Event-Radar: Event-driven Multi-View Learning for Multimodal Fake News Detection

Zihan Ma[†1,5] **Minnan Luo**[*1,2,3] **Hao Guo**[4] **Zhi Zeng**[1,2] **Yiran Hao**[1,3] **Xiang Zhao**[4]

[1]School of Computer Science and Technology, Xi'an Jiaotong University, Xi'an, 710049, China

[2]Ministry of Education Key Laboratory of Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an, 710049, China

[3]Shaanxi Province Key Laboratory of Big Data Knowledge Engineering, Xi'an Jiaotong University, Xi'an, 710049, China

[4]Laboratory for Big Data and Decision, Nation University of Defence Technology, Changsha, China

[5]SGIT AI Lab, State Grid Corporation of China

`mazihan880@stu.xjtu.edu.cn; minnluo@xjtu.edu.cn`

Abstract

The swift detection of multimedia fake news has emerged as a crucial task in combating malicious propaganda and safeguarding the security of the online environment. While existing methods have achieved commendable results in modeling entity-level inconsistency, addressing event-level inconsistency following the inherent subject-predicate logic of news and robustly learning news representations from poor-quality news samples remain two challenges. In this paper, we propose an **Event**-d**R**iven f**A**ke news **D**etection frAmewo**R**k (Event-Radar) based on multi-view learning, which integrates visual manipulation, textual emotion and multimodal inconsistency at event-level for fake news detection. Specifically, leveraging the capability of graph structures to capture interactions between events and parameters, Event-Radar captures event-level multimodal inconsistency by constructing an event graph that includes multimodal entity subject-predicate logic. Additionally, to mitigate the interference of poor-quality news, Event-Radar introduces a multi-view fusion mechanism, learning comprehensive and robust representations by computing the credibility of each view as a clue, thereby detecting fake news. Extensive experiments demonstrate that Event-Radar achieves outstanding performance on three large-scale fake news detection benchmarks. Our studies also confirm that Event-Radar exhibits strong robustness, providing a paradigm for detecting fake news from noisy news samples.

1 INTRODUCTION

Against the backdrop of the rapid expansion of social media, online platforms like Twitter have emerged as the primary channels for people to obtain information. Unfortunately, they have also become breeding grounds for the proliferation and dissemination of fake news. Fake news publishers exploit these platforms by spreading erroneous information, fueling societal divisions, fostering conspiracy theories, and posing threats to societal safety (Zhao et al., 2015; Lao et al., 2021). The "viral spread of information" during the 2016 US presidential elections (Fisher et al., 2016) and the COVID-19 pandemic (Naeem and Bhatti, 2020) vividly depict how fake news disrupts societal order. Typically, visual media like photos often trigger strong emotional reactions in readers, leading to higher engagement on social media, thereby serving as an ideal vehicle for fake news (Qi et al., 2019).

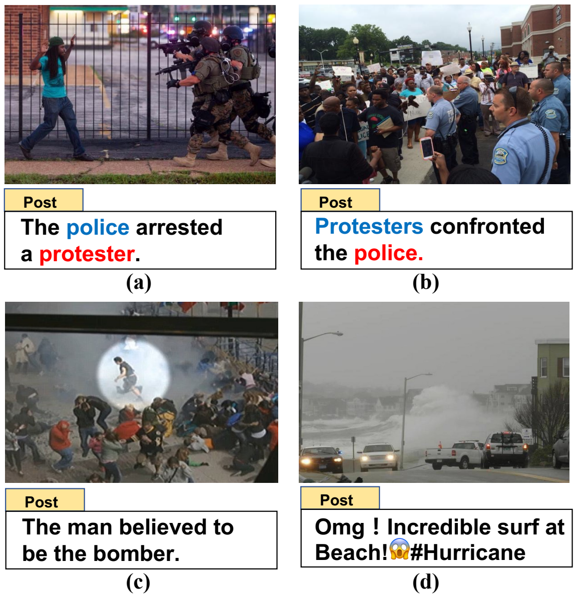

Figure 1: (a) and (b) demonstrate the limitations of element-level consistency detection. (c) and (d) demonstrate noise in the news on social media.

Some researchers argue that the inconsistency between posts and images is a key feature in judging the authenticity of news, and they have proposed methods to model this text-visual inconsis-

[*]Corresponding Author

[†]Part of the work was assisted by SGIT AI Lab, State Grid Corporation of China during the author's internship.

Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 5809–5821
August 11-16, 2024

tency (Chen et al., 2022; Zhou et al., 2023). In addition, news on social media is diverse, and inconsistency is not an absolute criterion for determining news authenticity (Ying et al., 2023). Detecting manipulated images (Cao et al., 2020) or provocative emotion in post (Zhang et al., 2021b) is also an effective view for detecting fake news. As a result, integrating as many available multimodal clues as possible becomes crucial for fake news detection, called multi-view learning (Ying et al., 2023; Wu et al., 2021; Zeng et al., 2023; Wan et al., 2024). Although methods based on inconsistency or multi-view learning have achieved many promising results, the lack of inconsistency checks at the event level still affects the accuracy of detection methods. Meanwhile, most existing methods overlook the impact of inherent noise in multimodal news data. Therefore, we summarize two main shortcomings of current methods:

- **Event-level multimodal inconsistency**: In the context of news being regarded as a collection of events, 89% of news images encompass events characterized by subjects, objects, and predicates (Li et al., 2022). As illustrated in Fig.1 (a) and (b), both images contain entities such as 'police' and 'protesters.' However, due to differing subject-verb relationships, they convey significantly different meanings. Although existing methods have achieved excellent results in modeling inconsistency at element-level, aligning subjects and objects in images, merely achieving alignment at the element level may not effectively measure the relationship between news posts and images. This limitation leads the model to learn features biased towards check the authenticity of the news.
- **Noise of multimodal samples**: With the rise of we-media, the casual composition of news has led to the proliferation of poor-quality news on social media. Some images undergo compression processing, making it almost impossible to recognize entities within them, while some news posts contain very few words. Additionally, certain multi-view methods incorporate pattern features to detect image manipulation in fake news. Some news publishers use image editing techniques to highlight key elements in news images, as shown in Figure 1(c), leading to biases in models relying on image manipulation for detection. On social media, certain platforms use symbols like "#" in the post for tagging or mentioning, as illustrated in Fig. 1(d), leading to misjudgments by models analyzing post content and emotion. These noise of multimodal news characterized by poor-quality and capable of causing cognitive bias in models, usually significantly impacts the generalization performance.

To tackle these challenges, we propose the **Event**-d**R**iven f**A**ke news **D**etection fr**A**mewo**R**k (Event-Radar) based on multi-view learning. The framework leverages statistical distributions to learn more robust news representations at the event-level. Specifically, we model individual news as a multimodal graph and extract subgraphs representing events present in both images and posts. Additionally, we utilize textual emotion and image pattern features as additional clues for multi-view learning, leveraging features from different views to enhance classification accuracy. However, this assumes that the quality or importance of these views is relatively stable across all samples. When feature from certain view is severely compromised, it can significantly impact the accuracy of classification (Wu et al., 2022). To address this issue, the beta distribution is utilized to estimate the credibility for each view, biasing the model towards trusting views with higher credibility. The contributions of this paper are three-folded:

- We propose a novel event-driven fake news detection framework that elucidates the inherent subject-verb logic in multimodal news.
- We attempt to address the issue of varying sample quality in news by estimating the credibility of each viewpoint using a Beta distribution. We determine the weight of feature fusion based on the magnitude of credibility, aiming to integrate features more heavily from views with higher credibility.
- Event-Radar not only outperforms all existing multi-view multimodal fake news detection frameworks but also provides a robust approach to resist the disturbance of noise samples, addressing the issue of model bias introduced by the complex data distribution in the real world.

2 RELATED WORK

2.1 Multimodal Fake News Detection

Traditional multimodal fake news detection extensively leverages latent information from both images and posts to obtain comprehensive multimodal news representations (Liu et al., 2023; Chen et al., 2023; Khattar et al., 2019). Approaches like Safe (Zhou et al., 2020) and BTIC (Zhang et al., 2021a) enhance multimodal representations by setting appropriate loss functions. Detecting fake news through modal inconsistency, measuring

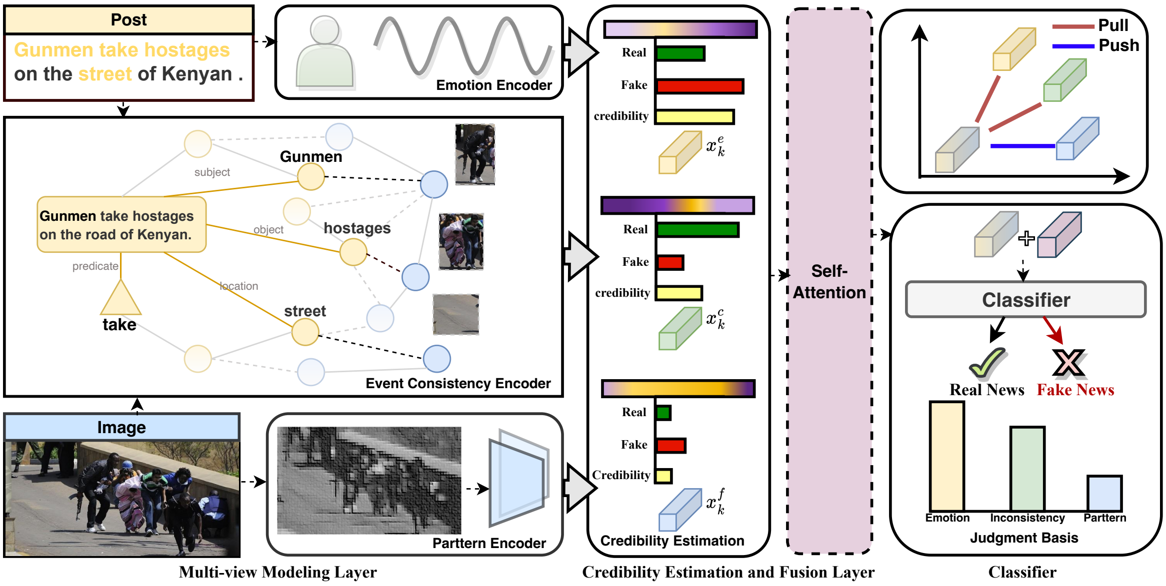

Figure 2: Overview of proposed Event-Radar.

the authenticity of news through entity alignment, is a prevailing method in current fake news research. CAFE (Chen et al., 2022) calculates the ambiguity between different modal elements using KL divergence, while FND-CLIP (Zhou et al., 2023) achieves excellent results through element-level semantic detection. However, entity-level inconsistency checks are relatively coarse and do not model the subject-predicate relationships between entities at the event level. Additionally, the challenge arises when fake news publishers employ image editing or deepfake techniques (Chen et al., 2022), rendering these methods ineffective on certain samples, requiring the integration of pattern features or provocative text emotion for multimodal learning.

2.2 Multi-view Learning

Leveraging multiple views to learn from data has proven to be effective in various tasks (Dang et al., 2023, 2024; Wu et al., 2024). Multi-view models based on CCA (Wang et al., 2016) are widely used for multi-view learning. MOE (Shazeer et al., 2017) based on the principle of divide and conquer introduces the mixed expert method by partitioning input samples into multiple subtasks and training an expert for each subtask. TMC (Han et al., 2021) uses the Dirichlet distribution to check class probabilities, parameterizing evidence from different views. In fake news detection, models like MVNN (Cao et al., 2020) and MCAN (Wu et al., 2021) incorporate pattern features as clues for multi-view learning, and BMR (Ying et al., 2023) introduces an enhanced multi-gate mixture expert network, demonstrating the advantages of multi-view learning in fake news detection.

However, multi-view methods suffer significant performance degradation when features from individual views are lost or contain a substantial amount of noise, leading to erroneous judgments. Hence, we propose a methodology that harnesses credibility to integrate multi-view features.

3 METHODOLOGY

Fig. 2 illustrates an overview of the Event-Radar framework, comprising a multi-view modeling layer and a credibility estimation layer. Specifically, we initially encode the events, emotions, and pattern information of the news to comprehensively assess the news representation from various views. To model multimodal news events, we introduce an event inconsistency measurement module based on event subgraphs. To obtain credible representations from each view, we employ Beta distribution to compute the credibility of each view and fuse modality features guided by this credibility. Subsequently, for information interaction, we employ a self-attention mechanism to fuse modal information from various views. Finally, we employ a classifier to perform fake news detection.

3.1 Event Inconsistency Encoder

Considering the intricate subject-predicate logic among entities in multimodal news events, our model is designed to capture the complex relationships within and between modalities. Motivated by multimodal learning (Li et al., 2022), we establish a cross-modal graph $\mathcal{G}_k$ as a representation for multimodal news. For the k-th post-image pair (P_k, I_k) in the dataset, we initially tokenize P_k into m tokens and extract n objects from the image I_k using Faster R-CNN (Chen et al., 2019). To obtain features in the same d-dim embedding space, we employ frozen CLIP (Radford et al., 2021) model to extract multimodal features T_k and V_k, *i.e*,

$$
\begin{aligned}
T_k &= CLIP(P_k) \\
&= [t_k^{CLS}, t_k^1, t_k^2, \cdots, t_k^m] \in \mathbb{R}^{(m+1)\times d}, \\
V_k &= CLIP(I_k) \\
&= [v_k^{CLS}, v_k^1, v_k^2, \cdots, v_k^n] \in \mathbb{R}^{(n+1)\times d},
\end{aligned}
$$

where t_k^{CLS} denotes the encoded representation of the [CLS] token, while v_k^{CLS} represents the encoding of the entire image.

We construct a multimodal graph $\mathcal{G}_k$ for (P_k, I_k) to leverage the initial representation of the multimodal graph using relationships between post tokens and image objects. Specifically, we consider the embeddings of the post tokens in T_k and the embeddings of objects in V_k as nodes in the graph $\mathcal{G}_k$. The node matrix is the concatenation of T_k and V_k, denoted as: $H_k = [T_k, V_k] \in \mathbb{R}^{(m+n+2)\times d}$. The edge weight coefficients are initialized by computing the similarity between nodes and then scaled to range between $[0, 1]$, *i.e*, $A_k^{i,j} = h_k^i \cdot h_k^j / 2\|h_k^i\| \|h_k^j\|$, where h_k^i and h_k^j are node features, $h_k^i, h_k^j \in H_k$. To extract event-specific subgraphs $\mathcal{G}_k^P$ and $\mathcal{G}_k^I$ corresponding to posts and images within the news graph $\mathcal{G}_k$, we employing the approach depicted in Fig.3. Specifically, we utilize both the Stanford NLP (Manning et al., 2014) and TextSmart NLP tools (Zhang et al.; Liu et al.) to perform NER on English and Chinese posts, obtaining the identification of subject t_k^s, object entity t_k^o, and location adverbial t_k^{loc} from T_k. These identified entities are then linked to t_k^{CLS} to form the post-event subgraph $\mathcal{G}_k^P$, with nodes represented as $H_k^P = [t_k^s, t_k^o, t_k^{loc}]$. To establish the mapping between textual entities and their corresponding image nodes within the $\mathcal{G}_k^I$ subgraph, we select the image object with the highest similarity as the representation of the textual entity nodes, denoted as $H_k^I = [v_k^s, v_k^o, v_k^{loc}]$.

In order to emphasize the events within the news, we apply weighting to the predicate entity and t_k^{CLS}, resulting in an enhanced representation of the post denoted as $t_k'^{CLS} = (t_k^{CLS} + t_k^p)/2$, where t_k^p represents the predicate entity within the post.

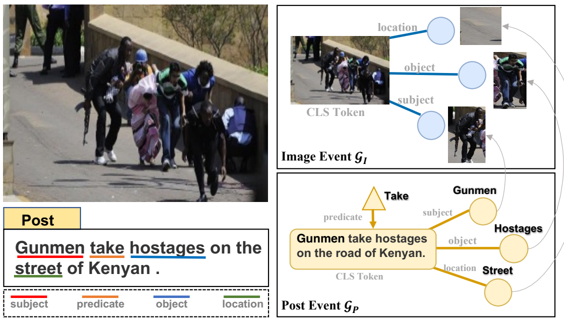

Figure 3: The process of constructing event subgraphs.

For any missing entities, we substitute them with zero-vectors of matching dimensions to denote the absence of critical entities within the news content. The initial weights of edges A_p^k and A_I^k within the subgraphs are set to 1. Subsequently, we employ an L-layer Graph Convolutional Network (GCN) (Kipf and Welling, 2016) for learning the multimodal graph. The features at the l-th layer are computed by:

$$H_k^l = \textbf{ReLU}(\tilde{A}_k^l H_k^{l-1} W^l), \tag{1}$$

where $\tilde{A}_k^l = D_k^{-\frac{1}{2}} A_k D_k^{-\frac{1}{2}}$ and D_k represent the degree matrix of the initial weights A_k. W^l denotes the learnable parameters. Subsequently, separate convolutions are applied to the event subgraphs of posts and images to obtain the node representations for the i-th layer of the event graph, *i.e*,

$$H_{P_k}^l = \textbf{ReLU}(\tilde{A}_{P_k}^l H_{P_k}^{l-1} W_p^l), \tag{2}$$

$$H_{I_k}^l = \textbf{ReLU}(\tilde{A}_{I_k}^l H_{I_k}^{l-1} W_i^l), \tag{3}$$

where $\tilde{A}_{P_k}^l$ and $\tilde{A}_{I_k}^l$ represent the normalized adjacency matrices for the event graphs of posts and images, respectively. W_p^l and W_i^l denote the learnable parameters for the event graphs associated with posts and images. Inspired by(Sheng et al., 2021), the edge weights among all entities are dynamically adjusted based on the latest representations, which aims to better reflect the relevance of entities within the events, *i.e*,

$$
\begin{aligned}
\Delta A_k^l &= \sigma\left(H_k^l W_a^l H_k^{l\,T}\right), \\
A_k^l &= \alpha A_k^{l-1} + (1-\alpha)\Delta A_k^l,
\end{aligned}
$$

where W_a^l denotes the learnable parameters, σ represents the sigmoid function, and α stands for the hyperparameter determining the update rate. Finally, we utilize a comparative function (Shen et al.,

2018) to perform graph-to-graph comparison between post events and image events, capturing the inconsistency of the event across modalities, denoted as x_k^c, *i.e*,

$$x_k^c = W_c[H_{P_k}^L, H_{I_k}^L, H_{P_K}^L - H_{I_k}^L, H_{P_k}^L \odot H_{I_k}^L] \in \mathbb{R}^d,$$

where W_c represents the learnable parameters, $\odot$ denotes the Hadamard product.

3.2 Emotion and Pattern Encoder

Emotion encoder. We follow (Zhang et al., 2021b) and extract the emotion feature of the news publisher from the original post P_k, *i.e*,

$$x_k^e = f_{emo}(P_k). \quad (4)$$

Pattern encoder. The general distribution of the image and the minute traces left by manipulating or compression are defined as image patterns.We employ Multi-Head Self-Attention (MHSA) network to encode image features transformed by Discrete Cosine Transform (Liu and Li, 2003), *i.e*,

$$x_k^f = \frac{1}{l_I} \sum_{j=0}^{l_I} \mathbf{MHSA}(DCT(f_j)) \in \mathbb{R}^d, \quad (5)$$

where $\mathbf{MHSA}(\cdot)$(Vaswani et al., 2017) represents the Multi-Head Self-Attention network;l_I represents the number of image patches; $DCT(\cdot)$ signifies the Discrete Cosine Transform (Liu and Li, 2003); f_j stands for the j-th patch of the image I_k.

3.3 Single view credibility calculation

The confidence levels of the three mentioned features vary for different multimodal fake news detection scenarios. Intuitively, calculating credibility and integrating them can enhance the detection performance. TMC (Han et al., 2021) has demonstrated that utilizing the Dirichlet distribution can effectively estimate the credibility of a single view. As the Beta distribution serves as a dimensionality reduction of the Dirichlet distribution and shares the same mathematical significance in binary classification scenarios, we interpret the output before the softmax operation of the classifier for the v-th view as the "evidence" e^v for inferring fake news. This "evidence" quantifies the support for the classification result gathered from the input and is employed to derive the parameters β^v of the Beta distribution, *i.e*,

$$e_r^v = \mathbf{Softplus}(o_r^v) \quad (6)$$
$$\beta_r^v = 1 + e_r^v \quad (7)$$

for $r \in \{0, 1\}$, where o_r^v represents the output of the final layer of the classifier model for the v-th view regarding the r-th classification result. Consequently, we infer the credible quality b_k^v of classifying the news into the k-th class, *i.e*, $b_r^v = e_r^v / S^v$, where $S_v = \beta_0^v + \beta_1^v$ represents the strength of the beta distribution. Beta distribution parameterizes the "evidence" as credible quality and serves as the conjugate prior for the classification distribution. We connect the parameters of the beta distribution to the uncertainty of the model's classification u^v using the Subjective Logic Theory framework (Jsang, 2018). Specifically, the sum of credible quality and uncertainty for the classification of real and fake news under a certain view is constrained to be 1, *i.e*, $u^v + b_0^v + b_1^v = 1$. q^v characterizes the credibility of a particular view clue, which will be directly used in the fusion process of multi-view features. Certainly, it's straightforward to understand that the credibility q^v of the v-th view can be inferred by subtracting the uncertainty from 1, *i.e*, $q^v = 1 - u^v = b_0^v + b_1^v$.

We concatenate the credibility of three views, *i.e*, modality inconsistency q_k^c, post emotion q_k^e, and image pattern q_k^f, to form the credibility vector, *i.e*,

$$Q_k = [q_k^c, q_k^e, q_k^f]. \quad (8)$$

3.4 Multi-view fusion layer

After obtaining the credibility estimates from individual views, the fusion of representations of inconsistency, emotions, and patterns with the corresponding credibility is achieved using a Multi-Head Self-Attention Network. This process enables modality interaction by multiplying the representations with their respective credibility, *i.e*,

$$x_k = \mathbf{MHSA}([x_k^c, x_k^e, x_k^f] \cdot Q_k^T). \quad (9)$$

Simultaneously, we evaluated the structural differences among representations from different views (Lei et al., 2022) to enhance the model's generalization performance, *i.e*,

$$\tilde{x}_k = flatten(sample(M)), \quad (10)$$

where M denotes the attention matrix, which undergoes downsampling after being flattened. The fused features derived from multiple views enable robust detection of intricate fake news samples within social networks.

Method	Twitter		Weibo		Pheme	
	Accuracy	F1 Score	Accuracy	F1 Score	Accuracy	F1 Score
EANN (Wang et al., 2018)	0.648	0.6385	0.782	0.780	0.681	0.721
SAFE (Zhou et al., 2020)	0.762	0.761	0.763	0.761	0.811	0.767
MVAE (Khattar et al., 2019)	0.745	0.744	0.824	0.823	0.852	0.827
CLIP+MLP (Radford et al., 2021)	0.857	0.853	0.887	0.886	0.870	0.845
CAFE+CLIP (Chen et al., 2022)	0.879	0.857	0.897	0.896	<u>0.882</u>	0.856
MCAN+CLIP (Wu et al., 2021)	<u>0.917</u>	<u>0.911</u>	0.900	0.899	<u>0.882</u>	<u>0.861</u>
FND-CLIP (Zhou et al., 2023)	0.902	0.896	<u>0.907</u>	<u>0.907</u>	0.875	0.857
BMR (Ying et al., 2023)	0.883	0.870	0.889	0.889	0.863	0.830
Event-Radar	**0.928**	**0.923**	**0.919**	**0.919**	**0.901**	**0.880**

Table 1: Fake news detection system's accuracy and binary F1 scores on three datasets. **Bold** indicates the best performance, while <u>underlined</u> denotes the second-best performance. Event-Radar demonstrates superior performance across all three datasets compared to all seven multimodal fake news detection baselines. The detailed classification results in each category will be provided in the appendix.

3.5 Training and Inference

After combining the features from multiple views, we connect these fused features with the structural difference features. Then, applying a linear transformation, we obtain the predicted results, *i.e*,

$$y_k = W_o \cdot [x_k, \tilde{x}_k] + b_o, \tag{11}$$

where W_o and b_o represent the learnable parameters. During the training process of Event-Radar, we refer to (Kiela et al., 2018) and utilized the the credible loss $\mathcal{L}_u$ for each view to ensure the model's judgments are more confident for each sample, *i.e*,

$$\mathcal{L}_u(x_k^v) = \sum_{k \in \mathcal{Y}} \hat{y}_k \cdot (\psi(S^v) - \psi(\beta_k^v)), \tag{12}$$

where $\mathcal{Y}$ is the annotated label set, $\hat{y}_k$ represents the ground truth label and $\psi(\cdot)$ denotes the digamma function. We also incorporate a contrastive learning loss $\mathcal{L}_c$ to encourage features to be as distant as possible from distributions with low credibility in the embedding space, *i.e*,

$$\mathcal{L}_c(s_k, q_{min}^t) = \frac{\sum_{i \neq t}^{(c,e,f)} s_k^i (1 - q_{min}^t) + s_k^t q_{min}^t}{\sum_i^{(c,e,f)} s_k^i},$$

where q_{min}^t signifies the value of the lowest credibility among the views, t denotes the view with the lowest credibility and $s_k = \left\{ s_k^c, s_k^e, s_k^f \right\}$ represents the set of similarities between the single-view representations and the fused representation. The overall loss function can be presented as:

Category	Ablation Settings	Accuracy	F1 Score
Full Model	Event-Radar	**0.928**	**0.923**
Fusion Method	Use MOE	0.919	0.913
	w/o L_c	0.911	0.907
	Only Concat	0.908	0.903
View	w/o Inconsistency	0.897	0.891
	w/o Emotion	0.905	0.893
	w/o Pattern	0.892	0.878

Table 2: Ablation study of Event-Radar. The test was conducted on Twitter. Other results are in the appendix.

$$\begin{aligned} \mathcal{L} = \sum_{k \in \mathcal{Y}} \hat{y}_k \log y_k + \lambda_1 \sum_{k \in \mathcal{Y}} \sum_{v}^{\{c,e,f\}} \mathcal{L}_u(x_k^v) \\ + \lambda_2 \sum_{k \in \mathcal{Y}} \mathcal{L}_c(s_k, u_{\min}^v), \end{aligned} \tag{13}$$

where λ_1 and λ_2 are hyperparameters used to balance these components.

4 Experiment

4.1 Experiment Settings

We evaluated the Event-Radar on three widely used benchmarks for fake news detection: Twitter (Boididou et al., 2015), Weibo (wei), and Pheme (Zubiaga et al., 2017). Twitter was released in 2015 at MediaEval, comprising 17673 news. Weibo is the most extensively used Chinese dataset with 9528 news exposing fake news. Pheme is designed for detecting fake news spread on social media and consists of five breaking news stories, encompassing a total of 3670 news. In all of our experiments, we used the division of the original dataset into training and test sets. We selected classic models EANN, SAFE, MVAE, CAFE, MCAN, FND-CLIP, and BMR as strong baselines.

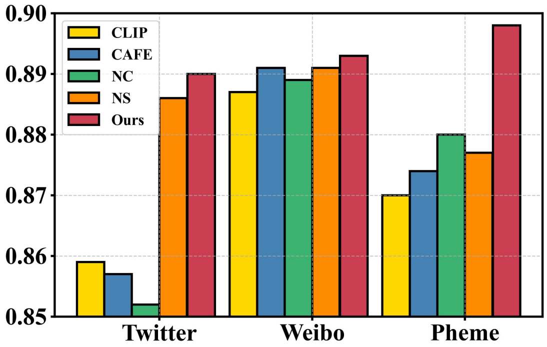

Figure 4: The results of inconsistency studies.

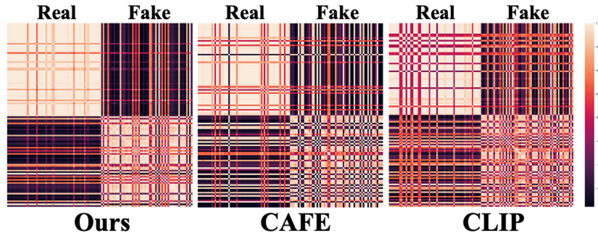

Figure 5: Heatmap visualization. Each cell in the heat maps represents the paired cosine similarity.

To ensure fairness, we replace the backbones of the latest strong baselines CAFE and MCAN, with CLIP having identical parameters. We also use CLIP+MLP as a comparative baseline. BMR proposed the use of MAE as a more suitable backbone for fake news detection; therefore, we did not alter its backbone. Meanwhile, BMR improved by removing poor-quality samples during data preparation, which, however, cannot address the complex data distribution in the real world. Therefore, during testing, we used the most original data distribution. More details of the implementation and baselines can be found in the appendix.

4.2 Main Result

We evaluated Event-Radar and eight representative baselines on three fake news detection benchmarks. Table.1 presents the results, indicating:

- The performance of Event-Radar consistently outperforms all baseline methods across the three datasets. On average, it achieves an 1.4% increase in accuracy and 1.43% increase in F1 scores compared to baslines on the three datasets.
- It is evident that leveraging the powerful multimodal representation capabilities, the CLIP+MLP method has achieved remarkably good detection performance. While methods like CAFE and MCAN show limited improvement on a CLIP-based backbone, Event-Radar demonstrates higher enhancement due to its ability to model at the event level and encode credibility across multiple views, compared to CLIP+MLP.
- Through multi-view feature modeling, MCAN and BMR have achieved excellent results in inferring fake news. However, subsequent experiments have shown that the ability of multi-view learning is highly sensitive to the quality of news samples. Event-Radar, through the adoption of more effective event modeling and credibility calculation, has demonstrated more promising inferential capabilities and accuracy. Its specific noise resistance will be further validated in robustness experiments.

4.3 Ablation Studies

We conduct further analysis to examine the roles of each module in our proposed model. The corresponding results are shown in Table. 2:

- To validate the reliability of our fusion approach, apart from simple concatenation of features and excluding the enhanced representation loss $\mathcal{L}_c$, we also employed the MOE used in BMR for multimodal fusion as a comparison to validate the effectiveness of our fusion strategy. We observe that while using the MOE fusion method yielded decent results, it fell 0.9% and 1% lower in accuracy and F1 Score respectively compared to Event-Radar. This also shows that our credibility-based fusion approach is effective.
- To assess the effectiveness of utilizing information from each modality, we removed event inconsistency, emotion, and pattern features, comparing these ablated versions with the complete model. The view ablation experiments in the table.2 show that removing any view leads to performance degradation, which emphasizes the advantage of learning fused representations from multiple views to better judge news veracity.

4.4 Inconsistency Studies

To demonstrate our superior ability in measuring multimodal inconsistency, we focused on the multimodal inconsistency view for fake news inference. For a fair comparison, we compared this view with CLIP and CAFE retaining only the inconsistency component. It is evident that our inconsistency modeling capability surpasses the current popular methods in Fig.4. Additionally, we constructed models that do not use event subgraphs and only perform convolution on the multimodal graph $\mathcal{G}$ (denoted as "NS"), and models that only construct event subgraphs without convolution (denoted as

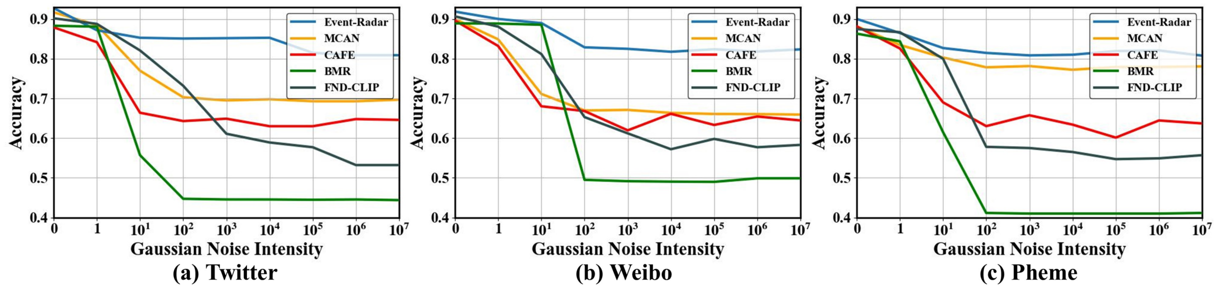

Figure 6: Classification results after adding Gaussian noise of different intensities.

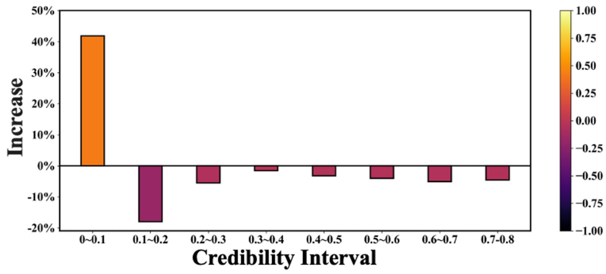

Figure 7: The changes in credibility distribution.

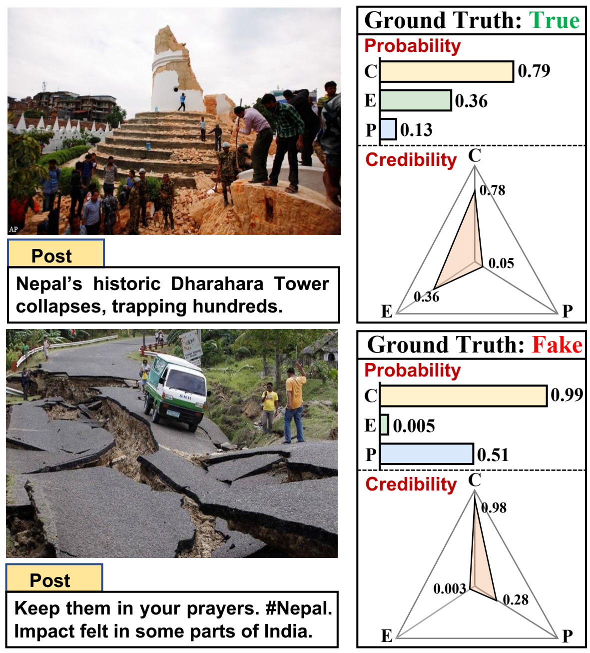

Figure 8: Case Study. We present several challenging instances along with their images and posts.

"NC"). We tested their inconsistency inference performance to validate the rationality of our inconsistency module design. As observed, our model achieved better accuracy, even outperforming most baselines in the main experiment. In Fig.5, we employed a heatmap visualization to measure the representational ability of the inconsistency module. We selected 60 true news and 60 fake news, calculating the paired similarity between incoming news. It is clear that our inconsistency modeling method exhibits significant intra-class similarity and inter-class differences, demonstrating strong discriminative capabilities.

4.5 Robustness Study

To further validate the ability of Event-Radar to learn from news samples with varying quality, we added Gaussian noise of different intensities to the features of each view in the test data. This was done to simulate information degradation or poor quality in a certain view, assessing the robustness of Event-Radar in integrating multimodal features. As shown in Fig.6, we tested the accuracy after adding noise to each modality and averaged the results to evaluate the performance of each model. Clearly, the performance of all models decreased to some extent after adding noise, while our model's accuracy dropped less and remained relatively stable. To further explore the underlying mechanisms, we plotted the change in the number of test samples in each credibility interval after adding Gaussian noise with an intensity of 10^2 in a certain view in Fig.7. It can be observed that after adding Gaussian noise, the number of samples in the credibility interval of 0 to 0.1 increased by 40%, while the number of high-credibility samples sharply decreased. This indicates that the model adjusts the fusion strategy by reducing the credibility in that view, thereby alleviating the performance decline.

4.6 Case Study

We present the probability and credibility associated with each view while classifying both real and fake news, thereby illustrating the classification process employed by Event-Radar. As shown in Fig. 8, it displays the contribution of specific views and the model's credibility in each view. We denote the inconsistency between post and image as "C", post emotion as "E", and pattern features as "P". In the first news, although the model classified it as fake news based on emotion and pattern features, the credibility for these modalities were low. The model chose to believe the judgment based on event inconsistency, resulting in the correct classification. Similarly, in the second example, the model provided correct credible judgments, leading to the correct classification result.

5 CONCLUSION

Event-Radar is a novel fake news detection framework that demonstrates exceptional event modeling capabilities and significant robustness, aimed at addressing the issue of maliciously crafted fake news. Extensive experiments validate that Event-Radar's classification performance surpasses all listed strong benchmarks. Further research confirms the effectiveness of our initial technical contributions, emphasizing Event-Radar's ability to resist interference from poor-quality news.

6 ACKNOWLEDGEMENT

This work was supported by the National Key Research and Development Program of China (No. 2022YFB3102600), the National Nature Science Foundation of China (No. 62192781, No. 62272374), the Natural Science Foundation of Shaanxi Province (2024JC-JCQN-62), the National Nature Science Foundation of China (No. 62202367, No. 62250009, No. 62137002), Project of China Knowledge Center for Engineering Science and Technology, and Project of Chinese academy of engineering "The Online and Offline Mixed Educational Service System for 'The Belt and Road' Training in MOOC China". We would like to express our gratitude for the support of K. C. Wong Education Foundation.

7 LIMITATIONS

Our work has two limitations that may impact the generalization ability of our proposed framework. While introducing event graphs has yielded promising results in fake news detection, we have yet to explore event representation learning from a causal relationship perspective. Additionally, the performance of the NER tools and object detection tools used can impact the structure of the event subgraph, thereby influencing the accuracy of event representation. Moreover, although the confidence-based fusion layer used in our work is effective in resisting interference from low-quality samples, extremely small confidence scores may result in an abundance of zero values in the classifier input, posing a risk of overfitting or gradient vanishing. We plan to address these limitations in future research.

8 ETHICS STATEMENT

This paper adheres to the ACM Code of Ethics and Professional Conduct. Firstly, the dataset utilized does not contain sensitive private information and poses no harm to society. Secondly, proper attribution is given to relevant papers and the sources of pre-trained models, along with detailed references to the toolkits used. Furthermore, our code will be released under the license of any artifacts used. Lastly, the proposed fake news detection method is designed to contribute to the safety and stability of the internet environment and public opinion.

References

Christina Boididou, Katerina Andreadou, Symeon Papadopoulos, Duc Tien Dang Nguyen, Giulia Boato, Michael Riegler, Yiannis Kompatsiaris, et al. 2015. Verifying multimedia use at mediaeval 2015. In *MediaEval 2015*, volume 1436. CEUR-WS.

Juan Cao, Peng Qi, Qiang Sheng, Tianyun Yang, Junbo Guo, and Jintao Li. 2020. Exploring the role of visual content in fake news detection. *Disinformation, Misinformation, and Fake News in Social Media: Emerging Research Challenges and Opportunities*, pages 141–161.

Kai Chen, Jiaqi Wang, Jiangmiao Pang, Yuhang Cao, Yu Xiong, Xiaoxiao Li, Shuyang Sun, Wansen Feng, Ziwei Liu, Jiarui Xu, Zheng Zhang, Dazhi Cheng, Chenchen Zhu, Tianheng Cheng, Qijie Zhao, Buyu Li, Xin Lu, Rui Zhu, Yue Wu, Jifeng Dai, Jingdong Wang, Jianping Shi, Wanli Ouyang, Chen Change Loy, and Dahua Lin. 2019. MMDetection: Open mmlab detection toolbox and benchmark. *arXiv preprint arXiv:1906.07155*.

Yixuan Chen, Dongsheng Li, Peng Zhang, Jie Sui, Qin Lv, Lu Tun, and Li Shang. 2022. Cross-modal ambiguity learning for multimodal fake news detection. In *Proceedings of the ACM Web Conference 2022*, pages 2897–2905.

Ziwei Chen, Linmei Hu, Weixin Li, Yingxia Shao, and Liqiang Nie. 2023. Causal intervention and counterfactual reasoning for multi-modal fake news detection. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2023, Toronto, Canada, July 9-14, 2023*, pages 627–638. Association for Computational Linguistics.

Zhuohang Dang, Minnan Luo, Chengyou Jia, Guang Dai, Xiaojun Chang, and Jingdong Wang. 2024.

Noisy correspondence learning with self-reinforcing errors mitigation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 1463–1471.

Zhuohang Dang, Minnan Luo, Chengyou Jia, Guang Dai, Jihong Wang, Xiaojun Chang, Jingdong Wang, and Qinghua Zheng. 2023. Disentangled representation learning with transmitted information bottleneck. *arXiv preprint arXiv:2311.01686*.

Matthias Fey and Jan Eric Lenssen. 2019. Fast graph representation learning with pytorch geometric. *arXiv preprint arXiv:1903.02428*.

Marc Fisher, John Woodrow Cox, and Peter Hermann. 2016. Pizzagate: From rumor, to hashtag, to gunfire in dc. *Washington Post*, 6:8410–8415.

Zongbo Han, Changqing Zhang, Huazhu Fu, and Joey Tianyi Zhou. 2021. Trusted multi-view classification. *arXiv preprint arXiv:2102.02051*.

Audun Jsang. 2018. *Subjective Logic: A formalism for reasoning under uncertainty*. Springer Publishing Company, Incorporated.

Dhruv Khattar, Jaipal Singh Goud, Manish Gupta, and Vasudeva Varma. 2019. Mvae: Multimodal variational autoencoder for fake news detection. In *The world wide web conference*, pages 2915–2921.

Douwe Kiela, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2018. Efficient large-scale multimodal classification. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32.

Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.

An Lao, Chongyang Shi, and Yayi Yang. 2021. Rumor detection with field of linear and non-linear propagation. In *Proceedings of the Web Conference 2021*, pages 3178–3187.

Zhenyu Lei, Herun Wan, Wenqian Zhang, Shangbin Feng, Zilong Chen, Jundong Li, Qinghua Zheng, and Minnan Luo. 2022. Bic: Twitter bot detection with text-graph interaction and semantic consistency. *arXiv preprint arXiv:2208.08320*.

Manling Li, Ruochen Xu, Shuohang Wang, Luowei Zhou, Xudong Lin, Chenguang Zhu, Michael Zeng, Heng Ji, and Shih-Fu Chang. 2022. Clip-event: Connecting text and images with event structures. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16420–16429.

Hui Liu, Wenya Wang, and Haoliang Li. 2023. Interpretable multimodal misinformation detection with logic reasoning. In *Findings of the Association for Computational Linguistics: ACL 2023, Toronto, Canada, July 9-14, 2023*, pages 9781–9796. Association for Computational Linguistics.

Lemao Liu, Haisong Zhang, Haiyun Jiang, Yangming Li, Enbo Zhao, Kun Xu, Linfeng Song, Suncong Zheng, Botong Zhou, Jianchen Zhu, Xiao Feng, Tao Chen, Tao Yang, Dong Yu, Feng Zhang, Zhanhui Kang, and Shuming Shi. Texsmart: A system for enhanced natural language understanding.

Yan Liu and Hong-Dong Li. 2003. Image and video processing techniques in the dct domain. *Journal of Image and Graphics*, 8(2):121–128.

Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. The Stanford CoreNLP natural language processing toolkit. In *Association for Computational Linguistics (ACL) System Demonstrations*, pages 55–60.

Salman Bin Naeem and Rubina Bhatti. 2020. The covid-19 'infodemic': a new front for information professionals. *Health Information & Libraries Journal*, 37(3):233–239.

Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. 2017. Automatic differentiation in pytorch.

Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830.

Peng Qi, Juan Cao, Tianyun Yang, Junbo Guo, and Jintao Li. 2019. Exploiting multi-domain visual information for fake news detection. In *2019 IEEE international conference on data mining (ICDM)*, pages 518–527. IEEE.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. 2021. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pages 8748–8763. PMLR.

Noam Shazeer, Azalia Mirhoseini, Krzysztof Maziarz, Andy Davis, Quoc Le, Geoffrey Hinton, and Jeff Dean. 2017. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. *arXiv preprint arXiv:1701.06538*.

Dinghan Shen, Xinyuan Zhang, Ricardo Henao, and Lawrence Carin. 2018. Improved semantic-aware network embedding with fine-grained word alignment. *arXiv preprint arXiv:1808.09633*.

Qiang Sheng, Xueyao Zhang, Juan Cao, and Lei Zhong. 2021. Integrating pattern- and fact-based fake news detection via model preference learning. In *CIKM '21: The 30th ACM International Conference on Information and Knowledge Management, Virtual Event, Queensland, Australia, November 1 - 5, 2021*, pages 1640–1650.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.

Herun Wan, Shangbin Feng, Zhaoxuan Tan, Heng Wang, Yulia Tsvetkov, and Minnan Luo. 2024. Dell: Generating reactions and explanations for llm-based misinformation detection. *arXiv preprint arXiv:2402.10426*.

Weiran Wang, Honglak Lee, and Karen Livescu. 2016. Deep variational canonical correlation analysis. *CoRR*, abs/1610.03454.

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.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. 2020. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*, pages 38–45.

Mingmin Wu, Yuxue Hu, Yongcheng Zhang, Zeng Zhi, Guixin Su, and Ying Sha. 2024. Mitigating idiom inconsistency: A multi-semantic contrastive learning method for chinese idiom reading comprehension. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 19243–19251.

Nan Wu, Stanislaw Jastrzebski, Kyunghyun Cho, and Krzysztof J Geras. 2022. Characterizing and overcoming the greedy nature of learning in multi-modal deep neural networks. In *International Conference on Machine Learning*, pages 24043–24055. PMLR.

Yang Wu, Pengwei Zhan, Yunjian Zhang, Liming Wang, and Zhen Xu. 2021. Multimodal fusion with co-attention networks for fake news detection. In *Findings of the association for computational linguistics: ACL-IJCNLP 2021*, pages 2560–2569.

An Yang, Junshu Pan, Junyang Lin, Rui Men, Yichang Zhang, Jingren Zhou, and Chang Zhou. 2022. Chinese clip: Contrastive vision-language pretraining in chinese. *arXiv preprint arXiv:2211.01335*.

Qichao Ying, Xiaoxiao Hu, Yangming Zhou, Zhenxing Qian, Dan Zeng, and Shiming Ge. 2023. Bootstrapping multi-view representations for fake news detection. In *Proceedings of the AAAI conference on Artificial Intelligence*, volume 37, pages 5384–5392.

Zhi Zeng, Mingmin Wu, Guodong Li, Xiang Li, Zhongqiang Huang, and Ying Sha. 2023. An explainable multi-view semantic fusion model for multimodal fake news detection. In *2023 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1235–1240. IEEE.

Haisong Zhang, Lemao Liu, Haiyun Jiang, Yangming Li, Enbo Zhao, Kun Xu, Linfeng Song, Suncong Zheng, Botong Zhou, Jianchen Zhu, Xiao Feng, Tao Chen, Tao Yang, Dong Yu, Feng Zhang, Zhanhui Kang, and Shuming Shi. Texsmart: A text understanding system for fine-grained ner and enhanced semantic analysis.

Wenjia Zhang, Lin Gui, and Yulan He. 2021a. Supervised contrastive learning for multimodal unreliable news detection in covid-19 pandemic. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, pages 3637–3641.

Xueyao Zhang, Juan Cao, Xirong Li, Qiang Sheng, Lei Zhong, and Kai Shu. 2021b. Mining dual emotion for fake news detection. In *Proceedings of the web conference 2021*, pages 3465–3476.

Zhe Zhao, Paul Resnick, and Qiaozhu Mei. 2015. Enquiring minds: Early detection of rumors in social media from enquiry posts. In *Proceedings of the 24th international conference on world wide web*, pages 1395–1405.

Xinyi Zhou, Jindi Wu, and Reza Zafarani. 2020. Safe: Similarity-aware multi-modal fake news detection. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 354–367. Springer.

Yangming Zhou, Yuzhou Yang, Qichao Ying, Zhenxing Qian, and Xinpeng Zhang. 2023. Multimodal fake news detection via clip-guided learning. In *2023 IEEE International Conference on Multimedia and Expo (ICME)*, pages 2825–2830.

Arkaitz Zubiaga, Maria Liakata, and Rob Procter. 2017. Exploiting context for rumour detection in social media. In *Social Informatics: 9th International Conference, SocInfo 2017, Oxford, UK, September 13-15, 2017, Proceedings, Part I 9*, pages 109–123. Springer.

A Implementation

We utilized PyTorch (Paszke et al., 2017), PyTorch Geometric (Fey and Lenssen, 2019), scikit-learn (Pedregosa et al., 2011), and Transformers (Wolf et al., 2020) to implement Event-Radar. Table.3 outlines the hyperparameter settings for easy replication of experimental results. For the model's backbone, "ViT-B/16" was employed for Twitter and Pheme datasets, while "Chinese-CLIP-ViT-B-16" provided by HuggingFace (Yang et al., 2022) was used for Weibo. The experiments were conducted on a Tesla A100 GPU.

Table 3: Hyperparameter settings of Event-Radar

Hyperparameter	Twitter	Weibo	Pheme
optimizer	Adam	Adam	Adam
learning rate	5e-4	5e-5	5e-5
credible loss coefficient	0.4	0.4	0.4
constractive loss coefficient	0.2	0.2	0.2
graph update rate α	0.6	0.6	0.6

B Baselines

To validate the effectiveness of our method, we applied the Event-Radar framework to the following seven strong baselines:

EANN (Wang et al., 2018) trained a fake news classifier based on extracting post events.

SAFE (Zhou et al., 2020) employs consistency as a loss function to optimize the task.

MVAE (Khattar et al., 2019) employs a variational autoencoder to model representations between text and images.

CLIP (Radford et al., 2021) exhibits strong multimodal representation capabilities. We concatenated CLIP with a two-layer MLP for our task.

CAFE (Chen et al., 2022) adaptively aggregates features based on the inherent cross-modal ambiguity, addressing misclassification issues arising from differences between different modalities.

MCAN (Wu et al., 2021) integrates pattern features into the co-attention network. It conducts detection by incorporating multiple views that fuse text, image semantics, and image pattern features.

FND-CLIP (Zhou et al., 2023) leverages the multimodal cognitive capabilities of clip by generating self-directed attentional weights to fuse features through modal similarity computed by CLIP.

BMR (Ying et al., 2023) models news features from multiple views through bootstrap multi-view representations. It utilizes the Mixture of Experts network for the fusion of multi-view features.

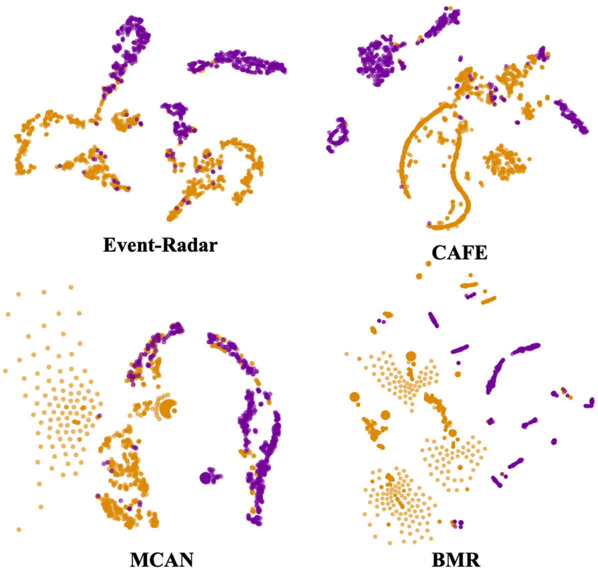

Figure 9: TSNE visualization of mined features on the test set. Dots with the same color are within the same label.

C Ablation Studies

The detailed abaltion study results in the ablation study are shown in Table4.

D Classification Result

The detailed classification results in the main experiment are shown in Table.5.

E Representation Study

In Fig.9, we present t-SNE visualizations of different model features learned by Event-Radar, CAFE, MCAN, and BMR on the Twitter test set. In Event-Radar, there is a relatively clear boundary between true and false news, and the clustering effect is good, with fewer outliers. This indicates that the features extracted by Event-Radar are more distinctive.

Table 4: Ablation Study Result.

Category	Ablation Settings	Twitter		Weibo		Pheme	
		Acc	F1	Acc	F1	Acc	F1
Full Model	Event-Radar	**0.928**	**0.923**	**0.919**	**0.919**	**0.901**	**0.880**
Fusion Method	Use MOE	0.919	0.913	0.912	0.911	0.894	0.876
	w/o L_c	0.911	0.907	0.891	0.891	0.900	0.876
	Only Concat	0.908	0.903	0.904	0.904	0.870	0.845
View	w/o Inconsistency	0.897	0.891	0.902	0.902	0.880	0.841
	w/o Emotion	0.905	0.893	0.887	0.886	0.887	0.854
	w/o Pattern	0.892	0.878	0.877	0.877	0.884	0.866

Table 5: Classification Result.

Dataset	Method	Accuracy	F1	Real News			Fake News		
				Precision	Recall	F1-score	Precision	Recall	F1-score
Twitter	EANN	0.648	0.639	0.584	0.759	0.660	0.810	0.498	0.617
	SAFE	0.762	0.761	0.695	0.811	0.748	0.831	0.724	0.774
	MVAE	0.745	0.744	0.689	0.777	0.730	0.801	0.719	0.758
	CLIP+MLP	0.857	0.853	<u>0.941</u>	0.824	0.879	0.755	<u>0.913</u>	0.827
	MCAN-CLIP	<u>0.917</u>	<u>0.911</u>	0.935	0.934	<u>0.934</u>	0.888	0.889	<u>0.888</u>
	FND-CLIP	0.902	0.896	0.935	0.907	0.921	0.851	0.894	0.872
	CAFE-CLIP	0.879	0.857	0.909	0.918	0.913	0.811	0.793	0.802
	BMR	0.883	0.870	0.865	**0.965**	0.912	**0.927**	0.746	0.827
	Event-Radar	**0.928**	**0.923**	**0.942**	<u>0.943</u>	**0.943**	<u>0.904</u>	**0.902**	**0.903**
Weibo	EANN	0.782	0.780	0.752	0.863	0.804	0.827	0.697	0.756
	SAFE	0.763	0.761	0.717	0.868	0.785	0.833	0.659	0.736
	MVAE	0.824	0.823	0.802	0.875	0.837	0.854	0.769	0.809
	CLIP+MLP	0.887	0.886	0.890	0.869	0.879	0.883	<u>0.903</u>	0.893
	MCAN-CLIP	0.900	0.899	0.915	0.869	0.892	0.887	0.827	0.907
	CAFE-CLIP	0.897	0.896	0.889	0.893	0.891	0.904	0.900	0.902
	FND-CLIP	<u>0.907</u>	<u>0.907</u>	<u>0.914</u>	<u>0.901</u>	<u>0.907</u>	<u>0.917</u>	0.901	<u>0.908</u>
	BMR	0.889	0.889	0.874	0.894	0.884	0.904	0.885	0.895
	Event-Radar	**0.919**	**0.919**	**0.924**	**0.905**	**0.914**	**0.932**	**0.915**	**0.924**
Pheme	EANN	0.681	0.721	0.701	0.750	0.747	0.685	0.664	0.694
	SAFE	0.811	0.767	0.806	**0.940**	0.866	<u>0.827</u>	0.559	0.667
	MVAE	0.852	0.827	0.871	0.917	0.893	0.806	0.719	0.760
	CLIP+MLP	0.870	0.845	0.899	0.917	0.908	0.800	0.763	0.781
	MCAN-CLIP	<u>0.882</u>	<u>0.861</u>	0.904	0.907	0.906	0.783	0.777	0.780
	CAFE-CLIP	<u>0.882</u>	0.856	<u>0.932</u>	0.902	<u>0.917</u>	0.765	<u>0.828</u>	0.795
	FND-CLIP	0.875	0.857	**0.937**	0.881	0.908	0.758	**0.862**	<u>0.807</u>
	BMR	0.863	0.830	0.879	0.834	0.905	0.820	0.700	0.755
	Event-Radar	**0.901**	**0.880**	0.925	<u>0.934</u>	**0.929**	**0.841**	0.822	**0.831**