Structure-adaptive Adversarial Contrastive Learning for Multi-Domain Fake News Detection

Lingwei Wei[1] and **Dou Hu**[1,2*] and **Wei Zhou**[1] and **Philip S. Yu**[3] and **Songlin Hu**[1,2*]
[1] Institute of Information Engineering, Chinese Academy of Sciences
[2] University of Chinese Academy of Sciences
[3] University of Illinois at Chicago
{weilingwei, hudou, zhouwei, husonglin}@iie.ac.cn
psyu@uic.edu

Abstract

The rapid proliferation of fake news across multiple domains poses significant threats to society. Existing multi-domain detection models typically capture domain-shared semantic features to achieve generalized detection. However, they often fail to generalize well due to poor adaptability, which limits their ability to provide complementary features for detection, especially in data-constrained conditions. To address these challenges, we investigate the propagation-adaptive multi-domain fake news detection paradigm. We propose a novel framework, Structure-adaptive Adversarial Contrastive Learning (StruACL), to adaptively enable structure knowledge transfer between multiple domains. Specifically, we first contrast representations between content-only and propagation-rich data to preserve structural patterns in the shared representation space. Additionally, we design a propagation-guided adversarial training strategy to enhance the diversity of representations. Under the StruACL objective, we leverage a unified Transformer-based and graph-based model to jointly learn transferable semantic and structural features for detection across multiple domains. Experiments on seven fake news datasets demonstrate that StruACL-TGN achieves better multi-domain detection performance on general and data-constrained scenarios, showing the effectiveness and better generalization of StruACL.

1 Introduction

Nowadays, mainstream social platforms have facilitated the news dissemination in a faster and cheaper way. Nevertheless, the ease has also caused the wide spread of fake news, which has had detrimental effects on individuals and society (Loomba et al., 2021). Triggered by the negative impact of fake news, fake news detection has become a pressing challenge due to its widespread impact across diverse platforms and domains.

News content and its corresponding user engagements (i.e., tree-structured propagation) are two key data types in detecting fake news. Content-based detection methods (Ma et al., 2016; Ruchansky et al., 2017; Karimi and Tang, 2019) capture intrinsic semantic or linguistic features of claim tweets to detect fake news. Propagation-based detection methods (Ma et al., 2018; Kumar and Carley, 2019; Ma and Gao, 2020; Hu et al., 2021; Bian et al., 2020; Wei et al., 2021; Lin et al., 2021; Wei et al., 2022a) are designed to integrate structural features to complement textual content for detection. Nevertheless, in real-world scenarios, labeled data for fake news is often scarce, particularly in specific low-resource domains or emerging topics, which hinders detection performance. Recently, multi-domain fake news detection has been widely studied to leverage and integrate knowledge from multi-domain data to improve target-domain detection (Zhu et al., 2023; Liang et al., 2022; Wang et al., 2018; Zhang et al., 2021; Nan et al., 2021; Li et al., 2024), alleviating the data limitation challenge to some extent.

However, the representations learned by most existing multi-domain detection paradigms fail to generalize well due to poor adaptability to the propagation structure. Firstly, as shown in Figure 1, some multi-domain approaches primarily focus on learning domain-invariant or domain-shared semantic features (Wang et al., 2018; Nan et al., 2021) on content-only training data. However, semantic features inherently differ from structural patterns, rendering these content-based methods inadequate for generalizing to samples that involve propagation. Furthermore, directly extending domain-specific propagation-based methods struggles to effectively adapt to detection scenarios lacking propagation structures, resulting in suboptimal detection performance for content-only samples (Wei et al., 2024). Therefore, a critical challenge lies in learning more robust representations by enhancing

*Corresponding author.

Findings of the Association for Computational Linguistics: ACL 2025, pages 9739–9752

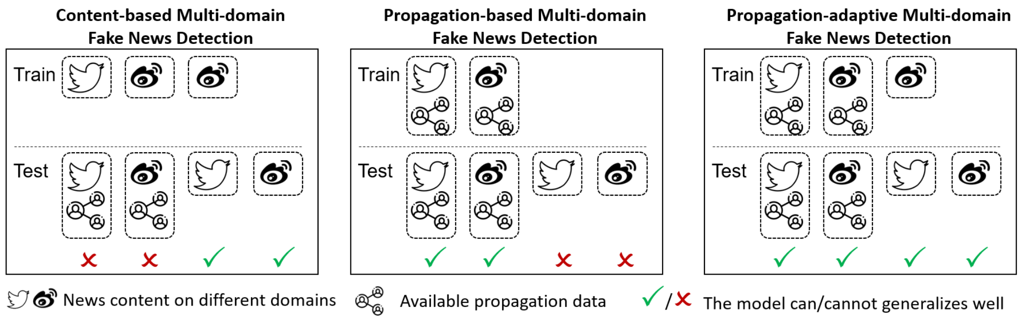


Figure 1: Difference between *propagation-adaptive multi-domain fake news detection* in this study and existing multi-domain fake news detection paradigms.

structure adaptability for multi-domain fake news detection.

In this paper, we study a novel propagation-adaptive multi-domain fake news detection paradigm, where the detection model is trained on both propagation-based data and content-only data. Our goal is to enhance generalization for both types of input.

To achieve this, we propose a new propagation structure-adaptive adversarial contrastive learning framework (StruACL) to adaptively learn generalized semantic and structural representations for multi-domain fake news detection. Specifically, we first design a new structure-aware contrastive learning (StruCL) objective to facilitate the adaptive transfer of structure knowledge during multi-domain training. With the guidance of structure label, StruCL leverages contrastive learning to differentiate representations between samples with and without propagation, effectively capturing and retaining structural knowledge in the shared representation space. By integrating this structural information, the learned representations become more informative, allowing the model to achieve enhanced performance in detecting fake news across both propagation-based and content-only domains.

Additionally, we design a propagation-guided adversarial training (PAT) strategy to enhance the diversity of representations under the data-constrained condition. PAT adaptively performs adversarial perturbations on original embeddings using the Fast Gradient Method (FGM) (Miyato et al., 2017) to generate worst-case samples for both content-only and propagation-based inputs. By jointly contrasting on both original and adversarial samples, the model can further effectively learn fine-grained semantic and structure knowledge via retaining the propagation-adaptive feature consistency. For the model architecture, we adopt a shared Transformer-based and graph-based network to jointly encode semantic and structural features from news content and available propagation across multiple domains, respectively. Under the proposed objective, our StruACL-TGN generalizes well across both content-only and propagation-structured domains.

We conduct experiments on seven fake news datasets with and without propagation. The experimental results demonstrate that our StruACL-TGN achieves superior performance in multi-domain fake news detection. Extensive experiments show the effectiveness of StruACL objective, particularly in data-limited application scenarios.

The main contributions are as follows: 1) We study a novel propagation-adaptive multi-domain fake news detection paradigm and develop a novel StruACL-TGN to learn generalized representations for detection on both domains with propagation data and content-only domains. 2) We design a new StruACL framework to learn more informative multi-domain representations. It contrasts semantic and structural representations to preserve and transfer structural knowledge, as well as introduces propagation-guided adversarial training to enhance the diversity of representations. 3) Experiments on seven fake news datasets demonstrate that StruACL-TGN achieves superior multi-domain detection performance. Extensive experiments further show that StruACL enhances the model's generalization capabilities in data-constraint applications.

2 Methodology

In this section, we first describe the problem definition of propagation-adaptive multi-domain fake news detection. Then, we propose a new StruACL-TGN to learn generalized representations on both domains with propagation data and content-only domains. The overall architecture is shown in

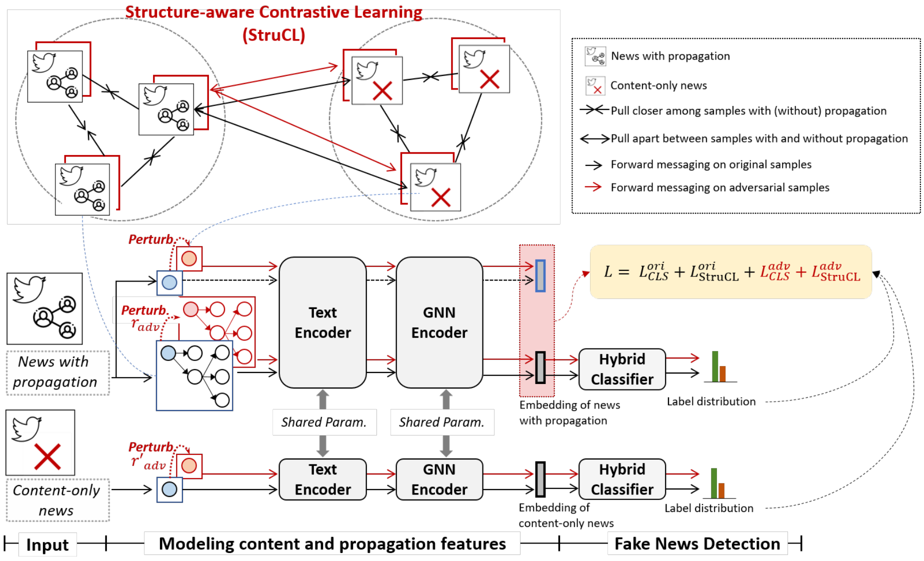


Figure 2: Overview of the proposed StruACL-TGN.

Figure 2. It adopts a shared Transformer-based and graph-based network to encode semantic and structural knowledge across multiple domains. For model training, we first propose a new StruCL to effectively utilize structure knowledge. Additionally, we design propagation-guide adversarial training (PAT) that generates worst-case samples to enhance the diversity of representations. Through applying PAT on original and adversarial samples, our StruACL can learn more informative multi-domain representations from domains with propagation data and domains with content-only data.

2.1 Problem Definition

Unlike existing multi-domain fake news detection tasks, propagation-adaptive multi-domain fake news detection aims to detect fake news across domains with heterogeneous data availability, where some domains have both propagation structures and content, while others only have content.

Formally, let $\mathcal{K}$ represent the set of all domains. Define $D^{(k)}$ as the dataset of each domain $k \in \mathcal{K}$. For each domain k that includes propagation data, $D^{(k)}$ is defined as $D^{(k)} = \{(x_i^{(k)}, G_i^{(k)}, y_i^{(k)})\}_{i=1}^{N_k}$, where $x_i^{(k)}$ is the content of the i-th news sample in domain k. $G_i^{(k)}$ is the propagation structure (e.g., a tree-like graph) associated with $x_i^{(k)}$. $y_i^{(k)} \in \{0, 1\}$ is the label indicating whether the news is fake or real. N_k is the number of samples in domain k. For each domain k' that excludes propagation data, $D^{(k)}$ is defined as $D^{(k')} = \{(x_i^{(k')}, y_i^{(k')})\}_{i=1}^{N_{k'}}$. where $x_i^{(k')}$ is the content of the i-th news sample in domain k'. $y_i^{(k')} \in \{0, 1\}$ is the fake news label. $N_{k'}$ is the number of samples in domain k'.

Propagation-adaptive multi-domain fake news detection aims to utilize both rich propagation structures and content-only samples to enhance detection performance across various domains. Fake news detection can be regarded as a binary classification task. Specifically, the objective is to learn a unified detection model $f(\cdot)$ that predicts the label $\hat{y}$ (e.g., fake or real) for a news item x (with or without propagation G across all domains: $\hat{y} = f(x, G)$, where $G = \emptyset$ for domains without propagation data.

2.2 Model Architecture

We adopt Transformer-graph network (TGN) as the model architecture. It consists of a shared Transformer-based semantic encoder, a graph-based structure encoder, and a hybrid classifier.

Transformer-based Semantic Encoder Considering multilingual settings, a pretrained multilingual BERT model (Conneau et al., 2020) on a monolingual corpus is utilized to facilitate language adaptation. Formally, given an input token sequence $x_{i1}, ..., x_{iT}$ where x_{ij} refers to j-th token in the i-th input sample, and T is the maximum sequence length, the model learns to generate the context representation of the input token sequences:

$$\mathbf{h}_i^s = \text{BERT}(\texttt{[CLS]}, x_{i1}, ..., x_{iN}, \texttt{[SEP]}), \quad (1)$$

where [CLS] and [SEP] are special tokens, typically placed at the beginning and end of each sequence, respectively. $\mathbf{h}_i^s$ indicates the hidden representation of the i-th input sample, computed by the representation of `[CLS]` token in the last layer of the encoder.

Graph-based Structure Encoder Based on the semantic representations, propagation-based models integrate structural features to enhance detection. Graph neural networks are widely applied to extract structural features through message-passing across nodes in the propagation graph. Given the input sample, which includes the textual content of the source news x and propagation trees G, existing models utilize various neural networks to extract high-level textual and structural features for detection. The formulation is defined as,

$$\mathbf{h}_i^g = \text{GNN}(x_i, G_i; \Theta), \tag{2}$$

where $\text{GNN}(\cdot)$ refers to graph-based encoders in propagation-based models (Bian et al., 2020; Wei et al., 2022b), and Θ refers to the corresponding trainable parameters. The input embedding of x_i and context c_i in G are initialized with the semantic embedding $\mathbf{h}_i^s$.

Hybrid Fake News Classifier We regard each domain as a task and employ a multi-task learning paradigm for multi-domain fake news detection. The paradigm using features of multiple tasks has shown promising performance across various natural language understanding tasks (Liu et al., 2019; Hu et al., 2025). To address the feature gap between different domains in distinguishing fake news, we design a hybrid fake news classifier to learn domain-specific and domain-shared discriminative features for detection.

Specifically, based on the final representation, domain-specific fake news classifiers are employed to predict the veracity label of each news content. For domain k, the initial prediction distribution is computed as,

$$\hat{\mathbf{y}'}^{(k)} = \mathbf{W}_c^{(k)}\mathbf{z} + \mathbf{b}_c^{(k)}, \tag{3}$$

where $\mathbf{W}_c^{(k)}$ and $\mathbf{b}_c^{(k)}$ are trainable parameters of domain k's classifier. Similarly, we apply a parameter-shared classifier to predict the veracity label for all domains, i.e.,

$$\hat{\mathbf{y}}^s = \mathbf{W}_c\mathbf{z} + \mathbf{b}_c. \tag{4}$$

Based on the above prediction, the final prediction for domain k is defined as,

$$\hat{\mathbf{y}}^{(k)} = Softmax(\frac{\hat{\mathbf{y}}^s}{2} + \frac{\hat{\mathbf{y}'}^{(k)}}{2}). \tag{5}$$

2.3 Optimization Process

2.3.1 Classification Objective

To achieve fake news detection, the model is trained by minimizing a joint loss function across all domains, considering both content and propagation data, i.e.,

$$\begin{aligned}\mathcal{L}_{\text{CLS}} = &\sum_{k \in \mathcal{K}_P} \frac{1}{N_k} \sum_{i=1}^{N_k} \ell(f(x_i^{(k)}, G_i^{(k)}), y_i^{(k)}) \\ &+ \sum_{k' \in \mathcal{K}_C} \frac{1}{N_{k'}} \sum_{i=1}^{N_{k'}} \ell(f(x_i^{(k')}, \emptyset), y_i^{(k')})\end{aligned}$$

where $\mathcal{K}_P \subseteq \mathcal{K}$ is the set of domains with propagation data, and $\mathcal{K}_C \subseteq \mathcal{K}$ is the set of domains without propagation data. $\ell(\cdot,\cdot)$ is the cross-entropy classification loss.

2.3.2 Structure-aware Contrastive Learning

Since the propagation-adaptive multi-domain fake news detection task involves domains with heterogeneous input types (i.e., some containing only news content while others include both content and propagation data), naively learning shared representations across domains would cause the model to converge a suboptimal compromise that fails to optimally capture either type's unique discriminative patterns for detection. To alleviate this issue, we design a new structure-aware contrastive learning (StruCL) objective to facilitate the adaptive transfer of structure knowledge during multi-domain training. It contrasts between representations with and without propagation structures. The objective of StruCL is defined as,

$$\begin{aligned}\mathcal{L}_{\text{StruCL}} &= \sum_{k \in \mathcal{K}_P} \mathcal{L}_{\text{StruCL}}^{(k)} \\ &= \sum_{k \in \mathcal{K}_P} -\frac{1}{N_k} \sum_{i=1}^{N_k} \left(\log \frac{e^{sim(\mathbf{z}_i^{(k)}, \mathbf{z}_{pos}^{(k)})/\tau}}{e^{sim(\mathbf{z}_i^{(k)}, \mathbf{z}_{pos}^{(k)})/\tau} + \sum_{j=1}^{N_k} \mathbb{1}_{g_i^{(k)} \neq g_j^{(k)}} e^{sim(\mathbf{z}_i^{(k)}, \mathbf{z}_j^{(k)})/\tau}} \right),\end{aligned} \tag{6}$$

where $g_i = 1$ refers to the hidden representation with structure information. The indicator function $\mathbb{1}_{g_i \neq g_j}$ equals 1 when the propagation structure label g_i and g_j are different, indicating a negative sample. $sim(\cdot,\cdot)$ is a pairwise similarity function, i.e., dot product. $\tau > 0$ is a scalar temperature parameter that controls the separation between the class with and without propagation structure.

StruCL explicitly encourage the separation between semantic and structural representations. It prevents the model from collapsing these fundamentally different feature spaces into an undifferentiated shared representation. Second, while main-

Benchmark		Dataset	Prop.	Resource	Lang.	Event	# Train	# Valid	# Test	# Total
CrossEval	CovidEval									
✓	×	Weibo21	×	Weibo	CN	Hybrid	5,751	1,918	1,923	9,592
×	✓	Covid19	×	Hybrid	EN	COVID19	6,420	2,140	2,140	10,700
✓	×	Twitter	✓	Twitter	EN	Hybrid	3,109	777	3,888	7,774
✓	✓	WeiboCovid19	✓	Weibo	CN	COVID19	163	40	208	411
×	✓	TwitterCovid19	✓	Twitter	EN	COVID19	159	39	202	400
×	✓	Arabic	✓	Twitter	AR	COVID19	136	36	184	356
×	✓	Cantonese	✓	Twitter	YUE	COVID19	577	143	724	1,444

Table 1: Statistics of 7 datasets for fake news detection. Prop. refers to whether the dataset contains propagation data. Lang. indicates language used in the dataset where CN, EN, AR, and YUE represent Chinese, English, Arabic, and Cantonese, respectively. Event summarizes the types of social events collected in the dataset. # Total refers to the total number of samples in each dataset.

taining this separation, StruCL ensures the coherence among representations of the same type. It reduces semantic or structural noise in representations, and forms compact representation of each type, allowing more effective knowledge transfer across domains.

2.3.3 Propagation-guided Adversarial Training

Considering the data-constrained condition, we further design a new propagation-guided adversarial training strategy (PAT) to enhance the diversity of representations. Different from previous methods (Hu et al., 2023) on language understanding applications, we apply an adversarial training strategy (e.g., FGM (Miyato et al., 2017)) to produce adversarial perturbations under a joint objective, i.e., structure-aware contrastive learning and multi-domain classification objectives. Specifically, the perturbations are put on the embedding layers of original samples, and then obtain adversarial samples. We then leverage the joint objective on these worst-case samples to maximize the consistency of transferable representations with or without propagation across multiple domains. Under the joint objective on both original and adversarial samples, our model can learn propagation-robust transferable features for multi-domain fake news detection.

The optimization objective for corresponding adversarial samples can be derived by following the calculation process for the original samples, denoted as, $\mathcal{L}_{\text{CLS}}^{\text{r-adv}} + \mathcal{L}_{\text{StruCL}}^{\text{r-adv}}$. Take the domain with propagation data as an example, the adversarial perturbation for content-only samples is defined as,

$$\min_{\theta} \mathbb{E}_{(x^{(k)},y^{(k)})\sim D^{(k)}} \max_{\|r_{\text{adv}}\|_q \leq \epsilon} (L_{\text{CLS}} + L_{\text{StruCL}}) \tag{7}$$

where $r_{\text{adv}} = -\epsilon \frac{g}{\|g\|_q}, g = \nabla \log p(y^{(k)}|x^{(k)}; \hat{\theta})$. PAT can generate structure-aware adversarial samples, which plays key role in fake news detection.

Totally, the overall loss of StruACL is defined as a sum of joint objective on both original and adversarial samples, i.e.,

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{CLS}} + \mathcal{L}_{\text{StruCL}} + \mathcal{L}_{\text{CLS}}^{\text{r-adv}} + \mathcal{L}_{\text{StruCL}}^{\text{r-adv}}. \tag{8}$$

3 Experiments

3.1 Experimental Setups

Datasets We conduct experiments on seven widely-used public datasets for fake news detection, where two content-based datasets including Weibo21, and Covid19, and five propagation-based datasets including Twitter, TwitterCovid19, WeiboCovid19, Arabic, and Cantonese. The statistics are shown in Table 1. **Weibo21** (Nan et al., 2021) collects Chinese tweets without propagation data on Weibo platform ranging from 2010-12-15 to 2021-03-31. Regarding to the breaking event COVID-19 pandemic, **Covid19** (Patwa et al., 2021) collects English textual tweets related to the topic of COVID-19 from public fact verification websites and social media (e.g., Facebook and Twitter[1]). **TwitterCovid19** (Kar et al., 2021; Lin et al., 2022), and **WeiboCovid19** (Lin et al., 2022) collect relevant textual tweets and propagation data on *Twitter* and Weibo, respectively. **Arabic** and **Cantonese**, originally collected by Alam et al. (2021) and Ke et al. (2020), contain textual claims in Arabic and Cantonese. Lin et al. (2023) further collect the propagation thread of each claim on both datasets.

Based on the above seven datasets, we build two major benchmarks to achieve different multi-domain fake news detection settings, each involving at least one content-based and propagation-based datasets to evaluate potential transferability of detection methods between semantic and propagation structure. We regard each dataset as

[1]Renamed *X* in 2023.

Methods	CrossEval $\|\mathcal{K}\| = 3$ Avg. Acc	Avg. F1	Δ_p F1	CovidEval $\|\mathcal{K}\| = 5$ Avg. Acc	Avg. F1	Δ_p F1
XLM-RoBERTa (Conneau et al., 2020)	87.33	87.15	0.00	77.97	74.91	0.00
GCNFN (Monti et al., 2019)	86.03	85.75	-1.59	79.39	76.49	+2.70
BiGCN (Bian et al., 2020)	86.48	86.28	-1.02	71.73	67.40	-9.71
EANN (Wang et al., 2018)	86.48	86.33	-0.93	79.30	77.36	+3.87
MDFEND (Nan et al., 2021)	88.49	88.22	+1.24	79.40	75.97	+1.90
M3FEND (Zhu et al., 2023)	87.87	87.62	+0.57	79.69	76.16	+2.50
FADED (Wei et al., 2025)	87.51	87.29	+0.15	79.56	77.14	+3.85
UCLR-TGN (Lin et al., 2024)	87.87	87.68	-0.40	78.82	76.76	+3.30
SAT-TGN (Wei et al., 2024)	87.60	87.32	+0.19	80.50	77.83	+4.34
StruACL-TGN (ours)	**89.44***	**89.25***	**+2.42**	**82.95***	**81.38***	**+10.12**

Table 2: Overall results of propagation-adaptive multi-domain fake news detection on *CrossEval* and *CovidEval* benchmarks. We highlight the best performance on each evaluation metric in bold. * refers to statistical significance over scores of the XLM-RoBERTa baseline under the t-test (p<0.05).

a domain and construct the two benchmarks using a multi-domain setting. Specifically, *CovidEval* includes five datasets related to the same event (i.e., COVID-19): Covid19, WeiboCovid19, TwitterCovid19, Arabic, and Cantonese. *CrossEval* includes three datasets from different social platforms and social events: Weibo21, Twitter, and WeiboCovid19. For each benchmark, we merge all domain-specific training/validation/test sets to form the train/valid/test set of each benchmark.

Evaluation Metrics Since fake news detection can be regarded as a binary classification, we adopt widely-used evaluation metrics for classification task, including accuracy (Acc), macro-averaged F1 score (F1). Additionally, to provide a more comprehensive evaluation of the model's overall performance across domains, we measure the average relative improvement (Δp) of F1 scores over the XLM-RoBERTa baseline on each domain.

Comparison Methods We compare with nine representative fake news detection methods. **XLM-RoBERTa** (Conneau et al., 2020) uses a PLM-based semantic encoder with a linear classifier for detection. **GCNFN** (Monti et al., 2019) uses the GCN to encode the propagation graph for detection. **BiGCN** (Bian et al., 2020) employs two GCNs to learn structural features from the propagation graph and dispersion graph. **EANN** (Wang et al., 2018) learns domain-invariant representations for detection. We re-implement by only considering the textual modality of news content across multiple domains. **MDFEND** Nan et al. (2021) uses a domain gate to aggregate multi-domain semantic representations. **M3FEND** (Zhu et al., 2023) introduces a domain adapter to extract domain-shared features from similar domains. **SAT** (Wei et al., 2024) learns structure-invariant features from samples with and without propagation for detection. We extend this framework for multi-domain detection with the same model architecture of our method, denoted as SAT-TGN. **UCLR** (Lin et al., 2024) uses unified contrastive transfer learning to enhance feature adaptation from well-resourced data to that of the low-resourced with only few-shot annotations. We implement this framework with the same model architecture of our method, denoted as UCLR-TGN. **FADED** (Wei et al., 2025) designs adversarial training from the perspective of document-level and entity-level. Here, we implement it via removing entity-level domain classifier due to the absence of explicit entity information in datasets.

Among them, XLM-RoBERTa, EANN, MDFEND, M3FEND, and FADED are content-based methods. BiGCN, GCNFN, SAT-TGN and UCLR-TGN are propagation-based methods. More details of the related works are listed in the Appendix.

Implementation Details All experiments are conducted on a single NVIDIA Tesla A100 80GB card. We use multilingual pretrained models (i.e., *xlm-roberta-base*) to extract textual features considering different languages across datasets, and fine-tune the semantic encoder during training. The dimension of hidden vectors is set to 64. The graph layers are set to 2. The Adamax optimizer is adopted for all methods with the learning rate initialized to 0.0001 and weight decay as 0. The temperature parameter is searched from $\{0.1, 1\}$. The perturbation radius is searched from $\{1, 5\}$ and the rate is set to 1. We run each model with 3 random seeds and report the average results of the test set for each method.

Methods	CrossEval								
	Weibo21$^\diamondsuit$		Twitter$^\spadesuit$		WeiboCovid19$^\spadesuit$		Avg.	Avg.	Δ_p
	Acc	F1	Acc	F1	Acc	F1	Acc	F1	F1
XLM-RoBERTa	88.46	88.33	87.42	87.40	86.11	85.71	87.33	87.15	0.00
MDFEND	90.80	90.75	85.78	85.77	88.89	88.15	88.49	88.22	+1.24
M3FEND	87.83	87.82	85.49	85.48	90.28	89.55	87.87	87.62	+0.57
FADED	89.91	89.90	85.80	85.80	86.81	86.16	87.51	87.29	+0.15
SAT-TGN	90.48	90.48	85.52	85.51	86.81	85.98	87.60	87.32	+0.19
StruACL-TGN (ours)	91.14*	91.09*	87.35	87.34	89.81	89.31	89.44*	89.25*	+2.42
StruACL-TGN* (ours)	**91.78***	**91.78***	**88.45***	**88.45***	**90.97***	**90.33***	**90.40***	**90.19***	**+3.50**

Table 3: Fine-grained experimental results of representative fake news detection methods on *CrossEval* benchmark, which involves fake news detection datasets across different social platforms. $^\diamondsuit$ refers to content-only datasets without propagation thread. $^\spadesuit$ indicates datasets with both textual content and propagation data. StruACL-TGN and StruACL-TGN* indicate the TGN model trained on the full benchmark and pair-wise benchmark under the proposed StruACL objective, respectively. Detailed results of pair-wise benchmark are listed in Table 5. * refers to statistical significance over scores of the XLM-RoBERTa baseline under the t-test (p<0.05).

Methods	CovidEval												
	Covid19$^\diamondsuit$		WeiboCovid19$^\spadesuit$		TwitterCovid19$^\spadesuit$		Arabic$^\spadesuit$		Cantonese$^\spadesuit$		Avg.	Avg.	Δp
	Acc	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1	F1
XLM-RoBERTa	96.92	96.90	85.42	84.80	63.37	60.07	77.17	71.52	66.99	61.23	77.97	74.91	0.00
GCNFN	93.93	93.87	84.72	84.12	69.80	63.30	81.52	78.60	66.99	62.56	79.39	76.49	+2.70
EANN	96.40	96.39	86.81	**86.40**	73.27	66.65	76.09	73.65	63.95	63.73	79.30	77.36	+3.87
MDFEND	96.40	96.39	84.72	84.03	67.82	59.17	75.54	70.67	**72.51**	**69.58**	79.4	75.97	+1.90
M3FEND	93.22	93.18	**86.81**	85.51	71.78	63.77	77.72	70.50	68.92	67.84	79.69	76.16	+2.50
FADED	**97.99**	**97.99**	82.64	82.26	76.24	72.07	73.37	70.96	67.54	62.43	79.56	77.14	+3.85
UCLR-TGN	97.38	97.37	78.47	78.39	69.31	63.93	80.98	77.42	67.96	66.67	78.82	76.76	+3.30
SAT-TGN	96.96	96.95	86.11	85.29	68.32	59.56	82.07	79.85	69.06	67.51	80.50	77.83	+4.34
StruACL-TGN (ours)	95.69	95.65	**86.81***	86.21*	**75.41***	**72.21***	**85.33***	**83.27***	71.50*	69.55*	**82.95***	**81.38***	**+10.12**

Table 4: Fine-grained experimental results of fake news detection on *CovidEval* benchmark, which involves fake news detection datasets related to the breaking event COVID-19. $^\diamondsuit$ refers to content-only datasets without propagation thread. $^\spadesuit$ indicates datasets with both textual content and propagation data. * refers to statistical significance over scores of the XLM-RoBERTa baseline under the t-test (p<0.05).

3.2 Overall Results

Comparison with Fake News Detection using Multi-domain Data The overall fake news detection results on CrossEval and CovidEval are listed in Table 2. We further show fine-grained results of representative methods with overall positive gains in Table 3 and Table 4. From results, our StruACL achieve the best overall detection performance based on all evaluation metrics on both benchmarks, showing the superiority of StruACL for multi-domain fake news detection. Specifically, compared with the XLM-RoBERTa, our StruACL achieves +2.42% and +10.12% Δp of F1 scores on CrossEval and CovidEval, respectively. First, directly apply existing propagation-based methods (e.g., BiGCN) to propagation-adaptive multi-domain fake news detection would get even negative performance due to poor generalization across domains. Additionally, EANN and FADED attempt to improve semantic feature adaptation; SAT solely learn structure-invariant features across multi-domain data; UCLR performs feature transfer in a straightforward hybrid way; MDFEND and M3FEND model domain-specific and domain-shared semantics with complex neural networks. These methods gains a certain but limited improvement since they ignore potential connections between semantics and structure. StruACL effectively performs knowledge transfer not only across multiple domains but also between semantics and structures, achieving the superior multi-domain detection performance.

Comparison with Fake News Detection using Pair-wise Domain Data We further evaluate on two specific domains, where the model is trained on the combined training sets of two specific domains and then tested on the combined test sets of the same two domains. This setting allows us to analyze how well the model generalizes across different domain combinations. Table 5 shows multi-domain detection results on pair-wise datasets where (a) and (b) indicate results trained on two heterogeneous datasets, one with propagation and another without; (c) and (d) indicates results trained on two propagation-based datasets and two content-based datasets with a homogeneous setting. From results, StruACL-TGN obtains superior average performance consistently, showing the

Methods	Weibo21$^{\diamond}$		WeiboCovid19$^{\blacklozenge}$		Avg.	
	Acc	F1	Acc	F1	Acc	F1
XLM-RoBERTa	86.17	86.12	87.50	86.85	86.83	86.48
EANN	80.56	80.09	90.43	90.43	85.49	85.24
MDFEND	88.92	88.84	89.58	89.00	89.25	88.92
SAT-TGN	87.50	87.08	85.28	85.18	86.39	86.13
StruACL-TGN	**89.18**	**89.15**	**90.97**	**90.33**	**90.08**	**89.74**
w/o StruACL	84.56	84.52	87.73	87.20	86.14	85.86

(a) Detection on Weibo21 and WeiboCovid19.

Methods	Weibo21$^{\diamond}$		Twitter$^{\blacklozenge}$		Avg.	
	Acc	F1	Acc	F1	Acc	F1
XLM-RoBERTa	90.74	90.71	86.27	86.26	88.50	88.49
EANN	89.08	89.05	86.29	86.29	87.69	87.67
MDFEND	88.66	88.62	83.69	83.69	86.18	86.16
SAT-TGN	88.82	88.82	82.79	82.70	85.81	85.76
StruACL-TGN	**91.78**	**91.78**	**88.45**	**88.45**	**90.12**	**90.12**
w/o StruACL	90.74	90.69	87.27	87.27	89.01	88.98

(b) Detection on Weibo21 and Twitter.

Methods	WeiboCovid19$^{\blacklozenge}$		TwitterCovid19$^{\blacklozenge}$		Avg.	
	Acc	F1	Acc	F1	Acc	F1
XLM-RoBERTa	86.81	86.16	76.24	73.94	81.52	80.05
EANN	87.50	86.56	74.26	73.51	80.88	80.04
MDFEND	82.64	80.57	71.78	64.60	77.21	72.58
SAT-TGN	88.19	87.54	76.24	73.40	82.22	80.47
StruACL-TGN	**88.89**	**88.38**	**78.22**	**76.40**	**83.55**	**82.39**
w/o StruACL	86.81	85.98	75.74	74.09	81.27	80.03

(c) Detection on TwitterCovid19 and WeiboCovid19.

Methods	Weibo21$^{\diamond}$		Covid19$^{\diamond}$		Avg.	
	Acc	F1	Acc	F1	Acc	F1
XLM-RoBERTa	90.74	90.73	97.52	97.52	94.13	94.12
EANN	89.24	89.23	97.34	97.33	93.29	93.28
MDFEND	88.46	88.46	97.29	97.28	92.87	92.87
SAT-TGN	90.95	90.95	97.57	97.56	94.26	94.26
StruACL-TGN	**92.04**	**92.01**	**98.18**	**98.17**	**95.11**	**95.09**
w/o StruACL	90.64	90.63	97.48	97.47	94.06	94.05

(d) Detection on Weibo21 and Covid19.

Table 5: Multi-domain detection results on pair-wise fake news datasets.

Methods	# Param	TwitterCovid19	WeiboCovid19	Arabic
XLM-RoBERTa	265MB×3	65.69	74.62	77.49
TextCNN	266MB×3	52.03	82.35	75.28
GCNFN	265MB×3	66.31	81.91	77.46
BiGCN	265MB×3	43.34	85.64	51.44
StruACL-TGN	265MB	**72.21**	**86.21**	**83.27**

Table 6: Comparison results of StruACL with single-task learning methods on three lower-resource datasets.

Methods	CrossEval			CovidEval		
	Avg. Acc	Avg. F1	Δp F1	Avg. Acc	Avg. F1	Δp F1
StruACL	**89.44**	**89.25**	**+2.42**	**82.95**	**81.38**	**+10.12**
w/o Adv	88.13	87.94	+0.91	81.94	80.13	+8.03
w/o StruCL	89.33	89.18	+2.33	81.93	80.41	+8.67
w/o StruACL	86.97	86.79	-0.40	81.49	79.84	+7.66
w/o Hybrid CLS	88.59	88.37	+1.42	81.75	80.11	+8.18
w/o $\text{Adv}_{\text{StruCL}}$	89.29	88.98	+2.11	79.86	77.45	+4.04
w/o Adv_{CLS}	87.15	86.95	-0.22	81.85	79.76	+7.23

Table 7: Overall results of the ablation studies on *CrossEval* and *CovidEval*. We report fine-grained results on each domain in Appendix.

effectiveness of StruACL on both heterogeneous and homogeneous settings. Additionally, by comparing Table 5 (a) and (b), which display results under two settings trained with Weibo21, we observe that our method shows a more significant improvement on WeiboCovid19,

Comparison with Fake News Detection on Lower-resource data Table 6 shows comparison results of single-task learning detection methods and our method on three lower-resource datasets (i.e., TwitterCovid19, WeiboCovid19, and Arabic). StruACL-TGN achieves better performance compared to detection methods trained on low resource training data with lighter network.

3.3 Ablation Study

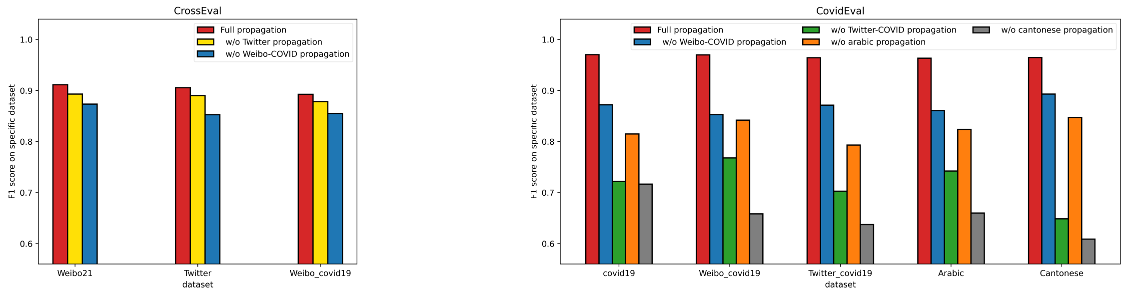
We further ablate the key components to evaluate the effectiveness of StruACL objective. **w/o Adv** refers to removing all adversarial perturbations during training. We also remove the perturbation based on structure-aware contrastive learning and cross-entropy classification objectives, respectively, denoted as **w/o $\textbf{Adv}_{\textbf{StruCL}}$**, and **w/o $\textbf{Adv}_{\textbf{CLS}}$**. **w/o StruCL** indicates removing the structure-aware contrastive learning, ignoring the transfer learning between structure and semantic. **w/o StruACL** is removing the full StruACL objective. **w/o Hybrid CLS** is removing the hybrid classifier and using a shared classifier for detection.

As shown in Table 7, the full model gains the best performance on both benchmarks consistently. The results demonstrate the effectiveness of each key component for detection. Additionally, for generating adversarial samples, eliminating the guidance of either structure-aware contrastive learning or task prediction gains (i.e., **w/o $\textbf{Adv}_{\textbf{StruCL}}$**, and **w/o $\textbf{Adv}_{\textbf{CLS}}$**) decreases performance to some extent, demonstrating the effectiveness of both objectives to generate structure-aware adversarial samples for boosting fake news detection.

3.4 Generalization Evaluation with Data-constrained Conditions

We evaluate the generalization under multi-domain data-constrained conditions. We vary the training set ratios to evaluate detection performance under limited data conditions. Specifically, for a predefined ratio (e.g., 20%), we randomly sampled subsets from the original training sets of all domains. All methods are tested using on the same sampled training subsets to ensure a fair comparison.

Figure 4 shows results of representative methods and our StruACL on CrossEval and CovidEval across various training set sizes. Our proposed Stru-



(a) Results on CrossEval.

(b) Results on CovidEval.

Figure 3: Results against removing domain-specific propagation in the training sets on CrossEval and CovidEval.

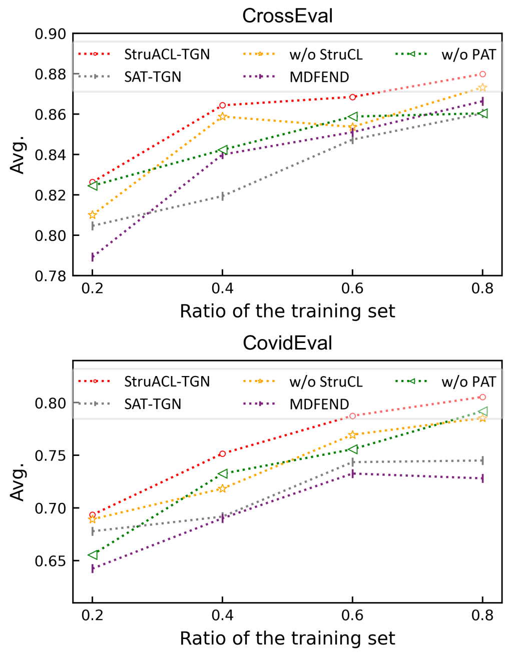


Figure 4: Results against different training set sizes. We report the average F1s of datasets on each benchmark.

ACL consistently achieves superior performance across all data-constrained settings, regardless of the training set ratio. This demonstrates the strong generalization capabilities of StruACL in scenarios with limited data. The performance improvement of StruACL is not only attributed to its ability to effectively learn semantic and structural features from multi-domain data but also to its capacity for transferring these learned semantic and structural representations across tasks. These advantages enable StruACL to efficiently utilize limited data and achieve generalized performance in data-constrained scenarios.

3.5 Effect of Training Propagation Structure

We analyze the effect of propagation structures in the training data during transferring between semantic and structure. We remove the propagation structure of the training set on the specific domain.

As shown in Figure 3, after removing propagation structures on Twitter and WeiboCovid19, the detection performance on all three datasets declined consistently. This indicates that propagation structures play a critical role in identifying fake news, as they provide complementary signals that enhance semantic analysis. Our StruACL can fully model interactions between semantic and structural features, thereby boosting multi-domain fake news detection. Interestingly, when removing the propagation data of WeiboCovid19, the detection performance on the Twitter declined more significantly compared to the performance drop observed for WeiboCovid19 itself. This may be because that StruACL leverages latent semantic associations related to Weibo, which facilitates the detection on WeiboCovid19 even in the absence of propagation features. In contrast, the performance gains on TwitterCovid19 are primarily driven by the transfer of propagation features from WeiboCovid19. This underscores the critical role of transferable propagation structure features in multi-domain detection.

4 Conclusion

This paper studies a propagation-adaptive multi-domain fake news detection paradigm. To achieve this, we develop a novel StruACL-TGN to learn generalized representations from propagation-based and content-only domains. StruACL contrasts semantic and structural representations to preserve and transfer structural knowledge, while introduces propagation-guided adversarial training to enhance the diversity of representations. Experiments on seven datasets show that StruACL-TGN achieves superior multi-domain detection on general and data-constraint settings, proving the effectiveness and generalization of StruACL.

Limitations

The current framework focuses on text-based semantic and propagation data. The study of multimodal inputs, such as images and videos will be left as future work to further enhance the robustness and versatility of fake news detection systems in increasingly complex and dynamic information ecosystems.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. U24A20335), the China Postdoctoral Science Foundation (No. 2024M753481), the Postdoctoral Fellowship Program of CPSF (No. GZC20232969), and Youth Innovation Promotion Association CAS. The authors thank the anonymous reviewers and the meta-reviewer for their helpful comments.

References

Firoj Alam, Shaden Shaar, Fahim Dalvi, Hassan Sajjad, Alex Nikolov, Hamdy Mubarak, Giovanni Da San Martino, Ahmed Abdelali, Nadir Durrani, Kareem Darwish, et al. 2021. Fighting the covid-19 infodemic: Modeling the perspective of journalists, fact-checkers, social media platforms, policy makers, and the society. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 611–649.

Tian Bian, Xi Xiao, Tingyang Xu, et al. 2020. Rumor detection on social media with bi-directional graph convolutional networks. In *AAAI*, pages 549–556.

Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. 2011. Information credibility on twitter. In *WWW*, pages 675–684.

Ming Chen, Zhewei Wei, Zengfeng Huang, Bolin Ding, and Yaliang Li. 2020. Simple and deep graph convolutional networks. In *ICML*, volume 119 of *Proceedings of Machine Learning Research*, pages 1725–1735. PMLR.

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. Unsupervised cross-lingual representation learning at scale. In *ACL*, pages 8440–8451. Association for Computational Linguistics.

Dou Hu, Yinan Bao, Lingwei Wei, Wei Zhou, and Songlin Hu. 2023. Supervised adversarial contrastive learning for emotion recognition in conversations. In *ACL*, pages 10835–10852.

Dou Hu, Lingwei Wei, Wei Zhou, and Songlin Hu. 2025. An information-theoretic multi-task representation learning framework for natural language understanding. In *AAAI*, pages 17276–17286. AAAI Press.

Dou Hu, Lingwei Wei, Wei Zhou, et al. 2021. A rumor detection approach based on multi-relational propagation tree. *Journal of Computer Research and Development*, 58(7):1395–1411.

S. Mo Jang, Tieming Geng, Jo-Yun Queenie Li, et al. 2018. A computational approach for examining the roots and spreading patterns of fake news: Evolution tree analysis. *Comput. Hum. Behav.*, 84:103–113.

Debanjana Kar, Mohit Bhardwaj, Suranjana Samanta, and Amar Prakash Azad. 2021. No rumours please! a multi-indic-lingual approach for covid fake-tweet detection. In *2021 grace hopper celebration India (GHCI)*, pages 1–5. IEEE.

Hamid Karimi and Jiliang Tang. 2019. Learning hierarchical discourse-level structure for fake news detection. In *NAACL-HLT*, pages 3432–3442.

Liang Ke, Xinyu Chen, Zhipeng Lu, Hanjian Su, and Haizhou Wang. 2020. A novel approach for cantonese rumor detection based on deep neural network. In *2020 IEEE International conference on systems, man, and cybernetics (SMC)*, pages 1610–1615. IEEE.

Ling Min Serena Khoo, Hai Leong Chieu, Zhong Qian, and Jing Jiang. 2020. Interpretable rumor detection in microblogs by attending to user interactions. In *AAAI*, pages 8783–8790.

Thomas N. Kipf and Max Welling. 2017. Semi-supervised classification with graph convolutional networks. In *ICLR (Poster)*.

Sumeet Kumar and Kathleen M. Carley. 2019. Tree lstms with convolution units to predict stance and rumor veracity in social media conversations. In *ACL*, pages 5047–5058.

Jiayang Li, Xuan Feng, Tianlong Gu, and Liang Chang. 2024. Dual-teacher de-biasing distillation framework for multi-domain fake news detection. In *ICDE*, pages 3627–3639. IEEE.

Chaoqi Liang, Yu Zhang, Xinyuan Li, Jinyu Zhang, and Yongqi Yu. 2022. Fudfend: fuzzy-domain for multi-domain fake news detection. In *CCF International Conference on Natural Language Processing and Chinese Computing*, pages 45–57. Springer.

Hongzhan Lin, Jing Ma, Liangliang Chen, Zhiwei Yang, Mingfei Cheng, and Chen Guang. 2022. Detect rumors in microblog posts for low-resource domains via adversarial contrastive learning. In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 2543–2556.

Hongzhan Lin, Jing Ma, Mingfei Cheng, et al. 2021. Rumor detection on twitter with claim-guided hierarchical graph attention networks. In *EMNLP*, pages 10035–10047.

Hongzhan Lin, Jing Ma, Ruichao Yang, Zhiwei Yang, and Mingfei Cheng. 2024. Towards low-resource rumor detection: Unified contrastive transfer with propagation structure. *Neurocomputing*, 578:127438.

Hongzhan Lin, Pengyao Yi, Jing Ma, Haiyun Jiang, Ziyang Luo, Shuming Shi, and Ruifang Liu. 2023. Zero-shot rumor detection with propagation structure via prompt learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 5213–5221.

Xiaodong Liu, Pengcheng He, Weizhu Chen, and Jianfeng Gao. 2019. Multi-task deep neural networks for natural language understanding. In *ACL*, pages 4487–4496.

Sahil Loomba, Alexandre de Figueiredo, Simon J Piatek, et al. 2021. Measuring the impact of covid-19 vaccine misinformation on vaccination intent in the uk and usa. *Nature human behaviour*, 5:337–348.

Jing Ma and Wei Gao. 2020. Debunking rumors on twitter with tree transformer. In *COLING*, pages 5455–5466.

Jing Ma, Wei Gao, Prasenjit Mitra, et al. 2016. Detecting rumors from microblogs with recurrent neural networks. In *IJCAI*, pages 3818–3824.

Jing Ma, Wei Gao, and Kam-Fai Wong. 2018. Rumor detection on twitter with tree-structured recursive neural networks. In *ACL*, pages 1980–1989.

Nikhil Mehta, Maria Leonor Pacheco, and Dan Goldwasser. 2022. Tackling fake news detection by continually improving social context representations using graph neural networks. In *ACL*, pages 1363–1380.

Takeru Miyato, Andrew M. Dai, and Ian J. Goodfellow. 2017. Adversarial training methods for semi-supervised text classification. In *ICLR (Poster)*.

Federico Monti, Fabrizio Frasca, Davide Eynard, et al. 2019. Fake news detection on social media using geometric deep learning. In *ICLR (Workshop)*.

Qiong Nan, Juan Cao, Yongchun Zhu, Yanyan Wang, and Jintao Li. 2021. MDFEND: multi-domain fake news detection. In *CIKM*, pages 3343–3347. ACM.

Parth Patwa, Shivam Sharma, Srinivas PYKL, Vineeth Guptha, Gitanjali Kumari, Md. Shad Akhtar, Asif Ekbal, Amitava Das, and Tanmoy Chakraborty. 2021. Fighting an infodemic: COVID-19 fake news dataset. In *CONSTRAINT@AAAI*, volume 1402 of *Communications in Computer and Information Science*, pages 21–29. Springer.

Kashyap Popat. 2017. Assessing the credibility of claims on the web. In *WWW (Companion Volume)*, pages 735–739.

Martin Potthast, Johannes Kiesel, Kevin Reinartz, et al. 2018. A stylometric inquiry into hyperpartisan and fake news. In *ACL*, pages 231–240.

Natali Ruchansky, Sungyong Seo, and Yan Liu. 2017. CSI: A hybrid deep model for fake news detection. In *CIKM*, pages 797–806.

Michael Sejr Schlichtkrull, Thomas N. Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov, and Max Welling. 2018. Modeling relational data with graph convolutional networks. In *ESWC*, volume 10843 of *Lecture Notes in Computer Science*, pages 593–607.

Kai Shu, Limeng Cui, Suhang Wang, et al. 2019. defend: Explainable fake news detection. In *KDD*, pages 395–405.

Yu Tong, Weihai Lu, Zhe Zhao, Song Lai, and Tong Shi. 2024. MMDFND: multi-modal multi-domain fake news detection. In *ACM Multimedia*, pages 1178–1186. ACM.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *NIPS*, pages 5998–6008.

Petar Velickovic, Guillem Cucurull, Arantxa Casanova, et al. 2018. Graph attention networks. In *ICLR (Poster)*.

Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science*, 359(6380):1146–1151.

Yaqing Wang, Fenglong Ma, Zhiwei Jin, Ye Yuan, Guangxu Xun, Kishlay Jha, Lu Su, and Jing Gao. 2018. Eann: Event adversarial neural networks for multi-modal fake news detection. In *Proceedings of the 24th acm sigkdd international conference on knowledge discovery & data mining*, pages 849–857.

Lingwei Wei, Dou Hu, Yantong Lai, et al. 2022a. A unified propagation forest-based framework for fake news detection. In *COLING*, pages 2769–2779.

Lingwei Wei, Dou Hu, Wei Zhou, and Songlin Hu. 2024. Transferring structure knowledge: A new task to fake news detection towards cold-start propagation. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8045–8049. IEEE.

Lingwei Wei, Dou Hu, Wei Zhou, et al. 2021. Towards propagation uncertainty: Edge-enhanced bayesian graph convolutional networks for rumor detection. In *ACL*, pages 3845–3854.

Lingwei Wei, Dou Hu, Wei Zhou, et al. 2022b. Uncertainty-aware propagation structure reconstruction for fake news detection. In *COLING*, pages 2759–2768.

Wenjie Wei, Yanyue Zhang, Jinyan Li, Panfei Liu, and Deyu Zhou. 2025. Cross-domain fake news detection based on dual-granularity adversarial training. In *COLING*, pages 9407–9417. Association for Computational Linguistics.

Ruichao Yang, Xiting Wang, Yiqiao Jin, Chaozhuo Li, Jianxun Lian, and Xing Xie. 2022. Reinforcement subgraph reasoning for fake news detection. In *KDD*, pages 2253–2262.

Huaiwen Zhang, Shengsheng Qian, Quan Fang, and Changsheng Xu. 2021. Multimodal disentangled domain adaption for social media event rumor detection. *IEEE Trans. Multim.*, 23:4441–4454.

Yongchun Zhu, Qiang Sheng, Juan Cao, Qiong Nan, Kai Shu, Minghui Wu, Jindong Wang, and Fuzhen Zhuang. 2023. Memory-guided multi-view multi-domain fake news detection. *IEEE Trans. Knowl. Data Eng.*, 35(7):7178–7191.

Overall of Appendix

In the appendix, we will provide related work on fake news detection, and fine-grained experimental results of ablation studies.

A Related Work

A.1 Fake News Detection

Fake news detection aims to automatically identify a news piece as real or fake.

Early works on **content-based fake news detection** rely on feature engineering to capture textual characteristics, e.g., topic features (Castillo et al., 2011), writing styles and consistency (Popat, 2017; Potthast et al., 2018). After the emergence of deep learning, some works (Ma et al., 2016; Ruchansky et al., 2017; Karimi and Tang, 2019) applied various neural networks to learn high-level linguistic features from the source news or combing its retweets.

Generally, users on social media share opinions, conjectures and evidence for checking fake news. Through their various interactive behaviors, a propagation tree describing the law of information transmission is formed and plays a significant role in fake news detection. Vosoughi et al. (2018); Jang et al. (2018) have empirically shown that compared with the truth, false news has deeper propagation structures, and reaches a wider audience. To leverage structure properties, **propagation-based fake news detection models** (Ma et al., 2016; Shu et al., 2019; Khoo et al., 2020) learn the sequential structure features in the propagation trees by RNN-based or attention-based modules. (Shu et al., 2019) jointly learned the sequential effect of comments and correlation between source news and the corresponding comments. To capture structural propagation patterns, (Ma et al., 2016) constructed a tree-structured neural network to model the propagation structure. (Khoo et al., 2020) adopted Transformer (Vaswani et al., 2017) to learn long-distance interactions. Recently, many researchers (Bian et al., 2020; Hu et al., 2021; Lin et al., 2021; Wei et al., 2021, 2022b; Mehta et al., 2022; Yang et al., 2022) regard the propagation tree as a graph, and employ various graph-based models (Kipf and Welling, 2017; Schlichtkrull et al., 2018; Chen et al., 2020; Velickovic et al., 2018) to capture topological structure features for detection. applied two graph convolutional networks (GCNs) (Kipf and Welling, 2017) to learn structural patterns from two distinct directed graphs. Hu et al. (2021); Lin et al. (2021) further learn multi-relational interactions in the propagation graph. Wei et al. (2024) study cold-start propagation and learn transferable features from samples with propagation for improving detection of content-only samples.

A.2 Fake News Detection across Multiple Domains

In real-world scenarios, fake news typically originates and propagate across various domains or platforms, due to real-time events, social trends, and other factors. Thus, multi-domain fake news detection has draw significant attention.

Most works aims to study domain-shared (Zhu et al., 2023; Liang et al., 2022) and domain-invariant semantic features (Wang et al., 2018; Zhang et al., 2021; Li et al., 2024; Lin et al., 2024; Wei et al., 2025) for detecting fake news across multiple domains. For example, Wang et al. (2018) learn event-invariant representations for multi-domain detection via considering the effect of event diversity. Nan et al. (2021) utilize domain gate to alleviate the domain shift issue for aggregation of multi-domain representations. Zhu et al. (2023) further develop a domain adapter to extract domain-shared features from similar domains. Liang et al. (2022) design a fuzzy domain label to capture multi-domain knowledge. Li et al. (2024) study the unbalanced multi-domain data issue and leverage two teacher models to mitigate the domain bias via knowledge distillation. Tong et al. (2024) design a progressive hierarchical extraction network to achieve domain-adaptive domain-related knowledge extraction. Lin et al. (2024) employ contrastive transfer learning to enhance feature adaptation from well-resourced data to that of the low-resourced with only few-shot annotations. Wei et al. (2025) perform semantic features adaptation via adversarial training from the perspective of document-level and entity-level.

However, most of the above methods focus on the news content across different domains. They typically fail to utilize propagation structure across multiple domains that provides positive complementary efforts on fake news detection. To fill the gap, we develop a novel StruACL framework to adaptively learn generalized semantic and structural representations for multi-domain fake news detection. StruACL can boost feature adaptation of structure knowledge across both content-only and propagation-structured domains.

Methods	Weibo21$^{\diamondsuit}$		Twitter$^{\blacklozenge}$		WeiboCovid19$^{\blacklozenge}$		Avg.	Avg.	Δp
	Acc	F1	Acc	F1	Acc	F1	Acc	F1	F1
StruACL	91.14	91.09	87.35	87.34	89.81	89.31	**89.44**	**89.25**	**+2.42**
w/o Adv	89.93	89.91	86.53	86.50	87.96	87.41	88.13	87.94	+0.91
w/o StruCL	91.21	91.20	87.20	87.20	89.58	89.13	89.33	89.18	+2.33
w/o StruACL	88.30	88.25	85.59	85.57	87.04	86.55	86.97	86.79	-0.40
w/o Hybrid CLS	89.29	89.22	87.60	87.59	88.89	88.31	88.59	88.37	+1.42
w/o $\text{Adv}_{\text{StruCL}}$	91.23	91.19	87.02	87.00	89.35	88.77	89.29	88.98	+2.11
w/o Adv_{CLS}	88.96	88.91	84.29	84.28	88.19	87.66	87.15	86.95	-0.22

Table 8: Ablation results on *CrossEval* benchmark.

Methods	Covid19$^{\diamondsuit}$		WeiboCovid19$^{\blacklozenge}$		TwitterCovid19$^{\blacklozenge}$		Arabic$^{\blacklozenge}$		Cantonese$^{\blacklozenge}$		Avg.	Avg.	Δp
	Acc	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1	F1
StruACL	95.69	95.65	86.81	86.21	75.41	72.21	85.33	83.27	71.50	69.55	**82.95**	**81.38**	**+10.12**
w/o Adv	96.21	96.21	88.89	88.31	76.24	72.32	82.61	78.80	65.76	65.00	81.94	80.13	+8.03
w/o StruCL	95.61	95.58	86.11	85.84	73.27	70.68	84.24	81.29	70.44	68.67	81.93	80.41	+8.67
w/o StruACL	95.79	95.78	86.81	86.46	73.27	68.76	83.33	81.06	68.23	67.16	81.49	79.84	+7.66
w/o Hybrid CLS	95.00	94.96	87.50	86.76	73.76	69.57	83.15	80.75	69.34	68.50	81.75	80.11	+8.18
w/o $\text{Adv}_{\text{StruCL}}$	97.88	97.88	82.18	81.91	68.32	62.60	80.62	77.22	70.30	67.62	79.86	77.45	+4.04
w/o Adv_{CLS}	96.73	96.72	88.89	88.23	72.28	66.48	83.70	81.12	67.68	66.27	81.85	79.76	+7.23

Table 9: Ablation results on *CovidEval* benchmark.

B Fine-grained Ablation Results

Table 8 and Table 9 report fine-grained results of ablation studies on CrossEval and CovidEval, respectively.