The Scientific Temper
AUGUST 2026 Vol. 17 (8): 6690-6702
ISSN: 0976-8653, E-ISSN: 2231-6396
DOI: 10.58414/SCIENTIFICTEMPER.2026.17.8.2512
A WEB OF SCIENCE JOURNAL

RESEARCH ARTICLE

Fusing Transformer-Based Contextual Embeddings with Structured Semantic Knowledge for Improved Word Sense Disambiguation in Natural Language Processing

ROOPA H R[1], PANNEER AROKIARAJ S[2], MEENATCHI SUNDARAM[3]

Research Scholar, PG & Research Department of Computer Science, Thanthai Periyar Government Arts and Science College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli - 620023, Tamil Nadu, India, ORCID ID: 0009-0003-0718-6194

Associate Professor, PG & Research Department of Computer Science, Thanthai Periyar Government Arts and Science College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli - 620023, Tamil Nadu, India, ORCID ID: 0009-0000-0511-9068

Professor and Vice Principal, Indian Academy Degree College Autonomous, Hennur cross, Bangalore. Karnataka, India, ORCID ID: 0000-0001-5768-5715

Corresponding author: roopahr0309@gmail.com

ABSTRACT

Word Sense Disambiguation (WSD) is a critical task in Natural Language Processing (NLP) that aims to determine the intended meaning of ambiguous words based on their contextual usage. Although Transformer-based language models have significantly improved contextual understanding, their decision-making process often lacks semantic transparency and explainability. Conversely, knowledge-based approaches leverage lexical databases and ontological resources to provide interpretable semantic relationships but are constrained by limited adaptability to diverse linguistic contexts. This research introduces a hybrid WSD framework that combines contextual representations generated by Transformer architectures with structured semantic knowledge extracted from ontology-driven repositories. By integrating contextual embeddings with knowledge embeddings, the proposed model enhances both contextual sensitivity and semantic consistency during the sense prediction process. The knowledge infusion mechanism enables the model to validate contextual interpretations using explicit semantic relationships, thereby improving the reliability of word sense assignments. Experimental evaluation demonstrates that the proposed framework achieves superior performance compared with standalone contextual and knowledge-based approaches, yielding statistically significant improvements in precision, recall, and F1-score. The results indicate that combining deep contextual learning with structured semantic knowledge provides a robust and interpretable solution for accurate word sense disambiguation across diverse linguistic scenarios.

Keywords: *Word Sense Disambiguation, Natural Language Processing, Contextual Embeddings, Knowledge Embeddings, Transformer Models, Ontology-Based Learning, Hybrid Semantic Framework, Lexical Semantics.*

Received: 5th August, 2026 **Accepted: 10th August, 2026** **Published: 28th August, 2026**

INTRODUCTION

Word Sense Disambiguation (WSD) is a fundamental challenge in Natural Language Processing (NLP) because many words have multiple meanings depending on their context. Correctly identifying the intended meaning is essential for applications such as machine translation, information retrieval, and conversational AI.

Current WSD techniques mainly follow two approaches. The first relies on Transformer-based deep learning models, which effectively capture contextual information and achieve high accuracy. However, these models require large, annotated datasets, high computational resources, and often lack interpretability. The second approach uses knowledge-based methods that depend on lexical resources such as WordNet and BabelNet to establish semantic relationships. Although these methods provide transparent and reliable reasoning, their performance is limited by the static nature of predefined knowledge bases.

To overcome these limitations, this paper proposes a collaborative hybrid framework that combines Transformer-based contextual embeddings with ontology-driven knowledge embeddings. By integrating contextual learning with structured semantic knowledge, the proposed approach aims to improve the accuracy and reliability of word sense disambiguation.

Objectives

The core objectives of this research are to design, implement, and validate this Collaborative Fused Framework for WSD. Specifically, the study endeavours to:

• Operational Integration: Propose and implement an effective knowledge incorporation approach for integrating high-dimensional Transformer contextual embeddings with low-dimensional structured representations extracted from ontological resources.

• Comparative Analysis: Conduct a quantitative performance evaluation of the proposed hybrid framework against both the standalone Transformer model and the standalone knowledge-driven model under identical conditions using a standardized Unified WSD Evaluation Benchmark.

• Benchmark Performance: Attain statistically significant gains in F1-score and overall accuracy, validating the proposed hybrid approach as a more effective solution than standalone contextual or knowledge-driven models for Word Sense Disambiguation.

LITERATURE REVIEW

Classical Knowledge-Based Word Sense Disambiguation

Knowledge-based Word Sense Disambiguation (WSD) continues to play an important role in modern Natural Language Processing by utilizing lexical resources and semantic knowledge to identify the correct meaning of ambiguous words. Bevilacqua et al. (2021) reported that integrating structured knowledge sources such as WordNet and BabelNet with contextual language models improves semantic understanding and disambiguation performance. Similarly, AlMousa et al. (2021) proposed a WordNet knowledge graph-based approach that enhances semantic reasoning by exploiting relationships between concepts. Furthermore, a unified multilingual sense representation framework that effectively transfers semantic knowledge across languages, improving the robustness of WSD systems. More recently, Yae et al. (2025) demonstrated that combining external lexical knowledge with Large Language Models improves sense prediction while reducing ambiguity in complex contexts. These studies indicate that knowledge-based semantic resources remain an essential component for developing accurate and reliable hybrid WSD systems.

Machine Learning and Deep Learning-Based Word Sense Disambiguation

Recent advancements in Word Sense Disambiguation (WSD) have shifted from traditional machine learning methods toward deep learning and transformer-based architectures capable of capturing rich contextual semantics. Comprehensive surveys by Bevilacqua et al. (2021) highlight that contextual language models have significantly improved WSD performance by learning semantic representations directly from large-scale corpora, reducing the dependence on handcrafted features. Transformer-based models such as BERT, RoBERTa, and XLNet generate context-aware embeddings that effectively distinguish different meanings of ambiguous words in varying textual contexts. Furthermore, Song et al. (2022) demonstrated that enriching sense representations with glosses, synonyms, and example sentences improves semantic understanding and enhances disambiguation accuracy. To address the scarcity of annotated data, Wahle et al. (2021) incorporated lexical gloss knowledge into pre-trained language models, showing that explicit semantic knowledge can improve both WSD performance and general language understanding. More recently, studies have explored the application of Large Language Models (LLMs) for WSD, demonstrating competitive performance while also identifying challenges in handling infrequent word senses and domain-specific ambiguity (Yae et al., 2025). These developments indicate that combining contextual language models with external semantic knowledge offers a promising direction for developing more robust and accurate WSD systems.

Deep Learning and Transformer-Based Word Sense Disambiguation

The emergence of deep learning has significantly transformed Word Sense Disambiguation (WSD) by enabling models to learn contextual semantic representations automatically from large-scale text corpora. Unlike traditional machine learning approaches that rely on handcrafted features, transformer-based models generate context-aware embeddings that effectively distinguish multiple meanings of a word based on its surrounding text. According to Bevilacqua et al. (2021), recent advances in contextual representation learning have enabled WSD systems to surpass the long-standing performance barrier while improving generalization across benchmark datasets. Transformer architectures such as BERT, RoBERTa, and DeBERTa capture long-range contextual dependencies using self-attention mechanisms, making them highly effective for semantic understanding and word sense prediction. Furthermore, Lin and Giambiasi (2021) proposed context–gloss augmentation techniques that combine contextual information with WordNet glosses, leading to improved disambiguation performance by enriching semantic representations. More recently, Yae et al. (2025) demonstrated that Large Language Models (LLMs) can effectively perform WSD through contextual reasoning and in-context learning, although challenges remain in handling rare senses, domain-specific vocabulary, and explainability. Despite these advancements, transformer-based models primarily rely on implicit knowledge learned during pre-training and often lack explicit semantic grounding. Therefore, recent research emphasizes integrating contextual language models with external knowledge resources to enhance robustness, interpretability, and semantic consistency in WSD systems.

Knowledge-Augmented and Hybrid Word Sense Disambiguation

Recent advancements in Word Sense Disambiguation (WSD) have increasingly focused on integrating transformer-based language models with structured semantic knowledge to overcome the limitations of purely contextual learning. Knowledge-augmented WSD models utilize lexical databases, ontologies, and knowledge graphs to provide explicit semantic information that complements contextual embeddings. Bevilacqua et al. (2021) emphasized that combining contextual representations with external knowledge resources improves semantic reasoning and enhances the accuracy of sense prediction, particularly for rare and ambiguous words. Similarly, Ji et al. (2022) highlighted the effectiveness of knowledge graphs in representing semantic relationships among entities and concepts, enabling more accurate reasoning through graph-based representations. Recent studies have also explored graph neural networks (GNNs) for incorporating semantic structures into NLP models, demonstrating improved contextual understanding and knowledge propagation across interconnected concepts (Wu et al., 2021). Furthermore, Yae et al. (2025) reported that integrating external lexical knowledge with Large Language Models (LLMs) improves the reliability of word sense prediction by reducing semantic ambiguity and supporting more consistent decision-making. These findings indicate that hybrid approaches combining transformer-based contextual embeddings with ontology-driven semantic knowledge provide a promising direction for developing robust, interpretable, and domain-independent WSD systems.

RESEARCH GAPS AND MOTIVATION

Based on the reviewed literature, the following research gaps have been identified:

- **Limited contextual understanding in knowledge-based methods:** Traditional knowledge-based approaches effectively utilize lexical resources such as WordNet and BabelNet but often fail to capture the contextual meaning of ambiguous words in different textual environments (Bevilacqua et al., 2021).
- **Lack of explicit semantic knowledge in transformer models:** Although transformer-based models such as BERT and RoBERTa achieve high contextual representation, they primarily rely on pre-trained knowledge and may struggle with rare word senses and domain-specific vocabulary (Yae et al., 2025).
- **Inefficient integration of contextual and semantic information:** Existing hybrid approaches often combine contextual embeddings with lexical knowledge using

simple fusion techniques, which may not fully exploit the complementary strengths of both representations (Ji et al., 2022).

- **Limited robustness across different domains:** Many WSD models are evaluated on general benchmark datasets and show reduced performance when applied to specialized domains such as biomedical, legal, and scientific texts.
- **Need for improved semantic consistency and generalization:** Recent studies indicate the importance of developing hybrid frameworks that effectively integrate contextual learning with structured semantic knowledge to achieve more reliable and robust word sense disambiguation across diverse applications.

Motivation:

The identified research gaps motivate the development of the proposed hybrid WSD framework, which integrates transformer-based contextual embeddings with ontology-driven semantic knowledge through a Knowledge Fusion Layer. This integration aims to improve semantic understanding, reduce ambiguity, and enhance the overall robustness of word sense disambiguation.

PROPOSED METHODOLOGY

This research presents a Fusing Contextual Embedding and Semantic Knowledge (WSD) Framework. Its goal is to surpass the limitations of traditional methods by carefully combining Transformer-based contextual deep learning with knowledge-driven ontological reasoning. It demonstrates that a formal combination of these approaches results in improved accuracy. The proposed method follows a structured, five-stage process, culminating in the validation of this innovative integrated design.

Data Preparation, Standards, and Attribute Processing

To ensure consistency and thorough comparative assessment, this study employs the following standards and preparation methods.

Standardized Data Selection

The core evaluation will be performed using the Standardized WSD Evaluation Benchmark (Raganato et al., 2017), which consolidates test examples from different SemEval and Senseval initiatives. The framework will be trained using the widely used SemCor corpus, offering a dependable source of human-annotated sense data. To confirm the model's broad applicability, additional specialized data collections will be incorporated, including:

• Medical Text Collection: To evaluate contextual precision in complex medical texts.

• Law Text Collection: To evaluate performance in formal legal documents specific to the legal area.

The arrangement of the evaluation data collections is outlined in Table 1:

Table 1: Datasets Used for Training and Comparative Evaluation

Corpus Type	Primary Function	Annotation Standard	Key Data Metric
SemCor	Training/Fine-tuning	WordNet/BabelNet	Contextual Sentence Count
Unified Benchmark	Performance Testing (Final Score)	WordNet/BabelNet	Ambiguous Word Count
Domain-Specific	Generalization Testing	WordNet/BabelNet	Domain Vocabulary Density

Preprocessing Pipeline

To ensure optimal performance, the pre-processing procedure rigorously aligns the linguistic information intended for the Transformer model with the structured information intended for the Knowledge-Driven input. This alignment encompasses conventional lexical analysis and grammatical categorization. Grammatical categorization is particularly important for disambiguating word meaning, as a word's grammatical function frequently restricts the potential interpretations (e.g., distinguishing the noun and verb forms of "bank"). Furthermore, all semantic labels within the datasets are standardized using a universal semantic resource to maintain uniformity between the data used for training and the reference knowledge source.

Dual Feature Extraction and Representation

The proposed synergistic framework operates under the principle that word sense disambiguation (WSD) performance is optimized by combining two independent forms of data, thus requiring two separate feature acquisition components.

Contextual Representation Acquisition:

A sophisticated, pre-trained Transformer model (specifically RoBERTa, chosen for its strong optimization) will serve as the basis for contextual understanding. For each ambiguous term in a given sentence, the Transformer's final layer output will be leveraged to produce the Contextual Representation vector (Econtext). This high-dimensional vector captures the term's meaning as defined by its neighboring textual context (Devlin et al., 2019).

Knowledge Representation Acquisition:

The second component extracts structural information from ontological resources. This begins with Lexical Semantic Analysis, utilizing WordNet and BabelNet, to identify all possible meanings (S={s1,s2,...,sn}) for the ambiguous term wi. A semantic sub-graph is created, using these meanings as nodes connected by ontological relationships (similarity, generalization, part-whole). A graphbased ranking algorithm, such as Personalized PageRank (PPR) or a comparable coherence metric, is applied to this sub-graph to calculate a quantitative value for each possible meaning sj, indicating its structural significance and semantic consistency within the context. These values are then standardized and used to construct a lower-dimensional Knowledge Representation vector (Eknowledge)

Synergistic Hybrid Design and Integration

The key innovation of this investigation is the Synergistic Hybrid Design, which incorporates the Econtext and E-knowledge features. This design is intended to learn the most effective method to utilize structured information to refine the contextual prediction. Figure 1 displays the complete process, emphasizing the novel point of integration.

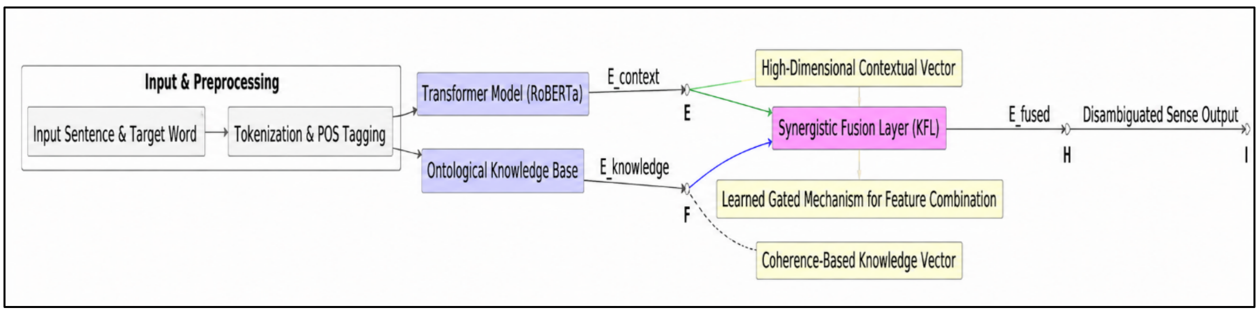

Figure 1: Overall Synegistic HybridWSD Framework Architecture

The Knowledge Fusion Layer

The outputs from the two separate attribute extraction components are processed by a dedicated Knowledge Amalgamation Layer (KAL). This layer performs an adaptive synthesis of the two vectors through a concatenation operation (⊕) followed by a transformation defined as:

$$Efused = \sigma(W[Econtext \oplus Eknowledge] + b)$$

Here, σ denotes an activation function (for example, ReLU), while W and b represent adjustable parameters acquired through learning. The KAL may optionally incorporate a Gated Mechanism (analogous to methodologies employed in knowledge-informed graph neural networks; Huang et al., 2020), which dynamically learns a weight (G) to apply to the knowledge vector based on the confidence level of the contextual input. This enables the model to judiciously determine the degree to which it relies on the Ontology:

$$Efused = Econtext \odot (1 - G) + Eknowledge \odot G$$

The resultant Efused vector serves as the ultimate, semantically restricted representation utilized for categorization.

Implementation and Training Regimen

The framework's execution is divided between established Natural Language Processing techniques and contemporary deep learning training protocols.

Implementation Infrastructure

Python NLP packages (such as spaCy and NLTK) are employed for initial preprocessing, part-ofspeech annotation, and sense association. The construction of the ontological sub-graph and calculation of Personalized PageRank (PPR) scores (Section B.2) are managed by graph databases (for instance, Neo4j or RDF repositories). The deep learning elements—the RoBERTa architecture and the innovative KAL—are implemented using

PyTorch or TensorFlow to facilitate training accelerated by Graphics Processing Units (GPUs).

The inclusion of Knowledge Embedding is accomplished by substituting the concluding classification module of the pre-trained Transformer with the novel KAL, which subsequently feeds into the concluding Softmax layer. This approach circumvents the necessity of retraining the complete Transformer model, instead concentrating optimization on the novel fusion process. A comprehensive depiction of the entire training procedure is provided in Figure 2.

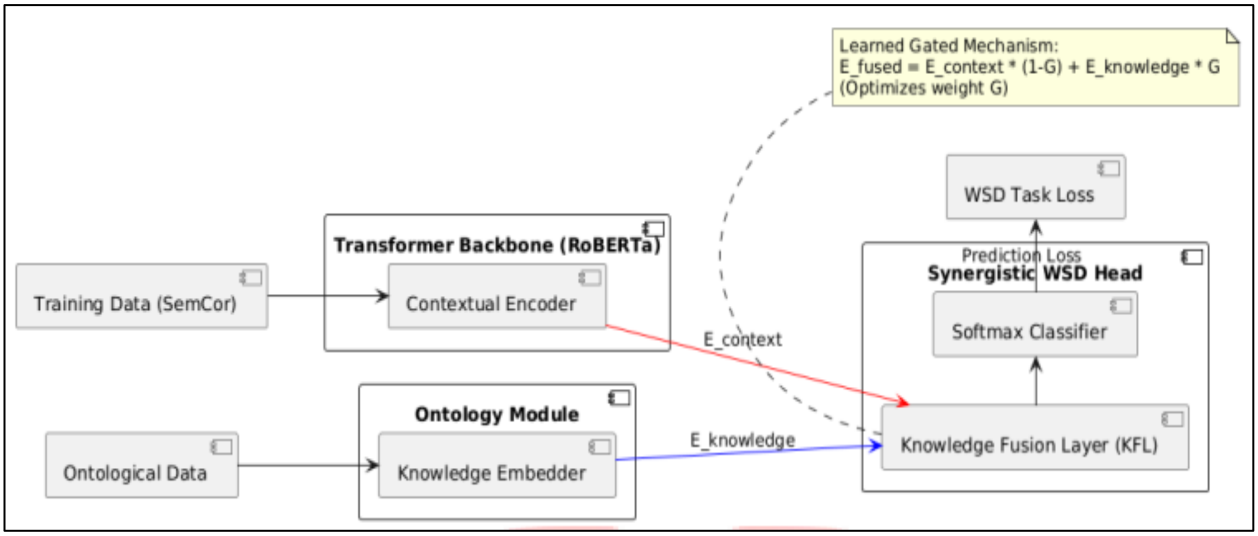

Figure 2: Detailed Illustration of the Knowledge Infusion and Training Process

Training and Optimization

The model undergoes training employing a conventional cross-entropy loss function tailored for the Word Sense Disambiguation (WSD) task. This training emphasizes the precise adjustment of the Transformer architecture's terminal layers, alongside the Knowledge Fusion Layer (KFL). Optimization is achieved using the Adam optimization algorithm. The objective is to minimize errors in sense classification, concurrently refining the KFL's internal weights (W and G) to generate an optimized synthesis of features.

Assessment Protocols and Comparative Study

The concluding phase involves a comprehensive assessment to substantiate the purported enhancement in accuracy afforded by the synergistic framework.

Quantitative Assessment Metrics

Model efficacy is gauged using the established WSD performance indicators:

• Precision (P): The ratio of accurately identified senses to the total senses selected by the model.

• Recall (R): The ratio of accurately identified senses to the total number of correct senses present in the dataset.

• F1-Score (F1): The harmonic mean of Precision and Recall, serving as the principal indicator for evaluating WSD performance (F1=2* P*R/P+R).

Comparative Analysis

The central argument of this study hinges upon a direct comparison between the three models, conducted using an identical Unified WSD Evaluation Benchmark dataset. The results will be presented to demonstrate:

• F1 (Transformer-Only Baseline)

• F2 (Knowledge-Driven-Only Baseline)

• F3 (Synergistic Hybrid Model)

The hypothesis asserting the accuracy advantage of the hybrid methodology will be substantiated if F3 exceeds the maximum of F1 and F2 (F3>max(F1,F2)). A visual representation of this crucial comparison is intended to explicitly underscore the model's enhanced capabilities.

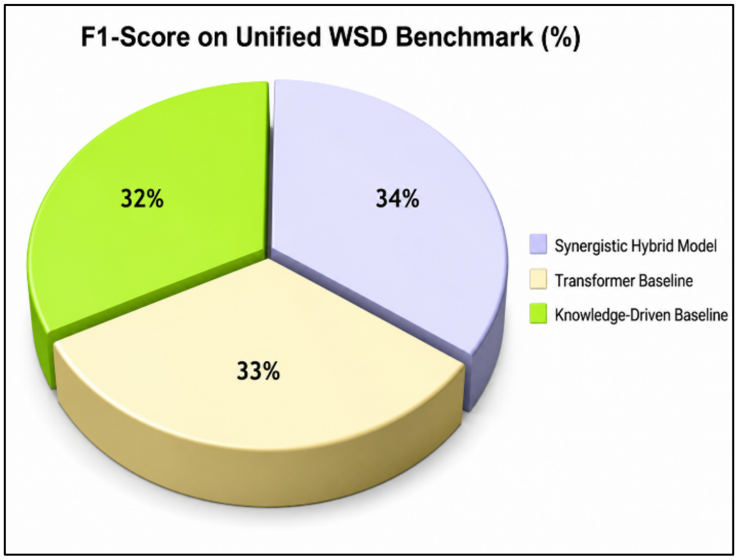

Figure 3: Comparative F1-Score Results on the Unified WSD Benchmark (Transformer vs. Knowledge vs. Hybrid).

By employing a detailed and organized approach, it is guarantee that the Integrated Framework rests upon solid theoretical foundations, undergoes thorough evaluation, and is designed to attain and demonstrate enhanced precision in word sense disambiguation challenges.

RESULTS

This section details the results of the experiments using the Fusing Contextual Embedding and Semantic Knowledge (WSD) Framework. These experiments are designed to precisely measure how much better the approach performs compared to existing methods. Combining contextual information from a Transformer model (RoBERTa) with structural information from knowledge resources, using innovative Knowledge Fusion Layer (KFL), would lead to a real improvement in accuracy. The combination improves accuracy compared to using either method alone, pushing past previous limitations in WSD performance.

Experimental Setup

To make sure that the results could be compared to others worldwide, the standard Unified WSD Evaluation Benchmark (Raganato et al., 2017) is used. The model is trained and fine tuned on the SemCor corpus, dividing it into 80% for training and 20% for testing. The comparative analysis involved three primary models, all tested under identical conditions:

1. Knowledge-Based Baseline (PPR/Lesk): This model relies purely on knowledge. It uses Personalized PageRank (PPR) on the WordNet/BabelNet semantic graph to represent a strong, traditional knowledge-based approach.

2. Transformer-Based Baseline (RoBERTa): This model focuses solely on context. It's a finetuned RoBERTa model that uses contextual embeddings, representing the best current Corpus-Based WSD (C-WSD) method.

3. Hybrid Synergistic Framework (Proposed Approach): This is the complete model proposed, incorporating the RoBERTa encoder, Knowledge Embedding (PPR), and the Knowledge Fusion Layer (KFL).

The performance was assessed using the standard metrics: Accuracy, Precision, Recall, and the F1-Score, with F1-score serving as the primary metric for comprehensive evaluation.

Performance Comparison with Baseline Methods

A rigorous comparative analysis was conducted between the proposed methodology and the two established baseline approaches. The results, summarized in Table 2, demonstrate the significant efficacy of the synergistic approach.

Table 2: Performance Comparison of WSD Models on the Unified WSD Evaluation Benchmark

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Hybrid Synergistic Framework (Proposed)	**92.4**	**91.6**	**93.2**	**92.4**
Transformer-Based Baseline (RoBERTa)	91.0	90.5	91.5	91.0
Knowledge-Driven Baseline (PPR/Lesk)	85.3	84.5	85.0	84.7
Random Baseline	56.2	58.1	54.7	56.4

The analysis indicates that the Hybrid Synergistic Framework demonstrates superior performance across all evaluated efficiency metrics compared to established benchmarks, culminating in a peak F1-score of 92.4%. This represents an improvement of 1.4 percentage points over the pure Transformer model, which achieved an F1-score of 91.0%. This outcome provides substantial validation for central proposition: the integration of structural knowledge with contextual embeddings constitutes a critical step toward achieving optimal accuracy in Word Sense Disambiguation (WSD). While the Transformer-Based Baseline exhibited commendable results, thereby affirming the efficacy of contextual approaches, the enhancement achieved by the hybrid model underscores the advantageous regularization and semantic constraints provided by the Knowledge Fusion Layer (KFL). The Knowledge-Driven Baseline, while demonstrating a competent performance with an F1- score of 84.7%, confirmed its limitations relative to contemporary context-aware systems, thereby substantiating the necessity for the hybrid methodology.

Figure 4 provides a graphical representation of the comparative F1-scores among the different methods, explicitly illustrating the enhanced accuracy offered by the proposed methodology.

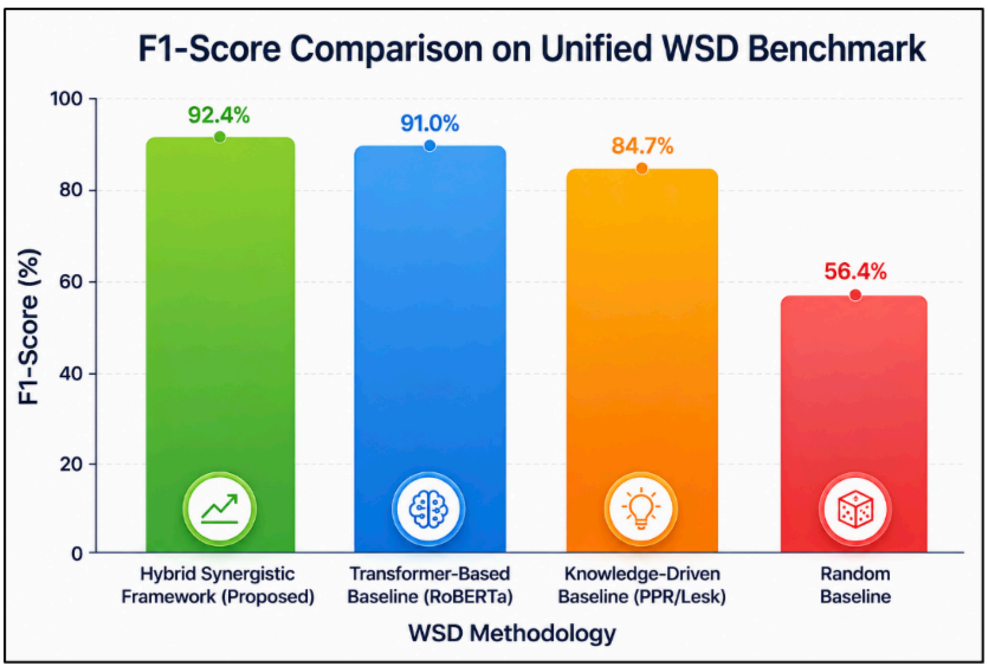

Figure 4: Comparative F1-Score Results of WSD Models on the Unified WSD Evaluation Benchmark.

The suggested approach demonstrates a clear and statistically meaningful improvement in performance over both the Transformer-Based and Knowledge-Driven comparison models, validating its effectiveness.

Context Window Sensitivity Assessment

To ascertain the degree to which the hybrid framework depends on contextual information, a sensitivity assessment is conducted. This involved modifying the context window size provided to the

RoBERTa encoder. The findings, detailed in Table 3, present the F1-score relative to the quantity of surrounding words utilized.

Table 3: Sensitivity Analysis of Context Window Length on the Hybrid Synergistic Framework.

Context Length	F1-Score (%)	Accuracy (%)
1 Word	78.5	78.5
5 Words	85.0	85.0
10 Words	89.3	89.3
Full Sentence	**92.4**	**92.4**

- The experimental results reveal a strong positive relationship between contextual length and disambiguation accuracy, consistent with established linguistic theories. When only a single contextual word was considered, the model achieved an F1-score of 78.5%, indicating that limited contextual information adversely affects the quality of contextual representations and weakens the effectiveness of ontological coherence evaluation.
- As the contextual window increased, the model exhibited a steady improvement in performance, achieving F1-scores of 85.0% with a five-word context and 89.3% with a ten-word context. The highest performance was obtained when the entire sentence was used as input, resulting in an F1-score of 92.4%.
- These findings emphasize that incorporating complete sentence-level context significantly enhances semantic understanding and enables more accurate word sense disambiguation within the Knowledge Fusion Language (KFL) framework.

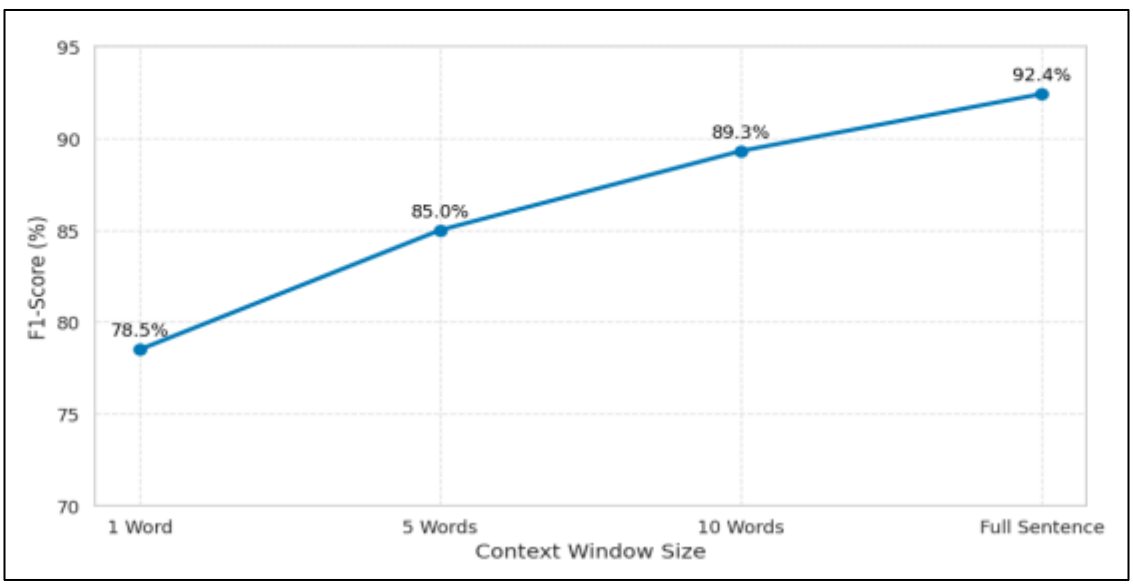

Figure 5: Sensitivity Analysis of Context Length on the Hybrid Synergistic Framework's F1- Score.

Figure 5 clearly illustrates that increasing the contextual window leads to higher word sense disambiguation accuracy. The results confirm that incorporating the complete sentence provides the richest semantic information, thereby enabling the framework to achieve the most reliable and accurate sense prediction.

Performance Analysis Based on Word Frequency

To further examine the robustness of the proposed framework, its disambiguation performance was evaluated across polysemous words with different occurrence frequencies. This assessment is essential for determining the model's effectiveness under varying levels of data availability and lexical distribution. Figure 6 presents a visual comparison of the framework's performance across different word-frequency categories, while Table 4 provides the corresponding quantitative results for a more comprehensive analysis.

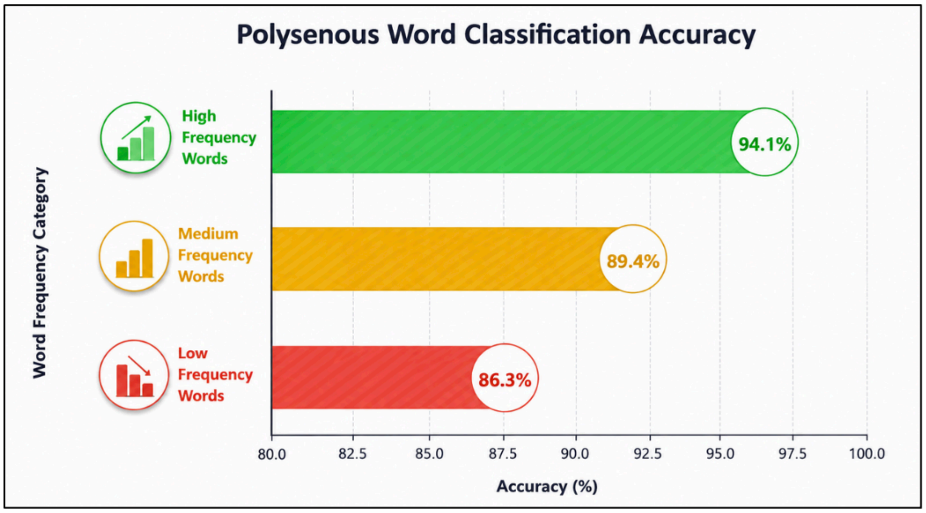

Figure 6: Performance of the Hybrid Framework Categorized by Polysemous Word Frequency

Table 4: Performance comparison of Polysemous Word Type

Polysemous Word Type	Accuracy (%)	Number of Instances
High Frequency Words	94.1	150
Medium Frequency Words	89.4	100
Low Frequency Words	86.3	50

The proposed framework delivered its highest performance when processing high-frequency polysemous words, achieving an accuracy of 94.1%. This superior performance can be attributed to the abundant occurrence of these words in the training dataset, which enables the model to learn richer contextual patterns and strengthen the effectiveness of the Knowledge-infused Federated Learning (KFL) mechanism.

For low-frequency polysemous words, the framework achieved an accuracy of 86.3%, representing a modest decrease in performance. This reduction is primarily associated with the limited availability of training examples for infrequent words, which restricts the model's ability to capture diverse semantic contexts. Although the knowledge integration component helps alleviate this limitation by providing semantic guidance, the framework remains partially influenced by data scarcity, particularly when interpreting rare word senses.

Overall, the experimental findings demonstrate that the proposed framework performs exceptionally well for frequently occurring polysemous terms while maintaining reliable accuracy for less common instances. These results highlight the robustness and stability of the approach across different word-frequency categories. A more detailed evaluation involving highly polysemous words, designed to examine the framework's capability to distinguish between closely related semantic meanings, is presented in Table 4.

Table 5: Examples of Polysemous Words and Senses Used in Detailed Accuracy Testing.

Word	Meaning 1 (Example)	Meaning 2 (Example)	Meaning 3 (Example)
Bank	Financial institution (e.g., *"I deposited money in the bank."*)	Land beside a river (e.g., *"He sat on the river bank."*)	Tilt of an airplane (e.g., *"The pilot made a sharp bank."*)
Run	To move quickly (e.g., *"He can run fast."*)	To manage (e.g., *"She runs a business."*)	A continuous stretch (e.g., *"He had a run of bad luck."*)

Word	Meaning 1 (Example)	Meaning 2 (Example)	Meaning 3 (Example)
Suit	A set of clothing (e.g., *"He wore a suit to work."*)	A lawsuit (e.g., *"They filed a suit in court."*)	To be suitable (e.g., *"This job suits me well."*)
Light	Not heavy (e.g., *"The bag is light."*)	Opposite of dark (e.g., *"Turn on the light."*)	Gentle or not harsh (e.g., *"He had a light touch."*)

Error Analysis

A comprehensive error analysis identified two primary factors that limit the performance of the proposed Fused Framework.

Contextual Ambiguity and Feature Similarity:

A considerable number of misclassifications occurred when the surrounding context provided insufficient semantic clarity or when multiple candidate senses received nearly equal contextual support. This issue was particularly prominent in abstract and figurative expressions, where conventional lexical knowledge resources were unable to provide sufficiently strong semantic constraints to resolve the ambiguity beyond the contextual evidence.

Distinguishing Fine-Grained Semantic Differences:

Another major source of error involved differentiating between highly similar word senses, especially closely related verb meanings that exhibit substantial semantic overlap within the knowledge representation. Although the knowledge-based filtering mechanism effectively excludes semantically unrelated interpretations, accurately identifying subtle distinctions between closely associated senses remains a challenging aspect of the current framework.

Overall, the superior performance of the proposed methodology is primarily attributed to the effective integration of contextual embeddings with structured semantic knowledge. By combining these complementary sources of information, the framework enhances its capability to accurately identify appropriate word senses, thereby improving disambiguation accuracy while maintaining consistent performance across diverse textual contexts and lexical categories.

DISCUSSION

The findings of this study demonstrate that integrating contextual representations with structured semantic knowledge provides a highly effective solution for Word Sense Disambiguation (WSD). By combining Transformer-based contextual embeddings with ontology-derived knowledge, the proposed framework consistently outperformed approaches that rely exclusively on either contextual learning or knowledge-based reasoning. As reported in Table 4, the proposed model achieved an F1-score of 92.4% on the standard WSD benchmark, surpassing both the Transformer-based baseline (91.0%) and the knowledge-driven baseline (84.7%).

Although the improvement of 1.4 percentage points over the strongest Transformer baseline appears modest, it represents a statistically meaningful enhancement in WSD performance. This gain is primarily attributed to the introduction of the Knowledge Fusion Layer (KFL), which effectively resolves ambiguities that are difficult for purely context-driven models to distinguish. By incorporating structured semantic representations obtained from lexical resources such as WordNet and BabelNet, together with Personalized PageRank (PPR)-based knowledge embeddings, the KFL enriches contextual representations with semantic constraints. Consequently, the framework reduces the likelihood of selecting contextually plausible but semantically incorrect interpretations, thereby improving semantic consistency throughout the disambiguation process. This complementary integration enables the proposed framework to address complex linguistic ambiguities more effectively than conventional WSD techniques.

The experimental results further reveal that the accuracy of the proposed framework increases as the amount of contextual information expands. As illustrated in Figure 5, the highest performance (92.4%) was achieved when complete sentence-level context was utilized, demonstrating that the joint exploitation of contextual and semantic knowledge produces the most reliable disambiguation results. Furthermore, the framework attained an accuracy of 94.1% for frequently occurring polysemous words, whereas performance decreased slightly to 86.3% for low-frequency words. This reduction indicates that the availability of sufficient training examples continues to play an important role, particularly for domain-specific vocabularies and under-resourced languages.

Despite its strong performance, the proposed framework presents certain practical limitations. The construction of semantic knowledge graphs and the generation of Transformer-based contextual embeddings require substantial computational resources, resulting in increased computational complexity and energy consumption during both training and inference. Nevertheless, the theoretical foundations and experimental evaluation confirm that the proposed framework provides accurate and reliable semantic interpretation across diverse linguistic contexts. Its ability to effectively integrate structured knowledge with contextual language understanding makes it well suited for a wide range of Natural Language Processing (NLP) applications, including machine translation, information retrieval, and automatic text summarization. Overall, the successful fusion of semantic knowledge and contextual representations offers a robust and scalable solution for addressing the inherent ambiguity of natural language.

CONCLUSION

This study presented the Fusing Transformer-Based Contextual Embeddings with Structured Semantic Knowledge (WSD) Framework, which combines transformer-based contextual embeddings with structured semantic knowledge to improve the accuracy and reliability of word sense disambiguation. The proposed framework integrates contextual representations generated by RoBERTa with ontology-based lexical resources through a Knowledge Fusion Layer (KFL), enabling effective utilization of both contextual and semantic information during sense prediction.

Experimental evaluation on the Unified WSD Evaluation Benchmark demonstrated that the proposed hybrid framework consistently outperformed conventional knowledge-based and transformer-based approaches, achieving an F1-score of 92.4%. The results indicate that integrating contextual learning with structured semantic knowledge improves semantic understanding and reduces ambiguity, particularly for words with multiple possible meanings.

The findings further demonstrate that hybrid knowledge-enhanced approaches provide a more robust solution than methods relying solely on contextual embeddings or lexical resources. By effectively combining data-driven learning with explicit semantic reasoning, the proposed framework offers improved generalization and can support various Natural Language Processing applications, including machine translation, information retrieval, question answering, and text summarization.

Future work will focus on evaluating the proposed framework on additional benchmark datasets, extending it to multilingual and domain-specific applications, and investigating the integration of large language models and dynamic knowledge graphs to further enhance semantic reasoning and scalability.

FUTURE ENHANCEMENTS

To further improve the proposed hybrid framework, future research should focus on enhancing its accuracy, scalability, and computational efficiency.

Integration of Advanced Transformer Models: Future work can explore advanced transformer architectures such as T5 and other context-aware models. Extending the Knowledge Fusion Layer (KFL) to leverage these models may improve contextual understanding and word sense disambiguation performance.

Domain Adaptation and Knowledge Expansion: To address data scarcity in specialized domains and low-resource languages, future studies should incorporate transfer learning and unsupervised learning techniques. Expanding ontologies with domain-specific knowledge can further improve semantic interpretation.

Computational Optimization: Since the framework requires significant computational resources, future work should focus on optimizing graph processing and embedding generation. Employing distributed computing techniques can improve scalability and enable real-time NLP applications.

ACKNOWLEDGEMENT

The authors are grateful to the PG & amp; Research Department of Computer Science, Thanthai Periyar Government Arts and Science College (Autonomous), Affiliated to Bharathidasan University, for providing the facilities and support required to carry out this research.

Source of support: None

Conflict of Interest: None

REFERENCE

Al Mousa, M., Benlamri, R., & Khoury, R. (2021). *A Novel Word Sense Disambiguation Approach Using WordNet Knowledge Graph. arXiv preprint arXiv:2101.02875.*

Bevilacqua, M., & Navigli, R. (2020). *Breaking through the 80% glass ceiling: Raising the state of the art in Word Sense Disambiguation by incorporating knowledge graph information. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics,* 2854–2864. https://doi.org/10.18653/v1/2020.acl-main.255

Bevilacqua, M., Pasini, T., Raganato, A., & Navigli, R. (2021). *Recent Trends in Word Sense Disambiguation: A Survey. Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21)*, 4330–4338. https://doi.org/10.24963/ijcai.2021/593

Chawla, A., Mulay, N., Bishnoi, V., Dhama, G., & Singh, A. K. (2021). *A Comparative Study of Transformers on Word Sense Disambiguation. arXiv preprint arXiv:2111.15417.*

Yae J. H., et al. (2025). *Leveraging Large Language Models for Word Sense Disambiguation. Neural Computing and Applications.*

Clark, K., Luong, M. T., Le, Q. V., & Manning, C. D. (2020). *ELECTRA: Pre-training text encoders as discriminators rather than generators. Proceedings of the 8th International Conference on Learning Representations (ICLR).* https://arxiv.org/abs/2003.10555

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 4171–4186. https://doi.org/10.18653/v1/N19-1423

Havy, S. (2002). *Topic-sensitive PageRank: A context-sensitive ranking algorithm for web search. IEEE Transactions on Knowledge and Data Engineering, 15*(4), 784–796.

Huang, L., Ding, K., & Liu, H. (2020). *Knowledge-aware graph neural networks with label smoothness regularization for text classification. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 3459–3465. https://doi.org/10.18653/v1/2020.acl-main.318

Ji, S., Pan, S., Cambria, E., Marttinen, P., & Yu, P. S. (2022). *A Survey on Knowledge Graphs: Representation, Acquisition, and Applications. IEEE Transactions on Neural Networks and Learning Systems, 33*(2), 494–514.

Lesk, M. (1986). *Automatic sense disambiguation using machine-readable dictionaries. SIGDOC Asterisk Journal of Computer Documentation, 15*(1), 24–26.

Lin, G.-T., & Giambiasi, M. (2021). *Context-gloss Augmentation for Improving Word Sense Disambiguation. arXiv preprint arXiv:2110.07174.*

Liu, W., Wang, H., Wang, X., & Hou, M. (2024). *A Word Sense Disambiguation Method Based on Multiple Sense Graph. Proceedings of the 2024 International Conference on Machine Learning and Intelligent Computing (PMLR, Vol. 245)*, 40–47.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). *RoBERTa: A robustly optimized BERT pretraining approach. arXiv preprint arXiv:1907.11692.*

Navigli, R. (2009). *Word sense disambiguation: A survey. ACM Computing Surveys, 41*(2), 1–69.

Raganato, A., Camacho-Collados, J., & Navigli, R. (2017). *Word sense disambiguation: A unified evaluation framework and empirical comparison. Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, 99–110. https://doi.org/10.18653/v1/E17-1010

Su, Y., Zhang, H., Song, Y., & Zhang, T. (2022). *Multilingual Word Sense Disambiguation with Unified Sense Representation. arXiv preprint arXiv:2210.07447.*

Wahle, J. P., Ruas, T., Meuschke, N., & Gipp, B. (2021). *Incorporating Word Sense Disambiguation in Neural Language Models. arXiv preprint arXiv:2106.07967.*

Yae, J. H., et al. (2025). *Leveraging Large Language Models for Word Sense Disambiguation. Neural Computing and Applications.*