Implicit Relation Extraction with Large Language Models via Implicit Information Augmentation

Yulin Zhang[1], Yahui Zhao[1*], Yiping Ren[2*], Guozhe Jin[1], Rongyi Cui[1]

[1]Intelligent Information Processing Inst, Yanbian University, Yanji, 133002, China

[2]Yanbian Vocational and Technical College, Yanji, 133002, China

***Abstract*—Relation Extraction (RE) is a core task in information extraction. Although existing approaches have achieved substantial progress in modeling explicit relations, their performance remains limited when confronted with implicit relations that require contextual reasoning or commonsense knowledge, particularly in complex scenarios involving multiple entities and multiple relations. To address this challenge, this paper proposes Implicit Relation Extraction with Large Language Models (IMP-RELLM). The proposed approach explicitly introduces implicit information generated by Large Language Models (LLMs) to enhance the model's ability to capture deep semantic representations and latent relational patterns in text. Specifically, IMP-RELLM adopts an instruction fine-tuning paradigm that organizes task descriptions, predefined relation sets, implicit information, and the original context into a unified structured prompt, guiding the model to attend to semantically valid entity relations that are not explicitly expressed in the text during generation. To reduce training costs and improve scalability, parameter-efficient fine-tuning is employed to adapt multiple mainstream LLMs. Extensive experiments demonstrate that the proposed method substantially outperforms baseline models without implicit information on implicit relation extraction tasks, achieving consistent improvements in both F1 score and recall. The advantages are particularly pronounced in high relational complexity settings and pure implicit relation scenarios, highlighting the critical role of implicit information augmentation in enhancing the reasoning capability and relational coverage of LLMs. Our code is available at https://github.com/IIP408/IMP-RELLM.**

***Index Terms*—implicit relation extraction, large language model, prompt augmentation**

I. Introduction

Entity–relation extraction aims to identify entities and their semantic relations from unstructured text and serves as a fundamental component for constructing structured knowledge representations. It plays a critical role in a wide range of applications, including knowledge graph (KG) construction, information retrieval, and question answering systems. With the advancement of deep learning (DL) techniques, substantial progress has been achieved in modeling explicit semantic relations. However, accurately identifying implicit relations in complex contexts remains a significant challenge[1].

In natural language corpora, relations between entities are not always explicitly expressed through surface linguistic patterns. A large proportion of semantic associations can only be correctly identified through contextual reasoning, commonsense completion, or cross-sentence understanding. Such implicit relations are often characterized by dispersed expresss weak semantic explicitness, and obscured reasoning paths, which makes them difficult to capture using traditional methods that rely on explicit features or localized contextual modeling.

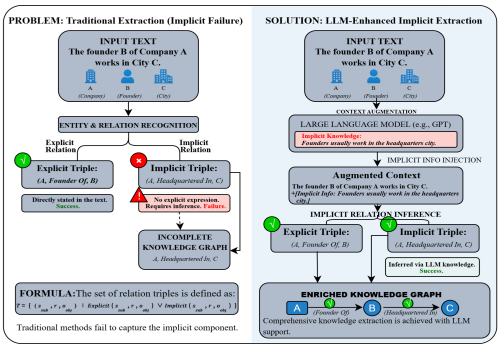

Fig. 1. LLM-Enhanced Implicit Relation Extraction Example.

Existing relation extraction methods primarily focus on structural modeling and representation learning, aiming to capture dependencies between entities and relations through various encoding and interaction mechanisms. However, these approaches typically rely on the implicit assumption that relational cues can be directly learned from the surface representations of the original text. As a result, semantic information that is not explicitly expressed in the text is often insufficiently modeled. This assumption becomes particularly problematic in scenarios involving multiple entities and multiple relations, leading to systematic omissions of implicit relations.

In recent years, LLMs have shown strong capabilities in language understanding and reasoning, offering new opportunities for relation extraction. Compared with traditional models, LLMs can leverage broader contextual information and commonsense knowledge for inference[2]. However, directly applying LLMs to relation extraction still faces challenges, including sensitivity to prompt design and the lack of explicit modeling of intermediate reasoning for implicit relations. As illustrated in Fig. 1, traditional methods can only extract the explicit relation (A, Founder Of, B) from the sentence "The founder B of Company A works in City C," while failing to infer the implicit relation (A, Headquartered In, C). In contrast, by incorporating LLM-based context augmentation, the proposed approach effectively recovers the missing relation and constructs a more complete knowledge graph.

To address the aforementioned challenges, this paper investigates implicit relation extraction from the perspective of explicit reasoning information modeling and proposes an LLM-based framework for implicit relation extraction. The

*Corresponding authors: Yahui Zhao: yhzhao@ybu.edu.cn, Yiping Ren: renyiping63@163.com

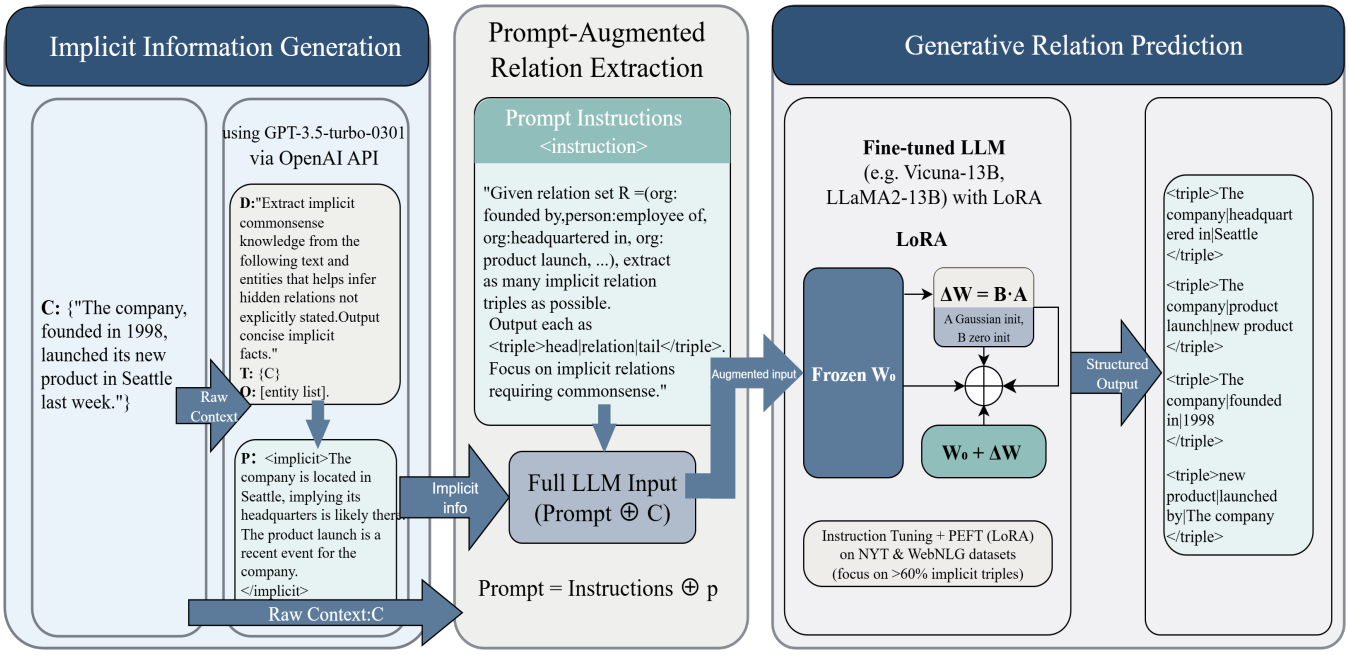

Fig. 2. The overall architecture of IMP-RELLM for relation extraction.

proposed framework explicitly generates latent semantic information relevant to relational reasoning and incorporates it as auxiliary context into the extraction process. In this way, the model is guided to perform relation prediction within a structured generation space constrained by a predefined relation schema.

The main contributions of this work are summarized as follows:

- The key challenges of implicit relation extraction are analyzed from a reasoning modeling perspective, reformulating implicit relation identification as a problem of reasoning context augmentation.
- An explicit implicit information generation and injection mechanism is proposed to provide LLMs with richer contextual support for relation reasoning.
- A unified generative relation extraction framework is developed to handle multi-entity and multi-relation scenarios, and its effectiveness on implicit relation extraction tasks is empirically validated.

II. Related Work

A. Relation Extraction Under Explicit and Generative Modeling Paradigms

Relation extraction has long focused on identifying semantic associations between entities in text. Early approaches primarily relied on handcrafted features combined with conventional classification models to determine relations between entity pairs [3]. With the advancement of deep learning, neural network–based methods have achieved substantial progress by automatically learning contextual representations for relation extraction [5],[6],[7],[8],[9]. These approaches typically adopt sequence modeling or attention mechanisms to predict relations from explicit textual representations and demonstrate stable performance in explicit relation identification scenarios.

To further enhance the modeling of complex semantic structures, some studies incorporate syntactic dependency information or graph-structured representations, using graph neural networks to aggregate contextual features across entities [10],[11],[12]. In addition, to mitigate error propagation between entity recognition and relation prediction, joint extraction strategies have been proposed, which simultaneously model entities and relations within a unified framework [4],[13],[14]. Although these methods exhibit advantages in multi-relation settings, their inference processes largely rely on observable textual cues, resulting in limited capability for modeling implicit relations.

In recent years, generative relation extraction approaches have attracted increasing attention. These methods reformulate relation extraction as a text generation problem, thereby improving flexibility in scenarios involving multiple triplet outputs [15],[19]. Despite their advantages in structural modeling, generative paradigms are often sensitive to input formulations and lack explicit modeling of latent reasoning information. Consequently, they still suffer from instability and insufficient coverage in implicit relation extraction scenarios.

B. Large Language Model–Driven Relation Extraction and Reasoning Augmentation

With the significant advances of LLMs in language understanding and reasoning, recent studies have begun to explore their application to relation extraction. Existing approaches typically employ prompt-based learning, in-context examples, or parameter-efficient fine-tuning to guide LLMs toward generating structured relation outputs [16],[17],[18],[19]. Compared to conventional models, LLMs exhibit stronger generalization capabilities in complex contexts and low-resource settings.

To further enhance reasoning performance, some studies introduce chain-of-thought reasoning, external knowledge augmentation, or retrieval-augmented mechanisms to guide models toward more coherent predictions [20],[21],[22],[23],[24]. However, in most of these ap-

proaches, the reasoning process remains implicitly embedded within the model and is not explicitly integrated into the relation extraction pipeline. As a result, generating stable and accurate multi-relation predictions within a constrained relation space remains challenging.

III. Conceptual Framework

A. Overview

Implicit relation extraction focuses on identifying semantic relations between entity pairs that are not explicitly stated in the surface text but can be inferred through contextual evidence or commonsense reasoning. Given an input text segment C containing multiple entities, the objective is to predict the complete set of entity–relation–entity triplets, which is formally defined as:

$$T = \{(s_i, r_i, o_i)\}_{\{i=1\}}^{N} \tag{1}$$

where s_i and o_i denote the subject and object entities, respectively, $r_i \in R$ represents a relation type selected from a predefined relation set R, and N indicates the total number of extracted triplets.

The IMP-RELLM is shown in Fig. 2 and consists of:(1) an Implicit Information Generation Module, which derives latent semantic cues relevant to relation reasoning from the original text. (2) a Prompt-Augmented Relation Extraction Module, which explicitly injects the generated implicit information together with task-level instructions into the LLM input. and (3) a Generative Relation Prediction Module, which produces structured entity–relation–entity triplets under the enhanced contextual conditions.

B. Implicit Information Generation

The GPT-3.5-turbo-0301 model, endowed with extensive commonsense knowledge acquired from large-scale text corpora, demonstrates strong capability in semantic parsing and understanding. Therefore, the GPT-3.5-turbo-0301 model is utilized via the OpenAI API, and task-specific prompt templates are designed to guide the LLM to uncover implicit knowledge that is not explicitly expressed in the input text. The generated implicit information is denoted as p_{imp}. This information is dynamically integrated into the prompts for implicit relation extraction, thereby providing more precise contextual support for entity–relation extraction.

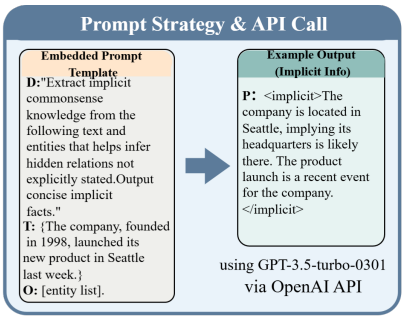

Fig. 3. Prompt template.

In this section, a dedicated instruction schema is designed for implicit information generation, denoted as E = {D, O, C}, where D represents the task description, O denotes the option tokens consisting of a predefined list of entity types, and C corresponds to the input text tokens. The prompt template for implicit information generation is illustrated in Fig. 3.

C. Prompt-Augmented Relation Extraction

After generating implicit information, IMP-RELLM explicitly integrates it into the relation extraction process via a prompt enhancement mechanism. To further improve task adaptability while maintaining parameter efficiency, LoRA[25] is employed to fine-tune the backbone LLM, such as LLaMA or Vicuna. The model input consists of three components: the task-level instruction prompt, the generated implicit information, and the original text.

The task-level instruction prompt specifies the description of the relation extraction task, provides a predefined set of relation classes R, and enforces output format constraints. These components are concatenated to form the complete model input, which can be expressed as:

$$P = \left[p_{ins}; p_{imp}\right] \tag{2}$$

$$Input = [P; C] \tag{3}$$

where p_{ins} denotes the task-level instruction prompt, P denotes the prompt text, and [;] indicates the concatenation operation.

Under this formulation, the Large Language Model is constrained to perform structured generation within the predefined relation space and explicitly leverages the injected implicit information during semantic reasoning. This design effectively reduces the risk of missing implicit relations.

D. Generative Relation Prediction

In the generative relation prediction layer, the model produces a sequence of relation triplets Y conditioned on the input context C and prompt P. The generated output sequence is defined as

$$Y = \left[y_1, y_2, \ldots, y_T\right] \tag{4}$$

where T denotes the sequence length in tokens and y_t denotes the token generated at time step t. The conditional probability of the output sequence is factorized in an autoregressive manner as

$$P(Y \mid Input) = \prod_{t=1}^{T} P(y_t \mid y_{<t}, Input) \tag{5}$$

where $y_{<t}$ represents the previously generated tokens before position t. This generative formulation naturally supports predicting a variable number of relation triplets, making it suitable for texts containing multiple entities and complex relational structures. The generated sequence is subsequently parsed into a set of structured entity–relation–entity triplets, which constitutes the final prediction.

To reduce computational and storage overhead while maintaining strong performance, LoRA is employed for parameter-efficient fine-tuning of the backbone Large Language Model. For a linear transformation with a pretrained weight matrix W_0, LoRA introduces a trainable low-rank update ΔW such that

$$W = W_0 + \Delta W \tag{6}$$

$$\Delta W = BA \tag{7}$$

where $W \in R^{d\times k}$ denotes the adapted weight matrix, $W_0 \in R^{d\times k}$ denotes the frozen pretrained weight matrix, and $\Delta W \in R^{d\times k}$ denotes the parameter update. $B \in R^{d\times r}$ and $A \in R^{r\times k}$ are trainable low-rank parameter matrices, and r is the rank hyperparameter satisfying $r \ll d$. This decomposition factorizes the original high-dimensional update into two low-dimensional matrices, Specifically, matrix A is initialized with a Gaussian distribution, while matrix B is initialized as a zero matrix. This design substantially reduces the number of trainable parameters while preserving the expressive capacity of the model.

To enable text-to-text generation by updating only specific model parameters, the LoRA method is employed to fine-tune the LLM on the target dataset. During fine-tuning, the LLM is trained using its built-in loss function:

$$L = -\sum_{t=1}^{T} \log P(y_t \mid y_{<t}, Input) \quad (8)$$

IV. Experiment

A. Experimental Setup

To evaluate the effectiveness of IMP-RELLM, experiments are conducted on two widely used benchmark datasets, NYT [26] and WebNLG [27]. Since this work focuses on implicit relation extraction, only samples in which implicit triplets account for more than 60% of all annotated relations are selected for evaluation. This filtering strategy enables a more accurate assessment of model performance under implicit relation scenarios.

Parameter-efficient fine-tuning is adopted to train multiple mainstream LLMs, including LLaMA2-7B [28], LLaMA2-13B [28], LLaMA3-8B [29], Vicuna-7B [30], and Vicuna-13B [30]. During fine-tuning, LoRA is employed to reduce the number of trainable parameters. The corresponding hyperparameter settings are reported in TABLE I.

All experiments are conducted on a CentOS Linux 8.2 environment using Python 3.9 and PyTorch 2.3.1. Training and evaluation are performed on eight NVIDIA A40 GPUs, each equipped with 48 GB of memory.

TABLE I.
LLM Parameters

	Hyperparameters	Settings
LLMs	Learning Rate	3e-5
	Batch size	6
	Epoch	30
	Weight Decay	0.001
LoRA	LoRA Alpha	32
	LoRA Dropout	0.05
	r	8

B. Performance Comparison Across Different LLM Backbones

TABLE II reports the performance of IMP-RELLM on the NYT and WebNLG datasets when different LLMs are used as backbone models. Precision, Recall, and F1 score are adopted as evaluation metrics.

As shown in TABLE II, several observations can be made: (1) Models with larger parameter scales consistently achieve superior performance on both datasets, confirming a positive correlation between model size and performance. (2) Vicuna-13B attains the highest F1 scores on both NYT and WebNLG, demonstrating a more balanced trade-off between precision and recall. In contrast, LLaMA2-13B yields slightly inferior results under the same parameter scale, indicating that pre-training strategies and data distributions also play a crucial role in implicit RE performance. (3) LLaMA3-8B exhibits relatively high recall but lower precision on the NYT dataset. Its performance degradation is more pronounced on WebNLG, suggesting limited generalization capability in more complex relational settings.

TABLE II.
Experiment on Comparing the Performance of Different LLMs in Extracting Implicit Relations

Model	NYT			WebNLG		
	Prec(%)	*Rec(%)*	*F1(%)*	*Prec(%)*	*Rec(%)*	*F1(%)*
IMP-RELLM(Llama2-7B)	79.16	92.77	85.42	68.96	80.67	74.36
IMP-RELLM(Llama2-13B)	80.87	93.19	86.60	68.25	83.09	74.94
IMP-RELLM(Llama3-8B)	72.01	92.81	81.09	68.89	74.74	71.7
IMP-RELLM(Vicuna-7B)	80.3	92.89	86.13	63.41	78.4	70.11
IMP-RELLM(Vicuna-13B)	**81.43**	**94.37**	**87.43**	**69.61**	**81.33**	**75.01**

C. Ablation Study

To examine the contribution of the implicit information module in IMP-RELLM, TABLE III presents an ablation study comparing three settings: the full model configuration, a variant without implicit information (w/o II), and a variant without any fine-tuning (w/o all).

TABLE III.
Results of Ablation Experiment

Model	NYT			WebNLG		
	Prec(%)	*Rec(%)*	*F1(%)*	*Prec(%)*	*Rec(%)*	*F1(%)*
IMP-RELLM(Llama2-7B)	79.16	92.77	85.42	68.96	80.67	74.36
w/o II	72.29	86.34	78.69	61.03	76.94	68.07
w/o all	64.12	49.97	56.16	56.05	57.90	56.96
IMP-RELLM(Llama2-13B)	80.87	93.19	86.60	68.25	83.09	74.94
w/o II	74.13	88.89	80.84	58.46	77.59	66.68
w/o all	64.82	46.58	54.21	55.09	62.88	58.73
IMP-RELLM(Llama3-8B)	72.01	92.81	81.09	68.89	74.74	71.70
w/o II	67.97	86.10	75.97	61.06	71.96	66.06
w/o all	46.29	29.70	36.18	48.85	40.48	44.27
IMP-RELLM(Vicuna-7B)	80.30	92.89	86.13	63.41	78.40	70.11
w/o II	69.29	88.97	77.91	55.28	73.21	62.99
w/o all	61.22	53.46	57.08	46.91	49.48	48.16
IMP-RELLM(Vicuna-13B)	**81.43**	**94.37**	**87.43**	**69.61**	**81.33**	**75.01**
w/o II	65.52	87.26	74.84	59.41	79.06	67.84
w/o all	44.57	32.24	37.41	54.87	61.86	58.15

The results show that incorporating implicit information consistently leads to substantial improvements in F1 score across all models and datasets. Compared to models fine-tuned with only basic prompts, the inclusion of implicit information yields particularly notable gains in recall, indicating that implicit information effectively assists the model in identifying latent relations that are not explicitly expressed in the text.

Moreover, when the fine-tuning process is completely removed (w/o all), performance drops significantly, demonstrating that parameter-efficient fine-tuning remains essential for implicit relation extraction. These ablation results validate the effectiveness and complementarity of the individual components within the IMP-RELLM framework.

D. Performance Under Different Relation Complexities

To further analyze the impact of implicit information under varying relational complexities, the NYT dataset is partitioned according to the number of triplets contained in each text instance. Vicuna-13B is selected as the representative backbone model for this analysis. The results are reported in TABLE IV and illustrated in Fig. 4.

TABLE IV.
COMPARISON EXPERIMENTS UNDER DIFFERENT RELATION COMPLEXITIES

Model	Number of Triples	Prec(%)	Recall(%)	F1(%)
vicuna-13B w/o II	Triple 1~2	45.24	38.59	41.65
	Triple 3~4	47.04	20.27	28.33
	Triple ≥ 5	50.00	15.22	23.33
vicuna-13B w/II	Triple 1~2	91.00	89.70	90.34
	Triple 3~4	93.82	71.41	81.09
	Triple ≥ 5	90.54	48.55	63.21

The results show that as relational complexity increases, models without implicit information exhibit a clear performance bottleneck in F1 score when the number of triplets is greater than or equal to five. In contrast, models augmented with implicit information maintain relatively high recall in high-complexity scenarios, suggesting that implicit information enables the model to more effectively capture latent relations and patterns, thereby improving recall.

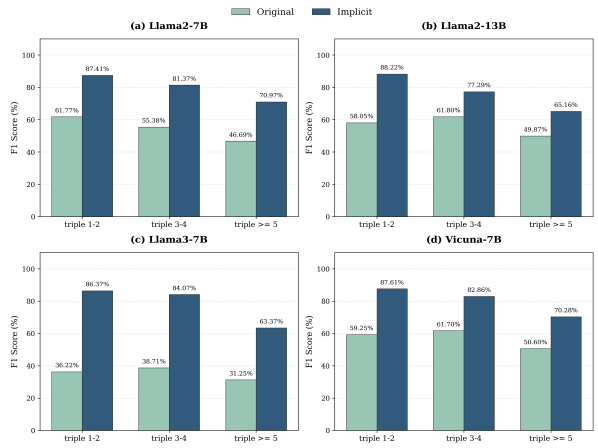

Fig. 4. Comparison of LLM utilizing implicit information under different relationship complexity conditions

The visualization in Fig. 4 further demonstrates that, across different levels of relational complexity, models incorporating implicit information consistently outperform those using only basic prompts in terms of F1 score, and this advantage becomes more pronounced as relational complexity increases.

E. Analysis of Pure Implicit Triplet Recognition

To further evaluate model performance in purely implicit relation scenarios, TABLE V reports the recognition results of pure implicit triplets on the NYT and WebNLG datasets using different LLM backbones. In this study, all entity–relation triplets in the text that are identified by the LLM as implicitly expressed are collectively referred to as purely implicit triplets.

TABLE V.
EXPERIMENT ON COMPARING THE PERFORMANCE OF DIFFERENT LLMS IN RECOGNIZING PURELY IMPLICIT TRIPLES

Model	NYT			WebNLG		
	Prec(%)	*Rec(%)*	*F1(%)*	*Prec(%)*	*Rec(%)*	*F1(%)*
llama2-7B w/o II	54.58	60.55	57.41	54.69	61.05	57.70
llama2-7B w/ II	82.40	**90.59**	**86.30**	65.53	64.97	65.25
llama2-13B w/o II	41.24	45.65	43.33	53.69	66.24	59.31
llama2-13B w/ II	78.72	86.37	82.37	68.70	**72.70**	**70.64**
llama3-8B w/o II	36.01	35.66	35.84	46.71	41.38	43.88
llama3-8B w/ II	79.58	82.98	81.25	60.29	48.04	53.47
vicuna-7B w/o II	36.85	49.10	42.13	45.42	52.49	48.70
vicuna-7B w/ II	70.52	74.95	72.67	58.00	61.38	59.64
vicuna-13B w/o II	46.10	41.56	43.71	52.65	64.23	57.86
vicuna-13B w/ II	**83.93**	79.56	81.69	**69.95**	71.21	70.57

The results indicate that, for purely implicit triplets, the incorporation of implicit information significantly improves model performance. Across multiple LLMs, introducing implicit information yields a minimum F1 improvement of 28.89% on NYT and 7.55% on WebNLG. Models augmented with implicit information generally achieve higher recall, indicating more comprehensive identification of latent relations.

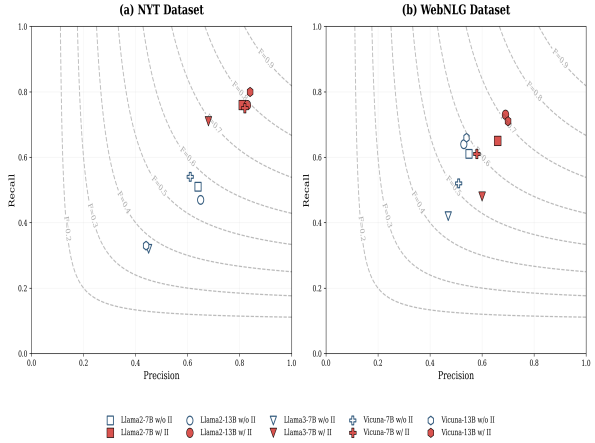

Fig. 5. Effects of different LLMs on the extraction of pure implicit condition

Fig. 5 provides a visualization in the precision–recall space that further corroborates these observations. Models with implicit information augmentation are consistently distributed in regions with higher F1 iso-contours, demonstrating stronger coverage capability in pure implicit relation extraction tasks.

V. CONCLUSION

This paper addresses the limitations of existing entity relation extraction approaches in modeling implicit relations and proposes IMP-RELLM, a LLM–based method for implicit relation extraction. By explicitly incorporating implicit information generated by LLMs, the proposed approach effectively compensates for the insufficiency of contextual and commonsense information in the original text,

thereby enhancing the model's ability to understand and reason about latent entity relations in complex contexts. Through parameter-efficient fine-tuning, IMP-RELLM achieves substantial performance gains while significantly reducing training and deployment costs, demonstrating strong scalability and practical applicability. Experimental results on the NYT and WebNLG datasets show that IMP-RELLM consistently outperforms baseline methods that do not incorporate implicit information on implicit relation extraction tasks, with particularly robust performance in scenarios involving multiple entities, multiple relations, and purely implicit relations.

ACKNOWLEDGMENT

This work is supported by Jilin Provincial Department of Science and Technology Development Plan Project [grant numbers: 20220203127SF], The school-enterprise cooperation project of Yanbian University [2024-10] and [2019-34 (continued)], National Natural Science Foundation of China [grant numbers 62162062].

REFERENCES

[1] H. Fei, B. Li, Q. Liu, et al., "Reasoning implicit sentiment with chain-of-thought prompting," *arXiv preprint arXiv:2305.11255,* 2023.

[2] X. Yang, X. Gong, B. Tang, et al., "CAG: A consistency-adaptive text-image alignment generation for joint multimodal entity-relation extraction," in *Proc. 33rd ACM Int. Conf. Inf. Knowl. Manage. (CIKM),* 2024, pp. 4183–4187.

[3] Y. Wang, B. Yu, Y. Zhang, et al., "TPLinker: Single-stage joint extraction of entities and relations through token pair linking," in *Proc. Int. Conf. Comput. Linguistics (COLING), Barcelona, Spain,* 2020, pp. 1572–1582.

[4] M. Shang, H. Huang, X. Sun, et al., "Relational triple extraction: One step is enough," in *Proc. 31st Int. Joint Conf. Artif. Intell. (IJCAI), Macao, China,* 2023, pp. 4360–4366.

[5] T. Mikolov, K. Chen, G. Corrado, et al., "Efficient estimation of word representations in vector space," in *Proc. Int. Conf. Learn. Represent. (ICLR), Online,* 2013.

[6] J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global vectors for word representation," in *Proc. Conf. Empir. Methods Nat. Lang. Process. (EMNLP), Doha, Qatar,* 2014, pp. 1532–1543.

[7] B. He, Y. Guan, and R. Dai, "Classifying medical relations in clinical text via convolutional neural networks," *Artif. Intell. Med., vol.* 93, pp. 43–49, 2019.

[8] Y. Luo, "Recurrent neural networks for classifying relations in clinical notes," J. Biomed. Inform., vol. 72, pp. 85–95, 2017.

[9] A. Vaswani, N. Shazeer, N. Parmar, et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), Long Beach, CA, USA,* 2017, pp. 5998–6008.

[10] T. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. Int. Conf. Learn. Represent. (ICLR), Toulon, France,* 2017, pp. 45–56.

[11] Y. Zhang, Z. Guo, and W. Lu, "Attention-guided graph convolutional networks for relation extraction," in *Proc. 57th Annu. Meet. Assoc. Comput. Linguistics (ACL), Florence, Italy,* 2019, pp. 5308–5317.

[12] Y. Zhang, P. Qi, and C. D. Manning, "Graph convolution over pruned dependency trees improves relation extraction," in *Proc. Conf. Empir. Methods Nat. Lang. Process. (EMNLP), Brussels, Belgium,* 2018, pp. 2205–2215.

[13] M. Eberts and A. Ulges, "Span-based joint entity and relation extraction with transformer pre-training," in *Proc. Eur. Conf. Artif. Intell. (ECAI), Santiago de Compostela, Spain,* 2020, pp. 2163–2170.

[14] J. Wang and W. Lu, "Two are better than one: Joint entity and relation extraction with table-sequence encoders," in *Proc. Conf. Empir. Methods Nat. Lang. Process. (EMNLP), Online,* 2020, pp. 1706–1721.

[15] G. Li, Z. Xu, Z. Shang, et al., "Empirical analysis of dialogue relation extraction with large language models," in *Proc. 33rd Int. Joint Conf. Artif. Intell. (IJCAI), Jeju, South Korea,* 2024, pp. 6359–6367.

[16] Z. Wan, F. Cheng, Z. Mao, et al., "GPT-RE: In-context learning for relation extraction using large language models," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI), Macao, China,* 2023, pp. 6359–6367.

[17] P. Rajpoot and A. Parikh, "GPT-FinRE: In-context learning for financial relation extraction using large language models," in *Proc. 6th Workshop Financial Technol. Nat. Lang. Process., Bali, Indonesia,* 2023, pp. 42–45.

[18] V. Zavarella, J. C. Gamero-Salinas, and S. Consoli, "A few-shot approach for relation extraction domain adaptation using large language models," *arXiv preprint arXiv:2408.02377,* 2024.

[19] Z. Shi and H. Luo, "CRE-LLM: A domain-specific Chinese relation extraction framework with fine-tuned large language model," *arXiv preprint arXiv:2404.18085,* 2024.

[20] X. Ma, J. Li, and M. Zhang, "Chain of thought with explicit evidence reasoning for few-shot relation extraction," in *Proc. Annu. Meet. Assoc. Comput. Linguistics (ACL), Toronto, Canada,* 2023, pp. 2334–2352.

[21] K. Zhang, B. J. Gutiérrez, and Y. Su, "Aligning instruction tasks unlocks large language models as zero-shot relation extractors," in *Proc. Annu. Meet. Assoc. Comput. Linguistics (ACL), Toronto, Canada,* 2023, pp. 794–812.

[22] J. Li, Z. Jia, and Z. Zheng, "Semi-automatic data enhancement for document-level relation extraction with distant supervision from large language models," in *Proc. Annu. Meet. Assoc. Comput. Linguistics (ACL), Toronto, Canada,* 2023, pp. 5495–5505.

[23] M. Lee, Q. Zhu, C. Mavromatis, et al., "HybGRAG: Hybrid retrieval-augmented generation on textual and relational knowledge bases," *arXiv preprint arXiv:2412.16311,* 2024.

[24] Z. Zhan, S. Zhou, M. Li, et al., "RAMIE: Retrieval-augmented multi-task information extraction with large language models on dietary supplements," *J. Amer. Med. Inform. Assoc., vol. 32, no. 3,* pp. 545–554, 2025.

[25] E. Hu, Y. Shen, P. Wallis, et al., "LoRA: Low-rank adaptation of large language models," in *Proc. Int. Conf. Learn. Represent. (ICLR),* 2021.

[26] S. Riedel, L. Yao, and A. McCallum, "Modeling relations and their mentions without labeled text," in *Proc. Eur. Conf. Mach. Learn. Knowl. Discov. Databases (ECML PKDD), Barcelona, Spain,* 2010, pp. 148–163.

[27] C. Gardent, A. Shimorina, S. Narayan, et al., "Creating training corpora for NLG micro-planning," in *Proc. Annu. Meet. Assoc. Comput. Linguistics (ACL), Vancouver, Canada,* 2017, pp. 179–188.

[28] H. Touvron, L. Martin, K. Stone, et al., "LLaMA 2: Open foundation and fine-tuned chat models," *arXiv preprint arXiv:2307.09288,* 2023.

[29] A. Grattafiori, A. Dubey, A. Jauhri, et al., "The LLaMA 3 herd of models," *arXiv preprint arXiv:2407.21783,* 2024.

[30] L. Zheng, W. L. Chiang, Y. Sheng, et al., "Judging LLM-as-a-judge with MT-Bench and Chatbot Arena," *Adv. Neural Inf. Process. Syst., vol. 36,* pp. 46595–46623, 2023.