Entity Pair-guided Relation Summarization and Retrieval in LLMs for Document-level Relation Extraction

Fu Zhang†, Hongsen Yu†, Jingwei Cheng*, Huangming Xu
School of Computer Science and Engineering, Northeastern University, China
{zhangfu,chengjingwei}@mail.neu.edu.cn; yuhongsen163@163.com

Abstract

Document-level relation extraction (DocRE) aims to extract relations between entities in a document. While previous research has primarily focused on traditional small models, recent studies have extended the scope to large language models (LLMs). Current LLM-based methods typically focus on filtering all potential relations (candidate relations) within a document at one time and then performing triplet fact extraction. However, most approaches for candidate relation filtering are based on the document level, which results in insufficient correlation between candidate relations and entity pairs. In addition, the data imbalance problem caused by a large amount of no-relation data (NA problem) is another important reason for the suboptimal performance of LLM-based methods. To address these issues, we propose an **e**ntity **p**air-guided **r**elation **s**ummarization and **r**etrieval model (**EP-RSR**) for DocRE, which introduces an innovative LLM-based document-level relation extraction paradigm, EPRF (Entity Pair-Relation-Fact), along with an entity pair-level candidate relation filtering method. Our approach first selects entity pairs that potentially contain relations and uses them to guide relation summarization and retrieval for extracting relation facts. This enhances the relevance between candidate relations and entity pairs while alleviating the issue of imbalanced NA data. Benchmark testing on three datasets demonstrates that our approach achieves state-of-the-art (SOTA) performance for LLM-based models[1].

1 Introduction

Document-level relation extraction (DocRE) is a crucial task in natural language processing (NLP), aimed at identifying and extracting semantic relations between entities within a given document. Compared to sentence-level relation extraction, the challenge of document-level approaches is that a large number of triplet facts need to be obtained through joint reasoning of multiple sentences (Yao et al., 2019), which places higher demands on the reasoning capabilities of the model.

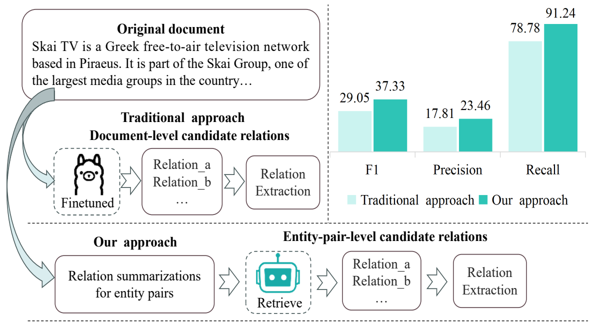

Figure 1: Differences between traditional approach and our approach in LLM-based document-level relation extraction. Additionally, we provide a preliminary comparison of the F1 scores between the document-level and entity-pair-level candidate relation filtering methods.

Traditional approaches for DocRE primarily relies on graph neural networks (GNNs) and pre-trained language models (PLMs). GNN-based models mainly perform explicit reasoning on graphs constructed from entities and sentences in a document, e.g., EoG (Christopoulou et al., 2019), GAIN (Zeng et al., 2020), and SIRE (Zeng et al., 2021). Approaches based on PLMs take the word sequence of a document as input and leverage the transformer (Vaswani, 2017) to implicitly capture the long-range contextual dependencies between entities, e.g., ATLOP (Zhou et al., 2021), EIDER (Xie et al., 2022), and DREEAM (Ma et al., 2023).

With the notable success of large language models (LLMs) like GPT in the field of NLP (Brown et al., 2020; Touvron et al., 2023), approaches for LLM-based DocRE have gradually gained traction. Unlike traditional approaches, LLM-based DocRE approaches primarily treat LLMs as black boxes, focusing on relation extraction paradigms

[1] Our code: https://github.com/LookingYu/EP-RSR.
†Equal contribution. *Corresponding author.

Findings of the Association for Computational Linguistics: NAACL 2025, pages 4022–4037
April 29 - May 4, 2025 ©2025 Association for Computational Linguistics

and prompt engineering (Gao et al., 2023; Li et al., 2023). Moreover, leveraging the strong generalization capabilities of LLMs, specific fine-tuned on DocRE tasks has become an essential process for enhancing relation extraction performance (Xue et al., 2024). However, existing research indicates that the performance of LLM-based DocRE still lags behind that of traditional small models at their state-of-the-art (SOTA) levels, underscoring the necessity of further advancing research in this area.

Given the wide variety of relation types between entities in DocRE (e.g., there are 96 types of relations in the Re-DocRED (Tan et al., 2022) dataset), the *filtering of candidate relations* between entities has become one of the key factors influencing the performance of current LLM-based DocRE methods. Current relation extraction paradigms in the sentence-level field generally favor the inclusion of candidate relations in prompts to enhance the effectiveness of triplet fact extraction, which often treat the entire set of relation types as the candidate relation set (Wang et al., 2023b). However, this approach is not suitable for DocRE with a large number of relations. Recently, Xue et al. (2024) propose a document-based candidate relation filtering method, which first selects all potentially existing relations within a given document at one time and then performs triplet fact extraction. However, this method relies heavily on the document itself for candidate relation selection and lacks sufficient correlation with the head and tail entities to be extracted, as illustrated in Figure 1.

Additionally, the *no-relation* (NA label) *data imbalance problem* (NA problem) in DocRE is another significant factor leading to suboptimal performance of LLM-based models. For example, in the DocRED (Yao et al., 2019) dataset, entity pairs without relation account for 97.17% of the dataset. The excessive amount of NA data poses a risk of false positives in relation prediction (Wan et al., 2023). Gao et al. (2023) addresses this by enhancing prediction capabilities through data programming combined with multiple weak supervision sources. However, the effectiveness of this method is often constrained in the absence of high quality sources of weak supervision.

To address the aforementioned issues, we propose an entity pair-guided relation summarization and retrieval model EP-RSR based on a novel LLM-based DocRE paradigm EPRF (Entity Pair-Relation-Fact). This paradigm EPRF first selects entity pairs, then filters candidate relations, and finally performs triplet fact extraction. Based on this paradigm, our approach first selects entity pairs that potentially contain relations from a given document, thereby alleviating the imbalance issue of NA data. Subsequently, we introduce entity-pair-level relation retrieval, which retrieves candidate relations for an entity pair from our constructed training data based on relation summarization of the entity pair. Finally, utilizing the candidate relations and their descriptions, we judge whether the entity pair has the candidate relations, thus achieving the goal of triplet fact extraction. Our contributions are summarized as follows:

- We propose an entity pair-guided relation summarization and retrieval model EP-RSR based on a novel paradigm EPRF for LLM-based document-level relation extraction. Our method effectively alleviates the issue of NA data imbalance and enhances the model's capacity to extract triplet facts.

- We propose a candidate relation filtering method based on entity-pair-level relation retrieval, which retrieves candidate relations at the entity-pair level based on relation summarizations of entity pairs, thereby enhancing the relevance between candidate relations and entity pairs.

- We implemented this method on the DocRED, Re-DocRED and DWIE datasets. Results demonstrate that our approach achieves significant performance improvements over competitive LLM-based baselines (+**7.42** F1 on DocRED test set, +**4.97** F1 on Re-DocRED dev set and +**18.00** F1 on DWIE test set).

2 Related Work

2.1 Traditional DocRE

Document-level relation extraction is a pivotal task in NLP, wherein over 40.7% of relations are dependent on cross-sentence joint reasoning (Yao et al., 2019). Early studies predominantly employ deep learning models such as convolutional neural networks (CNN) and long short-term memory networks (LSTM) for semantic representation learning (Zheng et al., 2018; Yao et al., 2019; Tang et al., 2020). As research progressed, the concept of graphs is increasingly integrated, exemplified by graph convolutional networks enhanced with global contextual information (Sahu et al., 2019), edge-oriented graph extraction techniques (Christopoulou et al., 2019), and graph aggregation-and-inference network which features a double graph design (Zeng et al., 2020). Moreover, the application of pre-trained language models BERT (Devlin et al., 2019) and RoBERTa (Liu et al., 2019) in this domain has proliferated, encompassing approaches like adaptive thresholding with local context pooling (Zhou et al., 2021), knowledge distillation strategies (Ma et al., 2023), and trainable memory module (Gao et al., 2024).

2.2 LLM-based DocRE

With the immense potential of LLMs increasingly evident across various domains (Wang et al., 2023a; Zhou et al., 2023; Xu et al., 2023), document-level relation extraction methods centered around LLMs have emerged, encompassing both fine-tuned and non-fine-tuned approaches.

For non-fine-tuned approaches, Gao et al. (2023) enhance model's relation extraction capabilities by combining prompting techniques with data programming. Li et al. (2023) enrich the DocRE dataset by integrating LLMs with a natural language inference module. Additionally, Ozyurt et al. (2023) propose the REPLM model, which is designed for few-shot relation extraction within a contextual framework using LLMs.

In terms of fine-tuned approaches, Li et al. (2024) propose a new fine-tuned LLM-based DocRE method, which adds relation sets and entity pairs to prompts for document-level relation extraction. Xue et al. (2024) propose a new RE paradigm model AutoRE. Although the author initially adopted the calculation method and goal of document-level relation triplet extraction (DocRTE), this method has achieved exceptional performance in LLM-based DocRE. Due to the large number of relations involved in DocRE, current LLM-based approaches generally avoid incorporating the entire relation list into the prompt template. To achieve this, AutoRE proposes a paradigm RHF (Relation-Head-Fact) that enhances relation extraction performance by filtering the relation list to obtain candidate relations and subsequently extracting triplet facts. However, this method primarily filters candidate relations based on the given document, which often has low correlation with individual entity pairs. Since the core objective of DocRE is to acquire triplets of entity pairs, candidate relations should be more strongly associated with specific entity pairs. Therefore, we propose an entity pair-guided relation summarization and retrieval model EP-RSR based on a novel LLM-based DocRE paradigm EPRF (Entity Pair-Relation-Fact), which focuses on entity-pair-level candidate relations, thereby enhancing the relevance between candidate relations and entity pairs to achieve better triplet fact extraction.

3 Methodology

3.1 Problem Definition

Given a document $D = \{s_i\}_{i=1}^{n_s}$, where each sentence $s_i = \{w_j\}_{j=1}^{n_w^i}$ contains n_w^i words, and an entity set $V = \{e_i\}_{i=1}^{n_e}$. The DocRE task is to predict the relation $r \in R \cup \{NA\}$ between entity pair (e_h, e_t), where $h, t \in \{1, \cdots, n_e\}$ and $h \neq t$. The set $R = \{r_i\}_{i=1}^{n_r}$ represents a predefined collection of relations, while NA signifies the absence of relation between the entity pairs.

3.2 Overview

Based on our new paradigm EPRF (Entity Pair-Relation-Fact), which first selects entity pairs, then filters candidate relations, and finally performs triplet fact extraction, our model EP-RSR comprises three key components: (1) **Entity information enhanced relation summarization** module, which first selects entity pairs that potentially contain relations from a given document. Then, it generates a relation summarization for each entity pair, which includes entity information and relation information related to the entity pair. (2) **Entity-pair-level relation retrieval** module, which retrieves candidate relations that exhibit a higher relevance to the entity pair based on the relation summarization. (3) **Triplet fact judgement** module, which achieves the goal of triplet fact extraction by judging whether the entity pair has the candidate relations. An illustration of the overall framework of our approach is shown in Figure 2.

3.3 Entity Information enhanced Relation Summarization

A significant challenge in DocRE is long contexts. Existing LLM-based methods typically identify all candidate relations from a given document without adequately considering the relevance of the contextual information carried by these candidate relations to the target head and tail entities. To obtain context information that is more pertinent to the entity pairs, we propose a novel entity information enhanced relation summarization approach.

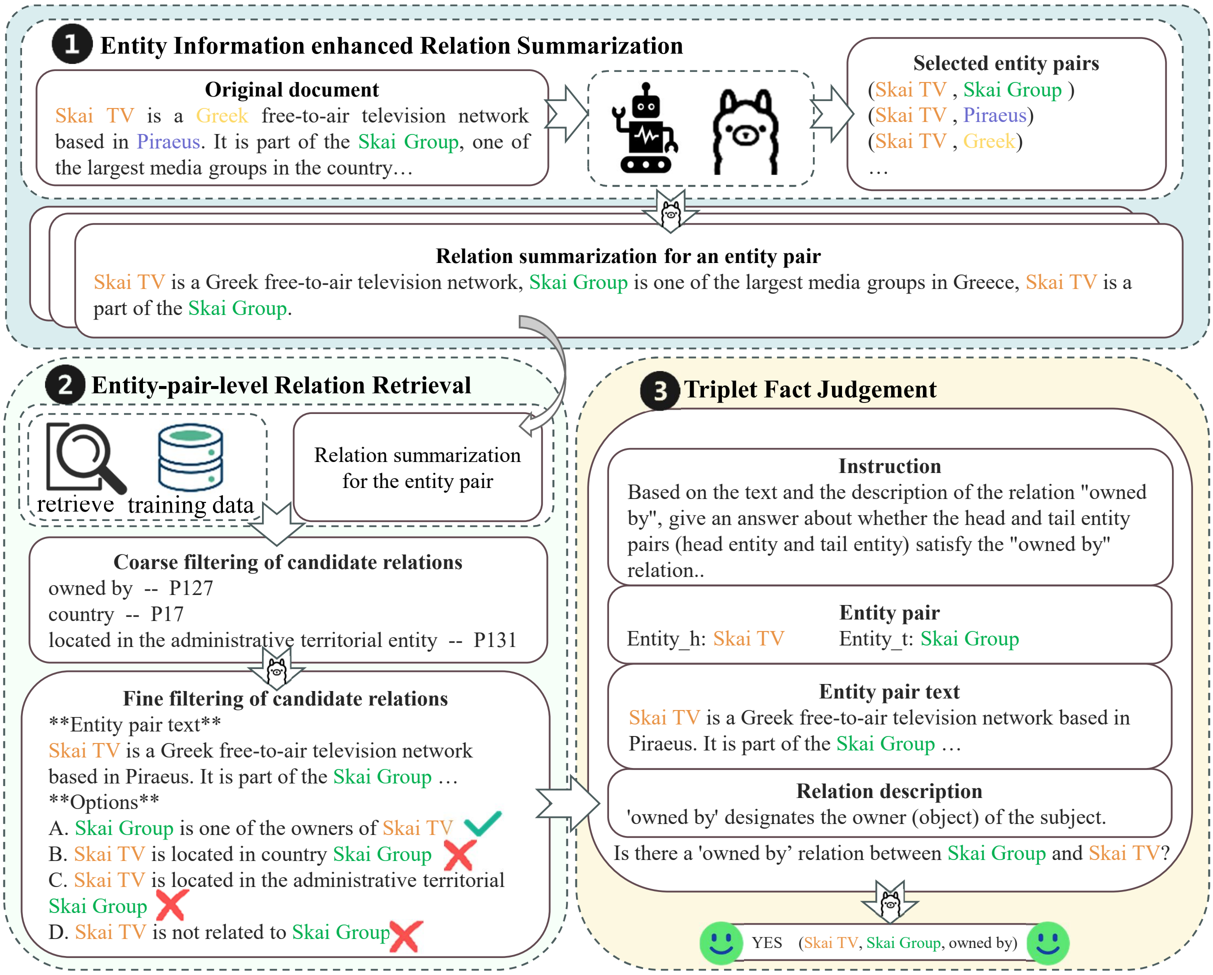

Figure 2: The overview of our model EP-RSR. It contains three key parts: (1) Select potential entity pairs in a document and obtain enhanced relation summarization for each entity pair. (2) Retrieve entity-pair-level candidate relations using the relation summarizations based on double filtering mechanisms. (3) Judge and extract triplet facts based on the candidate relations and their descriptions.

Specifically, we first *select entity pairs* that potentially contain relations from a given document. Although, considering that the huge number of entity pairs in DocRE poses a significant challenge, for instance, the DocRED dataset's dev set includes 26,141 entities, leading to 362,313 entity pairs. Getting enhanced relation summaries for all entity pairs would incur significant computational costs. In addition, the large number of entity pairs causes NA problem, which can lead to false positives in relation prediction (Gao et al., 2023). To address these issues, we propose using LLMs to conduct multi-sampling selection of entity pairs within the document, prior to relation summarization. This approach identifies the entity pairs that potentially contain relations, thereby reducing costs and effectively alleviating NA problem. In detail, we implement the following steps:

Multi-sampling. The multi-sampling process takes as input an instruction I, a document D, and a set of entities $V = \{e_i\}_{i=1}^{n_e}$. It employs LLMs to perform k random sampling, resulting in an entity pair set P_i in one sampling iteration.

$$P_i = LLM(I, D, V). \tag{1}$$

Selecting entity pairs that potentially contain relations. During each sampling iteration, entity pairs are filtered based on their cosine similarity to the true entities in the dataset. This filtering process uses a threshold T to obtain the set of valid entity pairs P_i':

$$P_i' = \{(e_h, e_t) \in P_i \,|\, sim(e_h, e_h') > T, \\ sim(e_t, e_t') > T\}, \tag{2}$$

where $e_h', e_t' \in V$, $sim(e_h, e_h')$ denotes the cosine similarity between the entity e_h and the entity e_h' (similar to $sim(e_t, e_t')$).

We finally merge the sets of valid entity pairs obtained from all k sampling iterations to obtain the set of entity pairs potentially containing relations, denoted as P_e:

$$P_e = \bigcup_{i=1}^{k} P_i'. \tag{3}$$

Further, by utilizing the summarization capabilities of LLMs, we can deduce the *relation summarization* R_s between an entity pair (e_h, e_t) in P_e. As an example, for the entity pair ("Skai TV", "Skai Group"), LLMs gets the relations summarization "Skai TV is a component of the Skai Group" based on the document D.

Additionally, considering that relation descriptions often impose implicit constraints on entity types, for instance, the "country" relation typically assumes that the head entity is non-human and the tail entity is a nation. Therefore, we integrate entity information $E(e_h)$ and $E(e_t)$, such as "Skai TV is a Greek free-to-air television network" and "Skai Group is one of the largest media groups in Greece" with the relation summarization R_s to construct a new enhanced relation summarization R_s'. The formula is defined as follows:

$$E(e_h) = LLM(e_h, D), \tag{4}$$

$$E(e_t) = LLM(e_t, D), \tag{5}$$

$$R_s^{(h,t)} = LLM(e_h, e_t, D), \tag{6}$$

$$R_s'(e_h, e_t) = E(e_h) + E(e_t) + R_s^{(h,t)}, \tag{7}$$

where $+$ represents the concatenation of strings and the entity pair $(e_h, e_t) \in P_e$.

3.4 Entity-pair-level Relation Retrieval

To obtain candidate relations, we propose an entity-pair-level relation retrieval method. The method retrieves the top-k candidate relations by double filtering from our pre-constructed training data $\mathcal{T_D}$ based on the relation summarization corresponding to an entity pair.

Pre-constructed training data $\mathcal{T_D}$: The training data is constructed from the train set $\mathcal{T}$ of the dataset. Specifically, for each entity pair $(e_h^{\mathcal{T}}, e_t^{\mathcal{T}})$ with relations in the train set of the dataset, a relation summarization $R_s'(e_h^{\mathcal{T}}, e_t^{\mathcal{T}})$ for it will be generated according to the approach in Section 3.3.

The relation summarization $R_s'(e_h^{\mathcal{T}}, e_t^{\mathcal{T}})$ is then converted into a vector representation $H(e_h^{\mathcal{T}}, e_t^{\mathcal{T}})$ using Sentence-transformers (Reimers and Gurevych, 2019):

$$H(e_h^{\mathcal{T}}, e_t^{\mathcal{T}}) = Encode(R_s'(e_h^{\mathcal{T}}, e_t^{\mathcal{T}})). \tag{8}$$

The vector representations of the entity pairs are stored as keys, with their true relation labels as values in the pre-defined training data $\mathcal{T_D}$:

$$\mathcal{T_D} = \{H(e_h^{\mathcal{T}}, e_t^{\mathcal{T}}) : Label(e_h^{\mathcal{T}}, e_t^{\mathcal{T}})\}. \tag{9}$$

Coarse filtering of candidate relations: Initially, the relation summarization R_s' for entity pair (e_h, e_t) in P_e is computed by Eq. (8) to obtain H.

Next, the cosine similarity between encoded representation H and each encoded representation $H^{\mathcal{T}}$ in training data $\mathcal{T_D}$ is computed using the formula:

$$sim(H, H^{\mathcal{T}}) = \frac{H \cdot H^{\mathcal{T}}}{\|H\|\|H^{\mathcal{T}}\|}. \tag{10}$$

All cosine similarity results are sorted, and top k similar data are selected. The corresponding relation labels form the candidate relations set R_C.

Fine filtering of candidate relations: By leveraging predefined relation templates RT, the coarse-filtered candidate relations R_C for entity pair (e_h, e_t) are transformed into natural language. This transformation is followed by a fine filtering using multiple-choice QA and prior knowledge of the entity pairs to yield the final set of candidate relations R_F. The steps are outlined as follows:

First, the candidate relations R_C are converted into a set of natural language options Q using the predefined relation templates RT and prior knowledge PK of the entity pairs:

$$Q = Conversion(R_C, RT, PK, e_h, e_t). \tag{11}$$

If the natural language option set Q is empty, it is considered that there is no relation between the entity pair (e_h, e_t). Otherwise, the natural language options set Q, along with the document D and the instruction I, are combined to form a multiple-choice question. We add "no relation" as an additional option to the question. This question is processed by LLMs, resulting in an answer $A(e_h, e_t)$:

$$A(e_h, e_t) = LLM(I, D, Q, e_h, e_t), \tag{12}$$

where $A(e_h, e_t)$ may contain multiple options.

Finally, the answer $A(e_h, e_t)$ is converted into corresponding relation labels, yielding final set of candidate relations R_F for the entity pair (e_h, e_t).

3.5 Triplet Fact Judgement

For each candidate relation $r \in R_F$, we construct a prompt and feed it into LLMs. Each prompt consists of the following components:

Instruction I_r: A tailored instruction is provided for each candidate relation r, detailing the DocRE task and specifying the output requirement. The model is required to output "YES" or "NO". If the entity pair contains the candidate relation, the model outputs "YES"; otherwise, it outputs "NO".

Relation Description RD_r: To enhance the model's understanding of the specific meaning of the candidate relation r, we include a detailed relation description in the prompt. This improves the model's ability to judge candidate relations.

Test Input x_{test}: The document D, along with the head entity e_h and the tail entity e_t, is provided to the model as input. The model generates the corresponding answer.

The entire process is as follows:

$$y_{test} = LLM(I_r, RD_r, x_{test}), \quad (13)$$

where $y_{test} \in \{YES, NO\}$. If the output is "YES", we infer that the given entity pair (e_h, e_t) has the relation r, and the model will output triplet (e_h, e_t, r). If the output is "NO", it is inferred that there is no relation between the entity pair (e_h, e_t), and the model does not output anything.

4 Experiments

4.1 Experimental Setup and Baselines

Dataset We conducted experiments on three DocRE datasets: DocRED (Yao et al., 2019), its revised version Re-DocRED (Tan et al., 2022), and a new gold-annotated dataset DWIE (Zaporojets et al., 2021). The detailed statistics of the datasets can be found in Appendix A.

Evaluation Metrics We employ F1 and Ign F1 as main evaluation metrics following (Yao et al., 2019). Ign F1 is employed to assess the F1 score while excluding relations shared between the training and test sets.

Implementation Settings We utilized Llama3-8B (Touvron et al., 2023) as primary model for our experimental framework, and employed the Sentence-Transformer model all-mpnet-base (Reimers and Gurevych, 2019). We use Llama-Factory (Zheng et al., 2024) framework to fine-tune LLM based on LoRA (Hu et al., 2022). Our experiments were predominantly conducted on three NVIDIA 3090 GPUs for fine-tuned, with inference operations carried out on a single 4090 GPU.

We only fine-tune the LLM for entity pair selection, multiple-choice QA in Section 3.4, and triple fact judgment stages. Fine-tuning parameters and inference parameter settings at each stage can be found in the Appendix B. The prompt templates for each stage are shown in the Appendix C.

In addition, the experimental analysis of the k in top-k candidate relations retrieval and the k-times multi-sampling in entity pair selection are detailed in Appendix D.

In order to better test the versatility of our approach, we also conducted experiments on the DocRTE task. The details and results of the experiments are shown in Appendix E.

Baseline Models In this study, the primary models for comparison include both non-fine-tuned and fine-tuned LLM models for DocRE. The non-fine-tuned models consist of PromptRE (Gao et al., 2023), DocGNRE (Li et al., 2023), and few-shot on ChatGPT (Han et al., 2023). For the fine-tuned model, we select LMRC (Li et al., 2024) and AutoRE (Xue et al., 2024) as a comparison baseline for DocRE. Additionally, we also incorporate the relation extraction paradigms of D-F (Document-facts) and D-R-F (Document-relation-facts) mentioned in AutoRE (Xue et al., 2024) as our comparison baselines.

In addition, to further demonstrate the performance of LLM-based and traditional models, we also select some traditional models introduced in Section 2.1 as baselines, including CNN (Yao et al., 2019), LSTM (Yao et al., 2019), BiLSTM (Yao et al., 2019), ATLOP (Zhou et al., 2021), DREEAM (Ma et al., 2023), and TTM-RE (Gao et al., 2024).

4.2 Main Results

All experimental results are presented in Table 1 and Table 2. Our model *consistently outperforms all LLM-based baselines on three datasets*. Moreover, we draw several interesting conclusions:

Our approach significantly outperforms current *fine-tuned* baseline AutoRE (Xue et al., 2024), achieving SOTA performance in the LLM-based DocRE methods. Specifically, compared to AutoRE, our model achieves an enhancement of **7.42** in F1 score on DocRED test set and an improvement of **18.00** in F1 score on DWIE test set. These significant improvements demonstrate the effectiveness of the proposed method.

Model	DocRED				Re-DocRED			
	Dev		Test		Dev		Test	
	F1	Ign F1	F1	Ign F1	F1	Ign F1	F1	Ign F1
Traditional models								
CNN (Yao et al., 2019)	43.45	41.58	42.26	40.33	-	-	-	-
LSTM (Yao et al., 2019)	50.68	48.44	50.07	47.71	-	-	-	-
BiLSTM (Yao et al., 2019)	50.94	48.87	51.06	48.78	-	-	-	-
ATLOP (Zhou et al., 2021)	61.09	59.22	61.30	59.31	77.63	76.88	77.73	76.94
DREEAM (Ma et al., 2023)	61.42	59.60	61.13	59.12	-	-	77.94	77.34
TTM-RE (Gao et al., 2024)	-	-	-	-	78.13	78.05	79.95	78.20
LLM-based models								
PromptRE (Gao et al., 2023)	-	-	-	-	10.55	9.03	-	-
DocGNRE (Li et al., 2023)*	13.84	13.65	13.93	13.67	11.18	11.10	11.12	11.04
ChatGPT (Han et al., 2023)	32.21	-	-	-	28.89	-	-	-
LMRC (Li et al., 2024)	39.25	38.62	38.66	38.09	52.56	52.29	52.45	52.15
D-F (Xue et al., 2024)*	46.38	44.77	47.08	45.30	54.22	53.48	53.33	52.50
D-R-F (Xue et al., 2024)*	45.77	44.32	47.50	45.98	56.58	56.10	54.84	54.35
AutoRE (Xue et al., 2024)*	47.17	45.58	47.15	45.45	60.17	59.25	59.29	58.33
Ours (EP-RSR)	**53.77**	**51.25**	**54.57**	**51.77**	**65.14**	**63.93**	**64.24**	**63.03**

Table 1: Experimental results on two public datasets for DocRE. Results with * are our reproduction using Llama3-8B. Bold indicates the best results among the LLM-based methods.

Model	DWIE			
	Dev		Test	
	F1	Ign F1	F1	Ign F1
ChatGPT (Han et al., 2023)	-	-	26.72	-
DocGNRE (Li et al., 2023)*	11.85	10.55	13.12	10.73
AutoRE (Xue et al., 2024)*	56.53	52.74	56.38	49.31
Ours (EP-RSR)	**70.23**	**66.32**	**74.38**	**69.56**

Table 2: Performance on the DWIE dataset for DocRE. Results with * are our reproduction using Llama3-8B.

Model	F1	Ign F1
Our Method	**53.77**	**51.25**
w/o Entity pair selection	21.83	19.12
w/o Entity information	51.51	49.25
w/o Entity-pair-level candidate relations	47.00	45.57
w/o Triplet fact judgement	47.65	44.49

Table 3: Ablation study on the DocRED.

Moreover, compared to *non-fine-tuned* methods, our model demonstrates a significant performance breakthrough, achieving an improvement of approximately **21.56** in F1 on DocRED dev set over the ChatGPT-based DocRE baseline (Han et al., 2023), thereby exhibiting a greater degree of competitiveness. Although the fine-tuned process necessitates additional computational resources and time, the experimental results indicate that these investments are worthwhile.

Additionally, our approach is better than some earlier traditional small models (e.g., BiLSTM (Yao et al., 2019)), but is still inferior to the latest traditional small models (e.g., DREEAM (Ma et al., 2023) and TTM-RE (Gao et al., 2024)). Our LLM-based results further narrows the performance gap with the latest traditional small models, making it a promising approach for future DocRE.

4.3 Ablation Study

To evaluate the contribution of each module to the model's performance, we conducted an ablation study on DocRED dev set, as illustrated in Table 3. Our observations include the following aspects:

Impact of entity pair selection Entity pair selection is very effective for LLM-based approach on DocRE task. Removing it leads to a sharp drop in performance. F1 and Ign F1 drop by 31.94 and 32.13 respectively. It indicates that selecting entity pairs that potentially contain relations further enhances the ability of LLM triplet fact extraction.

Impact of entity information for relation summarization Removing the entity information in the relation summarization, the F1 and Ign F1 of

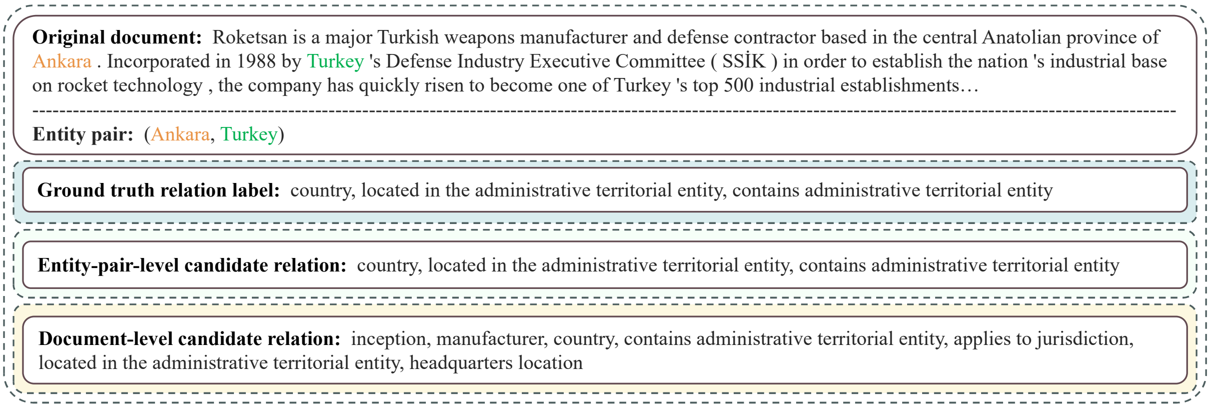

Figure 3: Case study for candidate relation selection.

our model decrease by 2.26 and 2 respectively, indicating that although the entity information is simple, adding it is very helpful for making the relation summarization contain more relation information.

Impact of entity-pair-level candidate relations Replacing the entit-pair-level candidate relations with the document-level candidate relations, the performance of our model experiences a significant drop, with F1 and Ign F1 dropping by 6.77 and 5.68 respectively, which indicates that our entity-pair-level method is more effective for selecting candidate relations.

Impact of triplet fact judgement Removing the triplet fact judgement, F1 and Ign F1 decreased by 6.12 and 6.76. It indicates that our method enables LLMs to better understand relations and facilitates more accurate relation extraction.

In addition, we conduct the ablation analysis on the multiple-choice QA and prior knowledge used in the fine filtering of candidate relations of Section 3.4, and the results are detailed in Appendix F.

4.4 Assessment of the effectiveness in alleviating NA problem

To further illustrate that our approach alleviates the NA problem, we perform data analysis based on the entity pair selection (denoted as Select) in Section 3.3. As shown in Table 4, compared with no selection, our method reduces the number of entity pairs by about 95.72%, reduces the percentage of NA data by about 45.32, and improves F1 by 31.94. It indicates that our method effectively alleviates NA problem in the LLM-based DocRE task.

Model	Number of entity pairs	F1	NA percentage
w Select	15480	53.77	51.85
w/o Select	362313	21.83	97.17

Table 4: Experimental analysis of the effectiveness of alleviating the NA problem.

4.5 Further Analysis of candidate relations at the entity-pair-level and document-level

The preceding ablation study has demonstrated the effectiveness of entity-pair-level candidate relations. To further investigate their significance, we replace the coarse filtering of candidate relations in Section 3.4 with document-level candidate relations and only calculate the F1 of this part as shown in Table 5. When using document-level candidate relations, the F1 score decreases by 8.28, accuracy decreases by 5.65, and recall decreases by 12.46, which indicates that the entity-pair-level candidate relation method is more effective.

Model	P	R	F1
Entity-pair-level	23.46	91.24	37.33
Document-level	17.81	78.78	29.05

Table 5: Results on different level candidate relations.

4.6 Case Study

We present a test case in Figure 3. In this case, compared with our entity-pair-level candidate relation method, the candidate relations obtained by the document-level method contain some correct answers, but the number of candidate relations is too large, and the correlation between many candidate relations and entity pairs is low. This case also illustrates the effectiveness of our entity-pair-level candidate relation method.

4.7 Time Cost Analysis

We perform an analysis of the time efficiency of our proposed model, EP-RSR, in terms of training and inference time, in comparison to the competitive baseline, AutoRE (Xue et al., 2024). Here, the reported training time corresponds to the duration required for fine-tuning LLMs, while the inference time refers to the time taken for LLMs to perform predictions.

The results, presented in Table 6, show that our approach *reduces* training time by 19,285 seconds and inference time by 2,814 seconds. Combining the results from Table 1 and Table 2, our model shows a significant improvement over AutoRE on the DocRE task. This further demonstrates that our model not only achieves superior performance but also maintains a relatively low time cost.

Model	Training time	Inference time
AutoRE	68672s	14877s
Our Method	49387s	12063s

Table 6: Time cost on the DocRED dataset.

5 Conclusion

In this paper, we introduce a novel LLM-based DocRE framework based on our proposed entity-pair-level candidate relations and a new LLM-based DocRE paradigm EPRF. Our model achieves SOTA results in LLM-based methods on DocRED, Re-DocRED and DWIE datasets, while also alleviating the NA problem in DocRE. Although our model is still inferior to the latest traditional small models, our LLM-based results further narrows the performance gap with the latest small models, making it a promising approach for future LLM-based DocRE.

Limitations

Error Propagation Issue The entire model is constrained by the initial entity pairs filtering step, which, while eliminating many irrelevant entity pairs, may also discard some genuinely relation-containing pairs. This can adversely affect the model's overall performance ceiling. To better address this issue, it may be beneficial to consider multiple entity pairs filtering sources or to employ alternative methods that alleviate the false positive problem commonly associated with LLMs.

Insufficient Training Data As shown in our experiment and analysis in Appendix G, the long-tail problem in train set of dataset results in a skewed distribution of relation instances, leading to a limited training effects for certain relations and subsequently deteriorating the prediction results for the corresponding relation triplets. To effectively tackle this challenge, leveraging external knowledge sources or generating relevant data using LLMs could help mitigate the long-tail issue.

Acknowledgments. The authors sincerely thank the reviewers for their valuable comments, which improved the paper. The work is supported by the National Natural Science Foundation of China (62276057).

References

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.

Fenia Christopoulou, Makoto Miwa, and Sophia Ananiadou. 2019. Connecting the dots: Document-level neural relation extraction with edge-oriented graphs. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 4925–4936.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL*, pages 4171–4186.

Chufan Gao, Xulin Fan, Jimeng Sun, and Xuan Wang. 2023. Promptre: Weakly-supervised document-level relation extraction via prompting-based data programming. *arXiv preprint arXiv:2310.09265*.

Chufan Gao, Xuan Wang, and Jimeng Sun. 2024. TTM-RE: Memory-augmented document-level relation extraction. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 443–458, Bangkok, Thailand. Association for Computational Linguistics.

Ridong Han, Tao Peng, Chaohao Yang, Benyou Wang, Lu Liu, and Xiang Wan. 2023. Is information extraction solved by chatgpt? an analysis of performance, evaluation criteria, robustness and errors. *arXiv preprint arXiv:2305.14450*.

Edward J Hu, yelong shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu

Chen. 2022. Lora: Low-rank adaptation of large language models. In *International Conference on Learning Representations*.

Junpeng Li, Zixia Jia, and Zilong Zheng. 2023. Semi-automatic data enhancement for document-level relation extraction with distant supervision from large language models. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 5495–5505.

Xingzuo Li, Kehai Chen, Yunfei Long, and Min Zhang. 2024. Llm with relation classifier for document-level relation extraction. *arXiv preprint arXiv:2408.13889*.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

Youmi Ma, An Wang, and Naoaki Okazaki. 2023. Dreeam: Guiding attention with evidence for improving document-level relation extraction. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, pages 1971–1983.

Yilmazcan Ozyurt, Stefan Feuerriegel, and Ce Zhang. 2023. Document-level in-context few-shot relation extraction via pre-trained language models. *arXiv preprint arXiv:2310.11085*.

Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Sunil Kumar Sahu, Fenia Christopoulou, Makoto Miwa, and Sophia Ananiadou. 2019. Inter-sentence relation extraction with document-level graph convolutional neural network. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4309–4316.

Qingyu Tan, Lu Xu, Lidong Bing, Hwee Tou Ng, and Sharifah Mahani Aljunied. 2022. Revisiting docred-addressing the false negative problem in relation extraction. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 8472–8487.

Hengzhu Tang, Yanan Cao, Zhenyu Zhang, Jiangxia Cao, Fang Fang, Shi Wang, and Pengfei Yin. 2020. Hin: Hierarchical inference network for document-level relation extraction. In *Advances in Knowledge Discovery and Data Mining: 24th Pacific-Asia Conference, PAKDD 2020, Singapore, May 11–14, 2020, Proceedings, Part I 24*, pages 197–209. Springer.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

A Vaswani. 2017. Attention is all you need. *Advances in Neural Information Processing Systems*.

Zhen Wan, Fei Cheng, Zhuoyuan Mao, Qianying Liu, Haiyue Song, Jiwei Li, and Sadao Kurohashi. 2023. Gpt-re: In-context learning for relation extraction using large language models. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 3534–3547.

Shuhe Wang, Xiaofei Sun, Xiaoya Li, Rongbin Ouyang, Fei Wu, Tianwei Zhang, Jiwei Li, and Guoyin Wang. 2023a. Gpt-ner: Named entity recognition via large language models. *arXiv preprint arXiv:2304.10428*.

Xiao Wang, Wei Zhou, Can Zu, Han Xia, Tianze Chen, Yuan Zhang, Rui Zheng, Junjie Ye, Qi Zhang, Tao Gui, Jihua Kang, J. Yang, Siyuan Li, and Chunsai Du. 2023b. Instructuie: Multi-task instruction tuning for unified information extraction. *ArXiv*, abs/2304.08085.

Yiqing Xie, Jiaming Shen, Sha Li, Yuning Mao, and Jiawei Han. 2022. Eider: Empowering document-level relation extraction with efficient evidence extraction and inference-stage fusion. In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 257–268.

Derong Xu, Wei Chen, Wenjun Peng, Chao Zhang, Tong Xu, Xiangyu Zhao, Xian Wu, Yefeng Zheng, and Enhong Chen. 2023. Large language models for generative information extraction: A survey. *arXiv preprint arXiv:2312.17617*.

Lilong Xue, Dan Zhang, Yuxiao Dong, and Jie Tang. 2024. Autore: Document-level relation extraction with large language models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)*, pages 211–220.

Yuan Yao, Deming Ye, Peng Li, Xu Han, Yankai Lin, Zhenghao Liu, Zhiyuan Liu, Lixin Huang, Jie Zhou, and Maosong Sun. 2019. Docred: A large-scale document-level relation extraction dataset. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 764–777.

Klim Zaporojets, Johannes Deleu, Chris Develder, and Thomas Demeester. 2021. Dwie: An entity-centric dataset for multi-task document-level information extraction. *Information Processing Management*, 58(4):102563.

Shuang Zeng, Yuting Wu, and Baobao Chang. 2021. Sire: Separate intra-and inter-sentential reasoning for document-level relation extraction. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 524–534.

Shuang Zeng, Runxin Xu, Baobao Chang, and Lei Li. 2020. Double graph based reasoning for document-level relation extraction. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1630–1640.

Wei Zheng, Hongfei Lin, Zhiheng Li, Xiaoxia Liu, Zhengguang Li, Bo Xu, Yijia Zhang, Zhihao Yang, and Jian Wang. 2018. An effective neural model extracting document level chemical-induced disease relations from biomedical literature. *Journal of biomedical informatics*, 83:1–9.

Yaowei Zheng, Richong Zhang, Junhao Zhang, Yanhan Ye, and Zheyan Luo. 2024. Llamafactory: Unified efficient fine-tuning of 100+ language models. *arXiv preprint arXiv:2403.13372*.

Huixue Zhou, Mingchen Li, Yongkang Xiao, Han Yang, and Rui Zhang. 2023. Llm instruction-example adaptive prompting (leap) framework for clinical relation extraction. *medRxiv*, pages 2023–12.

Wenxuan Zhou, Kevin Huang, Tengyu Ma, and Jing Huang. 2021. Document-level relation extraction with adaptive thresholding and localized context pooling. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 14612–14620.

A Details and Statistics of Datasets

We conducted our experiments on the DocRED (Yao et al., 2019), Re-DocRED (Tan et al., 2022), and DWIE (Zaporojets et al., 2021) datasets. Statistics for three datasets are reported in Table 7.

DocRED is a dataset designed for the task of relation extraction from multi-paragraph documents, comprising 132,375 entities, 1,829,756 entity pairs and 56,354 relation facts annotated across 5,053 Wikipedia articles. It comprises 3,053 train documents, 998 dev documents, and 1,000 test documents, encompassing 96 relation types. In contrast to previous relation extraction tasks that primarily focused on single sentences or short text corpora, DocRED incorporates more complex contextual information and inter-paragraph relations, necessitating models to possess enhanced understanding and reasoning capabilities.

Re-DocRED dataset serves as an improvement over DocRED, analyzing the causes and impacts of the false negative issues present in the original dataset, and re-annotating a total of 4,053 documents from DocRED to address these concerns. It comprises 3,053 train documents, 500 dev documents, and 500 test documents.

DWIE consists of 802 general news articles in English, randomly selected from a corpus collected from the German broadcaster Deutsche Welle between 2002 and 2018. It is annotated at the document level for NER, coreference resolution, RE, and entity linking. It comprises 602 train documents, 98 dev documents, and 99 test documents, encompassing 65 relation types.

Dataset	Split	#Doc.	#Rel.	#Ent.	#Facts.
DWIE	train	602	65	16,494	14,403
	dev	98		2,785	2,624
	test	99		2,623	2,495
DocRED	train	3,053	96	59,493	38,180
	dev	998		19,578	12,323
	test	1,000		19,539	-
Re-DocRED	train	3,053	96	59,359	85,932
	dev	500		9,684	17,284
	test	500		9,779	17,448

Table 7: Statistics on datasets, where Doc. (resp. Rel or Ent) abbreviates documents (resp. relations or entities).

B Parameter Settings

B.1 Fine-tuned Parameters

For the entity pair selection, multiple-choice QA, and triple fact judgment stages, the parameters of LLM fine-tuning are shown in the Table 12.

B.2 Inference Parameters

The parameter settings for LLMs inference at each stage are shown in the Table 8.

Stage	temperature	top_p
Entity information	0.9	0.9
Relation summarization	0.9	0.9
Entity pair selection	0.9	0.9
Multiple-choice QA	0.0001	0.9
Triplet fact judgment	0.1	0.9

Table 8: Inference parameters at each stage.

C Prompt Templates

We use different prompts for LLM in different experimental stages. For LLMs in the entity information and relation summarization stage, we adopt the non-fine-tuned approach. The specific prompts are shown in the Table 13. For LLMs in the entity pair selection, multiple-choice QA, and triplet fact judgment stages, we use fine-tuned. The specific instruct tuning template is shown in the Table 14. In the multiple-choice QA stage, the relation template that converts candidate relations into natural language sentences is shown in Table 16. The relation description of the candidate relations involved in the triplet fact judgment stage is shown in Table 17.

D Experimental Analysis of Parameters *k*

D.1 Impact of the parameter *k* on the top-k retrieved candidate relations

As illustrated in the Table 9, an increase in the parameter *k* corresponds to a rise in the number of true relation labels within the retrieved candidate relations. However, this augmentation is accompanied by a decline in the overall accuracy of the experimental results.

Parameter	P	R	F1
k=80	52.76	54.81	53.77
k=40	53.04	54.57	53.79
k=20	53.36	54.19	53.77
k=10	53.73	53.40	53.57
k=5	54.29	52.26	53.26
k=1	56.51	37.41	45.02

Table 9: The impact of the parameter *k* on the top-k retrieved candidate relations on the DocRE.

D.2 Analysis of the parameter *k* on multi-sampling

The choice of parameter *k* on multi-sampling in entity pair selection stage is crucial. Increased sampling improves estimation precision by reducing random errors. Larger *k* values provide better data distribution insights and enhance model robustness to outliers. However, higher sampling iterations also increase computational costs. The results, as shown in Figure 4, indicate that increased sampling improves recall but reduces accuracy, emphasizing the need for an optimal choice of *k*.

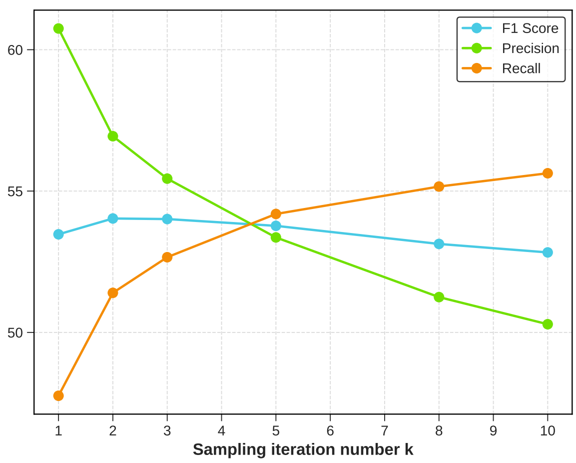

Figure 4: Analysis of parameter *k* on multi-sampling.

Model	Dev	Test
AutoRE(Mistral-7B)*	53.01	51.91
AutoRE(Llama3-8B)*	57.79	56.74
Our Method(Mistral-7B)	54.90	53.41
Our Method(Llama3-8B)	**58.77**	**56.81**

Table 10: Experimental F1 results on Re-DocRED for the DocRTE Task. Results marked with * are our reproductions.

E Scalability Analysis of EP-RSR on DocRTE

To validate the performance of our model on the DocRTE, we make a modification to the structure of our model EP-RSR:

(1) The prompt in the entity pair selection phase has been changed, and the input of the entity list has been deleted. The specific prompt is shown in Table 15. The subsequent selecting entity pairs that potentially contain relations part, which selects appropriate entities based on cosine similarity, has been deleted.

(2) Because our method uses prior knowledge related to entities, the specific type of the entity is required. Therefore, we added an additional step of predicting the entity type in the entire experiment to obtain the entity type of a given entity based on LLMs. This step only increases the additional reasoning time, and the training time remains unchanged. The specific additional reasoning time is 467s. The reasoning time of the entire experimental large model is 12530s, which is still lower than the reasoning time of the comparison model AutoRE of 14877s.

We compare it with AutoRE, using the same evaluation metrics as those used in AutoRE for DocRTE. The comparison results in Table 10 indicate that EP-RSR demonstrates strong competitiveness and generalization in DocRTE. Compared with AutoRE, our method shows an improvement of 1.89 in F1 score on the Re-DocRED dev set for the DocRTE task. Considering that our model is primarily designed for DocRE, the performance improvement in the DocRE task demonstrates that our method is clearly more advantageous in the DocRE task than in DocRTE.

F Analysis of Multiple-choice QA and Prior Knowledge

we conduct the ablation analysis on the multiple-choice QA and prior knowledge mentioned in Section 3.4, as illustrated in Table 11. Our observations include the following aspects:

Impact of multiple-choice QA Removing multiple-choice QA from the fine filtering of candidate relations causes the F1 and Ign F1 scores to drop by 13.07 and 11.71, which further illustrates that after the initial coarse filtering of candidate relations, further fine filtering is still required.

Impact of prior knowledge Removing prior knowledge in mentioned in Section 3.4, F1 and Ign F1 decreased 3.46 and 3.56. It indicates that prior knowledge is helpful in removing more insignificant candidate relations.

Model	F1	Ign F1
Our Method	**53.77**	**51.25**
w/o Multiple-choice QA	40.70	39.54
w/o Prior knowledge	50.31	47.69

Table 11: Analysis of Multiple-choice QA and Prior Knowledge.

G Analysis of Long-tail Problem in Train Set

There exists a long-tail problem in DocRE dataset, where many relations have a small number of associated labels. To further assess the impact of relation labels number in the train set on the model, we calculated the F1 score for each relation in the DocRED dev set. The results are shown in Figure 5. As the number of relation labels in the train set decreases, the overall F1 score for the relation also exhibits a downward trend. There is a clear positive correlation between the frequency of relations in the train set and their relation extraction performance. For example, the high-frequency relation P569 represents the “date of birth” relation with an F1 score of 88.82, while the low-frequency relation P39 represents the “position held” relation with an F1 score of 0. It indicates that the presence of the long-tail problem will degrade the triplet prediction results for some relations. To better address this issue, one might consider leveraging external knowledge sources or LLMs to generate corresponding data, thereby alleviating the long-tail problem.

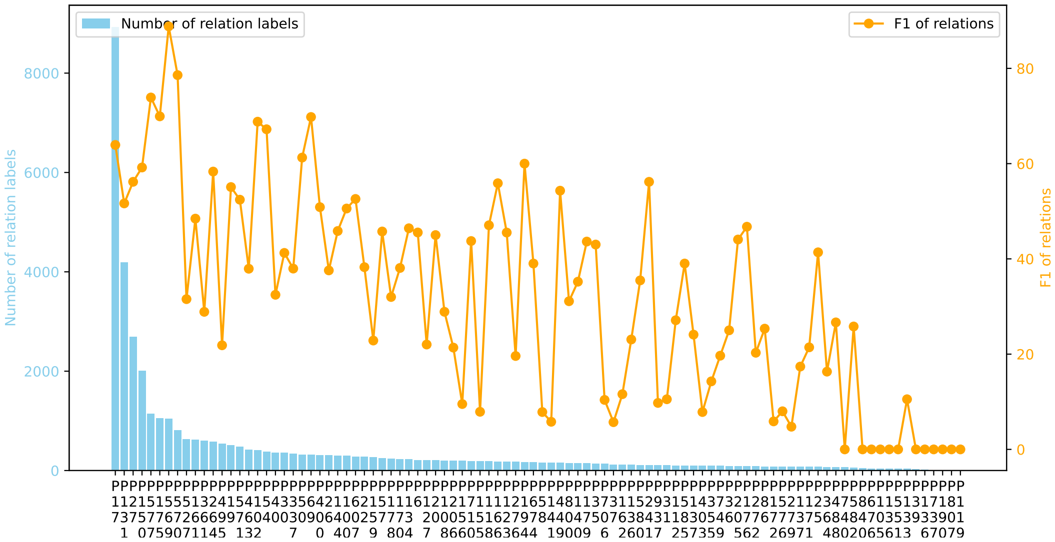

Figure 5: Analysis of long-tail problem in train set.

Stage	Parameter
Entity pair selection	lora_rank: 64, lora_alpha: 128, lora_dropout: 0.05, cutoff_len: 2048, learning_rate: 1.0e-4, num_train_epochs: 5.0, warmup_ratio: 0.1, bf16: true
Multiple-choice QA	lora_rank: 64, lora_alpha: 128, lora_dropout: 0.05, cutoff_len: 2048, learning_rate: 1.0e-4, num_train_epochs: 1.0, warmup_ratio: 0.1, bf16: true
Triplet fact judgment	lora_rank: 64, lora_alpha: 128, lora_dropout: 0.05, cutoff_len: 2048, learning_rate: 1.0e-4, num_train_epochs: 1.0, warmup_ratio: 0.1, bf16: true

Table 12: Fine-tuned parameters.

Stage	Prompt
Entity information	The text is as follows: {title} {doc} What is "{entity}"? (For example, US is a country and 12 is a number) Answer in one sentence. Only output answers without outputting anything else. The answer is:
Relation summarization	The text is as follows: {title} {doc} What is the relationship between {entity_h} and {entity_t}? Answer in one sentence. Only output answers without outputting anything else.

Table 13: Entity information prompt and Relation summarization prompt.

Stage	Instruct Tuning Template
Entity pair selection	Given a text and an entity list as input, list the entity pairs that can be identified as possibly containing a relation. ## Text: {doc_text} ## Entity list: {entity_list}
Multiple-choice QA	Determine which option can be inferred from the given text. ## Text: {doc_text} ## Options: {options}
Triplet fact judgment	Based on the text and the description of the relation "{rel}", give an answer about whether the head and tail entity pairs (head entity and tail entity) satisfy the "{rel}" relation. ## Relation description: {rel_description} ## The text to be extracted: {doc_text} ## Entity pair to be extracted: {extract_entity_pair}

Table 14: Instruct Tuning Template for EPRF.

Stage	Pormpt
Entity pair selection	Given a text as input, list the entity pairs that can be identified as possibly containing a relation. ## Text: {doc_text}

Table 15: Pormpt for DocRTE.

Relation	Relation Template
P6	`<tail> is the head of government of <head>`
P17	`<head> is located in country <tail>`
P19	`<tail> is the place of birth of <head>`
P20	`<tail> is the place of death of <head>`
P22	`<tail> is the father of <head>`
P25	`<tail> is the mother of <head>`
P26	`<tail> is the spouse of <head>`
P27	`<head> is a citizen of <tail>`
P30	`<head> is on the continent of <tail>`
P31	`<head> is the instance of <tail>`
P35	`<tail> is the head of state <head>`
P36	`<tail> is the capital of <head>`
P37	`<tail> is the official language of <head>`
...	...

Table 16: Relation Template.

Relation	Relation Description
country	For the 'country' relation, the subject pertains to a non-human entity, such as an organization, place, or event. The object signifies the sovereign state where the subject is based or occurs. Example: (Amazon Inc, country, United States).
country of citizenship	The 'country of citizenship' relation denotes that the subject, an individual, is recognized as a citizen by the object, a country. Example: (Elon Musk, country of citizenship, United States).
contains administrative territorial entity	The relation 'contains administrative territorial entity' involves a subject, an administrative territory, encompassing the object, a subdivision or part of this administrative territory. Example: (California, contains administrative territorial entity, Los Angeles).
has part	The 'has part' relation reflects that the subject, an entity or whole, comprises the object, a part or component of the subject. Example: (A car, has part, engine).
date of birth	In the 'date of birth' relation, the subject, a person, was born on the object, the specified date. Example: (John Doe, date of birth, January 1, 1990).
part of	In the 'part of' relation, the subject, a component or section, belongs to the object, a larger whole or aggregate. Example: (Engine, part of, a car).
...	...

Table 17: Relation Description.