BorderRE: An Open-Domain Relation Extraction Framework via Dynamic Boundaries

Anonymous ACL submission

Abstract

Zero-shot relation extraction (ZSRE) in open-domain scenarios requires models to not only accurately classify unseen relations but also robustly reject out-of-distribution (OOD) noise. Existing methods predominantly rely on static distance thresholds, which overlook the anisotropic nature of relation manifolds, inevitably leading to overconfidence and sub-optimal rejection performance. To address these limitations, we propose BorderRE, a novel framework driven by dynamic boundaries. Specifically, we adaptively generate polyhedral boundaries for each known relation and construct an absolute rejection buffer via boundary-aware mutually exclusive contrastive learning. This mechanism simultaneously enhances zero-shot classification accuracy and enables high-precision OOD interception. Extensive experiments on standard benchmarks demonstrate that BorderRE consistently outperforms state-of-the-art methods in both classification and open-set rejection tasks.

1 Introduction

Open-domain relation extraction is a critical technology for constructing and maintaining dynamic knowledge graphs (Pawar et al., 2017). Traditional supervised learning paradigms rely heavily on extensive annotated data, struggling to generalize to unseen relations (Wu and He, 2019). Consequently, zero-shot relation extraction (ZSRE) has emerged as a prominent research focus (Levy et al., 2017). In practical applications, such as technological intelligence analysis, systems are required not only to accurately identify known relations but also to exhibit robust rejection capabilities against out-of-distribution (OOD) noise.

Current ZSRE research typically fine-tunes small language models (SLMs) (Li et al., 2024b; Zhang et al., 2025; Zhao et al., 2023) to obtain task-specific embeddings, or leverages the in-context learning and retrieval capacities of large language models (LLMs) (Li et al., 2025). While LLMs demonstrate remarkable zero-shot deduction abilities, deploying them to process massive textual corpora incurs prohibitive API costs and latency, hindering large-scale industrial application. Consequently, prototype-based dual-tower architectures (Li et al., 2024b,a) powered by contrastive learning (Chen et al., 2020; Gao et al., 2021) remain the dominant paradigm due to their efficiency. These methods typically model relations as solitary prototype points in the feature space, relying on global static thresholds or nearest-neighbor searches to intercept unseen samples. However, in real-world inference, the actual distribution of relation instances is rarely symmetric; rather, it exhibits highly anisotropic, manifold-like structures. A unified static threshold fails to accommodate this geometric complexity—relations with narrow semantics cluster compactly, whereas those with broader semantics disperse sparsely. Relying on metric-based maximum similarity and static thresholds inevitably leads to overconfidence towards OOD noise. To mitigate this, we propose fitting relation manifolds with dynamic boundary vectors, which serve as explicit geometric dividing lines between intra-class and inter-class features. This mechanism effectively filters noise and tightly constrains the distribution of instances and relations.

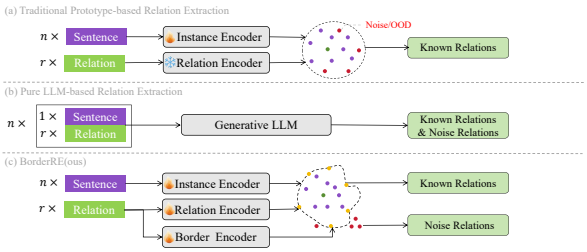


Figure 1: Comparison of relation extraction paradigms. BorderRE utilizes dynamic boundaries for efficient OOD rejection compared to static thresholds.

To resolve these challenges, we propose Bor-

derRE, a novel open-domain ZSRE framework driven by dynamic boundaries. As illustrated in Figure 1, BorderRE departs from traditional paradigms by employing optimized feature encoding as a foundation and dynamically generating physical boundaries to support high-precision extraction and robust OOD rejection.

In summary, our main contributions are threefold:

- **Anisotropic Dynamic Boundary Generation**: We propose a parameterized boundary generator that adaptively constructs polyhedral boundaries for each relation. Coupled with a novel boundary-aware mutually exclusive loss, this mechanism shatters the isotropic assumption, significantly enhancing both zero-shot classification accuracy and open-set OOD rejection.

- **Alignment-Enhanced Dual-Tower Architecture**: We design a co-training framework that aligns instance manifolds with relation prototypes, establishing a high-quality, geometrically structured representation foundation for boundary definition.

- **Two-Stage Progressive Training**: We introduce a decoupled training strategy that isolates semantic alignment from boundary optimization, effectively preventing the severe fluctuation of dynamic boundaries during early training phases and ensuring stable convergence.

2 Related Work

Existing ZSRE methods can be broadly categorized into representation-learning (SLM-based) and LLM-based approaches. Representation-learning methods, such as AlignRE (Li et al., 2024b) and EMMA (Li et al., 2024a), introduce contrastive learning to align the sentence and relation description embeddings in a dual-tower architecture. Building upon this, CE-DA (Zhang et al., 2025) proposes custom embedding and dynamic aggregation to strengthen the interaction between different types of relation side information. On the other hand, LLM-based methods leverage generative and retrieval capabilities to address zero-shot scenarios. For instance, QA4RE (Zhang et al., 2023) formulates ZSRE as a question-answering task, RelationPrompt (Chia et al., 2022) utilizes prompt tuning to generate synthetic data, while RE-GAR-AD (Li et al., 2025) redefines it as a generation-augmented retrieval problem using LLM embeddings to reduce token input overhead.

Despite improvements in prototype and sentence representations, these methods predominantly rely on static thresholds or nearest-neighbor searches for out-of-distribution rejection, lacking explicit modeling of cluster boundary distributions. In contrast, our approach relaxes the constraints of standard dual-tower joint training and focuses on constructing adaptive dynamic boundaries. By explicitly modeling the spatial limits of relation clusters, our method breaks the isotropic assumption, significantly optimizing open-set rejection performance while maintaining high fundamental classification capabilities.

3 Methodology

The proposed framework, BorderRE, consists of two progressive components, as illustrated in Figure 2: the underlying alignment-enhanced dual-tower feature encoding, and the core dynamic boundary generation with mutually exclusive constraints.

3.1 Alignment-Enhanced Dual-Tower Feature Encoding

Following the paradigm established by AlignRE (Li et al., 2024b), we adopt a dual-tower architecture to obtain high-quality fundamental representations. For the input instances, we utilize the Sentence Encoder (Devlin et al., 2019) with a prompt-tuning strategy to extract the hidden states of `[MASK]` tokens corresponding to the entities and relation slot. These states are then aggregated via a weighted pooling mechanism to form the instance representation v_{ins}:

$$v_{ins} = \alpha H_{[\text{MASK}]_h} + \beta H_{[\text{MASK}]_t} + \gamma H_{[\text{MASK}]_r} \quad (1)$$

On the relation side, we employ the Prototype Encoder to embed multiple side information (such as label names, descriptions, and aliases) into initial embeddings. These are subsequently fused through semantic alignment to obtain the comprehensive relation prototype P_r:

$$P_r = \sum_{j=1}^{n} W_j * p_r^j \quad (2)$$

While these encoding schema alignment techniques provide a robust feature foundation, our core contribution lies in how we dynamically define boundaries upon these representations. Joint training

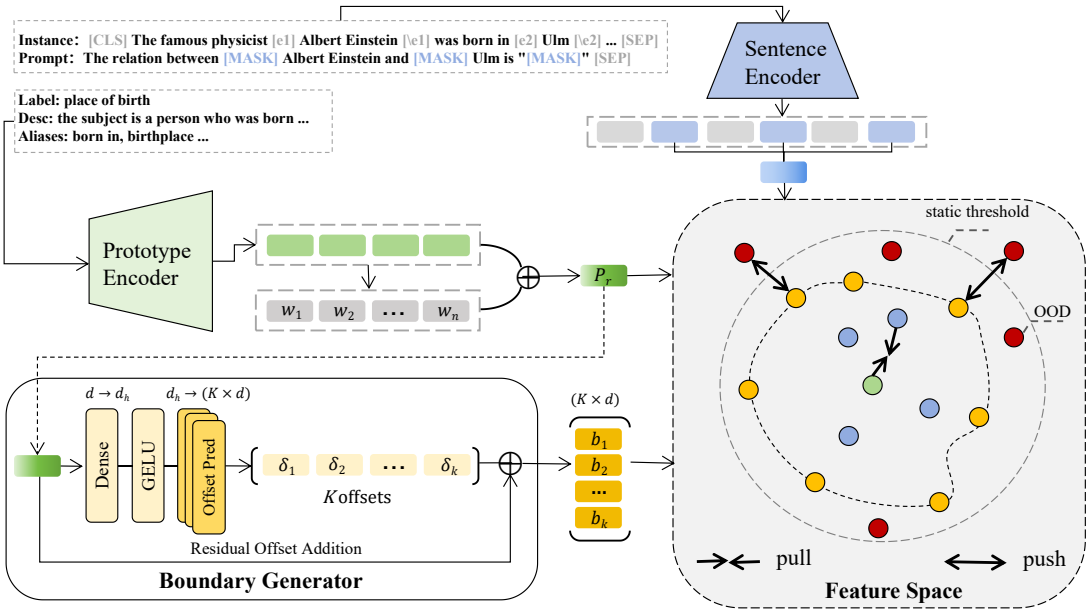


Figure 2: The overall framework of open-domain relation extraction via BorderRE based on dynamic boundaries.

allows static relation prototypes to approach the manifold of current instances, establishing a high-quality representation foundation for subsequent boundary definitions.

3.2 Dynamic Boundary Generation and Mutually Exclusive Constraints

In open-domain extraction, to overcome the defects of a single prototype static threshold, BorderRE uses a parameterized boundary generator to dynamically generate K polyhedral boundary vectors $B_r = \{b_1, \ldots, b_K\}$ for each relation prototype P_r:

$$b_k = P_r + \delta_k, \quad \delta_k = \text{MLP}(P_r) \tag{3}$$

This parameterized boundary generator is designed in accordance with three core principles: *lightweight efficiency*, *spatial correlation*, and *geometric flexibility*. First, rather than generating high-dimensional boundary vectors from scratch, the MLP solely predicts a spatial offset δ_k, which drastically reduces the parameter optimization space. Second, by integrating the offset δ_k with the original relation prototype P_r via a residual connection, we explicitly guarantee that the generated boundaries remain tightly anchored to their respective semantic cluster centers. Finally, this generative paradigm empowers the boundaries to dynamically adjust their shapes and radii based on the semantic density of specific relation clusters, seamlessly adapting to the anisotropic distribution inherent in open-domain feature spaces.

Specifically, the MLP is implemented as a two-layer feed-forward network activated by the Gaussian Error Linear Unit (GELU) (Hendrycks and Gimpel, 2016). Compared to standard ReLU, GELU provides a smoother, non-monotonic transformation in the negative domain, mitigating the dead neuron problem and facilitating refined gradient propagation during multidirectional spatial offset prediction. Furthermore, to prevent the generated boundaries from collapsing into a singular direction, we impose an orthogonal initialization constraint on the weights of the final offset prediction layer, ensuring a uniform distribution of boundary vectors across the spatial manifold.

Building upon this, we introduce a boundary-aware mutually exclusive contrastive learning loss (L_B) to impose bidirectional physical geometric constraints in the feature space. For positive sample constraints, acting as central attraction, we force the similarity between positive instances and their corresponding relation prototypes to be strictly higher than the dynamic threshold $T_p = \text{sim}(b_k, P_{y_i})$ defined by the boundary vectors. For negative sample constraints, acting as marginal repulsion, we force negative samples to be pushed to the out-of-bounds area formed by the polyhedral boundary vectors.

Specifically, the boundary-aware loss L_B consists of the positive sample boundary loss L_p and the negative sample boundary loss L_n, where α and β are scaling factors, and $\mathcal{N}_i$ is the negative

sample set:

$$L_p = \frac{1}{\alpha} \log \Big(1 + \sum_{k=1}^{K} \exp\big(-\alpha(\mathrm{sim}(x_i, P_{y_i}) - T_p)\big)\Big) \quad (4)$$

$$L_n = \frac{1}{\beta} \log \Big(1 + \sum_{x_{neg} \in \mathcal{N}_i} \exp\big(\beta(\mathrm{sim}(x_{neg}, P_{y_i}) - T_n)\big)\Big) \quad (5)$$

The positive and negative bidirectional constraints jointly construct a clear and absolute rejection buffer zone in the feature space.

During actual inference, the boundary vectors are utilized to formulate a definitive rejection rule. Specifically, the decision threshold $\bar{T}_r$ for a relation r is calculated as the average similarity between its relation prototype P_r and all corresponding K boundary vectors:

$$\bar{T}_r = \frac{1}{K} \sum_{k=1}^{K} \mathrm{sim}(b_k, P_r) \quad (6)$$

For an input instance x_i, if its maximum similarity to any known relation prototype P_r is lower than the corresponding average boundary threshold $\bar{T}_r$, x_i is rejected as an OOD sample. Otherwise, it is classified into the relation category that yields the highest similarity score.

3.3 Two-Stage Progressive Training

To decouple feature space alignment from boundary optimization, BorderRE employs a two-stage progressive training strategy. If the Sentence Encoder and boundary generator are trained jointly from scratch, the unaligned semantic space will cause the dynamic boundaries to fluctuate violently, severely disrupting the formation of fundamental relation clusters.

During the first phase (*Semantic Alignment*), we freeze the parameters of the boundary generator to concentrate on establishing a structured initial semantic space. Diverging from previous methods like AlignRE that strictly freeze the Prototype Encoder, we unfreeze it to perform collaborative training (co-training) alongside the Sentence Encoder. This paradigm enables static relation prototypes to flexibly converge towards the shifting manifold of instance embeddings. Following AlignRE, we adopt a margin-based contrastive loss L_M as the fundamental alignment objective, compelling instance vectors to form preliminary clusters around their corresponding relation prototypes:

$$L_M = \sum_{i=1}^{N} \max\Big(0, \max_{j \neq y_i}\big(\mathrm{sim}(v_{ins}^{(i)}, P_r^{(j)})\big) - \mathrm{sim}(v_{ins}^{(i)}, P_r^{(y_i)}) + \Delta\Big) \quad (7)$$

where N is the batch size, y_i is the ground-truth relation index for the i-th instance, and Δ is a predefined margin. This constructs a clearly structured semantic manifold and alleviates semantic drift.

During the second phase (*Boundary Collaboration*), we freeze the Prototype Encoder, unfreezing only the boundary generator and the Sentence Encoder. The imperative for freezing the Prototype Encoder lies in preserving the high-quality, stable relation anchors established in the preceding phase. Continuous updating of the Prototype Encoder would cause semantic centers to drift, depriving the boundary vectors of their structural anchors and triggering severe optimization instability. Consequently, anchoring the cluster centers allows us to fine-tune the Sentence Encoder with a minimal learning rate while the boundary generator focuses exclusively on deducing spatial offsets for OOD rejection. The joint optimization objective for this phase is the weighted sum of the margin loss and the boundary loss:

$$L_{tot} = L_M + \lambda L_B \quad (8)$$

Through decoupled training, the boundary generator focuses on learning the rejection mechanism, allowing boundary vectors and instance vectors to optimize their distribution around stable relation prototypes.

4 Experimental Results and Analysis

4.1 Experimental Setup

Datasets We evaluate our methods on two RE datasets: (1) FewRel (Han et al., 2018) and (2) Wiki-ZSL (Sorokin and Gurevych, 2017). Further details about the dataset preprocessing and data statistics are provided in the Appendix.

Implementation Details To ensure a fair and rigorous comparison, we adopt the same underlying sentence encoder (`bert-base-uncased` (Devlin et al., 2019)) as our primary baseline, AlignRE (Li et al., 2024b). During the Phase 1 collaborative training, both the instance encoder and the prototype encoder are optimized using a learning rate of 2e-6 for 5 epochs. In Phase 2, we freeze the prototype encoder to preserve the semantic anchors, while the instance encoder and the boundary generator are trained for another 5 epochs with learning rates of 1e-7 and 1e-4, respectively. We implement our framework using PyTorch and conduct all experiments on a single NVIDIA RTX 3090 GPU.

Model / Variant	Unseen=5			Unseen=10			Unseen=15		
	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1
ZS-BERT (Wu et al., 2025)*	64.91	64.50	64.70	58.76	57.94	58.35	54.02	53.25	53.63
AlignRE (Li et al., 2024b)†	81.33	80.89	81.11	73.43	72.80	73.11	67.92	67.05	67.48
QA4RE(GPT-3.5) (Zhang et al., 2023)*	70.21	65.44	67.74	61.15	58.90	60.00	53.42	55.12	54.26
REA(GPT-3.5) (Layegh et al., 2024)*	78.44	72.15	75.16	68.32	65.20	66.72	58.10	60.33	59.19
CE-DA (Zhang et al., 2025)*	84.55	83.90	84.22	76.12	75.50	75.81	71.05	70.22	70.63
RE-GAR-AD (Li et al., 2025)*	88.96	88.42	88.69	77.63	77.85	77.74	70.80	69.26	70.02
BorderRE (Phase 1)	89.88	89.45	89.66	83.13	81.80	82.46	77.03	75.24	76.12
BorderRE (Phase 2)	**91.14**	**90.48**	**90.81**	**85.70**	**84.81**	**85.25**	**77.53**	**75.61**	**76.56**

Table 1: Comparison of standard zero-shot relation classification performance on the Wiki-ZSL dataset. Baselines with * denote reference methods.

Model / Variant	Unseen=5			Unseen=10			Unseen=15		
	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1
ZS-BERT (Wu et al., 2025)*	88.02	87.33	87.67	76.54	75.80	76.17	67.24	66.26	66.75
AlignRE (Li et al., 2024b)†	93.30	92.90	93.10	86.41	85.14	85.77	77.63	77.00	77.31
QA4RE(GPT-3.5) (Zhang et al., 2023)*	86.20	80.51	83.26	75.13	72.66	73.87	60.12	62.33	61.21
REA(GPT-3.5) (Layegh et al., 2024)*	92.57	84.70	88.46	82.26	79.47	80.84	64.34	68.68	66.44
CE-DA (Zhang et al., 2025)*	95.26	95.08	95.17	88.61	87.60	88.10	84.03	82.60	83.31
RE-GAR-AD (Li et al., 2025)*	93.03	92.16	92.59	83.48	82.10	82.78	77.69	75.31	76.48
BorderRE (Phase 1)	94.44	93.65	94.04	87.73	86.07	86.89	83.52	82.22	82.86
BorderRE (Phase 2)	**95.91**	**95.68**	**95.79**	**90.62**	**89.93**	**90.27**	**84.31**	**83.52**	**83.91**

Table 2: Comparison of standard zero-shot relation classification performance on the FewRel dataset. Baselines with * denote reference methods.

The global quantity of boundary vectors is set to $K = 8$.

Baselines For zero-shot baselines on the FewRel and Wiki-ZSL datasets, our baseline models include ZS-BERT (Wu et al., 2025), AlignRE (Li et al., 2024b), and CE-DA (Zhang et al., 2025). For LLMs Baselines, in evaluating prompt-based LLM baselines, we selected QA4RE (Zhang et al., 2023), REA (Layegh et al., 2024), and the recent generation-augmented retrieval method RE-GAR-AD (Li et al., 2025) for comparison. We also present the performance using a vanilla prompt strategy (denoted as Vanilla), which involves directly prompting LLMs to deduce the relation within a sentence, absent any in-context demonstrations.

4.2 Standard Zero-Shot Classification Performance

To evaluate the fundamental classification capabilities of BorderRE, we first conduct standard zero-shot classification evaluations (without OOD rejection) on the FewRel and Wiki-ZSL benchmarks, varying the number of unseen categories ($m \in \{5, 10, 15\}$). As detailed in Table 2 and Table 1, under the most stringent setting ($Unseen = 15$), the complete BorderRE (Phase 2) achieves state-of-the-art F1 scores of 83.91% and 76.56%, respectively. These results comprehensively surpass strong baselines, including AlignRE and the GPT-3.5-backed QA4RE and REA, demonstrating that our dynamic boundaries and optimized feature encoding effectively enforce intra-class compactness.

Table 2 and Table 1 present the outputs of our progressive training model compared to baselines. The results of BorderRE (Phase 1) directly correspond to the output of the first phase (semantic alignment), where the model uses collaborative training to construct an initial semantic space. BorderRE (Phase 2) is the result after introducing dynamic boundary generation and mutually exclusive constraints.

Compared to the base model AlignRE, BorderRE (Phase 1) demonstrates significant improvements. By unfreezing the Prototype Encoder and performing collaborative training, Phase 1 allows the relation prototypes to dynamically approach the instance manifold, leading to a much better feature space alignment. For instance, on the Wiki-ZSL dataset (Unseen=15), Phase 1 improves the F1 score from 67.48% (AlignRE) to 76.12%. Furthermore, after introducing dynamic boundaries in Phase 2, the F1 score under FewRel (Unseen=15) increases from 82.86% (Phase 1) to 83.91%, and on Wiki-ZSL (Unseen=15) from 76.12% to 76.56%. This proves that dynamic boundaries effectively isolate overlapping clusters and draw similar samples closer, achieving the best comprehensive clas-

sification performance.

4.3 Open-Set Noise Rejection Validation

Beyond standard classification, we rigorously evaluate the model's robustness against open-set noise. To construct the noise distribution for open-set rejection, we partition the unseen relations into five folds: one fold is held out as out-of-distribution (OOD) relations (i.e., near-OOD noise), while the remaining folds serve as in-domain (ID) relations that the model is tasked to predict. We adopt three primary metrics for evaluating rejection performance: ID F1 (measures classification ability on known relations), OOD F1 (measures interception accuracy on unknown relations), and Harmonic F1 (H-score) which comprehensively reflects the overall open-set performance.

The in-domain and cross-domain rejection performance on FewRel is reported in Table 3, where Wiki-ZSL serves as the OOD source for the cross-domain setting. To ensure a rigorous and fair comparison, the optimal static threshold for the baselines is determined via a grid search (from 0.1 to 0.9 with a step size of 0.1) on the validation set, selecting the threshold that maximizes the H-score. Compared to the static threshold paradigm, BorderRE's dynamic boundaries yield a substantial improvement, elevating the in-domain OOD F1 score from 60.68% to 68.64% ($Unseen = 15$). When confronted with the starkly different cross-domain noise distribution, the OOD F1 score surges to 78.19%. These findings empirically validate that the anisotropic adaptability of dynamic boundaries successfully mitigates model overconfidence towards low-density OOD instances, enabling high-precision interception.

4.4 Ablation Study

To deeply analyze the contribution of each core component, we conduct an ablation study on the FewRel dataset.

As visually demonstrated in Figure 3 and detailed in Table 4, the complete BorderRE model achieves an F1 score of 83.91%. Substituting the boundary-aware mutually exclusive loss with standard contrastive loss (w/o Boundary Loss, functionally reverting to Phase 1) drops the F1 score to 82.86%, underscoring that dynamic physical boundaries effectively optimize feature space cluster aggregation over isotropic distance metrics. Furthermore, freezing the Prototype Encoder during the first stage (w/o Co-train) drops the F1 score to

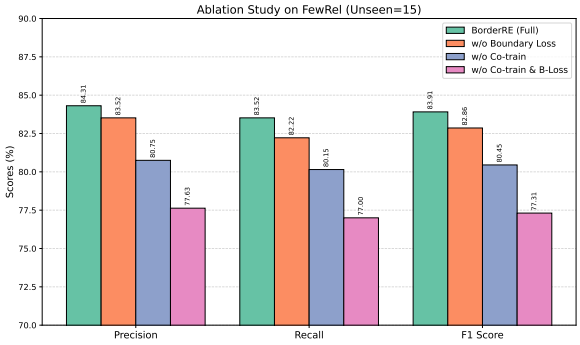


Figure 3: Ablation study of feature extraction and boundary components on FewRel (Unseen=15).

80.45%. This validates that collaborative training is crucial; it permits relation prototypes to dynamically approximate the instance manifold, establishing a high-quality representation foundation for boundary generation. Finally, omitting both collaborative training and boundary loss (w/o Co-train & Boundary Loss) degrades the model to the AlignRE baseline (77.31%), confirming the cumulative efficacy of our proposed mechanisms.

4.5 Qualitative Analysis and Visualization

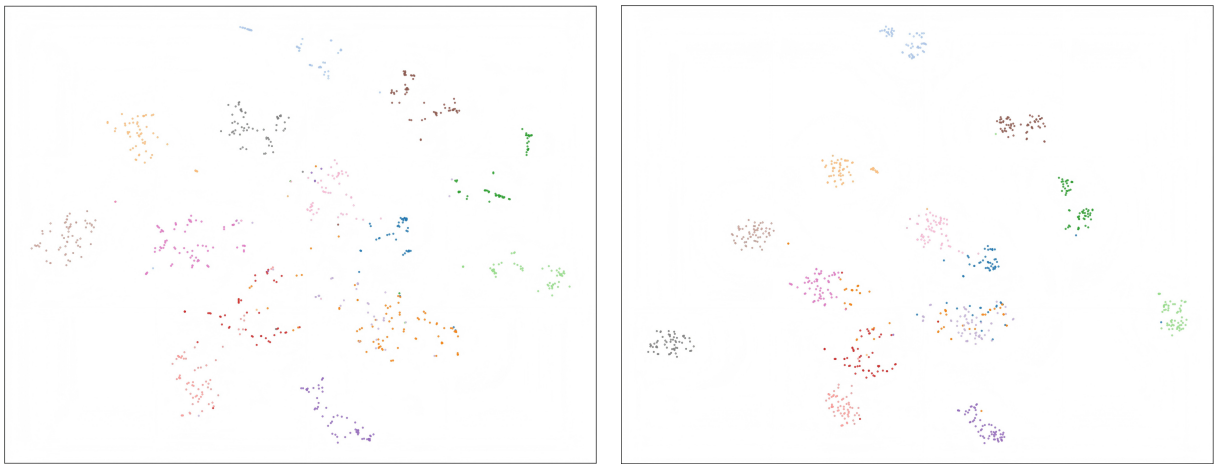
To intuitively illustrate the geometric advantages of our proposed BorderRE, we visualize the feature distributions of the Wiki-ZSL dataset (Unseen=15) using t-SNE. As shown in Figure 4, we compare the instance clustering effects between the baseline AlignRE and our fully trained BorderRE.

In Figure 4(a), we observe that while instances of the same relation tend to group together, the boundaries between different clusters remain extremely ambiguous and heavily overlapped (as highlighted by the red ellipses). This geometric entanglement makes the model highly susceptible to false positives when relying on a global static threshold, as instances from confusing relations easily cross into adjacent semantic territories.

Conversely, Figure 4(b) demonstrates the representation space optimized by BorderRE. Benefiting from the boundary-aware mutually exclusive contrastive loss, the geometric structure undergoes a profound transformation. The explicit constraints imposed by the dynamic boundaries act as a repulsive force, effectively pushing apart overlapping clusters to establish clear inter-class margins. Simultaneously, the attractive force of the central prototypes ensures that instances are tightly pulled within their respective polyhedral hulls. This visualization corroborates that BorderRE successfully

Setup	Unseen	Static Threshold			Dynamic Boundary Threshold		
		ID F1	OOD F1	H-score	ID F1	OOD F1	H-score
In-domain	5	85.72	72.67	78.57	**94.10**	**83.91**	**88.63**
	10	79.59	64.09	70.43	**84.04**	**73.32**	**77.97**
	15	76.55	60.68	66.99	**80.97**	**68.64**	**73.68**
Cross-domain	5	88.17	84.31	85.96	**93.12**	**90.09**	**91.53**
	10	75.71	79.72	77.35	**83.90**	**81.14**	**82.43**
	15	72.85	74.92	73.82	**76.50**	**78.19**	**77.25**

Table 3: Comparison of open-set in-domain and cross-domain rejection performance (%).



(a) AlignRE (b) BorderRE

Figure 4: t-SNE visualization of the instance feature distributions on Wiki-ZSL (Unseen=15). (a) Displays the clustering effect of AlignRE. (b) Shows the highly structured semantic space established by BorderRE.

Model Variant	Prec.	Rec.	F1
BorderRE (Full)	84.31	83.52	83.91
w/o Boundary Loss	83.52	82.22	82.86
w/o Co-train	80.75	80.15	80.45
w/o Co-train & Boundary Loss	77.63	77.00	77.31

Table 4: Ablation study of feature extraction and boundary components on FewRel (Unseen=15).

Boundary Vectors (K)	Precision	Recall	F1 Score
0 (Phase 1)	83.52	82.22	82.86
1	83.94	83.27	83.60
4	84.29	83.62	83.95
8	**84.62**	**83.87**	**84.24**
16	84.31	83.52	83.91
32	84.62	83.42	84.02

Table 5: Average classification performance under varying quantities of boundary vectors (K) on FewRel ($m = 15$).

shapes a highly discriminative and robust semantic space, perfectly aligning with our theoretical motivations for effective OOD rejection.

4.6 Analysis of Boundary Vector Quantity (K)

To investigate the impact of the boundary vector quantity (K) on rejection and classification performance, we analyze the F1 scores across various K values.

As depicted in Figure 5 and Table 5, setting $K = 0$ reverts the framework to Phase 1 (without explicit boundaries), yielding a baseline F1 score of 82.86%. When $K = 1$, the polyhedral boundary degrades into an isotropic hypersphere, which is prone to underfitting due to the anisotropic nature of the feature space. The performance peaks at $K = 8$ (84.24%), where the vectors optimally

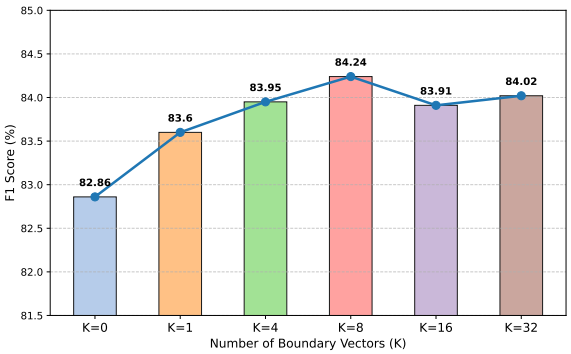


Figure 5: Impact of boundary vector quantity (K) on zero-shot classification performance (FewRel, $m = 15$).

delineate the geometric contours of the relation clusters. Theoretically, excessive boundary vectors ($K > 8$) induce parameter redundancy, rendering the model susceptible to overfitting local noise within specific clusters. Moreover, high-dimensional boundary intersections can compromise the clarity of the rejection zone. To ensure the BorderRE framework maintains consistent and robust generalization capabilities across massive relation categories without cumbersome hyperparameter tuning, we establish $K = 8$ as a globally shared configuration, achieving an optimal trade-off between geometric fitting precision and model stability.

5 Limitations

Currently, the quantity of dynamic boundaries (K) in our framework is fixed as a global hyperparameter. Future research could explore adaptive boundary generation strategies tailored to the intra-class variance of specific data manifolds. Furthermore, for relations with highly ambiguous semantic boundaries and exceedingly lengthy contextual dependencies, extracting robust prototype features remains challenging, occasionally constraining overall recall performance.

6 Conclusion

In this paper, we propose BorderRE, a novel framework designed to tackle the critical challenges of out-of-distribution noise and anisotropic instance distributions in open-domain zero-shot relation extraction. By generating parameterized polyhedral boundary vectors, we successfully substitute traditional rigid static thresholds with flexible geometric delineations. Accompanied by a two-stage collaborative training strategy and a mutually exclusive contrastive learning objective, BorderRE establishes highly structured semantic spaces characterized by strict intra-class compactness and clear inter-class separation. Extensive experimental evaluations demonstrate that our framework consistently sets new state-of-the-art benchmarks in both standard zero-shot classification and rigorous open-set rejection scenarios.

References

Alfred V Aho and Jeffrey D Ullman. 1972. *The theory of parsing, translation, and compiling*, volume 1. Prentice-Hall Englewood Cliffs, NJ.

American Psychological Association. 1983. *Publications Manual*. American Psychological Association, Washington, DC.

Rie Kubota Ando, Tong Zhang, and Peter Bartlett. 2005. A framework for learning predictive structures from multiple tasks and unlabeled data. *Journal of machine learning research*, 6(11).

Galen Andrew and Jianfeng Gao. 2007. Scalable training of l 1-regularized log-linear models. In *Proceedings of the 24th international conference on Machine learning*, pages 33–40.

Ashok K. Chandra, Dexter C. Kozen, and Larry J. Stockmeyer. 1981. Alternation. *Journal of the Association for Computing Machinery*, 28(1):114–133.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. 2020. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pages 1597–1607. PmLR.

Yew Ken Chia, Lidong Bing, Soujanya Poria, and Luo Si. 2022. Relationprompt: Leveraging prompts to generate synthetic data for zero-shot relation triplet extraction. In *Findings of the association for computational linguistics: ACL 2022*, pages 45–57.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pages 4171–4186.

Tianyu Gao, Xingcheng Yao, and Danqi Chen. 2021. Simcse: Simple contrastive learning of sentence embeddings. In *Proceedings of the 2021 conference on empirical methods in natural language processing*, pages 6894–6910.

Dan Gusfield. 1997. *Algorithms on Strings, Trees and Sequences*. Cambridge University Press, Cambridge, UK.

Xu Han, Hao Zhu, Pengfei Yu, Ziyun Wang, Yuan Yao, Zhiyuan Liu, and Maosong Sun. 2018. Fewrel: A large-scale supervised few-shot relation classification dataset with state-of-the-art evaluation. In *Proceedings of the 2018 conference on empirical methods in natural language processing*, pages 4803–4809.

Iris Hendrickx, Su Nam Kim, Zornitsa Kozareva, Preslav Nakov, Diarmuid O Séaghdha, Sebastian Padó, Marco Pennacchiotti, Lorenza Romano, and Stan Szpakowicz. 2010. Semeval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals. In *Proceedings of the 5th international workshop on semantic evaluation*, pages 33–38.

Dan Hendrycks and Kevin Gimpel. 2016. Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415*.

Amirhossein Layegh, Amir H Payberah, and Mihhail Matskin. 2024. Rea: Refine-estimate-answer prompting for zero-shot relation extraction. pages 301–316.

Omer Levy, Minjoon Seo, Eunsol Choi, and Luke Zettlemoyer. 2017. Zero-shot relation extraction via reading comprehension. In *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)*, pages 333–342.

Shilong Li, Ge Bai, Zhang Zhang, Ying Liu, Chenji Lu, Daichi Guo, Ruifang Liu, and Sun Yong. 2024a. Fusion makes perfection: An efficient multi-grained matching approach for zero-shot relation extraction. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 2: Short Papers)*, pages 79–85.

Zehan Li, Fu Zhang, and Jingwei Cheng. 2024b. Alignre: An encoding and semantic alignment approach for zero-shot relation extraction. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 2957–2966.

Zehan Li, Fu Zhang, Tianyue Peng, He Liu, and Jingwei Cheng. 2025. Generation-augmented retrieval: Rethinking the role of large language models in zero-shot relation extraction. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 17928–17941.

Sachin Pawar, Girish K Palshikar, and Pushpak Bhattacharyya. 2017. Relation extraction: A survey. *arXiv preprint arXiv:1712.05191*.

Mohammad Sadegh Rasooli and Joel Tetreault. 2015. Yara parser: A fast and accurate dependency parser. *arXiv preprint arXiv:1503.06733*.

Daniil Sorokin and Iryna Gurevych. 2017. Context-aware representations for knowledge base relation extraction. In *Proceedings of the 2017 conference on empirical methods in natural language processing*, pages 1784–1789.

Shanchan Wu and Yifan He. 2019. Enriching pre-trained language model with entity information for relation classification. In *Proceedings of the 28th ACM international conference on information and knowledge management*, pages 2361–2364.

Siye Wu, Jing Cao, Xiran Su, and Qiyuan Tian. 2025. Zero-shot knowledge extraction with hierarchical attention and an entity-relationship transformer. In *2025 5th International Conference on Sensors and Information Technology*, pages 356–360. IEEE.

Fu Zhang, He Liu, Zehan Li, and Jingwei Cheng. 2025. Ce-da: Custom embedding and dynamic aggregation for zero-shot relation extraction. In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 9814–9823.

Kai Zhang, Bernal Jiménez Gutiérrez, and Yu Su. 2023. Aligning instruction tasks unlocks large language models as zero-shot relation extractors. pages 794–812.

Yuhao Zhang, Victor Zhong, Danqi Chen, Gabor Angeli, and Christopher D Manning. 2017. Position-aware attention and supervised data improve slot filling. In *Proceedings of the 2017 conference on empirical methods in natural language processing*, pages 35–45.

Jun Zhao, WenYu Zhan, Wayne Xin Zhao, Qi Zhang, Tao Gui, Zhongyu Wei, Junzhe Wang, Minlong Peng, and Mingming Sun. 2023. Re-matching: A fine-grained semantic matching method for zero-shot relation extraction. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6680–6691.

A Case Study and Error Analysis

To intuitively demonstrate the geometric advantages of dynamic boundaries over static thresholds, we present a case study on the FewRel dataset in Table 6.

OOD Interception vs. ID Recall: As shown in Table 6, the baseline relies on a rigid static threshold ($T = 0.20$), resulting in false positives for confusing OOD samples (Good Case 1) and false negatives for implicit ID samples (Good Case 2). In contrast, BorderRE dynamically adjusts the threshold $\bar{T}_r$. It successfully tightens the boundary ($\bar{T}_r = 0.26$) to reject noise and relaxes it ($\bar{T}_r = 0.18$) to capture sparse representations, demonstrating superior geometric flexibility.

Error Analysis: Despite overall improvements, BorderRE struggles with extreme cases. In Bad Case 1, the lack of explicit context causes the instance to fall outside the stringent dynamic boundary. Conversely, in Bad Case 2, broad relation categories (e.g., *part of*) inherently generate overly relaxed boundaries, inadvertently encapsulating highly confusing OOD noise. Addressing these overlapping edge cases via finer-grained contextual modeling remains a focus for future work.

Category	Instance Context & Predictions	Scores & Thresholds	Analysis
Good Case 1: Successful OOD Interception	**Sentence:** It was the only bridge to span the [E2] Irrawaddy River [/E2] until recent times, when a spate of bridge construction has been carried out by the government including the new [E1] Irrawaddy Bridge [/E1]. **Ground Truth:** OOD (*crosses*) **Matched Prototype:** *located in or next to body of water* **Baseline (Static):** Accepted (Wrong) **BorderRE (Ours): Rejected as OOD (Correct)**	Max Sim: 0.2110 Static T: 0.2000 **Dynamic $\bar{T}_r$: 0.2624**	The static threshold is too loose, causing a false positive due to lexical overlap. The dynamic boundary tightly constraints the prototype, correctly rejecting the noise.
Good Case 2: Successful ID Recall	**Sentence:** His father, [E2] Leopold III [/E2], was the Margrave of Austria, while his mother, [E1] Agnes of Waiblingen [/E1], was the daughter of the Emperor Henry IV. **Ground Truth:** *spouse* **Matched Prototype:** *spouse* **Baseline (Static):** Rejected as OOD (Wrong) **BorderRE (Ours): Accepted as *spouse* (Correct)**	Max Sim: 0.1874 Static T: 0.2000 **Dynamic $\bar{T}_r$: 0.1860**	The relation representation is implicit, causing low similarity. The static threshold rigidly rejects it, while the dynamic boundary adapts to the sparse manifold, successfully recalling it.
Bad Case 1: False Negative (Error Reject)	**Sentence:** The [E1] Point Lowly Lighthouse [/E1] is a popular icon of the upper [E2] Spencer Gulf [/E2] region. **Ground Truth:** *located in or next to body of water* **Matched Prototype:** *located in or next to body of water* **Baseline (Static):** Accepted (Correct) **BorderRE (Ours): Rejected as OOD (Wrong)**	Max Sim: 0.2130 Static T: 0.2000 **Dynamic $\bar{T}_r$: 0.2626**	The instance lacks strong explicit relational indicators. The dynamic boundary for this cluster is overly stringent, causing the instance feature to fall outside the polyhedral hull.
Bad Case 2: False Positive (Error Accept)	**Sentence:** The name was revived in August 2009 by [E1] Midway Games [/E1]'s former European subsidiaries which rebranded themselves as [E2] Tradewest [/E2]. **Ground Truth:** OOD (*follows*) **Matched Prototype:** *part of* **Baseline (Static):** Rejected as OOD (Correct) **BorderRE (Ours): Accepted as *part of* (Wrong)**	Max Sim: 0.1924 Static T: 0.2000 **Dynamic $\bar{T}_r$: 0.1786**	The OOD semantic space heavily overlaps with the *part of* relation. The wide dynamic boundary for *part of* mistakenly encompasses this highly confusing noise instance.

Table 6: Case study and error analysis comparing the static threshold paradigm with our dynamic boundaries.