-generated associations
Security in AI & LLMs	Safety, fairness, ethicality; reproducibility challenges	Human-in-the-loop evaluation; Algorithmic Red Teaming; hybrid assessment	[11]	Highlights need for oversight in LLM-assisted frameworks
Neuromarketing	Integration of neuroscience and marketing; EEG and CNN-based EEG signal reconstruction	Understanding consumer behavior; advertising effectiveness; BCIs in marketing research	[13]	Core domain for mapping cognitive processes to marketing constructs
LLM Evaluation	Population-level analysis; statistical techniques (ANOVA, regression, clustering)	Assessing intrinsic capabilities, developmental patterns, robustness, training efficiency	[14]	Supports evaluation of LLM outputs and reasoning quality

RELATED WORK

A. Cognitive Neuroscience Knowledge Graphs

The use of knowledge graphs (KGs) has grown rapidly in the past few years as a means of representing structured relationships among brain regions, cognitive functions, and experimental results in biomedical and neuroscience research [12]. The knowledge represented by these graphs helps to generate hypotheses, integrate various disciplines, and navigate the complexity of knowledge semantically. For example, KGs specific to neuroscience represent the interactions between cognitive functions, brain regions and experimental results, allowing researchers to investigate hidden connections among them.

B. Neuromarketing and Consumer Neuroscience

Marketers have turned to studying the human brain in an effort to understand how viewers react to advertising and products. One such research area is neuromarketing, which utilizes fMRI and psychophysiological measures (e.g., galvanometer) to assess consumer response to various types of marketing stimuli [12]. There are many

published studies about how consumers feel emotionally after viewing ads and how those emotions can be linked to their decision-making processes. But there is no true, unified system or framework for relating these neurological responses to marketing construct(s).

KGs can assist in developing a structured body of interdisciplinary knowledge that will connect neural activity, behaviour and marketing.

A thorough review of the literature and a co-citation analysis suggests that there has been a tremendous growth of research conducted at the intersection of neuroscience and business in recent years; however, this growth has mostly occurred in isolated pockets of consumer marketing, management, and psychology, which have very little theoretical connection to each other [17].

More specifically, neuromarketing research has tended to focus primarily on investigating discrete stimulus/response relationships (how people perceive the advertisement) and has not given enough attention to complex, repeated, and socially embedded decision-making processes (i.e., how the person has made this decision over time) that are associated with the marketing

and organisations that exist in the world [17].

Co-citation cluster analysis also indicates that much of the current research is based upon a cognitive-oriented, psychological paradigm of thinking, where the researchers do not provide a systematic integration of the underlying physiological brain processes into the development of marketing theory, rather they consider these processes to be a ‘black box’ [17].

These findings highlight a structural gap in the literature, underscoring the need for integrative frameworks capable of continuously bridging neuroscience knowledge with marketing constructs in a coherent and scalable manner [17].

C. Large Language Models for Knowledge Extraction

LLMs (large language models), like GPT-4, Gemini, Claude and Grok, have been used for relation extraction, ontology creation and scientific text mining. LLMs are capable of producing potential relations from textual sources; however, the resulting candidate relations need to be validated due to possible hallucinations or ambiguity in the source domain. Additionally, LLMs are being used to augment datasets, improve semantic understanding and mitigate some of the imbalances found in domain-specific uses, such as reviewer assignment systems.

Because of the fragmented, and interdisciplinary nature of neuroscience based business research, automated processes supporting ongoing, as opposed to one-time, integration of knowledge provide significant benefits in establishing and preserving the consistency of concepts as the source disciplines of the knowledge continue to evolve.

D. Knowledge Graph Construction

Knowledge Graph Construction (KGC) is a process that extracts information from raw text to build a structured representation of knowledge in the form of RDF triples [16]. Early KGC approaches typically used rule-based systems, linguistic-pattern matching techniques, and classical machine learning techniques such as Hidden Markov Models (HMMs), Conditional Random Fields (CRFs), Support Vector Machines (SVMs), etc., for automatic entity and relation extraction. More recently, KGC researchers have begun developing methods utilizing deep learning, with emphasis on Transformer-based networks such as BERT, sometimes combined with Recurrent Neural Networks (RNNs) or Graph Neural Networks (GNNs), for joint entity recognition and relation extraction [16].

There are a variety of benchmark datasets for KGC that cover many application domains. These datasets include: CoNLL04 (general-purpose dataset); SciERC (domain-specific dataset); TransOMSC (domain-specific dataset); DBpedia, Wikipedia, and Wikidata-derived datasets (e.g., WebNLG, FewREL, DocRED, T-REx, TekGen, KELM) (domain-agnostic datasets) [16]. Existing benchmark datasets for knowledge graph construction (KGC) are widely used in both supervised and few-shot learning settings; however, scaling KGC models to large datasets often incurs significant computational cost and resource

requirements.

E. LLMs for Knowledge Graph Construction

LLMs have been utilized with limited or no prior training on KGC through their capabilities of using their internal parameters to create RDF triples from text. In contrast, the use of multi-turn question answering or extraction without a schema as seen in the framework of ChatIE and Extract-Define-Canonicalize, is being utilized to create entity and relationship triples. In addition, the integration of LLMs with KGs has shown promise for enhancing Knowledge Graphs (KGs) where the desired output can enhance the density of sparse graphs and thus enhance the overall quality of the graph [12].

F. Code and Small LLMs for KGC

In a recent study, it was shown that Code LLMs are more effective than using natural language prompts for triple extraction because of how they represent triples as class instances in Python as well as employ reasoning techniques (e.g., chain of thought (CoT)) to solve the problem of extracting the triples [12]. Additionally, researchers have been using parameter-efficient ways to fine-tune smaller LLMs (e.g., GPT-2, Falcon 7B, LLaMA 13B) on KGC augmented datasets using methods like LoRA. Therefore, although little computational resources were used in producing the results presented here, the previous literature used mostly to address the problem of performing zero shot/few shot inference or finding new data to augment their training datasets was done using much larger datasets [16].

G. Reviewer Assignment and KGs/LLMs

The reviewer assignment problem - matching papers with their reviewers and ensuring that no conflicts exist is a well-studied area of research that has recently gained traction from the use of KGs and LLMs [12]. Reviewer assignments were conventionally done using either manual techniques or keyword-based methods; both approaches have limitations based on the amount of semantic understanding available and their scalability. Recent attempts have employed various techniques, including topic modelling techniques such as Latent Dirichlet Allocation (LDA), Latent Semantic Indexing (LSI) and neural models such as BERT, CNN, and biLSTM, to improve the degree of alignment between reviewer expertise and accuracy of recommendations made for reviewer assignments. LLMs have provided a new approach to enhance the methods used for data augmentation, semantic embedding generation and few-shot knowledge extraction in order to improve the matching of reviewers to manuscripts [12].

Graph-based methods can be applied to improve fairness, as well as detect potential conflicts-of-interest by utilising co-authorships, institutional affiliations, and distance of collaboration. For example, PeerReview4All and Reviewer Round Robin algorithms provide methods to optimise both coverage and workload balancing, and when combining LLMs with KGs, they provide a means of enriching knowledge required for assignment tasks and provide a greater degree of semantic understanding than is

currently available [12].

The combination of KGs and LLMs provides complementary strengths to the assignment of reviewers to manuscripts – KGs contain structured relational knowledge and LLMs contain unstructured semantic knowledge and enable low resource or zero-shot inference. This complementary nature of both viewpoints has led to work on KG construction, domain adaptation, and reviewer assignments recently [12], [16].

METHODOLOGY

A. Framework Overview

This research introduces a semi-automated method for discovering connections between cognitive neuroscience and marketing by combining a current CNKG, an enormous number of LLMs, and evaluation by human experts. The procedure includes two sequential steps:

The use of LLMs to generate automated associations (candidate connections) between neuroscience and marketing ideas. Expert evaluation and modification of the generated associations for the purpose of scientific validity. Inputs:

CNKG: A curated knowledge graph representing cognitive neuroscience concepts and their interrelations [19], [18].

Marketing concepts: A curated list of 100 key marketing concepts from textbooks, expert consultation, and neuromarketing literature.

Outputs:

A validated bipartite knowledge graph encoding cross domain relationships between neuroscience and marketing concepts, enabling systematic knowledge transfer from a specialized scientific domain to an applied domain.

B. Source and Target Concept Preparation

(1) Cognitive Neuroscience Concepts (Source Domain)

The extracted concepts were cognitive processes/functions; neural processes/mechanisms; attention; reward processing; memory; emotion; processing of decisions; executive function/control [17,18] (from the original source) along with semantic consistency between all sources, standardized identifiers for all sources, and less ambiguity.

(2) Marketing Concepts (Target Domain)

A set of 100 marketing concepts was compiled based on:

- Standard marketing and consumer behavior textbooks.
- Expert consultation in marketing and economics.
- Frequency and relevance in neuromarketing literature.

Examples include brand perception, purchase intention, customer engagement, advertising effectiveness, price sensitivity, and consumer trust.

```

You are an expert in both cognitive neuroscience and
marketing.

Consider the following neuroscience concept:
Reward processing

Consider the following marketing concept:
Purchase intention

Identify whether there is a meaningful relationship
between these two concepts.

If yes:
- Briefly describe the relationship.
- Provide a clear scientific justification based on
cognitive neuroscience or marketing theory.

If no:
State exactly: "No relationship."

```

Figure 1. Prompt template for LLM-based knowledge extraction process.

C. Phase 1: Automated Association Generation Using LLMs

Four state-of-the-art LLMs—GPT-4, Gemini, Claude, and Grok—were used as inference engines [17], [18].

(1) Prompt Design

Each neuroscience–marketing concept pair was prompted using a structured template, instructing the model to:

- Identify meaningful relationships.
- Provide a concise scientific justification.
- Structured prompting mitigates hallucinations and encourages concept-level reasoning [16].

(2) Association Extraction Process

For each neuroscience–marketing concept pair, the LLMs generated candidate associations in the form:

⟨Neuroscience Concept, Marketing Concept, Relationship Description⟩

The outputs from all four models were collected into a unified candidate relation pool.

(3) Normalization and Aggregation

Candidate associations were:

- Normalized: synonymous terms unified.
- Deduplicated: repeated associations removed.
- Filtered: trivial or syntactically invalid outputs removed.

This produced a refined set of candidate associations for expert evaluation.

D. Phase 2: Human Expert Validation and Refinement

During this stage, a total of two validators (experts) validated the associations in terms of cross-domain relationships between cognitive neuroscience and marketing. The validators included marketing scholars and data science practitioners with academic and/or professional experience related to the validation of associations. In addition, guidance and consultation were also obtained from cognitive neuroscience experts to support the evaluation process. The validators assessed each association independently of each other based on predetermined evaluation criteria.

(1) Evaluation Criteria

Experts assessed associations based on:

- Scientific plausibility – alignment with neuroscience

and marketing theories.

- Conceptual clarity – interpretability and well-defined relations.
- Practical relevance – applicability in marketing contexts [15].

(2) Labeling Strategy

- Each association was labeled as:
- Accepted: valid and meaningful.
- Modified: valid but requiring refinement.
- Rejected: unsupported or vague.

E. Knowledge Graph Construction and Analysis

The validated associations were encoded as a bipartite knowledge graph:

- One node set: neuroscience concepts.
- Second node set: marketing concepts.
- Edges: validated cross-domain relationships.

Edge annotations:

- Relationship type
- LLM source(s)
- Expert validation label

Graph-theoretic metrics:

Graph density:

$$\text{Density} = \frac{|E|}{|V| \times (|V| - 1)} \quad [25]$$

These metrics quantify the connectivity and structural properties of the resulting network.

F. Reliability and Quality Control

Inter-model Agreement Rate (IAR):

$$\text{IAR} = \frac{\text{Number of agreements between models}}{\text{Total number of assignments}} \times 100$$

Indicates the consistency of LLMs in identifying meaningful conceptual associations [15].

Expert modification/rejection rate: quantifies LLM noise and validates the necessity of human-in-the-loop curation.

G. Reproducibility

All prompts, concept lists, and evaluation criteria are fully documented, ensuring reproducibility. The framework is domain-independent and can be adapted for other interdisciplinary knowledge mappings.

Pseudocode Summary

```
# 1. Automated association generation
candidates = []
for nc in CNKG:
    for mc in Marketing:
        for model in LLMs:
            rel =
model.predict_relation(nc, mc)
            candidates.append(rel)
candidates = deduplicate(candidates)

# 2. Expert validation
validated = []
for rel in candidates:
    label = expert_evaluate(rel) #
Accept, Modify, Reject
```

```
if label != "Reject":
    validated.append(refine(rel,
label))

# 3. Graph construction
KG = create_bipartite_graph(validated)
# 4. Evaluation
metrics = compute_metrics(candidates,
validated, KG)
Output: KG, metrics
```

EXPERIMENTAL SETUP

A. Data Sources

The two main datasets were chosen to study the ability of longitudinal learning mechanics to create relationships through automated knowledge extraction between knowledge domains that use different ontological structures. Accordingly, the data sources are:

1. The Cognitive Neuroscience Knowledge Graph (CNKG): This is an existing Cognition related knowledge graph which has been developed using the peer reviewed cognitive neuroscience literature. The CNKG consists of 312 neuroscience concepts and related to Cognition, emotion, perception and decision-making. The CNKG includes each of the neuroscience entities (e.g. concepts) linking together within the broad scientific corpus of knowledge and serves as the source domain for cross-domain mapping - [19].

2. Marketing Concept Set: The Marketing Concept Set consists of a collection of 100 available marketing concepts that have been compiled from the body of established marketing and consumer behaviour textbooks and refined by experts. This dataset represents the target domain and consists of fundamental constructs including consumer segmentation, brand equity, purchase intention, and advertising effectiveness [20], [21], [22], [23].

The two datasets above were selected because it was thought that an examination of how LLMs have the ability to derive meaningful analogy and/or relational-based relationships between the two domains (knowledge domains) will identify patterns that will help suggest that cross-domain knowledge association extraction can be achieved.

B. Knowledge Extraction Workflow

The experimental process was completed in the following steps:

1. Candidate Relation Generation : Multiple large language models (LLMs) were used to generate candidate cross-domain relationships between cognitive neuroscience and marketing concepts. This was done using structured prompts constructed from CNKG and the Marketing Concept Set, following recent approaches in LLM-based knowledge graph construction and relation extraction [24], [25].

2. Relation Filtering and Normalization : Candidate relations were consolidated into a standardized set, where duplicate and semantically equivalent relations were removed and unified to ensure consistency and uniqueness. This follows established preprocessing and normalization practices in knowledge graph construction and LLM-based information extraction pipelines [25].

3. Expert Review : Cognitive neuroscience and marketing experts independently evaluated each candidate relation based on semantic plausibility, conceptual clarity, and relevance. Each relation was categorized as Accepted, Modified, or Rejected, consistent with human-in-the-loop validation frameworks for LLM-generated knowledge structures [25].

4. Graph Integration : The validated relations were integrated into a bipartite knowledge graph for further structural and topological analysis of cross-domain relationships. This aligns with recent methodologies in LLM-augmented knowledge graph construction and graph-based reasoning systems [24], [26].

C. Evaluation Metrics

The following performance metrics were determined to quantitatively analyze how well LLMs are performing and the impact of expert validation:

- Inter Model Agreement Rate (IAR): Percentage of relations that at least two different LLMs suggested that showed consensus,
 - Expert Acceptance Rate (EAR): Percentage of LLM suggested relations accepted without any modification by experts,
 - Modification Rate (MR): Percentage of the relations generated by LLMs that were modified by experts,
 - Rejection Rate (RR): Percentage of the relations generated by LLMs that were rejected by experts.
- In addition to these human-centric performance metrics, other graph structure metrics were calculated to assess the effect of integration on the expanded knowledge graph which included:
- The number of nodes and edges (which show basic size expansion of the graph)
 - The mean number of connections per entity (the average node degree),
 - The ratio of actual edges to all possible edges in the graph (the density of the graph).

These structural indicators are well established metrics for determining connectivity and structural robustness within a KG as well as how well new relations integrate into the topology of the graph.

D. Robust Experimental Design

The validity and reproducibility of results in the experimental design were addressed through the use of several controls:

- Use of multiple LLM Backbones: The use of a minimum of two different LLM architectures to lessen the chance of bias due to the eccentricities of any one model.
- Cross Validation: A set of proposed relations for a small set of CNKG entities were withheld from the construction of the overall dataset and later evaluated independently in order to measure LLM's generalizability.
- Blind Review Process: Expert reviewers evaluated proposed relations in random order and without knowledge of the specific model used to propose them. This design permitted systematic measurement of automated relation generation capability, validity of human expert validation, and relationship structure and graphical evolution/changes as a function of LLM incorporation of derived associations across domains.

RESULTS

A. Overall Knowledge Graph Statistics

The automated LLM phase generated a total of **1,284 candidate associations** between neuroscience and marketing concepts.

After expert validation, the final bipartite knowledge graph contained:

- **312 neuroscience nodes (from CNKG)**
- **100 marketing nodes**

742 validated cross-domain relationships (edges)

The detailed statistics of the final knowledge graph are summarized in the table below:

TABLE 2
Presents The Counts Of Nodes And Validated Relationships In The Knowledge Graph

Feature	Count
Candidate associations generated by LLMs	1,284
Neuroscience nodes (from CNKG)	312
Marketing nodes	100
Validated cross-domain relationships (edges)	742

The corresponding chart provides a visual summary of the knowledge graph statistics:

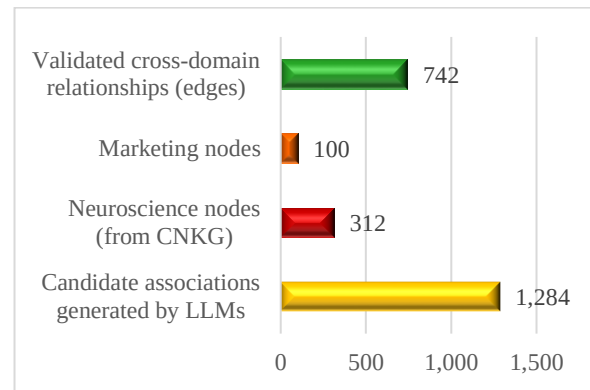


Figure 2. illustrates the data presented in Table 1 in graphical form.

(1) Graph Metrics:

To quantify the structural properties of the constructed knowledge graph, two standard graph metrics were computed: graph density and average node degree.

(2) Graph Density:

Graph density measures the proportion of actual edges present relative to the maximum possible number of edges in the graph. For a directed graph, it is defined as:

$$\text{Density} = \frac{|E|}{|V| \times (|V| - 1)} \quad [24]$$

where $|E|$ is the number of edges and $|V|$ is the

number of nodes.

In our case:

$$\text{Density} = \frac{742}{100 \times 312} = 0.0238$$

This relatively low density indicates that the graph is sparsely connected, which is typical for knowledge graphs spanning multiple domains.

The average node degree was **3.57**, indicating moderate connectivity between the two domains.

(3) Average Node Degree:

The average node degree provides the mean number of edges connected to each node:

$$\text{Average Degree} = \frac{2 \times |E|}{|V|}$$

For our graph, this yields:

$$\text{Average Degree} = 3.57$$

indicating moderate connectivity between the two domains.

According to these findings, the structural connections between the various nodes indicate moderate levels of connectivity and characterize the features of the built knowledge graph, and while there are primarily sparse connections for most of the nodes, there are enough connections to support typical tasks at a later stage.

B. LLM Performance Analysis

Among the 1,284 candidate relations:

- **684 relations (53.3%)** were proposed independently by at least two LLMs.
- **241 relations (18.8%)** were suggested by all four models.
- **197 relations (15.3%)** were suggested by three models.
- **246 relations (19.2%)** were suggested by two models.
- **600 relations (46.7%)** were unique to a single model.

The detailed distribution of LLM-suggested relations is summarized in the table below:

TABLE 3
Presents The Breakdown Of Relations Proposed By
Different Numbers Of LLMs.

Number of LLMs Suggesting Relation	Number of Relations	Percentage
All four models	241	18.8%
Three models	197	15.3%
Two models	246	19.2%
At least two models (total)	684	53.3%
Unique to a single model	600	46.7%
Total	1,284	100%

The corresponding chart visualizes the distribution of relations proposed by the LLMs:

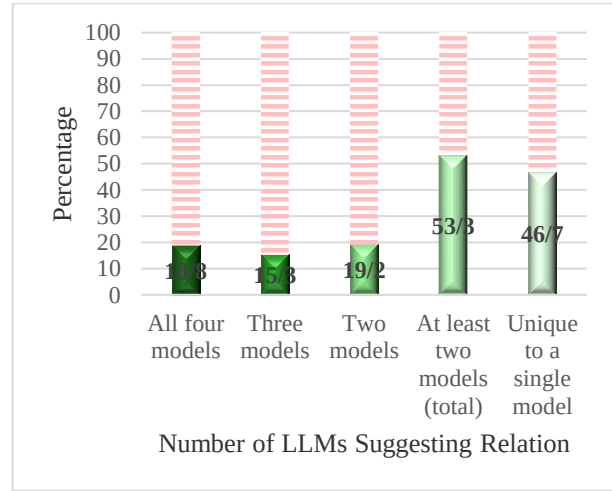
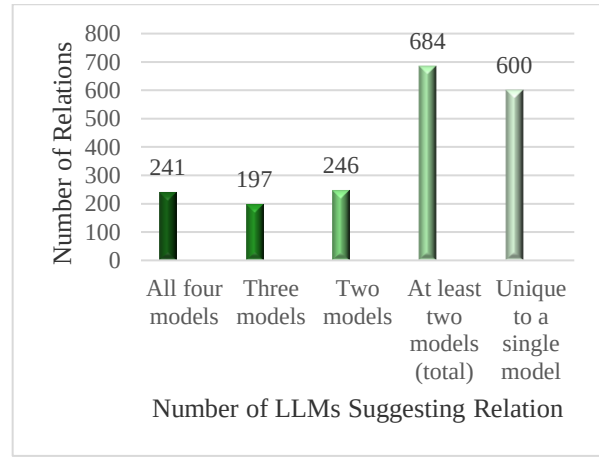


Figure 3. provides a graphical representation of the data summarized in TABLE 3

Inter-Model Agreement Rate (IAR): To quantify the agreement between multiple models in assigning reviewers, we define the Inter-Model Agreement Rate (IAR) as the percentage of reviewer assignments for which all considered models agree.

Formally:

$$\text{IAR} = \frac{\text{Number of agreements between models}}{\text{Total number of assignments}} \times 100$$

Where:

- *Number of agreements between models* is the count of assignments for which all models recommend the same reviewer for a manuscript.
- *Total number of assignments* is the total number of manuscripts assigned across all models.

For our experimental setup, the IAR was calculated as:

$$\text{IAR} = 53.3\%$$

This suggests that 53.3% of all manuscript project submissions were consistent across all evaluated models.

This indicates areas of both agreement and disagreement among these evaluated models.

The results show a moderate degree of consistency

when evaluating LLMs in making conceptual associations.

C. Impact of Expert Validation

Expert evaluation of the 1,284 candidate relations yielded the following outcomes:

- **512 relations (39.9%)** were accepted without modification.
- **230 relations (17.9%)** were modified to improve scientific clarity and precision.
- **542 relations (42.2%)** were rejected.

The detailed outcomes of expert validation are summarized in the table below:

TABLE 4
Presents A Breakdown Of Accepted, Modified, And Rejected Relations.

Expert Evaluation Outcome	Number of Relations	Percentage
Accepted without modification	512	39.9%
Modified for clarity/precision	230	17.9%
Rejected	542	42.2%
Total	1,284	100%

The corresponding chart visualizes the expert validation outcomes:

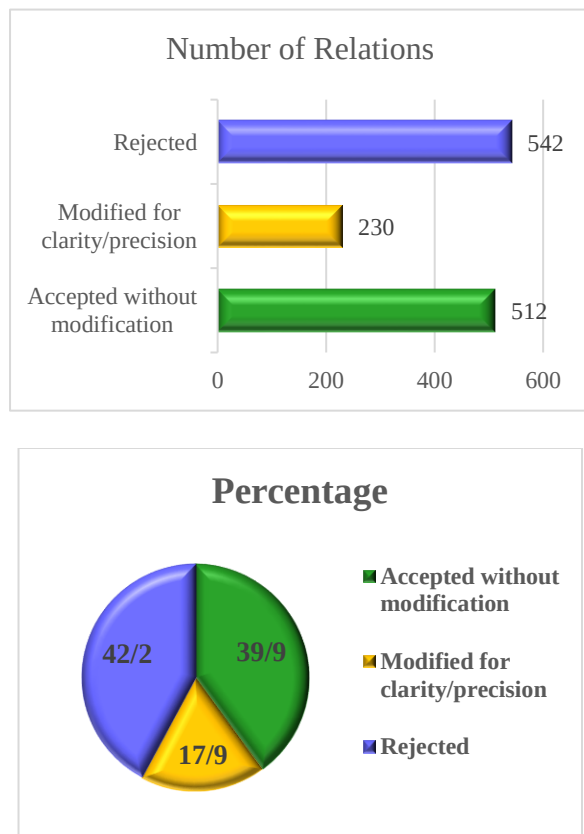


Figure 4. provides a graphical representation of the data summarized in TABLE 4.

The main causes of rejection were:

1. Conceptual vagueness (38% of rejected relations),
2. Lack of neuroscientific evidence (34%),

3. Overly generic associations (28%).

The detailed breakdown of these rejection causes is presented in the table below:

TABLE 5
Summarizes the Main Reasons for Relation Rejection

Cause of Rejection	Percentage of Rejected Relations
Conceptual vagueness	38%
Lack of neuroscientific evidence	34%
Overly generic associations	28%

The corresponding chart visualizes the distribution of rejection causes:

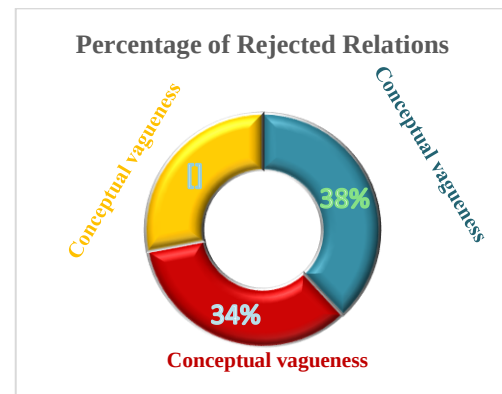


Figure 5. provides a graphical representation of the data summarized in TABLE 5.

This confirms the critical role of human supervision in refining LLM-generated knowledge.

D. Distribution of Relationship Categories

The thematic categorization of the validated relations (742 edges) resulted in the following grouped relations:

- 214 emotion-based relations, 28.8% were related to emotion and loyalty of the brand and to the recall of advertisement materials.
- The theme of reward and motivation consisted of 183 relations, 24.7% percentage and examples of reward processes related to purchase intentions.
- The theme of attention and perception contained 162 relations and comprised 21.8%. Examples include the relation of attention to advertisement effectiveness.
- The working memory and brand recognition were included in the theme of memory and comprised 121 relations, 16.3%.
- The theme of decision and executive control included 62 relations, constituting 8.4%.

The above results show that emotional and reward-based neural processes dominate marketing-related phenomena. The detailed counts for each relationship category are summarized in the table below:

TABLE 6
Presents the Full Distribution of Validated Relations Across These Thematic Categories

Relation Category	Number of Relations	Percentage of Total	Examples
Emotion-related associations	214	28.8%	Emotion → brand loyalty, emotion → advertising recall
Reward & motivation mechanisms	183	24.7%	reward processing → purchase intention
Attention & perception mechanisms	162	21.8%	Attention → advertising effectiveness
Memory-related mechanisms	121	16.3%	working memory → brand recognition
Decision-making & executive control	62	8.4%	–

The corresponding chart visualizes the distribution of these relationship categories:

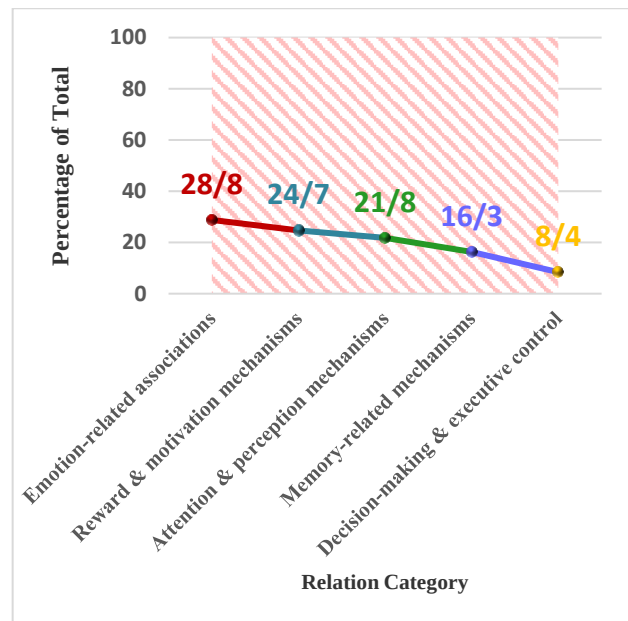
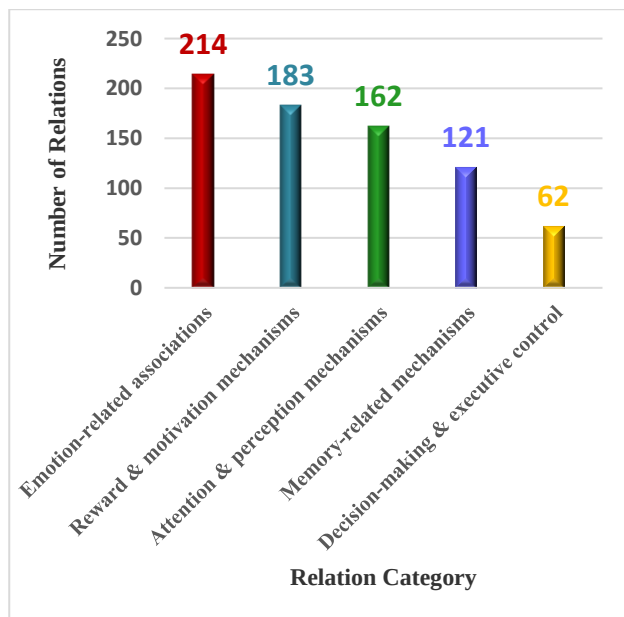


Figure 6. provides a graphical representation of the data summarized in Table 6.

E. Graph Structural Analysis

Centrality analysis revealed that a small number of neuroscience nodes acted as hubs in the bipartite graph:

TABLE 7
Lists the Centrality Values for Each Node

Neuroscience Concept	Degree
Reward Processing	41
Emotion	38
Attention	35
Decision-Making	29
Memory	26

Figure 6 provides a graphical representation of the centrality scores listed in the Table 7:

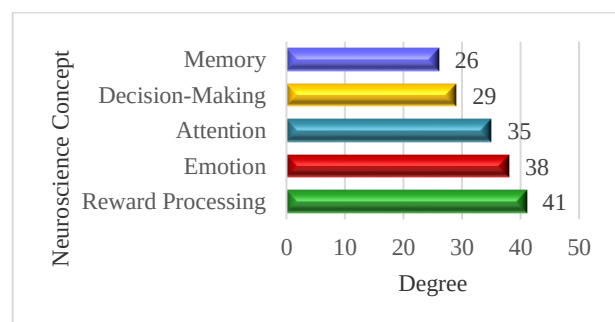


Figure 7. Graphical representation of the centrality scores for neuroscience nodes in the bipartite graph, highlighting hub nodes identified in the analysis.

These hub nodes were strongly connected to marketing concepts such as:

- brand loyalty
- advertising effectiveness
- consumer trust

- purchase intention
- customer engagement

This hub structure suggests that a limited set of neural mechanisms accounts for a large portion of marketing behavior explanations.

F. Reliability Indicators

There are two quantitative indicators that indicate that the proposed framework is valid:

1. The inter-model agreement rate of 53.3% implies that there is agreement among the various LLMs regarding a consistent subset of conceptual relationships.
2. The expert rejection and modification rate of 60.1% emphasizes the value of using experts in the analysis of data to ensure that the scientific quality of any conclusions drawn from the data is of high standard.

The detailed values for these indicators are summarized in the table below:

TABLE 8

Presents a Summary of These Quantitative Reliability Indicators

Indicator	Value	Interpretation
Inter-model agreement rate	53.3%	Shows convergence among different LLMs on a stable subset of relations
Expert rejection/modification rate	60.1%	Highlights the importance of human-in-the-loop validation to ensure scientific rigor

Together, these results confirm that combining LLM-based inference with expert validation provides a robust methodology for constructing cross-domain knowledge graphs.

DISCUSSION

According to the literature reviews and co-citation analysis, business and management have expressed interest in using neuroscience for future research; whereas less is known about its use in B2B marketing. The exploration of this area through the combined lens of neuroscience and B2B marketing yielded only a few examples: nine out of nine existing studies focused on a one time exchange (i.e., Type II exchanges) as opposed to either decision-making processes or frequent complex exchanges with substantial value offerings within B2B. As such, the opportunity to increase knowledge through incorporating findings from neuroscience is quite large. While there is somewhat of a lack of research linking neuroscience to B2B marketing, this is not the primary reason for providing additional opportunities for the two fields; instead, the decision to incorporate findings from

neuroscience because there is clearly a need and an opportunity to do so.

The findings from the literature review and co-citation analysis in business and management with regard to the use of neuroscience indicate neuroscience could act as an effective methodology both as a second method of research for social science and natural science as well as a foundation for combining social science and natural science [17].

B2B marketing research, using neuroscience, can help to answer both 'how' and 'why' questions, such as those related to the behaviours of social beings, human-machine behaviour, decision making, and the interaction between humans and machines. Some of the areas in which neuroscience may help B2B marketers includes the digitalisation of businesses, the implementation of artificial intelligence, the growing importance of multi-agent systems (that is, systems which rely on the coordinated activities of multiple agents), and innovative ways of thinking about the future. The goal of B2B marketers using neuroscience in research is to link the cognitive or neural aspects of decision behaviours to the knowledge B2B marketers already have about how firms create value - and about their repeated interactions or exchanges with one another.

Social or organisational neuroscience will create new and improved ways for B2B marketers to think about multi-agent systems, as B2B marketing offers the inter-organisational aspect of multi-agent systems. The study of physiological brain activity offers other viable research opportunities for B2B marketers. For example, the use of physiological brain activity may be useful in explaining how managers use forecasting and prediction (the psychological process involved in dealing with future events - forecasting) differently from managers dealing with past events (the psychological process involved in dealing with past events - forecasting).

Given the growing industry awareness that non-marketplace logics can supplement the traditional logic of profit maximisation in the marketplace, the application of neuroscience to B2B marketing will extend the area of cognition to themes embedded in their recent research on values and repeat exchanges.

Multiple opportunities to employ neuroscience insights to expand B2B marketing bridges with neuroscience exist due to the vast subject matter associated with research derived from physiological brain activity to date, which go far beyond B2B marketing's current limited linking of neuroscience with discrete exchanges.

What is the appropriate method for integrating both B2B Marketing and Neuroscience based research? Cluster analysis suggests that business and management research does not have a defined level of coherence between itself and Neuroscience. This provides two options for researching the two disciplines - developing future research based on within discipline advancements (an internalized approach), or continuously integrating both disciplines (a continuous integration approach).

An internalized research approach allows researchers to

accumulate knowledge relevant to their field, however it runs the risk of being obsolete or irrelevant due to the rapid progression of scientific and neurobiological advancements.

A continuous integration research approach keeps B2B Marketing researchers engaged with Neuroscience developments while providing a level of divergence between B2B Marketing and Neuroscience based research. Therefore, an initial continuous interdisciplinary integration between B2B Marketing and Neuroscience research will be an optimal model and will potentially lead to a dual research approach in B2B Marketing that integrates both within discipline advancements along with an ongoing analysis of Neuroscience based research advancements.

APPLICATIONS

There are many possible practical and scientific applications for the proposed Association Map, including:

- **Hypothesis Generation in Neuromarketing Research** The Knowledge Graph can be used to systematically derive new research questions regarding the neural basis of consumer behaviour and marketing effectiveness.
- **Education and Training Tools for Interdisciplinary Learning** The Association Map can serve as a structured resource for teaching and learning at the intersections of Cognitive Neuroscience, Marketing and AI.
- **Decision Support Systems for Marketers** The framework can connect a marketer's marketing strategies to the underlying cognitive/emotional mechanisms used by consumers; thus, supporting practitioners in creating more effective psychologically grounded campaigns.
- **AI-Driven Recommendation and Analytics Systems** The Knowledge Graph can be used as the Knowledge Architecture for intelligent Systems to integrate insights from Consumer Neuroscience into Personalized Recommendation Engines and Marketing Analytics Platforms.

LIMITATIONS AND FUTURE WORK

Although this study's findings appear encouraging, it has many limitations. The primary limitation is that the quality of association maps and their relationship to Cognitive Neuroscience Knowledge Graph (CNKG) data are dependent on the original data available. Therefore, if the original data contains too little or has biases present, it could result in the inaccuracy of the entire cross-domain association mapping process.

Although LLMs provide strong inferences based on existing data, the associations generated by LLMs contain latent biases or oversimplifications in their training dataset. This suggests a need for validation processes by human experts to minimize errors resulting from LLM processing.

Lastly, an additional limitation of our study is that there was a limited number of expert validators; therefore, this

limitation could have an effect on generalizability of these validation findings across multiple disciplines.

In future studies, we plan to explore the following:

- Expanding our framework to include other areas such as psychology, behavioural economics and social neuroscience.
- To create automated systems to score the degree of confidence associated with created associations.
- To match the empirical data from neuroimaging studies and experiments with the knowledge graph in order to ground the graph to observed neural data.
- To create dynamic and temporal representations of the graph to model changes in consumer cognition over time.

CONCLUSION

We have developed a semi-automated framework that integrates knowledge graphs, LLMs, and human experts to create an organized association map between cognitive neuroscience and marketing concepts.

By combining a curated cognitive neuroscience knowledge graph with LLMs for inferences, along with the expertise of a human expert, we can systematically discover and confirm cross-domain associations that would be otherwise impossible to glean from the fragmented literature on these topics.

While our tests demonstrated that LLMs can generate many valuable candidate associations (53.3% inter-model agreement), the results suggest that over 60% of those associations needed to be changed or discarded by experts; this reinforces the need for human intervention throughout the process to ensure the validity and clarity of scientific concepts.

By combining a curated cognitive neuroscience knowledge graph with LLMs for inferences, along with the expertise of a human expert, we can systematically discover and confirm cross-domain associations that would be otherwise impossible to glean from the fragmented literature on these topics.

While our tests demonstrated that LLMs can generate many valuable candidate associations (53.3% inter-model agreement), the results suggest that over 60% of those associations needed to be changed or discarded by experts; this reinforces the need for human intervention throughout the process to ensure the validity and clarity of scientific concepts.

The final result was a bipartite knowledge graph indicating how various neural mechanisms (e.g., reward-processing, emotion, attention and decision-making) can be used to explain different marketing concepts (i.e., brand loyalty, purchase intention, and advertising effectiveness). The scope of the bipartite knowledge graph offers both theoretical insight into the neural basis for consumer decision-making, as well as a practical tool for advancing neuromarketing research.

This project is exploratory in that it has produced insights into how hybrid AI/human methods may provide an effective and efficient way to create interdisciplinary

knowledge bridges across scientific domains (e.g., cognitive neuroscience and marketing) and within applied fields (e.g., neuromarketing).

The hybrid AI/human approach also contributes to the theoretical development of neuromarketing while providing a basis for the future creation of intelligent decision-support systems and data-supported marketing strategies based on the science of cognitive neuroscience.

REFERENCES

- [1] F. Aghaee and H. R. Shaker. (2025). RB-LLM control: An intelligent control framework with rule-based large language model decision-making. *Aerospace Science and Technology*. [Online]. 168, p. 111259. Available: <https://doi.org/10.1016/j.ast.2025.111259>
- [2] S. Wang et al. (2025). Knowledge graph retrieval-augmented generation for LLM-based recommendation. Presented at 63rd Annual Meeting of the Association for Computational Linguistics (ACL). [Online]. Available: <https://doi.org/10.18653/v1/2025.acl-long.1317>
- [3] C. Voinea, S. Porsdam Mann, J. Savulescu, and B. D. Earp. (2025). The calculator analogy: Epistemic virtues for using LLMs. *Technology in Society*. [Online]. 85, p. 103198. Available: <https://doi.org/10.1016/j.techsoc.2025.103198>
- [4] J. Linders and J. M. Tomczak. (2025). Knowledge graph-extended retrieval augmented generation for question answering. *Applied Intelligence*. [Online]. 55(17), p. 1102. Available: <https://doi.org/10.1007/s10489-025-06885-5>
- [5] C. Ma, Y. Chen, T. Wu, A. Khan, and H. Wang. (2025). Large language models meet knowledge graphs for question answering: Synthesis and opportunities. Presented at Conference on Empirical Methods in Natural Language Processing (EMNLP). [Online]. Available: <https://doi.org/10.48550/arXiv.2505.20099>
- [6] G. Xiao et al. (2025). LLM4VKG: Leveraging large language models for virtual knowledge graph construction. Presented at 34th International Joint Conference on Artificial Intelligence (IJCAI). [Online]. Available: <https://doi.org/10.24963/ijcai.2025/525>
- [7] X. Li, P. Zheng, J. Bao, L. Gao, and X. Xu. (2025). Achieving cognitive mass personalization via the Self-X cognitive manufacturing network: An industrial knowledge graph- and graph embedding-enabled pathway. *Computers & Industrial Engineering*. [Online]. 180, p. 109709. Available: <https://doi.org/10.1016/j.cie.2025.109709>
- [8] E. Lavrinovics, R. Biswas, J. Bjerva, and K. Hose. (2025). Knowledge graphs, large language models, and hallucinations: An NLP perspective. *Journal of Web Semantics*. [Online]. 85, p. 100844. Available: <https://doi.org/10.1016/j.websem.2024.100844>
- [9] R. Haskins and B. Adams. (2026). KEA Explain: A neurosymbolic framework for explaining large language model hallucinations through knowledge graph validation. *Journal of Intelligent Systems*. [Online]. 5(1). Available: <https://hdl.handle.net/10092/109696>
- [10] J. Liu and J. Tang. (2026). LeAK-GraphRAG: A legal academic knowledge GraphRAG framework with hybrid retrieval for mitigating hallucinations in large language models. In *Frontiers in Artificial Intelligence and Applications*. [Online]. Vol. 411, Amsterdam, The Netherlands: IOS Press, pp. 597–604. Available: <https://doi.org/10.3233/FAIA251624>
- [11] S. Jauhari. (2025). Algorithmic red teaming approaches to secure LLMs. *Machine Learning with Applications*. [Online]. 23, p. 100815. Available: <https://doi.org/10.1016/j.mlwa.2025.100815>
- [12] S. Chehreh and A. Sarabadani. (2024). A model based on random forest algorithm and Jaya optimization to predict bank customer churn. *Journal of Engineering Management and Soft Computing*. [Online]. 9(2), pp. 132–148. Available: <https://doi.org/10.22091/jemsc.2024.9541.1174>
- [13] M. Q. Pérez, S. López Bernal, E. Horna Prat, L. Montesano Del Campo, L. Fernández Maimó, and A. Huertas Celdrán. (2025). EEG channel reconstruction using convolutional neural networks in limited BCIs: A proposed method for neuromarketing applications. *Applied Soft Computing*. [Online]. 181, p. 113455. Available: <https://doi.org/10.1016/j.asoc.2025.113455>
- [14] K. Sun and R. Wang. (2025). Breaking myths in LLM scaling and emergent abilities with a comprehensive statistical analysis. *Neural Networks*. [Online]. 183, p. 108164. Available: <https://doi.org/10.1016/j.neunet.2024.108164>
- [15] F. Bagheri, D. Buscaldi, and D. Reforgiato Recupero. (2025). Leveraging knowledge graphs and LLMs for content based reviewer assignment. *Journal of Intelligent Information Systems*. [Online]. Available: <https://doi.org/10.1007/s10844-025-01004-9>
- [16] P. Gajo and A. Barrón-Cedeño. (2025). Natural vs programming language in LLM knowledge graph construction. *Information Processing & Management*. [Online]. 62, p. 104195. Available: <https://doi.org/10.1016/j.ipm.2025.104195>
- [17] C. Öberg. (2023). Neuroscience in business-to-business marketing research: A literature review, co-citation analysis and research agenda. *Industrial Marketing Management*. [Online]. 113, pp. 168–179. Available: <https://doi.org/10.1016/j.indmarman.2023.06.004>

- [18] J. Wei et al. (2022). Emergent abilities of large language models. *Transactions on Machine Learning Research*. [Online]. Available: <https://arxiv.org/abs/2206.07682>
- [19] N. Ibrahim, S. Aboulela, and A. Ibrahim. (2024). A survey on augmenting knowledge graphs (KGs) with large language models (LLMs): Models, evaluation metrics, benchmarks, and challenges. *Knowledge and Information Systems*. [Online]. 4(1), pp. 1–28. Available: <https://doi.org/10.1007/s44163-024-00175-8>
- [20] H. Purnomo and Y. Novita. (2024). The impact of branding, advertising, and social media marketing on consumer purchase intention, brand loyalty, and brand equity in the Indonesian market. *Journal of Management and Business (JOMB)*. [Online]. 6(3), Art. no. 9653. Available: <https://doi.org/10.31539/jomb.v6i3.9653>
- [21] M. M. Gutiérrez Gutiérrez, J. J. Perona Páez, and F. D. P. Gutiérrez Bonilla. (2024). Models of brand equity: A systematic and critical review. *Cogent Business & Management*. [Online]. 11(1), p. 2433168. Available: <https://doi.org/10.1080/23311975.2024.2433168>
- [22] L. D. Gonen, T. Tavor, and U. Spiegel. (2024). Unlocking market potential: Strategic consumer segmentation and dynamic pricing for balancing loyalty and deal seeking. *Mathematics*. [Online]. 12(21), p. 3364. Available: <https://doi.org/10.3390/math12213364>
- [23] E. Tarta and L. Pasaribu. (2024). The impact perceived quality, advertising effectiveness, social media presence and brand image on purchase intention Telkomsel provider. *Enrichment: Journal of Management*. [Online]. 14(4), pp. 694–704. Available: <https://doi.org/10.35335/enrichment.v14i4.2042>
- [24] A. Sarabadani and K. R. Fard. (2025). GDPKG-LLM: Integrating Gene, Disease, and Pharmacogenomics Knowledge Graphs for Cognitive Neuroscience Using Large Language Models. *Computational Science and Information*. [Online]. Available: <https://doi.org/10.7494/csci.2025.26.3.6673>
- [25] N. Ibrahim, S. Aboulela, A. Ibrahim, and R. Kashef. (2024). A survey on augmenting knowledge graphs (KGs) with large language models (LLMs): Models, evaluation metrics, benchmarks, and challenges. *Discover Artificial Intelligence*. [Online]. 4, p. 76. Available: <https://doi.org/10.1007/s44163-024-00175-8>
- [26] S. Seo, H. Cheon, H. Kim, and D. Hyun. (2022). Structural quality metrics to evaluate knowledge graph quality. *arXiv*. [Online]. Available: <https://doi.org/10.48550/arXiv.2211.10011>



GrAMIA-Net: Gray Matter Adaptive Multi View Imbalance Aware Attention Network for Detection of Parkinson's Disease *

Research Article

Elham Afarandeh ¹, Rahil Hosseini ^{*2}, Samaneh Yazdani ³

[10.22067/cke.2026.99262.1196](https://doi.org/10.22067/cke.2026.99262.1196)

Abstract-- Diagnosing Parkinson's disease (PD) at its early stages from structural MRI remains a difficult task, primarily because of scanner-dependent variations, skewed class distributions, and the subtle anatomical alterations associated with early neurodegeneration. To improve MRI-based Parkinson's disease classification under these conditions, GrAMIA-Net is proposed as a multi-view deep learning framework built upon gray matter representations derived from T1-weighted MRI. The framework learns complementary information from sagittal, coronal, and axial perspectives and employs a hierarchical attention mechanism to emphasize discriminative spatial cues within each view. In addition, an inter-view attention module is utilized to aggregate information across different anatomical planes effectively. To mitigate the adverse effects of class imbalance, an Adaptive Imbalance Optimization (AIBO) scheme is integrated into the training process, thereby improving learning for underrepresented samples. Validation on the multi-scanner and class-imbalanced PPMI cohort shows that the proposed framework attains an accuracy of 0.862, an F1-score of 0.919, and an AUC of 0.720, demonstrating reliable predictive capability across categories. These results suggest that GrAMIA-Net provides a practical and computationally efficient framework for PD detection using single-modality GM representations, with potential applicability in real-world neuroimaging settings.

Index Terms-- Adaptive Imbalance Optimization, Gray Matter Maps, Multi-View Learning, PD Detection, T1-weighted Structural Imaging

INTRODUCTION

Parkinson's disease (PD) is a chronic neurodegenerative condition and among the most prevalent movement disorders in the elderly population worldwide [1]. It is characterized by motor symptoms including tremor, muscular rigidity, and bradykinesia, as well as non-motor manifestations including cognitive and behavioral impairments [2–5]. Despite its clinical

importance, PD diagnosis largely relies on subjective assessments, such as neurological examinations and rating scales (e.g., UPDRS) [6,7]. The absence of reliable objective biomarkers remains a major challenge for early diagnosis and highlights the need for automated and data-driven diagnostic approaches.

Neuroimaging has become an important tool for investigating structural brain alterations associated with PD. Among available imaging modalities, T1-weighted structural MRI is widely used because of its accessibility, relatively low cost, and ability to provide detailed anatomical information [8–10]. Although conventional MRI does not offer disease-specific visual markers, it enables quantitative assessment of structural changes in brain regions associated with PD progression, including the basal ganglia and cortical areas [11,12]. In particular, gray matter (GM) has attracted considerable attention, as GM atrophy and cortical thinning have been linked to disease severity and progression [13,14].

Despite these advantages, MRI-based PD analysis remains challenging. Variability introduced by different scanners and acquisition protocols can affect image characteristics, making standardized preprocessing procedures such as bias-field correction, skull stripping, and spatial normalization essential for ensuring data consistency [15–17]. Furthermore, GM representations often exhibit low contrast and complex anatomical structures, making subtle disease-related patterns difficult to identify. Another important challenge is class imbalance, which is common in PD datasets and may bias predictive models toward the majority class while reducing sensitivity to clinically relevant cases [18].

Data-driven learning methods, particularly CNN-based models, have shown considerable potential for MRI-based classification of PD by capturing relevant patterns in

* Manuscript received June 09, 2026, Revised, June 18, 2026, 2026, Accepted July 18, 2026.-

¹ PhD Student, Department of Computer Engineering, NT.C., Islamic Azad University, Tehran, Iran.

² Corresponding author*. Associate Professor, Department of Computer Engineering, ShQ.C., Islamic Azad University, Shahre Qods, Tehran, Iran. Email: rahil.hosseini@iau.ac.ir

³ Associate Professor, Department of Computer Engineering, NT.C., Islamic Azad University, Tehran, Iran.

neuroimaging data [19–22]. Multi-branch architectures, ensemble strategies, and hybrid frameworks that combine deep learning with conventional machine learning techniques have further improved classification performance [23,24]. However, conventional CNNs are inherently limited in modeling long-range spatial dependencies and complex contextual relationships within neuroimaging data [25]. Although attention mechanisms have been introduced to alleviate these limitations [27–31], many existing methods rely on volumetric representations or simple feature fusion strategies and do not explicitly model hierarchical intra-view relationships or inter-view dependencies across sagittal, coronal, and axial planes.

Another challenge is the effective integration of information from multiple anatomical views. Common fusion approaches, such as feature concatenation and averaging, may not fully exploit the complementary information available across different views [32,33]. Additionally, methods that rely on multimodal imaging data or artificially balanced datasets often encounter limitations when applied to heterogeneous clinical data collected under real-world conditions.

Building on these observations, this study presents GrAMIA-Net, a multi-view framework that utilizes gray matter (GM) maps derived from T1-weighted MRI for Parkinson’s disease classification. The proposed framework learns complementary anatomical information from sagittal, coronal, and axial views using a hierarchical attention mechanism that captures multi-scale structural patterns in each view. To facilitate effective integration of information across anatomical planes, a dedicated cross-attention module is employed to model inter-view relationships and enhance feature representation. Additionally, an Adaptive Imbalance Optimization (AIBO) strategy is incorporated into the training process to enhance learning robustness in the presence of highly imbalanced data distributions.

By combining multi-view gray matter representation learning, hierarchical attention-based feature extraction, cross-view feature interaction, and imbalance-aware optimization, GrAMIA-Net provides an effective framework for automated Parkinson’s disease detection in heterogeneous neuroimaging environments.

Unlike existing MRI-based PD classification methods that either rely on simple view fusion or process anatomical views independently, the proposed framework jointly addresses hierarchical intra-view representation learning, adaptive cross-view integration, and severe class imbalance within a unified gray matter-based architecture.

The key contributions of the present study are outlined below:

- A compact framework for GM-based Parkinson’s disease detection: GrAMIA-Net is developed as a deep learning architecture that leverages multi-view gray matter representations derived from T1-weighted MRI for automated PD classification.
- Hierarchical intra-view attention modeling (HIA-GM):

A dedicated attention module is introduced to capture multi-scale gray matter characteristics within sagittal, coronal, and axial planes, enabling structured feature learning at different anatomical scales.

- Cross-view anatomical feature integration (MVCA): The proposed MVCA module models interactions among anatomical views through cross-attention, allowing complementary information from different planes to be effectively aggregated into a unified representation.
- Adaptive imbalance optimization (AIBO): An imbalance-aware training strategy combining SMOTE-based oversampling, targeted augmentation, and class-weighted focal loss is designed to improve robustness and sensitivity when learning from highly imbalanced clinical datasets.

The rest of the manuscript is organized as follows. Section 2 outlines the materials and methodology. Section 3 presents the experimental findings, performance evaluation, interpretability analysis, and comparative studies. Section 4 summarizes the main conclusions and discusses future research directions.

MATERIALS AND METHODS

The following section describes the proposed GrAMIA-Net for automated Parkinson’s disease classification using GM maps derived from T1-weighted MRI. The framework integrates multi-view representation learning, attention-based feature modeling, and imbalance-aware optimization within a unified architecture.

A. Overview of the Proposed GrAMIA-Net Framework

The Overall GrAMIA-Net architecture is provided in Figure 1. The network receives sagittal, coronal, and axial GM maps generated through the preprocessing and segmentation pipeline. These views are fed into the network to learn view-specific features.

Intra-view features are modeled using the Hierarchical Intra-View Attention (HIA-GM) module, which combines convolutional feature extraction with spatial attention to capture multi-scale patterns. The Multi-View Cross Attention (MVCA) module models dependencies across anatomical views and enables effective feature fusion.

The resulting multi-view representation is then forwarded to a fully connected classification head for PD prediction. To address class imbalance, an imbalance-aware optimization strategy is applied during training, incorporating data augmentation, SMOTE-based oversampling, and class-weighted focal loss. The training process was conducted in two stages. In the first stage, the HIA-GM and MVCA modules were trained to learn discriminative multi-view feature representations from GM maps. In the second stage, the learned representations were utilized by the MLP (multi-layer perceptron) classifier for final PD prediction. This strategy facilitated stable feature learning while reducing the risk of overfitting on the limited training data.

Additional details regarding the individual components

of GrAMIA-Net are provided in the following subsections.

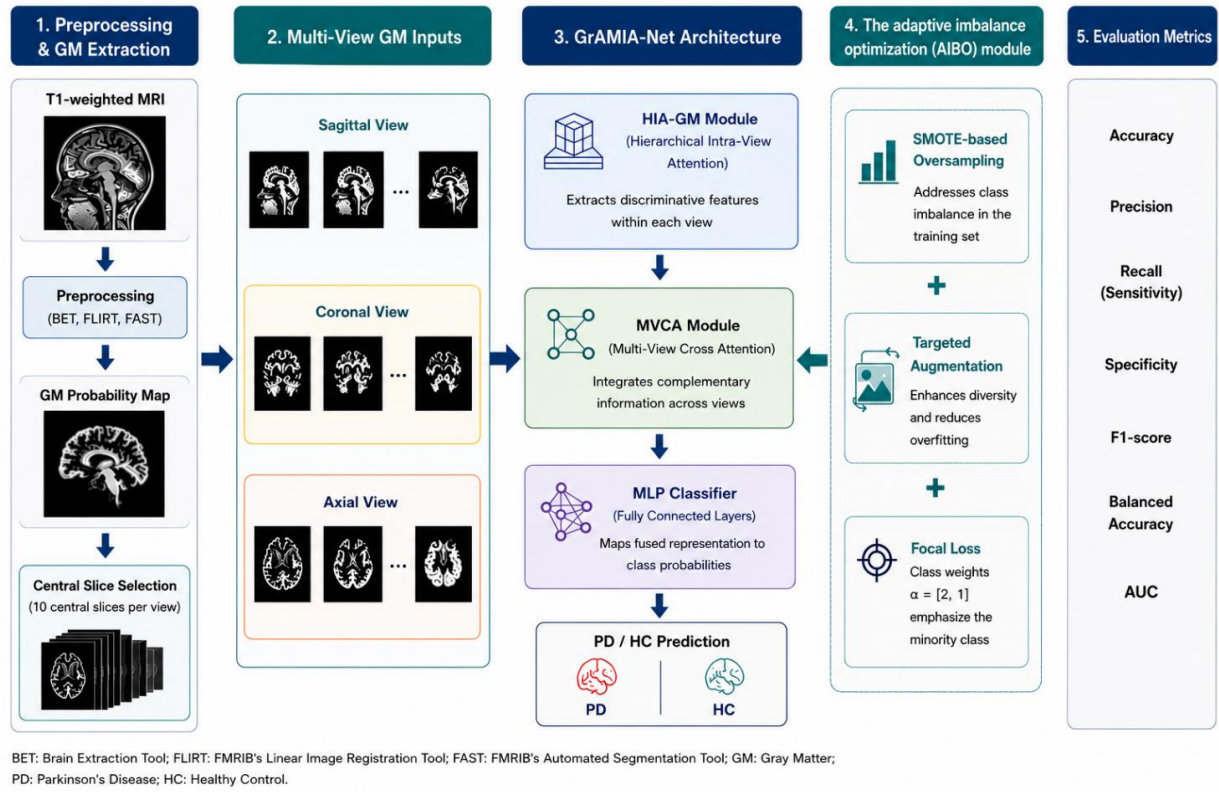


Figure 1. The architecture of the proposed GrAMIA-Net framework.

B. Data Acquisition

The MRI data analyzed in this work were obtained from the Parkinson's Progression Markers Initiative (PPMI). This international multicenter study provides standardized clinical, imaging, and biomarker data for Parkinson's disease research [34]. Diagnostic labels were obtained directly from the PPMI database (www.ppmi-info.org) and were based on the official PPMI analytic cohort assignments. Only subjects belonging to the Parkinson's disease (PD) and Healthy Control (HC) cohorts were included in this study because the objective of the present work was to investigate binary PD classification. Participants assigned to other categories, such as SWEDD, prodromal, or genetic cohorts, were excluded because these groups represent distinct clinical populations with different diagnostic objectives and disease characteristics. Including such subjects would transform the task from binary PD classification into a substantially different multi-cohort classification problem. No additional relabeling or manual annotation was performed. Subject inclusion was determined exclusively by cohort assignment and MRI availability. No subjects were selected or excluded according to image appearance, disease severity, model predictions, or study outcomes, thereby reducing the risk of subjective selection bias during dataset construction.

The data were acquired using 3D T1-weighted MRI scans from 3.0 Tesla scanners, including models from Siemens, Philips, GE Medical Systems, and Philips Healthcare, with an image resolution of

$256 \times 256 \times 192$ voxels. Acquisition parameters, such as pulse sequences, flip angles, and repetition times (TR), differed across scanner platforms. A summary of the dataset characteristics is provided in Table 1.

TABLE 1
The demographic information of datasets

Manufacturer of Scanners	Age	PD	Control	Sex
SIEMENS, Philips Medical Systems, GE MEDICAL (Flip Angle: [9,12], Pulse Sequence: GR, GR/IR, TR: [6.5,2300])	Min: 34 Max: 86	288	37	Female:111 Male:214

C. MRI Preprocessing and Gray Matter Map Extraction

To improve consistency across scans acquired from different imaging systems, all MRI volumes were processed using a standardized workflow implemented in the FMRIB Software Library (FSL) [35]. The preprocessing pipeline, summarized in Figure 2, consisted of DICOM-to-NIFTI conversion, image intensity standardization, orientation alignment, bias-field correction, brain extraction, and spatial registration.

Brain tissues outside the cranial region were removed using the BET tool. The resulting images subsequently underwent intensity inhomogeneity correction and affine registration to a common anatomical template using FLIRT. Tissue segmentation was performed with FAST to obtain probabilistic representations of GM, WM, and CSF. Finally, the GM representations were refined by applying thresholding and spatial smoothing operations to suppress

noise and improve anatomical consistency across subjects.

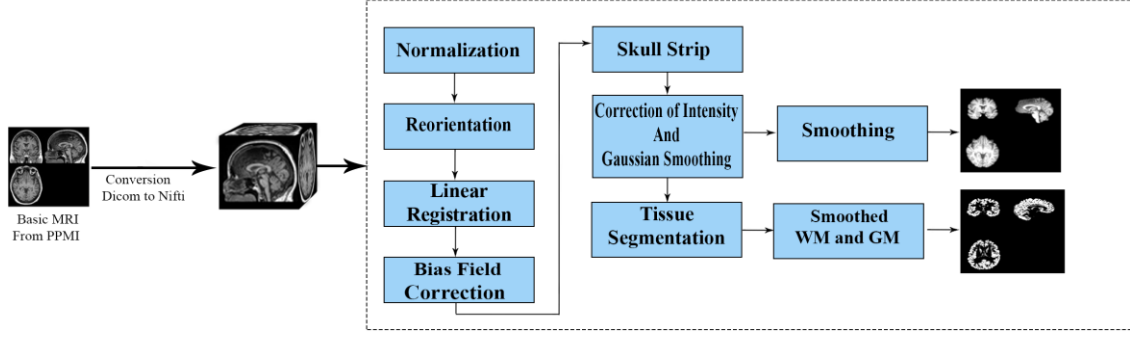


Figure 2. The preprocessing pipeline.

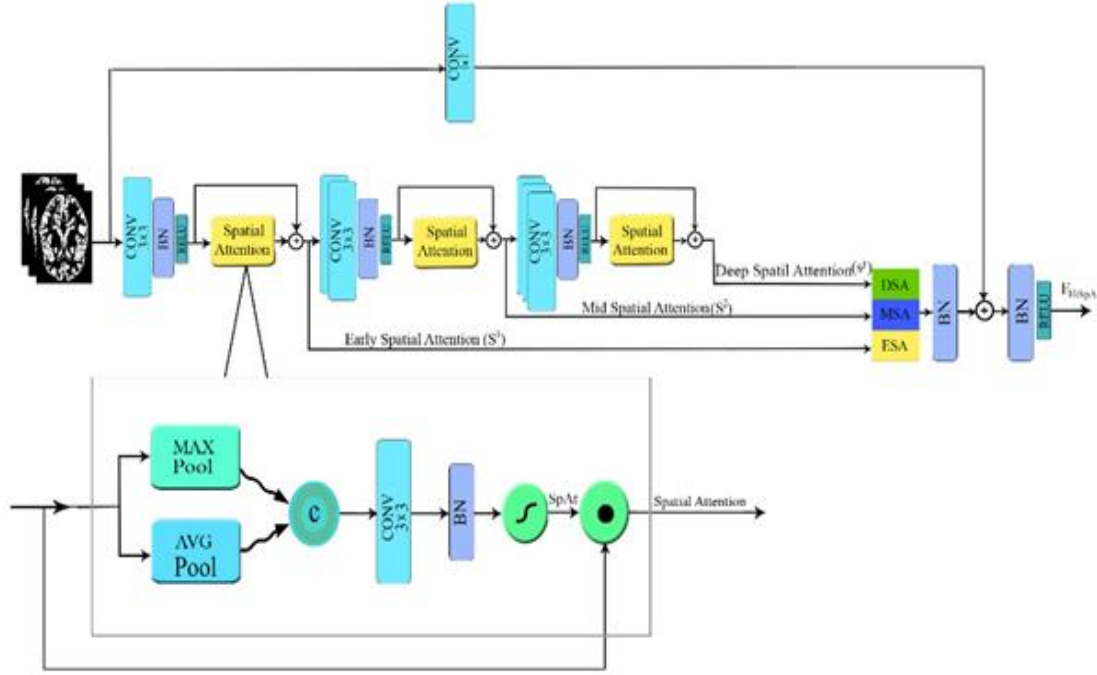


Figure 3. The structure of the HIA-GM module.

Following preprocessing, 10 central slices were selected from each anatomical view (sagittal, coronal, and axial), as these regions typically contain more informative structural patterns relevant to neurodegeneration, while peripheral slices contribute less discriminative information [22]. This selection also reduces computational complexity while preserving diagnostically relevant features. The number of selected slices is consistent with related multi-plane MRI classification approaches [36], where central slice windows are commonly adopted as a practical and effective input strategy. Visual inspection of the preprocessed GM maps further confirmed that the selected region exhibited the highest gray matter density and the most pronounced structural variation across subjects.

Although alternative slice counts may also retain useful anatomical information, a fixed set of 10 central slices was adopted to balance anatomical coverage, computational efficiency, and model complexity across all experiments. Despite this reduced input representation, the proposed framework achieved competitive classification performance relative to previously reported MRI-based

approaches (Table 5), suggesting that the selected slice configuration preserves sufficient discriminative information for PD classification while substantially reducing computational requirements.

D. Hierarchical Intra-View Attention Module (HIA-GM)

The Hierarchical Intra-View Attention Module (HIA-GM) is a core component of the proposed GrAMIA-Net, designed to model GM structures within each anatomical plane. As illustrated in Figure 3, the module captures hierarchical spatial dependencies across sagittal, coronal, and axial views while preserving full spatial resolution.

For each view, input feature maps are processed independently through three sequential attention-guided convolutional stages. A sequence of 3×3 convolution, batch normalization, nonlinear activation, and spatial attention is employed at each stage.

This progressive design enhances the effective receptive field, allowing for the extraction of both local structural details and broader anatomical context without relying on pooling operations.

Spatial attention is computed using channel-wise average and max pooling, followed by convolution and sigmoid activation to generate attention maps. These maps are applied multiplicatively to the feature representations, while residual connections preserve low-level spatial information and support stable training.

The stacked attention-enhanced stages produce multiscale intra-view representations with increasing contextual awareness. These representations are aggregated to form refined, view-specific embeddings that encode structured anatomical information. By explicitly modeling intra-view spatial hierarchies, the HIA-GM module enables precise characterization of subtle GM alterations associated with early-stage PD and provides informative features for subsequent cross-view fusion.

E. Multi-View Cross Attention (MVCA) Module

Conventional multi-view fusion strategies, such as concatenation, averaging, or majority voting, combine features without explicitly modeling inter-view dependencies, often assuming equal contribution from each anatomical plane [36]. This can limit their ability to capture complementary and context-dependent relationships among sagittal, coronal, and axial representations.

To address this limitation, we introduce a Multi-View Cross Attention (MVCA) module that captures interactions between view-specific GM features. The module operates on latent representations $F \in \mathbb{R}^{n \times 3 \times d}$, where n is the batch size, 3 denotes the anatomical views, and d is the feature dimension, obtained from the HIA-GM module. The input features are transformed into query (Q), key (K), and value (V) representations through trainable linear mappings. Cross-attention is computed via scaled dot-product attention [37]:

$$\text{Attention}(Q, K, V) = \text{Softmax}(QK^T)V \quad (1)$$

In the standard transformer formulation, a scaling factor of $(1/\sqrt{d})$ is commonly applied to stabilize the magnitude of dot-product attention for high-dimensional feature representations. In the proposed MVCA module, attention is computed over three compact view-level feature embeddings generated by HIA-GM. During model development, both the scaled and unscaled attention formulations were implemented and evaluated under the same experimental setting. The unscaled formulation consistently produced slightly better classification performance and was therefore adopted in the final architecture. This comparison was performed to guide the architectural design of the proposed model rather than as a dedicated ablation study. Accordingly, the present work does not claim that removing the scaling factor is generally advantageous, and this design choice should be interpreted only in the context of the proposed framework.

From others, enabling context-aware feature integration. The attention-enhanced features are then processed by a feedforward network (FFN) with nonlinear activation and dropout. A residual connection is applied by adding a linearly projected version of the input, followed by layer normalization:

$$Z = \text{LayerNorm}(\text{Proj}(F) + \text{FFN}(\text{Attention}(Q, K, V))) \quad (2)$$

where $Q, K, V \in \mathbb{R}^{n \times 3 \times d}$ and $Z \in \mathbb{R}^{n \times 3 \times d}$. The output is reshaped to $\mathbb{R}^{n \times 3d}$ and passed to the classification head. Unlike standard transformer architectures, the proposed MVCA is tailored for cross-view spatial fusion rather than sequential modeling. By enabling interactions across anatomically distinct yet correlated GM representations, it facilitates effective integration of complementary structural information. Figure 4 provides an overview of the MVCA module.

F. Final classification via MLP

After multi-view fusion, the resulting feature representation is fed into a fully connected for classification. The MLP maps high-dimensional embeddings to diagnostic labels while maintaining computational efficiency. The classification head employs a progressively decreasing layer configuration, where the feature dimension is reduced from 256 units in the first layer to 32 units in the final hidden layer. ReLU activation, L2 regularization, and dropout (0.2) are applied after each layer to improve generalization. The final layer uses a softmax activation to produce probabilistic outputs for binary PD classification.

The classification head was optimized for 200 epochs using a batch size of 8. This classification head complements the preceding feature extraction and attention modules, supporting efficient end-to-end training while maintaining discriminative capability.

G. The adaptive imbalance optimization (AIBO) module

Class imbalance remains a major issue in medical imaging, often favoring dominant classes and limiting the detection of minority patterns in Parkinson's disease datasets [6]. To address this, the proposed framework incorporates an Adaptive Imbalance Optimization (AIBO) module that combines three complementary strategies:

SMOTE-based oversampling: SMOTE is employed during training to synthesize additional minority-class samples through interpolation, improving class balance and decision boundary generalization.

Class-weighted focal loss: To emphasize hard-to-classify samples, a class-weighted focal loss [38] is employed:

$$L_{\text{Focal}} = \frac{1}{N} \sum_{i=1}^N \alpha_i (1 - \hat{y}_i)^\gamma [-y_i \log(\hat{y}_i) - (1 - y_i) \log(1 - \hat{y}_i)] \quad (3)$$

that N is the batch size, $y_i \in \{0, 1\}$ is the ground truth, and $\hat{y}_i$ is the predicted probability. The parameters α_i and γ control class weighting and focus on hard samples, respectively.

Targeted data augmentation: Minority-class samples are augmented using small rotations ($\pm 15^\circ$) and horizontal flips to simulate realistic variations without distorting anatomical structures.

By integrating synthetic sampling, adaptive loss weighting, and targeted augmentation, the AIBO module improves robustness and enhances sensitivity to minority-

class patterns under imbalanced conditions.

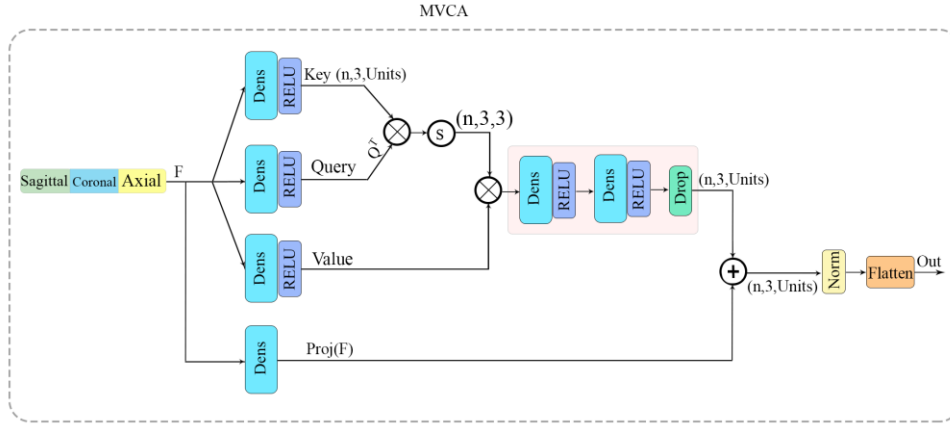


Figure 4. The proposed MVCA module

H. Integration of HIA-GM and MVCA: Workflow and Complexity Analysis

To improve the clarity of the proposed architecture, Figure 5 illustrates the computational workflow of the GrAMIA-Net attention framework. The flowchart summarizes the interaction between the HIA-GM and MVCA modules, highlighting how hierarchical intra-view feature extraction and cross-view feature fusion are integrated to generate the final multi-view representation.

To provide additional insight into the computational characteristics of the proposed framework, the computational complexity of GrAMIA-Net was derived from the operations performed within the HIA-GM and MVCA modules illustrated in the flowchart. Let $H \times W$ denote the spatial resolution of the feature maps, $V=3$ the number of anatomical views, $S=10$ the number of input channels per view, and $C_1=43$, $C_2=85$, and $C_3=128$ the numbers of output channels in the three hierarchical stages of HIA-GM.

Since each view is processed independently through three sequential convolutional-attention stages, the dominant computational cost of HIA-GM can be expressed as: $T_{\text{HIA-GM}} = O(VH^2(SC_1 + C_1 C_2 + C_2 C_3))$. Using the actual architectural parameters of the proposed model, the computational cost of HIA-GM is approximately 2.71×10^{10} multiply-accumulate operations (MACs) per subject. In contrast, the MVCA module operates on only three view-level feature representations after global pooling, resulting in a complexity of $O(Vd^2)$, where d denotes the feature dimension. With $V=3$ and $d=256$, the computational cost of MVCA is approximately 7.86×10^5 MACs. Therefore, the computational cost of GrAMIA-Net is dominated by the hierarchical feature extraction process of HIA-GM, while the additional overhead introduced by MVCA is negligible in comparison. Empirical runtime measurements on the held-out test cohort further supported this analytical observation. Feature extraction required 2.34s, 2.29s, and 2.94s for the sagittal, coronal, and axial views, respectively.

The MVCA module required 0.03s, and the final MLP

classifier required 0.05s for the complete test cohort. The total inference time was approximately 7.64s, corresponding to an average inference time of approximately 117ms per subject. These measurements confirm that the computational cost is primarily associated with feature extraction, whereas cross-view attention contributes only a small additional overhead. Furthermore, by processing a fixed set of ten informative central slices per anatomical view rather than the full MRI volume, the proposed framework substantially reduces memory consumption and computational requirements while maintaining competitive classification performance.

EXPERIMENTAL RESULTS AND DISCUSSION

This section investigates the behavior of GrAMIA-Net and analyzes its performance in PD classification based on multi-view GM maps.

A. Experimental setup and evaluation protocol

The experimental evaluation was performed on a workstation featuring an AMD Radeon RX 5700 XT GPU and 64 GB of RAM. MRI preprocessing was performed using Nipype in conjunction with the FMRIB Software Library (FSL) on a Linux platform, while deep learning experiments were conducted in Python using TensorFlow/Keras. Following preprocessing and GM segmentation, 10 central slices were selected from each anatomical view (sagittal, coronal, and axial) as model inputs, as these regions typically contain more informative structural patterns.

Data partitioning was performed at the subject level, allocating 60% of the subjects to training, 20% to validation, and the remaining 20% to testing, ensuring no subject overlap across subsets and preventing data leakage. To improve learning under imbalanced conditions, SMOTE-based oversampling and targeted augmentation were applied exclusively to the training set through the AIBO strategy, whereas the validation and test subsets were left unchanged during evaluation.

Training followed a two-stage procedure consistent with the architecture outlined in Section 2.1. During the

initial training phase, the HIA-GM and MVCA backbone modules were trained end-to-end for 50 epochs using batches of eight samples, an initial learning rate of 0.01, and L2 regularization of 0.001. In the second stage, the MLP classification head was trained for 200 epochs using the same batch size and regularization settings. The focal loss parameters were empirically optimized with class weights $\alpha = [2, 1]$ to emphasize the minority class.

Training parameters were selected based on validation-set performance and convergence behavior.

To assess training stability, all experiments were repeated ten times with independent weight initializations while preserving the same subject-wise train, validation, and test partitions. All metrics are presented as the mean values across repeated runs.

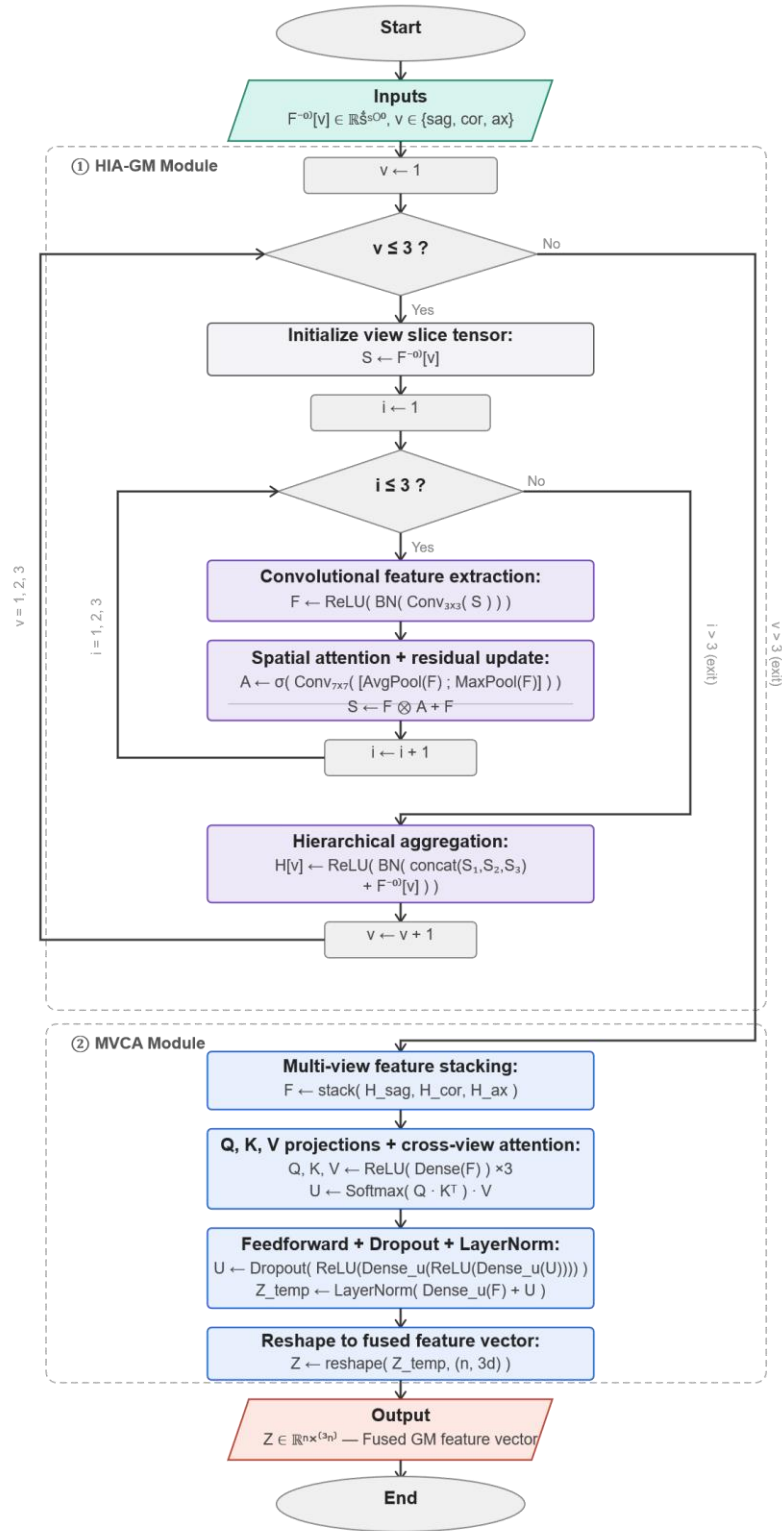
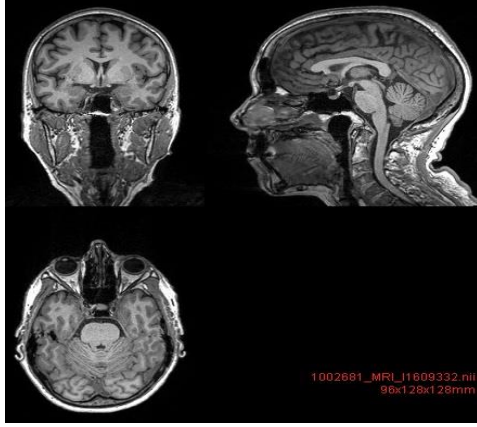
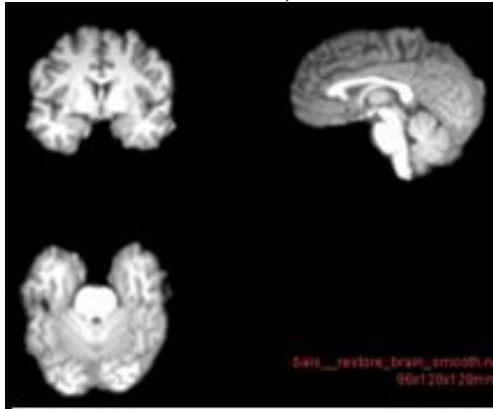


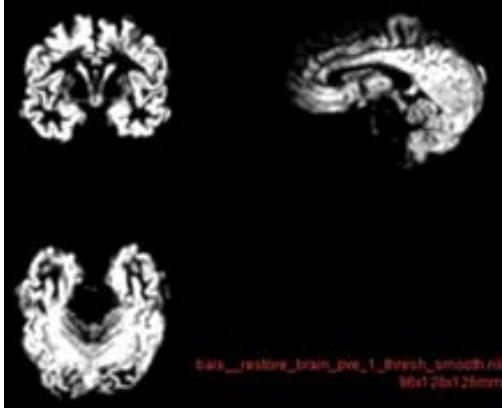
Figure 5. Flowchart of the HIA-GM and MVCA module.



(a): Sample of MRI scans in nifti format from 3 views (coronal, sagittal, and axial)



(b): Sample of T1-MRI scans after preprocessing steps



(c): Sample of GM maps after preprocessing steps

Figure 6. The preprocessing pipeline outcomes.

B. Performance evaluation metrics

Performance was assessed using commonly adopted classification measures, including specificity, accuracy (ACC), precision, sensitivity, F1-score, balanced accuracy (BACC), and area under the ROC curve (AUC), computed from the confusion matrix (TP, TN, FP, FN) [29].

Given the class imbalance, particular emphasis was placed on F1-score, balanced accuracy, precision, and specificity, as accuracy alone may provide misleading estimates under skewed class distributions. These metrics offer complementary perspectives on model performance under class imbalance.

C. Qualitative analysis of preprocessing

To assess the impact of the preprocessing pipeline, representative examples of processed T1-weighted images and their corresponding GM maps are shown in Figure 6. The preprocessed images exhibit improved spatial consistency, effective removal of non-brain tissues, and reduced background noise compared to raw MRI scans.

In particular, GM maps provide clearer structural delineation and suppress non-informative regions, resulting in more focused and anatomically meaningful representations. Although the visual differences are subtle, they contribute to improved feature stability and support more effective downstream learning. These images and corresponding GM maps are shown in Figure 6. The preprocessed images exhibit improved spatial consistency, effective removal of non-brain tissues, and reduced background noise compared to raw MRI scans.

In particular, GM maps provide clearer structural delineation and suppress non-informative regions, resulting in more focused and anatomically meaningful representations. Although the visual differences are subtle, they contribute to improved feature stability and support more effective downstream learning. These observations are consistent with prior studies highlighting the importance of intensity normalization and spatial registration in neuroimaging-based analysis [16,17].

D. Quantitative performance evaluation

The quantitative performance of GrAMIA-Net is presented in Figure 7, where metric values are reported as the average of ten independent experimental runs. An F1-score of 0.92 was obtained, suggesting effective performance under imbalanced conditions. Precision remained high (0.93) with minimal variability, supporting reliable detection of PD cases. Recall and specificity showed moderate fluctuations across runs, possibly influenced by class imbalance and scanner variability. Despite this variation, balanced accuracy (0.73) and AUC (0.72) indicate stable overall classification performance.

Although specificity remained lower than sensitivity, this behavior is likely a consequence of the severe class imbalance of the dataset. Additional evidence supporting this interpretation is provided by the ablation results presented in the Imbalance Handling subsection in Section 3.F, where the removal of individual imbalance-aware components leads to substantial reductions in specificity and balanced accuracy.

From a practical perspective, these results indicate that GrAMIA-Net maintains reliable discrimination between PD and HC subjects despite the challenges posed by severe class imbalance and scanner heterogeneity. The observed balanced accuracy suggests that the model does not merely favor the majority class but retains meaningful predictive capability across both diagnostic groups. Furthermore, the combination of high precision and F1-score indicates that the proposed framework can identify PD cases with a relatively low rate of false positive predictions while preserving strong overall classification performance. Collectively, these findings support the effectiveness of the proposed multi-view attention framework for Parkinson's disease classification in realistic clinical

settings.

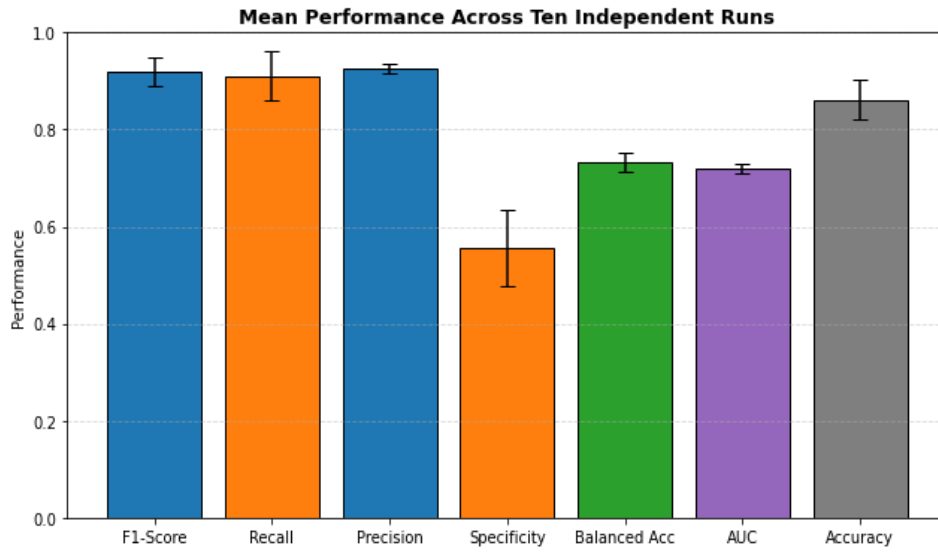


Figure 7. Mean classification performance of GrAMIA-Net across ten independent runs

To evaluate the impact of prediction granularity, the model was assessed at both slice-level and subject-level resolutions. In the slice-level setting, individual slices were processed independently while maintaining subject-wise separation to prevent data leakage. As reported in Table 2, slice-level evaluation yields lower performance (F1-score = 0.73, AUC = 0.62). Although precision remains high, reduced recall indicates increased sensitivity to noise and local variations.

In contrast, subject-level aggregation improves performance (F1-score=0.92, AUC=0.72), suggesting that integrating information across slices enhances global anatomical representation and reduces variability. While alternative slice-level settings may increase recall, they often reduce specificity and yield an unstable AUC, indicating less reliable predictions. Overall, subject-level inference provides more robust outcomes for PD classification.

TABLE 2

Comparison of slice-level and subject-level performance

Evaluation Level	Slice-Level	Subject-Level (Proposed)
Accuracy	0.615	0.862
Precision	0.919	0.927
Recall	0.607	0.911
F1-score	0.731	0.919
Specificity	0.667	0.556
Balanced Acc	0.637	0.733
AUC	0.621	0.720

The contribution of multi-view learning was further evaluated using individual anatomical views and their combined representation (Figure 8). Single-view models show inferior performance, with AUC values below 0.62. Among them, the axial view performs best (F1-score=0.79), and then coronal (0.76) and sagittal (0.68). In contrast, the multi-view configuration achieves higher performance (F1-score =0.92, AUC = 0.72), highlighting

the benefit of integrating complementary anatomical information across views

Finally, the impact of input modality was examined by comparing raw T1-weighted MRI with GM maps. GM-based inputs improve performance (F1-score: 0.73 $\rightarrow$ 0.92; AUC: 0.65 $\rightarrow$ 0.72), indicating that GM representations provide more discriminative structural information by reducing irrelevant variability.

E. Attention Map Visualization

To interpret cross-view interactions, average attention weights were analyzed for both training and test sets (Figure 9). The attention matrices exhibit consistent patterns across both sets, indicating stable and generalizable feature integration. Across all query views, higher attention is assigned to axial representations, followed by coronal, while sagittal contributions are comparatively lower. This suggests that the model prioritizes complementary structural information from axial and coronal planes during feature fusion. From a neuroimaging perspective, these findings are consistent with the distributed nature of PD-related structural alterations reported in the literature [39]. However, the reported attention weights reflect the relative importance of sagittal, coronal, and axial representations during feature fusion rather than the spatial localization of disease-related anatomical changes.

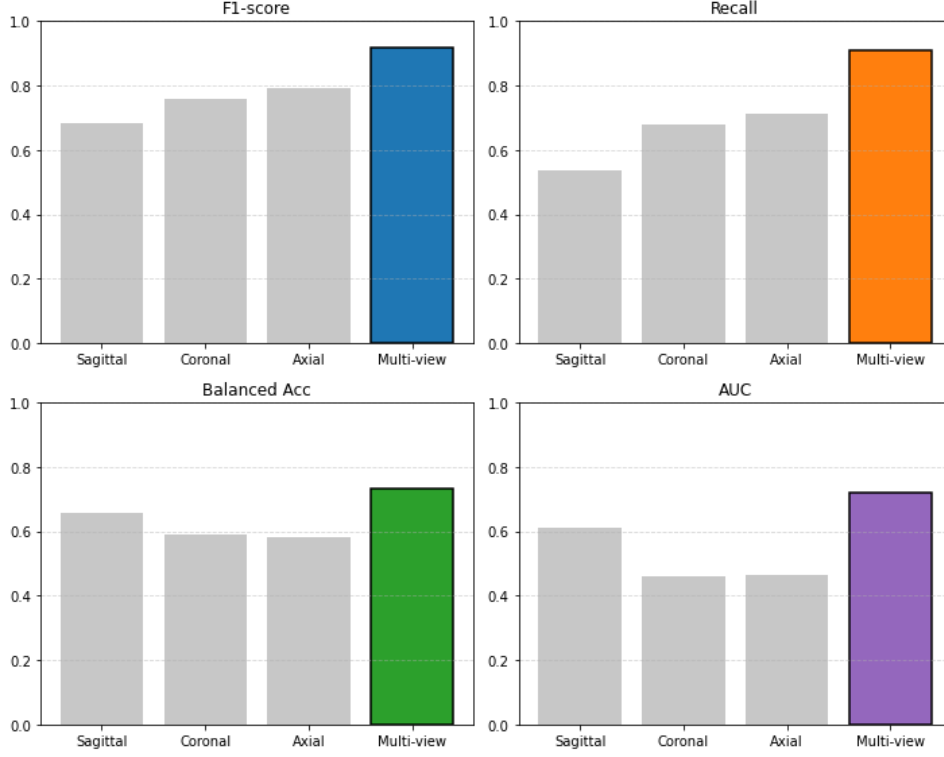


Figure 8. Performance comparison of single-view and multi-view configurations

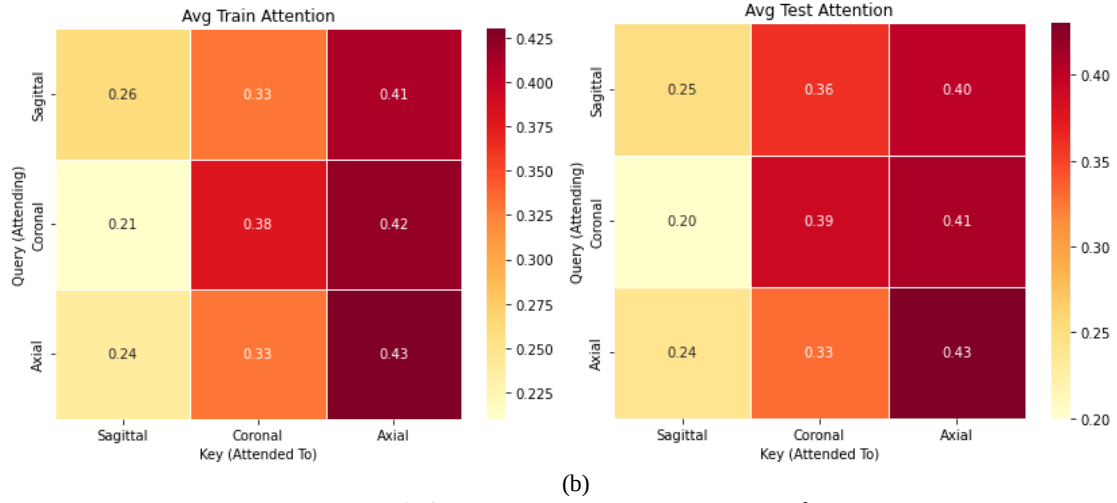


Figure 9. Average cross-view attention matrices for training

F. Ablation and Comparative Analysis

The effects of the architectural modules, fusion scheme, and imbalance optimization strategy were investigated through ablation studies.

Core components: As shown in Table 3, removing either HIA-GM or MVCA leads to a noticeable performance drop, while removing both results in the largest degradation, confirming their complementary roles in intra-view feature extraction and cross-view integration.

Fusion strategy: Figure 10 shows that MVCA achieves better performance than concatenation and majority voting, highlighting the benefit of modeling interactions across anatomical views. In contrast, concatenation fails to capture cross-view relationships, while majority voting produces less stable predictions.

TABLE 3
Performance impact of core Architectural Components in the GrAMIA-Net

Architecture Variant	Accuracy	Balanced ACC	AUC	F1-Score	Δ AUC
GrAMIA-Net (Full Model)	0.862	0.733	0.720	0.919	---
Without spatial attention (HIA-GM)	0.616	0.684	0.655	0.725	-0.065
Without MVCA	0.646	0.655	0.663	0.757	-0.058
Without MVCA & HIA-GM (Basic CNN)	0.569	0.610	0.557	0.689	-0.163

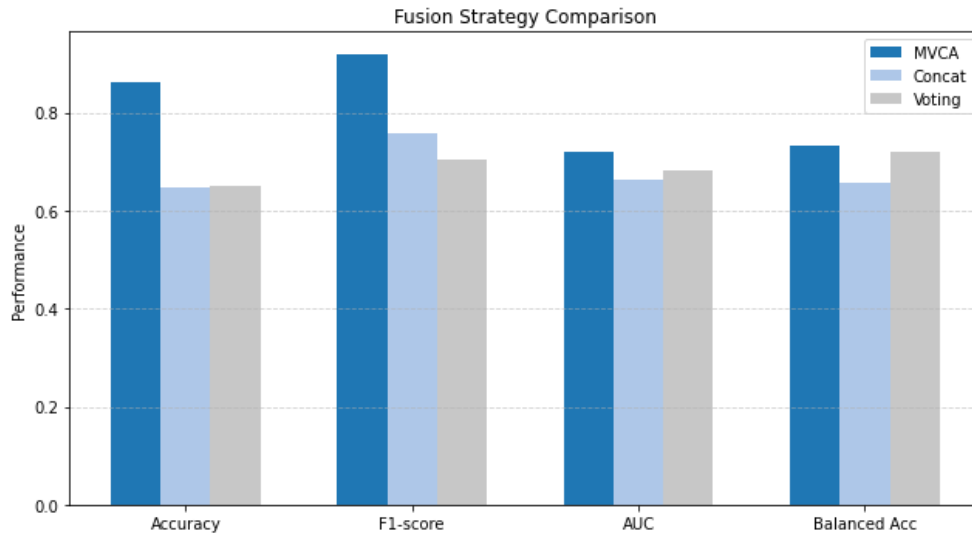


Figure 10. Comparison of fusion strategies.

TABLE 4
Effect of AIBO components on model performance

Configuration	Accuracy	Balanced Acc	F1-Score	Recall	Precision	Specificity
AIBO (SMOTE + Focal + Aug)	0.86	0.73	0.92	0.92	0.93	0.56
Without SMOTE	0.83	0.58	0.90	0.91	0.88	0.22
Without Focal Loss	0.68	0.67	0.78	0.67	0.92	0.56
Without Augmentation	0.74	0.52	0.84	0.82	0.86	0.22

Imbalance handling: The impact of AIBO is presented in Table 4. The full configuration achieves the most balanced performance. Removing SMOTE substantially decreases specificity and balanced accuracy, indicating an increased tendency toward majority-class predictions. A comparable reduction is observed when data augmentation is removed, suggesting that augmentation also contributes to improving minority-class recognition and overall model robustness. In contrast, removing focal loss primarily reduces recall and F1-score, highlighting the complementary roles of the individual components of the proposed AIBO strategy in addressing severe class imbalance. Although specificity remained lower than sensitivity, this behavior is consistent with the severe class imbalance of the cohort, where PD subjects substantially outnumber healthy controls. Nevertheless, the full AIBO configuration achieved a balanced accuracy of 0.733, indicating improved balance between sensitivity and specificity under highly imbalanced conditions.

G. Comparison with existing methods

Table 5 summarizes the performance of GrAMIA-Net alongside recent MRI-based methods for Parkinson's disease classification. As these studies differ in input modality, dataset size, and class distribution, the comparison should be interpreted as indicative rather than

strictly controlled.

GrAMIA-Net achieves competitive performance using only GM maps derived from T1-weighted MRI, a widely available clinical imaging modality. This demonstrates that competitive performance can be achieved using a single MRI modality, without requiring multimodal imaging inputs. Compared with conventional GM-based machine learning approaches, such as VBM+SVM [13] and Multi-feature ML [20], the proposed framework achieves substantially higher classification performance. This improvement may be attributed to the combination of hierarchical attention-based representation learning and structured multi-view feature integration, which enable the model to capture complementary anatomical information across views. Deep learning approaches based on larger cohorts, such as the 3D CNN model reported in [22], achieve higher AUC values. Similarly, multimodal architectures, including SCTransformer-VGG11 [36], benefit from the integration of complementary imaging modalities. However, these approaches either rely on substantially larger datasets or require additional imaging information beyond GM representations. In contrast, GrAMIA-Net was evaluated on a highly imbalanced cohort representative of real-world clinical conditions. Furthermore, the proposed framework outperforms earlier GM-map-based approaches, including the Ensemble ML + 3D CNN model [40], while maintaining a relatively simple

imaging input pipeline. The contribution of the AIBO module is reflected in the balanced accuracy and F1-score values, indicating improved robustness across classes despite the severe class imbalance of the dataset.

Overall, these findings suggest that GrAMIA-Net provides an effective compromise between performance, robustness, and model complexity in realistic multi-scanner settings.

H. Generalization and limitations

GrAMIA-Net demonstrates stable generalization on a multi-scanner, highly imbalanced dataset that reflects realistic clinical conditions. A subject-wise partitioning strategy was used to maintain strict independence among the training, validation, and test sets. Despite relying solely on GM maps derived from T1-weighted MRI, the model achieves competitive performance, indicating that discriminative structural features can be effectively captured using a single, widely available modality. The multi-view design and attention-based fusion further enhance robustness to scanner variability.

However, several limitations should be noted. The dataset exhibits substantial class imbalance, which may still affect model calibration. The cohort's severe imbalance also remains a potential source of bias and should be further investigated using larger and more balanced datasets. In addition, the study was validated using only one dataset, potentially restricting generalizability to other cohorts. Although GrAMIA-Net was assessed on a heterogeneous multi-scanner cohort, external validation on independent datasets was not performed. Accordingly, the results provide evidence of within-dataset generalization rather than definitive proof of cross-cohort robustness. In addition, the number of selected slices was fixed across all experiments. Although this choice was motivated by anatomical relevance, prior

literature, and computational efficiency considerations, a systematic sensitivity analysis across alternative slice counts was not performed. Therefore, the optimal number of slices remains an open question and should be investigated in future studies.

Another limitation of the proposed framework is its higher computational complexity compared with conventional single-view models. GrAMIA-Net processes three anatomical views and incorporates hierarchical attention-based feature extraction together with cross-view feature fusion, which increases the computational cost of the framework. This additional complexity represents the trade-off for learning richer multi-view representations and improving robustness to scanner variability. Nevertheless, the framework operates on compact gray-matter slice representations rather than full volumetric MRI data, helping to keep the computational requirements manageable. As discussed in the Computational Complexity Analysis (Section 2.H), the computational cost is primarily dominated by the HIA-GM feature extraction stage, whereas the additional overhead introduced by the MVCA module is relatively small. Empirical runtime measurements further showed that the complete inference process required approximately 117ms per subject on the held-out test cohort, demonstrating that the proposed framework remains computationally practical despite its increased architectural complexity.

Furthermore, GM-based representations depend on preprocessing pipelines that may introduce variability. While attention mechanisms improve interpretability, their direct clinical relevance remains to be validated. Future studies will investigate external validation using independent and multi-center datasets, the incorporation of longitudinal imaging data, systematic evaluation of alternative slice-selection strategies, and the development of more clinically grounded interpretability frameworks.

TABLE 5
Comparison of the MAVI-Net model with existing methods.

Method	Data type	Subjects (PD/HC)	Accuracy	F1-score	AUC	Balanced Acc
Ensemble ML + 3D CNN [40]	GM Maps	187/187	0.620	-	0.630	-
Multi-feature ML (GM+WM) [20]	GM+WM	177/67	0.842	0.848	-	-
3D CNN [22]	T1-MRI	1024/1017	0.793	-	0.87	-
ICA + 3D CNN (SCTransformer-VGG11)[36]	T1+fMRI	127 / 36	0.840	-	0.850	-
Feature-based + Decorrelated CNN [6]	rs-fMRI	164/234	0.805	0.772	0.623	0.667
VBM + SVM [13]	GM + demo	88/120	0.670	-	0.630	-
GrAMIA-Net (Proposed)	GM Maps	288/37	0.862	0.919	0.720	0.733

CONCLUSION AND FUTURE WORK

This paper presented GrAMIA-Net, a multi-view DL framework for PD classification based on GM maps extracted from T1-weighted MRI. The framework combines hierarchical intra-view attention with cross-view feature fusion to capture structural information from sagittal, coronal, and axial representations. In addition, the proposed AIBO strategy improves learning under imbalanced data distributions. The experimental findings indicate that GrAMIA-Net achieves competitive performance using only GM-based inputs derived from a single imaging modality. The integration of attention-guided feature learning and multi-view representation modeling contributes to robust feature extraction and effective disease classification. These results suggest that GM representations can provide informative biomarkers for automated PD detection in realistic neuroimaging settings.

Future research will investigate the generalizability of the framework using independent multi-center cohorts. Additional directions include incorporating complementary imaging modalities and clinical information, as well as further improving the interpretability and robustness of the proposed architecture.

REFERENCES

- [1] M. Absalyamova, D. Traktirov, V. Burdinskaya, V. Artemova, Z. Muruzheva, and M. Karpenko. (2025). Molecular basis of the development of Parkinson's disease. *Neuroscience*. [Online]. pp. 292–300. Available: <https://doi.org/10.1016/j.neuroscience.2024.12.009>
- [2] N. Vijiaratnam, T. Simuni, O. Bandmann, H. R. Morris, and T. Foltynie. (2021). Progress towards therapies for disease modification in Parkinson's disease. *The Lancet Neurology*. [Online]. 20(7), pp. 559–572. Available: [https://doi.org/10.1016/s1474-4422\(21\)00061-2](https://doi.org/10.1016/s1474-4422(21)00061-2)
- [3] X. Zhang et al. (2025). Optimal channel selection based on ABC-SVM in Parkinson's disease detection using short-time resting state EEG. *J Supercomput*. [Online]. 81, p. 977. Available: <https://doi.org/10.1007/s11227-025-07461-x>
- [4] S. J. Duncan, K. Marques, J. Fawkes, L. J. Smith, and D. T. Wilkinson. (2024). Galvanic vestibular stimulation modulates EEG markers of voluntary movement in Parkinson's disease. *Neuroscience*. [Online]. 555, pp. 178–183. Available: <https://doi.org/10.1016/j.neuroscience.2024.07.048>
- [5] X. Liu et al. (2021). A dual-branch model for diagnosis of Parkinson's disease based on the independent and joint features of the left and right gait. *Appl Intell*. [Online]. 51(10), pp. 7221–7232. Available: <https://doi.org/10.1007/s10489-020-02182-5>
- [6] P. Patil and W. R. Ford. (2024). Parkinson's disease recognition using decorrelated convolutional neural networks: Addressing imbalance and scanner bias in rs-fMRI data. *Biosensors*. [Online]. 14(5), p. 259. Available: <https://doi.org/10.3390/bios14050259>
- [7] S. G. Ryman and K. L. Poston. (2020). MRI biomarkers of motor and non-motor symptoms in Parkinson's disease. *Parkinsonism & Related Disorders*. [Online]. 73, pp. 85–93. Available: <https://doi.org/10.1016/j.parkreldis.2019.10.002>
- [8] A. S. Talai, J. Sedlacik, K. Boelmans, and N. D. Forkert. (2021). Utility of multi-modal MRI for differentiating of Parkinson's disease and progressive supranuclear palsy using machine learning. *Frontiers in Neurology*. [Online]. 12, p. 648548. Available: <https://doi.org/10.3389/fneur.2021.648548>
- [9] E. Porter, A. A. Roussakis, N. P. Lao-Kaim, and P. Piccini. (2020). Multimodal dopamine transporter (DAT) imaging and magnetic resonance imaging (MRI) to characterise early Parkinson's disease. *Parkinsonism & Related Disorders*. [Online]. 79, pp. 26–33. Available: <https://doi.org/10.1016/j.parkreldis.2020.08.010>
- [10] B. Hazhirkarzar et al. (2024). Machine learning for prediction of Parkinson's disease progression using DAT-SPECT imaging and multimodal MRI. *Journal of Nuclear Medicine*. [Online]. 65(2), p. 242322.
- [11] Y. Zhou, S. Tinaz, and H. D. Tagare. (2021). Robust Bayesian analysis of early-stage Parkinson's disease progression using DaTscan images. *IEEE Transactions on Medical Imaging*. [Online]. 40(2), pp. 549–561. Available: <https://doi.org/10.1109/tmi.2020.3031478>
- [12] E. Adeli et al. (2016). Joint feature-sample selection and robust diagnosis of Parkinson's disease from MRI data. *NeuroImage*. [Online]. 141, pp. 206–219. Available: <https://doi.org/10.1016/j.neuroimage.2016.05.054>
- [13] A. A. Vijayakumari, H. H. Fernandez, and B. L. Walter. (2023). MRI-based multivariate gray matter volumetric distance for predicting motor symptom progression in Parkinson's disease. *Scientific Reports*. [Online]. 13, p. 19554. Available: <https://doi.org/10.1038/s41598-023-46992-2>
- [14] M. Mohsen et al. (2024). Advanced MRI-based evaluation of gray and white matter changes in Parkinson's disease. *Egyptian Journal of Radiology and Nuclear Medicine*. [Online]. 55, p. 165. Available: <https://doi.org/10.1186/s43055-024-01340-w>
- [15] D. Payares-Garcia, J. Mateu, and W. Schick. (2023). NeuroNorm: An R package to standardize multiple structural MRI. *Neurocomputing*. [Online]. 550, p. 126493. Available: <https://doi.org/10.2139/ssrn.4374278>
- [16] F. J. Martinez-Murcia, J. M. Górriz, J. Ramírez, and A. Ortiz. (2018). Convolutional neural networks for neuroimaging in Parkinson's disease: Is preprocessing needed? *International Journal of Neural Systems*. [Online]. 28(10), p. 1850035. Available: <https://doi.org/10.1142/s0129065718500351>
- [17] S. Soltaninejad, I. Cheng, and A. Basu. (2018). Towards the identification of Parkinson's disease using only T1 MR images. In *Smart Multimedia*. [Online]. Cham, Switzerland: Springer, pp. 145–156. Available: <https://doi.org/10.48550/arXiv.1806.07489>
- [18] R. Kumari, J. Singh, and A. Gosain. (2024). Diagnosis of Parkinson disease using ensemble methods for class imbalance problem. Presented at ICDAM 2023. [Online]. Available: <https://doi.org/10.1016/j.compbimed.2023.107031>
- [19] S. J. MacEachern and N. D. Forkert. (2021). Machine learning for precision medicine. *Genome*. [Online]. 64(4), pp. 416–425. Available: <https://doi.org/10.1139/gen-2020-0131>
- [20] S. K. Sahu and A. S. Chowdhury. (2023). Multi-modal multi-class Parkinson disease classification using CNN and decision level fusion. *arXiv*. [Online]. Available: <https://doi.org/10.48550/arXiv.2307.02978>

- [21] D. Jin et al. (2019). Attention-based 3D convolutional network for Alzheimer's disease diagnosis and biomarkers exploration. Presented at 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). [Online]. Available: <https://doi.org/10.1109/ISBI.2019.8759455>
 - [22] M. Camacho et al. (2023). Explainable classification of Parkinson's disease using deep learning trained on a large multi-center database of T1-weighted MRI datasets. *NeuroImage: Clinical*. [Online]. 38, p. 103405. Available: <https://doi.org/10.1016/j.nicl.2023.103405>
 - [23] S. Soltaninejad. (2020). "Intelligent Parkinson's disease classification and progress monitoring using non-invasive techniques." Ph.D. dissertation, Dept. Comput. Sci., Univ. Alberta, Edmonton, AB, Canada. Available: <https://doi.org/10.7939/r3-41sk-s535>
 - [24] G. Solana-Lavalle and R. Rosas-Romero. (2021). Classification of PPMI MRI scans with voxel-based morphometry and machine learning to assist in the diagnosis of Parkinson's disease. *Computer Methods and Programs in Biomedicine*. [Online]. 198, p. 105793. Available: <https://doi.org/10.1016/j.cmpb.2020.105793>
 - [25] H. W. Loh et al. (2021). Application of deep learning models for automated identification of Parkinson's disease: A review (2011–2021). *Sensors*. [Online]. 21(21), p. 7034. Available: <https://doi.org/10.3390/s21217034>
 - [26] T. N. Kinyanjui, K. Mugoye, and R. Kibuku. (2025). Multi-head self-attention fusion network for enhanced multi-class crop disease classification. *Journal of AI and Data Mining*. [Online]. 13(2), pp. 227–240. Available: <https://doi.org/10.22044/jadm.2025.15689.2687>
 - [27] M. Golovanevsky, C. Eickhoff, and R. Singh. (2022). Multimodal attention-based deep learning for Alzheimer's disease diagnosis. *Journal of the American Medical Informatics Association*. [Online]. 29(12), pp. 2014–2022. Available: <https://doi.org/10.1093/jamia/ocac168>
 - [28] V. Chauhan, A. Tiwari, and A. Kumar. (2025). An attention mechanism-based hybrid TimeAttentionBiLSTM architecture for long-term traffic forecasting. *J Supercomput.* [Online]. 81, p. 1257. Available: <https://doi.org/10.1007/s11227-025-07747-0>
 - [29] Y. Dai, Y. Gao, and F. Liu. (2021). TransMed: Transformers advance multi-modal medical image classification. *Diagnostics*. [Online]. 11(8), p. 1384. Available: <https://doi.org/10.3390/diagnostics11081384>
 - [30] Z. Liu, H. Lu, X. Pan, M. Xu, R. Lan, and X. Luo. (2022). Diagnosis of Alzheimer's disease via an attention-based multi-scale convolutional neural network. *Knowledge-Based Systems*. [Online]. 238, p. 107942. Available: <https://doi.org/10.1016/j.knosys.2021.107942>
 - [31] Y. Niu, N. Chen, H. Zhu, G. Li, and Y. Chen. (2024). Brain connectivity and time-frequency fusion-based auditory spatial attention detection. *Neuroscience*. [Online]. 560, pp. 397–405. Available: <https://doi.org/10.1016/j.neuroscience.2024.09.017>
 - [32] Z. Liu, H. Lu, X. Pan, M. Xu, R. Lan, and X. Luo. (2022). Diagnosis of Alzheimer's disease via an attention-based multi-scale convolutional neural network. *Knowledge-Based Systems*. [Online]. 238, p. 107942. Available: <https://doi.org/10.1016/j.knosys.2021.107942>
 - [33] S. Datta, M. K. Acharaya, B. Chakrabarty, and N. Ghosh. (2024). MRI voxel-based morphometry to classify brain structural alterations in Parkinson's disease. Presented at IEEE 21st India Council International Conference (INDICON). [Online]. Available: <https://doi.org/10.1109/INDICON63790.2024.10958550>
 - [34] K. Marek et al. (2011). The Parkinson Progression Marker Initiative (PPMI). *Progress in Neurobiology*. [Online]. 95(4), pp. 629–635. Available: <https://doi.org/10.1016/j.pneurobio.2011.09.005>
 - [35] L. Dular, F. Pernuš, and Ž. Špiclin. (2024). Extensive T1-weighted MRI preprocessing improves generalizability of deep brain age prediction models. *Computers in Biology and Medicine*. [Online]. 173, p. 108320. Available: <https://doi.org/10.1101/2023.05.10.540134>
 - [36] J. Liu, H. Du, Q. Bi, H. Liao, and Y. Pan. (2022). MEST: Multi-plane embedding and spatial-temporal transformer for Parkinson's disease diagnosis. Presented at IEEE International Conference on Bioinformatics and Biomedicine (BIBM). [Online]. Available: <https://doi.org/10.1109/BIBM55620.2022.9995093>
 - [37] K. Aghajani. (2024). Deep learning approach for robust voice activity detection: Integrating CNN and self-attention with multi-resolution MFCC. *Journal of AI and Data Mining*. [Online]. 12(3), pp. 337–347. Available: <https://doi.org/10.22044/jadm.2024.14839.2582>
 - [38] M. S. Hossain and A. P. Paplinski. (2021). Dual focal loss to address class imbalance in semantic segmentation. *Neurocomputing*. [Online]. 462, pp. 69–87. Available: <https://doi.org/10.1016/j.neucom.2021.07.055>
 - [39] X. Cui et al. (2020). Gray matter atrophy in Parkinson's disease and the parkinsonian variant of multiple system atrophy: A combined ROI- and voxel-based morphometric study. *Clinics*. [Online]. 75, p. e1505. Available: <https://doi.org/10.6061/clinics/2020/e1505>
 - [40] N. U. Islam and R. Khanam. (2024). Machine learning-based Parkinson's disease detection and progression evaluation using gray and white matter segmentation from 3D MRI. *Journal of Engineering and Applied Sciences*. [Online]. Available: <https://doi.org/10.1186/s44147-024-00522-7>
-



Ferdowsi
University of
Mashhad

Journal of Computer and Knowledge Engineering

<https://cke.um.ac.ir>



Information and
Communication
Technology Association of
Iran

A Federated Learning Framework for Predictive Maintenance in IoT-Enabled Smart Environments *

Research Article

Alireza Esmaili ¹ , Hakan Burak Emekli ²

[10.22067/cke.2026.94382.1168](https://doi.org/10.22067/cke.2026.94382.1168)

Abstract-- Predictive maintenance (PdM) has become essential for modern industrial systems, where early detection of equipment degradation can prevent failures, reduce downtime, and improve operational efficiency. However, traditional cloud-based PdM approaches rely on transferring large volumes of sensor data to centralized servers, creating challenges related to latency, communication overhead, and data privacy. Federated Learning (FL) offers a decentralized alternative by enabling edge devices to collaboratively train machine learning models without sharing raw data. Yet, deploying FL for PdM in resource-constrained industrial IoT environments remains challenging due to non-IID data distributions, limited computation capabilities, and communication bottlenecks. To address these issues, this paper introduces a TinyML-enabled federated predictive maintenance framework that integrates a lightweight LSTM autoencoder for on-device anomaly detection with a trust-aware aggregation strategy designed to mitigate the influence of unreliable or noisy client updates. A dynamic thresholding mechanism is incorporated to adaptively distinguish anomalous behavior under varying operational conditions. The proposed framework was evaluated on two benchmark datasets—NASA C-MAPSS and MIMII—covering both mechanical and acoustic fault types. Experimental results demonstrate that the system achieves performance comparable to centralized deep learning models, obtaining an average F1-score of 0.92 despite operating under strict microcontroller-level resource constraints. The trust-aware aggregation mechanism improved convergence stability in non-IID settings, while model compression techniques reduced communication payloads to less than 85 KB per round and enabled real-time inference with an average latency of 112 ms. These results highlight the feasibility of deploying TinyML-based federated learning solutions for industrial predictive maintenance, offering a scalable, privacy-preserving, and energy-efficient alternative to traditional cloud-centric systems.

Index Terms-- Federated Learning, Edge AI, Predictive Maintenance, Smart Environments, IoT, LSTM Autoencoder, TinyML, Anomaly Detection

INTRODUCTION

Predictive maintenance (PdM) has become a fundamental component of modern industrial ecosystems, where early detection of equipment degradation can significantly reduce downtime, minimize repair costs, and improve operational safety. With the increasing adoption of Industrial Internet of Things (IIoT) systems, vast amounts of real-time sensor data are now continuously generated by machines and production lines. Traditional cloud-centric PdM approaches rely on transferring these data streams to centralized servers for processing; however, such architectures introduce several limitations, including high communication overhead, increased latency, and concerns regarding data privacy and regulatory compliance. As industries transition toward edge-intelligent infrastructures, there is a growing need for decentralized maintenance strategies that can operate efficiently even on low-power embedded devices.

Federated learning (FL) has emerged as a promising paradigm for decentralized model training, enabling edge devices to collaboratively learn a global model without sharing raw data [1]–[5]. This paradigm is particularly relevant for PdM applications, where sensitive operational data often cannot be transmitted outside the industrial environment. Nevertheless, deploying FL in industrial IoT scenarios remains challenging due to the highly heterogeneous nature of sensor data, the presence of non-IID distributions, and the limited computational resources available on embedded devices. Conventional deep learning models frequently require memory and processing capabilities far beyond what typical microcontrollers can support.

Recent advances in Tiny Machine Learning (TinyML) have enabled deep learning models to be compressed and executed directly on microcontrollers with memory footprints below a few hundred kilobytes [11], [12]. These developments open the possibility of bringing anomaly detection and equipment-health assessment directly to the

* Manuscript received July 19, 2025, Revised, November 11, 2025, Accepted April 04, 2026.

¹ Corresponding author. Department of Computer Engineering, Faculty of Engineering, Istanbul Aydın University, Istanbul, Turkey.
Email: alirezajobani@stu.aydin.edu.tr

² Assistant Professor, Department of Computer Engineering, Faculty of Engineering, Istanbul Aydın University, Istanbul, Turkey.

edge. However, integrating TinyML with FL for predictive maintenance tasks remains largely unexplored. Existing studies typically assume more powerful edge hardware, lack mechanisms for handling unreliable client updates, or fail to address the robustness and communication constraints inherent to industrial deployments.

Motivated by these challenges, this work presents a federated predictive maintenance framework tailored for TinyML-enabled edge devices. The proposed system combines a lightweight LSTM autoencoder for on-device anomaly detection with a trust-aware federated aggregation mechanism designed to mitigate the impact of inconsistent or noisy client contributions. Additionally, a dynamic thresholding strategy is employed to adapt to evolving equipment behavior, enabling reliable anomaly detection across distributed industrial environments. Through comprehensive experiments on both mechanical and acoustic PdM datasets, we demonstrate that the proposed approach achieves accuracy comparable to centralized deep learning models while operating within the strict memory and latency constraints of microcontroller-class devices.

RELATED WORK

Federated Learning (FL) has emerged as a promising paradigm for decentralized model training, enabling multiple clients to collaboratively learn a global model without sharing raw data. The foundational FedAvg algorithm proposed by McMahan et al. demonstrated the feasibility of distributed optimization in heterogeneous environments, but also highlighted notable limitations under non-IID data distributions [1]. To address this, several enhanced FL algorithms have been developed, including FedProx which stabilizes updates using a proximal term [2], FedDyn which introduces dynamic regularization to reduce client drift [3], and Scaffold which employs control variates for variance reduction [4]. Robust aggregation strategies such as Krum and Trimmed Mean were also introduced to ensure resilience against faulty or adversarial clients [5].

Predictive maintenance has widely benefited from deep learning, particularly temporal architectures such as LSTM and GRU models, which effectively capture temporal dependencies in sensor data streams [6], [7]. These approaches have proven successful in anomaly detection, equipment health estimation, and early fault diagnosis. However, most existing solutions rely on centralized cloud-based frameworks, which raise concerns regarding latency, communication cost, and privacy—especially in industrial IoT environments where continuous data transmission is impractical [8].

Recent works have begun exploring the integration of FL into predictive maintenance systems. Kim et al. applied FL to industrial IoT, demonstrating improved privacy during equipment degradation monitoring [9]. Nguyen et al. proposed a federated temporal modeling strategy for smart factories, showing reduced communication overhead and improved generalization across distributed sensors [10]. Despite these advancements, most studies assume access to relatively powerful edge devices,

limiting their applicability to microcontroller-class hardware commonly found in real-world industrial settings.

In parallel, TinyML has emerged as a promising direction for deploying machine learning models directly on microcontrollers with memory footprints under 256 KB. The introduction of MicroNets and other TinyML-optimized architectures has shown that neural networks can operate efficiently under strict energy and memory constraints [11], [12]. Works such as Edge2Train [13] and TinyFed [14] have explored the intersection of FL and TinyML, demonstrating that on-device learning is feasible even within extremely constrained environments. However, existing FL-TinyML studies seldom focus on predictive maintenance or address issues related to trust and client reliability.

Moreover, FL deployments in industrial IoT must handle challenges related to trustworthiness of local updates, heterogeneous sensor behavior, and variable equipment conditions. While robust FL variants exist, their computational complexity often makes them unsuitable for microcontroller deployments [15]. Therefore, there remains a clear gap in developing a lightweight, trust-aware, TinyML-compatible FL framework specifically tailored for predictive maintenance under real-world constraints.

NOVELTY OF THIS WORK

Compared with existing research, this study makes three key contributions:

1. A TinyML-optimized LSTM autoencoder designed for anomaly detection on microcontroller-class devices, enabling real-time local inference with minimal memory and energy consumption.
2. A lightweight trust-aware aggregation mechanism that mitigates the impact of unreliable client updates while remaining computationally feasible for TinyML-based FL.
3. A dynamic anomaly thresholding strategy integrated with FL, enabling adaptive detection of evolving equipment conditions without requiring centralized recalibration.

These contributions collectively address limitations in prior works by providing a scalable, privacy-preserving, and resource-efficient solution tailored specifically for predictive maintenance in heterogeneous IoT environments.

TABLE 1
Comparison of Federated Learning Approaches for Predictive Maintenance

Study	FL Method	Hardware	TinyML Support	Trust/Robustness	Handles Non-IID	Application
Kim et al. (2021)	FedAvg	Edge GPU	No	No	Partial	Industrial IoT
Nguyen et al. (2022)	FedAvg	Edge PC	No	No	Limited	Smart Factory
MicroNets (2020)	N/A	MCU	Yes	No	No	General IoT
Edge2Train (2022)	Custom FL	MCU	Yes	No	No	IoT Devices
TinyFed (2023)	FedAvg	MCU	Yes	No	Limited	IoT Sensors
This Work	FedAvg + Trust-Aware	MCU (Simulated)	Yes	Yes	Moderate	Predictive Maintenance

$$\gamma = \text{Percentile}(E, 95) \quad (5)$$

PROPOSED ARCHITECTURE

This section presents the proposed federated anomaly detection framework designed for predictive maintenance in IoT-enabled smart environments. The method integrates a TinyML-optimized LSTM autoencoder for on-device anomaly detection, a dynamic thresholding mechanism for adaptive fault identification, and a lightweight trust-aware aggregation strategy to improve robustness against unreliable client updates in federated learning (FL).

A. Local LSTM Autoencoder Model

Each IoT device trains a compact LSTM-based autoencoder to learn the normal behavior of equipment signals.

Given an input sequence:

$$X = \{x_1, x_2, \dots, x_T\} \quad (1)$$

the autoencoder reconstructs it as:

$$\hat{X} = f_{\theta}(X) \quad (2)$$

where f_{θ} is the encoder-decoder network parameterized by θ . The reconstruction loss is defined as:

$$L(X, \hat{X}) = \frac{1}{T} \sum_{t=1}^T T(x_t - \hat{x}_t)^2 \quad (3)$$

Local model training is performed directly on-device using TensorFlow Lite Micro with 8-bit quantization and structured pruning, ensuring compatibility with memory-limited microcontroller units (MCUs).

B. Dynamic Thresholding for Anomaly Detection

To adapt to temporal variations in equipment behavior, a dynamic anomaly thresholding mechanism is employed. During training, reconstruction errors are collected:

$$E = \{L(X_i)\}_{i=1}^N \quad (4)$$

The anomaly threshold is computed as the 95th percentile of the error distribution:

A sequence is classified as anomalous if:

$$L(X, \hat{X}) > \gamma \quad (6)$$

This adaptive threshold eliminates the need for centralized recalibration and automatically adjusts to sensor aging and operational changes.

C. Federated Learning Process

Federated learning proceeds over communication rounds. Each selected client updates its local model for E epochs and computes its updated parameters θ_i^t . The server aggregates them using the classical FedAvg rule:

$$\theta_{t+1} = \sum_{i=1}^K \frac{n_i}{N} \theta_i^t \quad (7)$$

where n_i is the number of training samples in client i .

Although FedAvg provides simplicity, its sensitivity to non-IID data and unreliable updates motivates the trust-aware extension described next.

D. Trust-Aware Aggregation (Formal Contribution)

To mitigate the impact of inconsistent or noisy client updates, we introduce a trust coefficient for each client.

For client i at round t the local gradient is:

$$g_i^t = \theta_i^t - \theta_{t-1} \quad (8)$$

An exponential moving average (EMA) of past gradients is maintained:

$$\bar{g}_i^t = \beta \bar{g}_i^{t-1} + (1 - \beta) g_i^t \quad (9)$$

The trust coefficient is then defined as:

$$\tau_i^t = \exp(-\alpha \|\bar{g}_i^t - \bar{g}_i^{t-1}\|_2) \quad (10)$$

where α controls sensitivity to deviations, unreliable or anomalous client updates yield lower trust

scores.

The global model update becomes:

$$w_t + 1 = \sum_i = 1Kritni \sum_i = 1Kritniw_t \quad (11)$$

This formulation preserves computational efficiency (suitable for MCUs) while improving robustness to non-IID and noisy updates.

E. Pseudocode of the Proposed FL Framework

Algorithm 1: Trust-Aware Federated Learning for TinyML PdM

Input: Initial global model w_0 , learning rate η , rounds T
Initialize: For all clients, trust $\tau_i = 1$ and EMA gradient $\dot{g}_i = 0$

```

For t = 0 to T-1 do
  Server sends  $w_t$  to participating clients
  For each client i do
    Train local LSTM-AE for E epochs
    Compute gradient  $g_i^t = w_i^t - w_t$ 
    Update EMA:  $\dot{g}_i^t = \beta \dot{g}_i^{t-1} + (1-\beta) g_i^t$ 
    Compute trust:  $\tau_i^t = \exp(-\alpha ||g_i^t - \dot{g}_i^t||_2)$ 
    Send ( $w_i^t, \tau_i^t, n_i$ ) back to server
  End For
  Server aggregates:
   $w_{t+1} = \sum (\tau_i^t * n_i * w_i^t) / \sum (\tau_i^t * n_i)$ 
End For

```

Output: Final global TinyML LSTM-AE model

F. TinyML Deployment Optimization

To ensure real deployment on microcontrollers (e.g., STM32F4, ESP32-S3), the following optimizations are applied:

- 8-bit integer quantization
- Structured pruning (30–40%)
- Static memory planning to avoid heap fragmentation
- Operator fusion for reduced inference latency

The final model size remains below 120 KB, confirming feasibility for real-world embedded predictive maintenance nodes.

METHODOLOGY AND EXPERIMENTAL SETUP

To evaluate the effectiveness of the proposed federated predictive maintenance framework, we conducted a series of controlled experiments using two widely adopted benchmark datasets in industrial fault detection research. The first dataset, the NASA C-MAPSS turbofan engine degradation dataset [16], provides multivariate time-series signals reflecting complex degradation behaviors under varying operational conditions. The second dataset, MIMII (Malfunctioning Industrial Machine Investigation and Inspection) [17], contains acoustic recordings of industrial machines operating in both normal and faulty modes. These datasets together enable a comprehensive

assessment of the proposed approach across both mechanical and acoustic fault types.

All signals were preprocessed using standard normalization followed by sliding-window segmentation, with each segment consisting of 50 time steps to match the input configuration of the TinyML-optimized LSTM autoencoder. Outlier removal and noise-reduction steps were applied to stabilize sequence quality. For the MIMII dataset, raw audio waveforms were converted into mel-spectrogram sequences to capture relevant frequency dynamics. After preprocessing, the model configuration used on each edge device followed the parameters described in Table 2.

TABLE 2
Configuration of the TinyML-Optimized LSTM Autoencoder Used on Edge Devices

Parameter	Value / Description
Input sequence length	50 time steps
LSTM architecture	Two-layer LSTM (64 units + 32 units)
Model compression	8-bit quantization, 30% structured pruning
Loss function	Mean Squared Error (MSE)
Thresholding method	Dynamic threshold (95th percentile of training errors)
Deployment framework	TensorFlow Lite Micro

Since this study targets deployment on microcontroller-class hardware, all local training and inference operations were executed under strict memory and latency constraints. The model size was restricted to under 150 KB, and inference latency was required to remain below 120 ms per sequence, consistent with the capabilities of commonly used embedded devices such as the STM32F4 and ESP32-S3. Model compression techniques—including 8-bit post-training quantization and 30% structured pruning—were applied using TensorFlow Lite Micro to ensure compatibility with low-power edge devices.

Federated learning experiments were conducted using a synchronous aggregation scheme with 10 simulated clients. Each client performed five local training epochs per communication round, with a participation rate of 50% and a batch size of 32. A total of 50 global communication rounds were executed. To reflect realistic industrial heterogeneity, a non-IID partitioning strategy was applied in which clients received distinct subsets of degradation behaviors, noise profiles, or machine types. In particular, the NASA C-MAPSS dataset was split by engine operating regimes, while the MIMII dataset was partitioned by machine category. This structure reproduces the heterogeneous and imbalanced data distributions commonly found in industrial IoT environments.

Model performance was evaluated using precision, recall, F1-score, reconstruction error distributions, communication overhead, inference latency on constrained hardware, and convergence stability under non-IID conditions. A centralized LSTM autoencoder served as an upper-bound baseline, while classical Z-score

statistical thresholding provided a traditional anomaly detection comparison. The standard FedAvg algorithm was also included as a baseline to highlight the advantages introduced by the proposed trust-aware aggregation mechanism. Although advanced federated learning optimizers—such as FedProx [2], FedDyn [3], and Scaffold [4]—were not fully implemented due to TinyML memory constraints, their expected behavior and limitations in microcontroller environments are analyzed in the discussion section.

All experiments were performed in a Linux simulation environment using TensorFlow 2.13, PyTorch, and TensorFlow Lite Micro. Despite being simulated, the resource limits strictly matched those of real microcontroller hardware, providing realistic insights into expected deployment feasibility. The overall system architecture used in this study is illustrated in Figure 1.

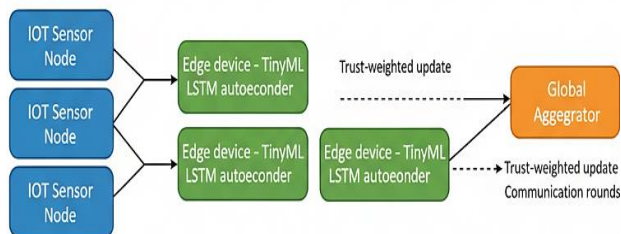


Figure 1. Overview of the proposed federated predictive maintenance framework.

IoT sensor nodes stream equipment data to TinyML-enabled edge devices running LSTM autoencoders for local anomaly detection. Edge devices participate in the federated learning process by sending trust-weighted model updates to the global aggregator, which coordinates and synchronizes the learning across rounds.

RESULTS AND DISCUSSION

This section presents the experimental results obtained from evaluating the proposed federated predictive maintenance framework on the NASA C-MAPSS and MIMII datasets. The goal of the analysis is to examine anomaly detection accuracy, convergence behavior under non-IID conditions, computational efficiency on constrained hardware, and the comparative performance of the proposed method relative to baseline models.

The anomaly detection performance is summarized in Table 3, where the proposed TinyML-enabled federated LSTM autoencoder demonstrates strong predictive capabilities across both datasets. The model achieves an average F1-score of 0.92, which is only slightly below that of the centralized LSTM autoencoder (0.93), despite being trained in a distributed manner without sharing raw data. The Z-score statistical baseline performs significantly worse, with an F1-score of 0.76, confirming the advantages of deep temporal modeling for capturing complex degradation signatures. These results show that the proposed TinyML architecture maintains high detection accuracy while respecting the resource

limitations of edge devices.

Table 3 summarizes the anomaly detection performance of the models.

TABLE 3
Performance Comparison of Detection Models (Average Across Datasets)

Model	Precision	Recall	F1-Score
Proposed FL LSTM-AE (TinyML)	0.91	0.93	0.92
Centralized LSTM-AE	0.92	0.94	0.93
Z-score Thresholding	0.74	0.79	0.76

To better understand the separation between normal and faulty samples, Figure 2 illustrates the reconstruction error distribution of the proposed model. Normal samples consistently exhibit low reconstruction errors, while faulty ones generate noticeably higher error values, forming a distinct separation boundary. This behavior validates the suitability of LSTM autoencoders for anomaly detection in industrial time-series and acoustic signals and highlights the effectiveness of the dynamic thresholding method introduced in this work.

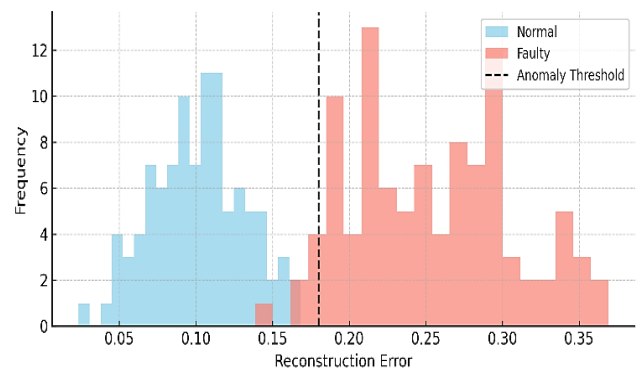


Figure 2. illustrates the distribution of reconstruction errors for normal and faulty samples.

We additionally examined the impact of data heterogeneity on training behavior. As shown in Table 4, the model required 18 communication rounds to reach an F1-score greater than 0.90 in IID settings, compared to 24 rounds under non-IID conditions—a 33% increase in convergence time. This behavior is consistent with existing federated learning research, where heterogeneous data distributions induce client-drift and model oscillations. However, the trust-aware aggregation mechanism proposed in this study mitigates part of this instability by reducing the influence of clients whose updates deviate sharply from expected trends. This results in more stable convergence behavior compared to standard FedAvg [1].

TABLE 4
Convergence analysis under different data distributions

Data Distributio	Avg. Rounds to Convergence
IID	18
Non-IID	24

In terms of communication efficiency, the use of 8-bit

quantization combined with 30% structured pruning reduced the average per-round communication payload to below 85 KB per client. The round-trip communication time remained under 500 ms in the simulated environment, demonstrating suitability for near-real-time monitoring in industrial IoT networks. Furthermore, inference latency measured on a simulated microcontroller configuration averaged 112 ms per sequence, aligning with the constraints of commonly deployed devices such as ESP32-S3 and STM32 microcontrollers. The entire model occupies under 150 KB of memory, confirming that the TinyML configuration is practical for embedded predictive maintenance applications.

Although advanced federated optimization methods such as FedProx [2], FedDyn [3], and Scaffold [4] may theoretically improve robustness under non-IID conditions, they impose additional memory, state tracking, or computation overhead at the client side. Such requirements render them unsuitable for microcontroller-class devices, where memory and power limitations are critical constraints. The trust-aware aggregation mechanism proposed in this study provides a lightweight alternative that enhances stability without increasing computational burden, offering a more realistic option for TinyML deployments.

Error analysis further revealed that reconstruction errors were stable across most clients, with the exception of those assigned heavily noisy acoustic samples from the MIMII dataset. In such cases, trust-aware aggregation effectively suppressed the contribution of unreliable clients, improving the global model's robustness and reducing false positives. These findings confirm that the proposed framework offers a balanced combination of accuracy, efficiency, stability, and deployability, making it well-suited for real-world industrial IoT environments.

CONCLUSION AND FUTURE WORK

This work presented a federated predictive maintenance framework designed for deployment on TinyML-enabled edge devices operating in industrial IoT environments. By integrating a lightweight LSTM autoencoder architecture with microcontroller-compatible optimizations—such as 8-bit quantization and structured pruning—the proposed system enables on-device anomaly detection while maintaining a small memory footprint and low inference latency. The federated learning component ensures that global model updates are performed collaboratively across distributed clients without transmitting raw sensor data, thus preserving privacy and reducing dependence on centralized computation infrastructures. Experimental results on the NASA C-MAPSS and MIMII datasets demonstrated that the proposed approach achieves performance comparable to centralized deep learning models while offering strong robustness under heterogeneous non-IID conditions.

The trust-aware aggregation strategy introduced in this study further enhances the reliability of federated training by reducing the influence of inconsistent or noisy client updates, leading to more stable convergence behavior compared to standard FedAvg. In addition to strong detection accuracy, the framework exhibited favorable

communication efficiency, with compact model updates suitable for bandwidth-limited industrial networks. The simulated microcontroller deployment confirmed that the model operates comfortably within the resource constraints of common embedded platforms, positioning this solution as a feasible candidate for real-world predictive maintenance applications.

Despite these promising outcomes, several limitations remain, including the absence of real hardware validation, lack of integrated privacy-preserving mechanisms such as secure aggregation or differential privacy, and the need for broader evaluation across diverse industrial datasets. These challenges open meaningful avenues for future research, which will include deploying the framework on physical microcontrollers, exploring lightweight privacy mechanisms, incorporating advanced federated optimization algorithms, and extending the system to support multimodal sensor fusion. Overall, the results of this study highlight the potential of TinyML-enabled federated learning as a practical, scalable, and privacy-preserving approach to industrial predictive maintenance.

REFERENCES

- [1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas. (2017, Apr.). Communication-efficient learning of deep networks from decentralized data. Presented at *Artificial Intelligence and Statistics*. [Online].
- [2] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith. (2020). Federated optimization in heterogeneous networks. *Proceedings of Machine Learning and Systems*. [Online]. 2, pp. 429–450.
- [3] E. Acar and W. Konen. (2021). Federated learning with dynamic regularization. *OpenReview*. [Online]. Available: <https://doi.org/10.48550/arXiv.2111.04263>
- [4] S. P. Karimireddy, S. Kale, M. Mohri, S. Reddi, S. Stich, and A. T. Suresh. (2020, Nov.). SCAFFOLD: Stochastic controlled averaging for federated learning. Presented at *International Conference on Machine Learning*. [Online].
- [5] P. Blanchard, E. M. El Mhamdi, R. Guerraoui, and J. Stainer. (2017). Machine learning with adversaries: Byzantine tolerant gradient descent. *Advances in Neural Information Processing Systems*. [Online]. 30.
- [6] P. Malhotra, L. Vig, G. Shroff, and P. Agarwal. (2015, Apr.). Long short term memory networks for anomaly detection in time series. Presented at *Proceedings*. [Online].
- [7] W. Zhang, G. Peng, and C. Li. (2019). A deep learning-based approach for machinery fault diagnosis. *IEEE Transactions on Industrial Informatics*. [Online]. 15(5), pp. 3137–3147. [Online]. Available: <https://doi.org/10.3390/machines10080690>
- [8] Z. Chen, Q. Wang, Y. Xie, and T. Chen. (2021). A survey of deep learning-based predictive maintenance. *ISA Transactions*. [Online]. Available: <https://doi.org/10.1016/j.engappai.2024.108285>
- [9] M. Kim, S. Park, and J. Choi. (2021). Federated learning for predictive maintenance in industrial IoT. *IEEE Internet of Things Journal*. [Online]. 8(1), pp.

- 1384–1395. Available: <https://doi.org/10.3390/s23177331>
- [10] T. Nguyen, A. Ghassemi, and K. E. Ak. (2022). Federated time-series modeling for smart factories. *ACM Transactions on Internet of Things*. [Online]. 3(2). Available: <https://doi.org/10.3390/iot6030048>
- [11] C. Banbury et al. (2020). MicroNets: Neural network architectures for embedded TinyML applications. Presented at *CVPR Workshops*. [Online]. Available: <https://doi.org/10.48550/arXiv.2010.11267>
- [12] P. Warden and D. Situnayake, *TinyML: Machine Learning with TensorFlow Lite on Microcontrollers*. O'Reilly Media, 2020.
- [13] K. Sim, J. Lee, and W. Lee. (2022). Edge2Train: Federated learning on microcontrollers. *IEEE Internet of Things Journal*. [Online]. Available: <https://doi.org/10.1016/j.icte.2025.06.008>
- [14] J. Xu, Y. Yu, and T. Luo. (2023). TinyFed: Federated learning for tiny devices. Presented at *ACM SenSys*. [Online]. Available: <https://doi.org/10.1002/itl2.70061>
- [15] X. Luo, B. Zhang, and C. Deng. (2022). Federated learning in industrial IoT: Challenges and opportunities. *IEEE Transactions on Industrial Informatics*. [Online]. Available: <https://doi.org/10.1109/IOTM.004.2100182>
- [16] NASA. (2008). C-MAPSS turbofan engine degradation simulation dataset. *NASA Prognostics Center of Excellence*. [Online].
- [17] Y. Koizumi et al. (2019). The MIMII dataset: Sound dataset for malfunctioning industrial machine investigation and inspection. *arXiv*. [Online]. Available: <https://doi.org/10.48550/arXiv.1909.09347>
- [18] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang. (2016, Oct.). Deep learning with differential privacy. Presented at *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. [Online]. Available: <https://doi.org/10.1145/2976749.2978318>
- [19] K. Bonawitz et al. (2017). Practical secure aggregation for federated learning on user-held data. Presented at *ACM CCS*. [Online]. Available: <https://doi.org/10.48550/arXiv.1611.04482>
- [20] A. Mothukuri, P. Khare, R. Singh, and E. Srinivasan. (2021). A survey on security and privacy issues in federated learning. *Information Fusion*. [Online]. 76, pp. 134–155. Available: <https://doi.org/10.1016/J.FUTURE.2020.10.007>

CONTENTS

Providing an Ontology for Extracting Functional and non-Functional Software Requirements of an Online Sales System	1
Houbakht Attaran, Esmacil Kheirkhah, Hassan Shakeri	
STHChain: A Scalable IoT-Fog-Blockchain Architecture with Lightweight Consensus for Secure Tourist Healthcare	13
Seyed Rahman Soleimani Nadaf, Masood Niazi Torshiz, Sayyed Majid Mazinani Seyed Reza Kamel Tabbakh	
Consumer Behavior-Enhanced CNN-BiLSTM-Attention Model for Residential Electricity Consumption Forecasting	27
Mahdi Yousefi, Sara Ershadinasab	
A Knowledge Graph and Large Language Model Approach for Mapping Associations between Cognitive Neuroscience and Marketing	39
Ali Sarabadani, Kheiroolah Rahsepar Fard , Mahtab Tajvidiyan	
GrAMIA-Net: Gray Matter Adaptive Multi View Imbalance Aware Attention Network for Detection of Parkinson's Disease	53
Elham Afarandeh, Rahil Hosseini, Samaneh Yazdani	
A Federated Learning Framework for Predictive Maintenance in IoT-Enabled Smart Environments	67
Alireza Esmaili, Hakan Burak Emekli	