

基于社交图与标签优化的多模态虚假新闻检测^①



王春东^{1,2,3}, 陈星宇^{1,2}, 林浩^{1,2}

¹(天津理工大学 计算机科学与工程学院, 天津 300384)

²(天津理工大学 计算机病毒防治技术国家工程实验室, 天津 300384)

³(天津公安警官职业学院, 天津 300382)

通信作者: 王春东, E-mail: michael3769@163.com

摘要: 现有虚假新闻检测方法存在明显不足. 这些方法未能充分挖掘多模态关联, 忽视关键信息导致冗余问题, 且未考虑相似新闻的正负关联影响检测精度. 此外, 能协同利用文本、图像和社交图这 3 种模态的方法较为稀缺. 针对这些问题, 提出了一种基于社交图与标签优化的多模态虚假新闻检测模型. 该方法同步提取文本、图像和社交图特征, 通过全局动态加权机制强化社交图关键特征并通过图像描述增强视觉表征, 利用跨模态交叉注意力实现模态间优势互补, 并结合对比学习进行跨模态对齐. 本文模型在标签优化器中通过双向标签注意力捕捉相似新闻间的正负相关性优化预测标签. 在 Weibo 和 PHEME 数据集上的实验表明, 本文方法较基线方法准确率提升 1.63% 和 3.01%, *F1* 分数提升 2.46% 和 3.67%.

关键词: 虚假新闻检测; 社交图; 多模态融合; 对比学习; 注意力机制; 标签优化

引用格式: 王春东, 陈星宇, 林浩. 基于社交图与标签优化的多模态虚假新闻检测. 计算机系统应用. <http://www.c-s-a.org.cn/1003-3254/10117.html>

Multimodal Fake News Detection Based on Social Graph and Label Optimization

WANG Chun-Dong^{1,2,3}, CHEN Xing-Yu^{1,2}, LIN Hao^{1,2}

¹(School of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China)

²(National Engineering Laboratory for Computer Virus Prevention and Control Technology, Tianjin University of Technology, Tianjin 300384, China)

³(Tianjin Public Security Police Profession College, Tianjin 300382, China)

Abstract: Existing fake news detection methods suffer from notable limitations. They typically fail to adequately explore multimodal correlations, overlook critical information, thereby introducing redundancy, and disregard the positive or negative influence of similar news articles, all of which impair detection accuracy. Moreover, approaches capable of synergistically utilizing textual, visual, and social-graph modalities remain scarce. To address these issues, this study proposes a multimodal fake-news detection model integrating social graphs and label optimization. The method simultaneously extracts features from text, images, and social graphs. Key social graph signals are enhanced via a global dynamic weighting mechanism, while visual representations are strengthened using image descriptions. Cross-modal complementary advantages are achieved through cross-modal cross-attention, and contrastive learning is applied for cross-modal alignment. In the label optimizer, bidirectional label attention is used to capture positive and negative correlations among semantically similar news articles, thereby refining the predicted labels. Experimental results on Weibo and PHEME show that the proposed model improves accuracy by 1.63 % and 3.01 %, and *F1*-score by 2.46 % and 3.67 %, respectively, compared to baseline methods.

Key words: fake news detection; social graph; multimodal fusion; contrastive learning; attention mechanism; label optimization

① 基金项目: 国家重点研发计划“区块链”重点专项 (2023YFB2703900); 天津市科委重大专项 (15ZXDSGX00030)

收稿时间: 2025-09-22; 修改时间: 2025-10-27; 采用时间: 2025-11-07; csa 在线出版时间: 2026-02-06

自从人类步入网络信息时代,网络平台与社交媒体已成为新闻传播的主要渠道.新闻创作和传播方式也发生了巨大的变革,用户不仅是信息消费者,更是信息创造者.用户在社交媒体上通过多元化的形式(如图片、文本、视频、音频等)发布消息,虚假新闻也在这一过程中大量滋生.根据在2025年1月由中国互联网络信息中心发布的第55次《中国互联网络发展状况统计报告》^[1]中的数据显示,截至2024年底,我国网民规模已超过11亿,互联网普及率稳步提升,这也使得虚假新闻日益增多.虚假新闻指的是刻意编造并传播的新闻,其动机可能涉及误导大众、操控舆论、政治干预或经济利益^[2].新冠疫情期间,有人将疫情传播归咎于5G网络,导致英国多起信号塔被纵火,严重影响市民生活^[3].因此,及时识别和阻断虚假新闻传播对于减少社会负面影响、维护社会稳定具有重要意义.

传统虚假信息检测方法通常采用单模态,如基于文本的关键词过滤或基于图像的特征识别.部分改进方法尝试浅层融合两种模态,但这些方法忽略了模态间的潜在关联性,且无法充分捕捉图文模态的深层语义关联,限制了检测性能^[4].多模态信息融合技术涉及表示、转换、对齐和融合,但缺乏捕捉相似新闻之间的正负关系.本研究提出MFND-SGLO模型,将社交图和标签优化器结合到多模态流程中,采用协同特征提取策略,增强文本、视觉和社交图特征的表达,提升多模态融合效果.通过跨模态交叉注意力和对比学习框架,实现模态间语义对齐,特别是标签优化器通过双向标签注意力挖掘相似新闻间的正负关联.

本文主要贡献如下.

(1) 引入图像描述模型生成图像描述,并与视觉特征融合,实现了对视觉特征的数据增强,提升多模态信息融合效果.同时,设计多模态对比交叉熵损失,增强了多模态特征的对齐和分类判别能力.

(2) 设计了增强的图注意力网络GAT-SE,该网络能够有效抑制噪声干扰和冗余特征,从而增强对全局信息的捕捉能力,提高模型区分虚假新闻的能力.

(3) 结合全局标签优化器捕捉社交图中相似帖子节点间的正负关系,解决了传统多模态方法未考虑相似新闻间正负关系的问题,减少了依赖单条新闻而导致的误判.

(4) 实验结果表明,设计的MFND-SGLO在多数场景优于基线模型.具体而言,在PHEME数据集上准确率与F1分数分别提升3.01%、3.67%.在Weibo数据集

上,准确率与F1分数分别提升1.63%、2.46%.

1 相关工作

1.1 基于单模态的虚假新闻检测方法

王振琦等人^[5]提出一种基于交叉特征感知融合的虚假新闻检测方法,通过捕捉新闻在文本、情感及风格特征上的领域差异和跨领域关联,提升了模型检测性能.李铭伟等人^[6]针对文本语义稀疏及多源信息关联的难题,提出了基于自监督异质子图和子图注意力网络的虚假新闻检测方法,有效融合文本、党派、主题等信息,提升了检测的准确率.然而,这些研究在模态的充分利用方面仍有不足. Agarwal等人^[7]通过融合细粒度与粗粒度特征,深入捕捉句子中词语的复杂相互依赖关系及语义信息,以实现虚假新闻的检测. Luvembe等人^[8]通过双重情感特征和优化的随机森林进行虚假新闻检测.

1.2 基于多模态的虚假新闻检测方法

多模态技术融合图片、文本等多种模态,弥补单模态方法的局限,进而做出更精准的判断.

例如, Wu等人^[9]提出了多模态协同注意力网络(MCAN),通过VGG模型捕捉图像的空间特征,并借助CNN提取图像的频率特征.再利用共同注意力机制将这两种特征进行有效融合,从而获得更为精准的图像表示. HMCAN^[10]把图像特征和文本特征共同输入到一个多模态上下文注意力网络中,通过捕捉不同模态之间和相同模态内部的关系,实现了更高效的融合. SAFE^[11]融合多模态特征与跨模态相似性,共同学习新闻的特征表示. Khattar等人^[12]提出了多模态变分自动编码器(MVAE),该模型采用双模态变分自编码器和二进制分类器,以识别虚假新闻,解决了现有方法难以高效整合文本与图像信息的问题.借助学习视觉和文本的共同表示,模型能够更精准地判断新闻的真实性. Chen等人^[13]提出的CAFE方法通过计算单模态特征分布之间的KL散度来量化不同模态间的模糊性,并将这种模糊性融入跨模态融合过程中,从而实现更有效的自适应特征聚合. MCNN^[14]不仅融合了文本的语义特征和视觉篡改特征,还考虑了文本与视觉信息之间的相似性. Yang等人^[15]提出的MRAN通过多层次编码网络提取文本特征并通过VGG19提取视觉特征紧接着设计关系感知注意力网络融合多模态信息. Lv等人^[16]提出的MDF-FND利用多头注意力对文本和图

像分别进行不确定性建模,然后基于 Dempster-Shafer 证据理论设计动态融合网络,按样本的置信度自适应地调整两种模态的权重.然而这些方法仅依赖单条数据内容而忽视社交背景(如传播路径)在虚假新闻检测中的作用.

基于上述方法的不足,引入图结构的方法通过社交网络分析增强了多模态特征的表达能力. Zheng 等人^[17]提出的 MFAN 模型在一个统一的框架中融合了文本、视觉和社交图谱等多种特征. Sun 等人^[18]针对

多模态虚假新闻检测中信息交互不足和弱模态优化不平衡的问题提出了 MGIN-AG 框架,该方法利用细粒度图结构与跨模态图注意力机制,并结合动态权重调节策略平衡不同模态的优化.

1.3 主流方法分析

如表 1 所示,本文详细对比了近年来主流的虚假新闻检测方法,包括但不限于本文提出的 MFND-SGLO 模型和引用的 10 个基线模型.对比维度涵盖了模态、图结构、标签建模、损失函数这 4 个指标.

表 1 主流虚假新闻检测方法比较

方法(年份)	模态	图结构	标签建模	损失函数
MVAE ^[12] (2019)	文本、图像	无	二进制分类器	VAE损失
SAFE ^[11] (2020)	文本、图像	无	标准分类	跨模态相似性学习
MCNN ^[14] (2021)	文本、图像	无	标准分类	多模态一致性损失
MCAN ^[9] (2021)	文本、图像	无	标准分类	协同注意力损失
HMCAN ^[10] (2021)	文本、图像	无	标准分类	多模态上下文注意力
CAFE ^[13] (2022)	文本、图像	无	标准分类	KL散度
MFAN ^[17] (2022)	文本、图像、社交图	基础社交图	标准分类	多模态注意力损失
MGIN-AG ^[18] (2023)	文本、图像、社交图	细粒度图注意力	标准分类	自适应梯度优化
MRAN ^[15] (2024)	文本、图像	无	标准分类	关系感知注意力损失
文献[7](2024)	文本	无	标准分类	—
MDF-FND ^[16] (2025)	文本、图像	无	不确定性建模	动态权重自适应损失
MFND-SGLO(Ours)	文本、图像、社交图	全局动态加权的社交图	双向标签注意力	多模态对比交叉熵损失

表 1 从模态、图结构、标签建模与损失函数这 4 个维度对比了主流虚假新闻检测方法.整体而言,早期方法(如 MVAE、SAFE 等)多局限于文本和图像的双模态融合,未引入图结构且标签建模较为简单;后续方法(如 MFAN、MGIN-AG 等)虽开始融入基础或细粒度的社交图结构,但在标签关系建模上仍显不足.相比之下,本文所提的 MFND-SGLO 模型融合了文本、图像与社交图这 3 种模态,采用全局动态加权的图结构,并引入双向标签注意力机制进行标签优化,同时使用多模态对比交叉熵损失,在模型结构与标签建模方面均体现出更全面的设计.

2 方法

2.1 方法概述

本方法使用多模态新闻数据(图像-文本对)及其真实性标签训练虚假新闻检测模型.模型架构如图 1 所示,包含原始数据、特征提取与交互和分类这 3 个模块.原始数据包括文本、图像和社交图,作为模型的输入.特征提取与交互模块利用 CLIP^[19]和 BERT^[20]

提取视觉和文本特征,并通过 BLIP^[21]生成图像描述,随后将该描述输入到 BERT 得到图像描述特征,并将其与视觉特征拼接得到联合视觉特征.然后文本与联合视觉特征经过跨模态交叉注意力得到联合视觉增强的文本特征和文本增强的联合视觉特征.而对输入的社交图进行稀疏性缓解捕捉节点间的潜在连接后,借助 GAT-SE 聚合社交图信息并对其全局动态加权,得到社交图特征.社交图特征会与联合视觉增强的文本特征通过对比学习对齐,得到对齐的文本特征和对齐的社交图特征.二者与文本增强的联合视觉特征再经跨模态交叉注意力生成 6 个相互增强的特征,拼接得到跨模态全局特征.全局标签优化器的输入不仅包括对齐的社交图特征,还包括跨模态全局特征经过 MLP 后得到的初步预测标签.全局标签优化器通过双向标签注意力捕捉社交图特征中相似帖子节点间的正负关系,据此更新新闻帖子的预测标签,最终输出预测结果.

2.1.1 模块间数据流与相互依赖关系

为了清晰地展示模型内部各模块的协同工作机制,

本节将详细阐述模块间的数据流路径与相互依赖关系, 以增强整体模型设计的结构透明度. 模型的数据处理

流程遵循一个从特征提取、特征交互到最终分类优化的逻辑顺序.

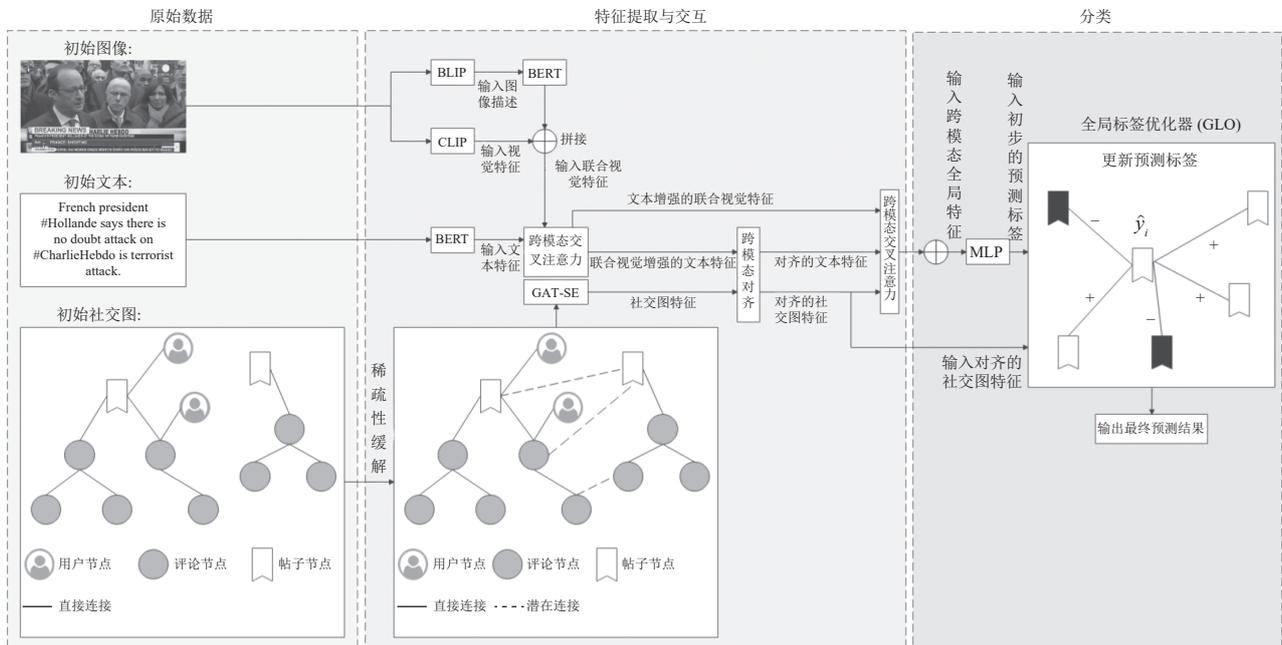


图1 MFND-SGLO 模型整体框架

模型以原始的多模态新闻数据 (文本、图像) 及其社交图作为输入, 数据流首先进入并行的特征提取阶段. 对于文本与视觉模态, 文本内容通过 BERT 模型提取深层语义特征, 图像则由 CLIP 模型提取视觉特征, 同时通过 BLIP 模型生成图像的文本描述, 该描述再经由 BERT 编码为图像描述特征. 这两种特征被拼接为联合视觉特征, 实现了对图像内容的多维度表征. 对于社交图模态, 基于新闻的传播和交互信息构建初始社交图. 为了解决社交图的稀疏性问题, 模型首先进行边的补全, 然后将处理后的社交图输入到增强的图注意力网络 GAT-SE 中. GAT-SE 模块通过聚合社交图的信息并进行全局动态加权, 有效捕捉关键的全局信息并抑制噪声, 最终生成社交图特征.

特征提取完成后, 数据流进入跨模态交互与对齐环节. 首先, 文本特征与联合视觉特征进入跨模态交叉注意力模块, 进行第 1 次深度融合. 此步骤的目的是让两种模态信息相互补充, 生成视觉增强的文本特征和文本增强的联合视觉特征, 这两种特征为后续的模态对齐和最终分类提供了更丰富的语义信息. 随后, 模型利用对比学习机制对联合视觉增强的文本特征与社交图特征进行对齐. 该步骤通过在特征空间中拉近正样

本对并推远负样本对, 实现了不同来源特征在语义上的一致性, 生成了对齐的文本特征与对齐的社交图特征. 接下来, 对齐的文本特征、对齐的社交图特征以及第 1 次融合产生的文本增强的联合视觉特征进行跨模态交叉注意力融合. 此过程生成了 6 个相互增强的特征, 它们被拼接成最终的跨模态全局特征, 该特征是整個模型对新闻内容最全面的表示.

最后, 数据流进入分类与优化阶段. 跨模态全局特征首先通过一个 MLP, 生成一个初步的分类预测标签. 该初步预测标签与对齐的社交图特征一同被送入全局标签优化器. 优化器内部的双向标签注意力机制是关键, 它通过分析社交图中相似新闻节点间的正负相关性, 对初步预测进行更新. 例如, 如果两条内容相似的新闻在社交图上表现出负相关, 该机制会据此调整它们的预测标签, 从而减少因孤立判断导致的错误. 经过标签优化器处理后, 模型输出最终的、经过上下文关系修正的虚假新闻预测结果.

综上所述, 模型的数据流设计体现了从独立特征表征到多层次、跨模态深度融合, 再到结合社交上下文进行全局优化的精密逻辑. GAT-SE 模块为模型提供了高质量的社交图信息, 跨模态对齐确保了不同模态

特征的语义一致性,而双向标签优化则在最终决策阶段引入了全局视角,三者环环相扣,共同提升了虚假新闻检测的准确性。

2.2 特征提取与交互模块

首先,设集合 $X = \{x_1, x_2, \dots, x_n\}$ 为社交平台上的多模态帖子集合。其中, x_n 代表多模态帖子集合中的第 n 条帖子。对于每条多模态帖子 $x_i \in X$, 有 $x_i = \{t_i, p_i, u_i, c_i\}$ 。其中, t_i 代表文本内容, p_i 代表图像内容, u_i 代表发帖的用户, c_i 代表评论集合。对于每一个评论集合 c_i , 有 $c_i = \{c_i^1, c_i^2, \dots, c_i^j\}$ 。其中, c_i^j 代表第 j 条评论。 u_i 代表发布第 i 条帖子 x_i 的用户。本文的目标是训练一个模型 $f: x_i \rightarrow Y$, 其中, $x_i \in X$ 。得到每个帖子的类别概率标签 $\hat{y}_i = (\hat{y}_i^r, \hat{y}_i^f)$, 其中, \hat{y}_i^r 和 \hat{y}_i^f 为该帖子为真或假的概率。

2.2.1 文本特征提取

对于文本模态,使用 BERT 提取帖子 x_i 中的文本 t_i 和评论 c_i 的特征,分别得到帖子 t_i 的特征向量 $b_i \in \mathbb{R}^{512}$ 以及评论特征向量集 $b_i^c = \{b_{i1}^c, b_{i2}^c, \dots, b_{ij}^c\}$ 。

2.2.2 视觉特征提取

对于视觉模态,使用 CLIP 和 BLIP 分别提取图像的视觉特征 $V_i \in \mathbb{R}^{512}$ 和图像描述。然后通过 BERT 处理图像描述得到图像描述特征 $b_i^d \in \mathbb{R}^{512}$, 接下来把图像描述特征与视觉特征进行拼接得到联合视觉特征。将图像描述特征与视觉特征拼接,旨在融合低层次视觉信息和高层次语义描述,形成联合视觉特征并通过全连接层降至 512 维得到 $V_i^d \in \mathbb{R}^{512}$ 。这既增强了视觉特征的表达能力,避免了直接与文本拼接可能导致的语义不一致,也为后续的跨模态交叉注意力融合提供了更全面的图像理解。

2.2.3 社交图特征提取

基于提取的文本特征构建社交图 $G = \{N, A, E\}$ 。其中, E 代表边, N 代表节点集合, $A \in \{0, 1\}^{|N| \times |N|}$ 表示节点间交互关系的邻接矩阵。社交图中包含帖子节点、评论节点和用户节点。用户节点的特征为其发布的所有帖子的文本特征的平均值。社交图中的边通过 3 种方式直接连接: 用户发布帖子、用户评论帖子、用户转发帖子。

由于数据隐私等限制,社交图在构建时面临稀疏问题。所以本文首先对社交图和社交图的邻接矩阵 A 进行稀疏性缓解。为了识别节点间的潜在关联,本文通过计算节点间的相似度度量 β_{ij} 作为两节点相似性的衡量

标准。如果没有直接连接的节点之间的余弦相似度大于 0.6, 则认为两者之间存在潜在连接,从而在两个节点之间构建边,记为 $e_{ij} = 1$, e_{ij} 代表潜在连接。

接下来,我们通过融合原始邻接矩阵 $A \in \{0, 1\}^{|N| \times |N|}$ 和潜在连接,得到新的邻接矩阵 $A' \in \{0, 1\}^{|N| \times |N|}$ 。我们定义 $a_{ij} = 1$ 表示两个节点存在直接连接,而 $a_{ij} = 0$ 则表示没有连接。若原始矩阵中存在直接连接 ($a_{ij} = 1$) 或存在潜在连接 ($e_{ij} = 1$), 则新邻接矩阵中对应位置 a'_{ij} 的取值为 1, 否则取值为 0。

基于此,本文构建出一个融合潜在关系的全新图结构。在进行社交图特征提取时,传统的图注意力网络^[22]虽然能够聚合特征,却过度依赖局部邻居信息,难以捕捉全局特征,容易受到冗余信息干扰。受 Zhao 等人^[23]的启发,本文设计了具有动态全局特征捕捉能力的 GAT-SE, 通过聚合邻居信息并动态调整特征权重,有效强化重要特征并抑制冗余特征。首先,对于节点 n_i , 利用新的邻接矩阵 A' 确定其邻居节点集合 $N_i = \{n_i^1, n_i^2, n_i^3, \dots, n_i^j\}$, 其中, j 表示邻居节点的数量, n_i^j 是第 j 个邻居节点。然后计算该节点与所有邻居节点的注意力权重 $\eta_i = \{\eta_i^1, \eta_i^2, \eta_i^3, \dots, \eta_i^j\}$ 。注意力权重的计算公式如式 (1):

$$\eta_i^j = \text{LeakyReLU}\left(P\left[Wb_i \oplus Wb_j\right]\right) \quad (1)$$

其中, $W \in \mathbb{R}^{300 \times 512}$ 和 $P \in \mathbb{R}^{1 \times 600}$ 是可学习的变换矩阵。 \oplus 代表拼接操作, LeakyReLU 是激活函数。随后我们借助 Softmax 函数对注意力权重完成归一化操作,进而求得各邻居节点的最终注意力系数 δ_i^j , 所有邻居节点的注意力权重之和为 1, 公式如式 (2):

$$\delta_i^j = \text{Softmax}\left(\eta_i^j\right) = \frac{\exp\left(\eta_i^j\right)}{\sum_{\eta_i^k \in \eta_i} \exp\left(\eta_i^k\right)} \quad (2)$$

在获取到归一化的注意力系数 δ_i^j 后,我们即可对邻居节点的特征进行加权求和,从而得到节点 n_i 的初步社交图特征 $G_i \in \mathbb{R}^{512}$, 公式如式 (3):

$$G_i = \delta \left(\sum_{j \in N_i} \eta_i^j T b_j \right) \quad (3)$$

其中, δ 代表 ReLU 激活函数, $T \in \mathbb{R}^{512}$ 是学习到的变换矩阵。接下来对初步社交图特征 G_i 进行全局性的动态加权,得到每个通道的权重大小 $S \in \mathbb{R}^{512}$, 公式如式 (4):

$$S = F_{\text{ex}}(G_i, U) = \sigma(U_2 \delta(U_1 G_i)) \quad (4)$$

其中, $U_1 \in \mathbb{R}^{W_0 \times 512}$ 和 $U_2 \in \mathbb{R}^{512 \times W_0}$ 是全连接层的权重矩阵, W_0 为更小的维度, σ 是 Sigmoid 函数, F_{ex} 代表激励操作. 然后进行缩放操作, 把获得的权重大小 S 分配给初步的社交图特征 G_i , 从而获得最终的社交图特征 $G_{ie} \in \mathbb{R}^{512}$, 公式如式 (5):

$$G_{ie} = S \cdot G_i \quad (5)$$

其中, \cdot 代表按通道相乘.

2.2.4 跨模态融合

本文利用跨模态交叉注意力机制对文本特征、联合视觉特征和社交图特征展开融合, 旨在学习不同模态间的注意力权重, 进而强化多模态特征的信息表达. 具体而言, 首先通过计算文本特征与联合视觉特征的查询矩阵、键矩阵与值矩阵用于增强它们的模态内特征表示. 例如, 对于文本特征 $b_i \in \mathbb{R}^{512}$ 我们通过学习到的线性变换矩阵 $W_Q^b, W_K^b, W_V^b \in \mathbb{R}^{512 \times 512}$ 得到对应的查询、键与值矩阵 $Q_i^b, K_i^b, V_i^b \in \mathbb{R}^{1 \times 512}$.

对于联合视觉特征, 根据学习到的线性变换矩阵 $W_Q^V, W_K^V, W_V^V \in \mathbb{R}^{512 \times 512}$ 得到联合视觉特征 $V_i^V \in \mathbb{R}^{512}$ 的查询、键和值矩阵 $Q_i^V, K_i^V, V_i^V \in \mathbb{R}^{1 \times 512}$. 随后通过式 (6) 计算文本模态的自注意力特征表示为 $F_i^b \in \mathbb{R}^{512}$, 其中, $W_b^o \in \mathbb{R}^{512 \times 512}$ 是线性变换矩阵.

$$F_i^b = \left(\text{Softmax} \left(\frac{Q_i^b K_i^{bT}}{\sqrt{d}} \right) V_i^b \right) W_b^o \quad (6)$$

同理, 随后我们通过 Q_i^V, K_i^V, V_i^V 和线性变换矩阵 $W_b^o \in \mathbb{R}^{512 \times 512}$ 得到联合视觉模态的自注意力特征表示 F_i^V . 接下来我们对 F_i^V 和 F_i^b 进行跨模态交互. 例如, 我们通过 F_i^V 得到查询矩阵 Q_i^V , 通过 F_i^b 得到键矩阵和值矩阵 K_i^b 和 V_i^b . 然后通过式 (7) 计算出联合视觉增强的文本特征 F_i^{Vb} , 其中, $W_{Vb}^o \in \mathbb{R}^{d \times d}$ 是线性变换矩阵:

$$F_i^{Vb} = \left(\text{Softmax} \left(\frac{Q_i^V K_i^{bT}}{\sqrt{d}} \right) V_i^b \right) W_{Vb}^o \quad (7)$$

同理, 随后通过 Q_i^b, K_i^V 和 V_i^V 得到文本增强的联合视觉特征 F_i^{bV} .

2.3 结合跨模态对齐的对比学习

新闻帖子自身的文本特征表示应与其社交图特征表示紧密相连, 本文引入对比损失^[24]用于对齐文本特征与社交图特征. 对比学习旨在使相似样本在特征空

间中更接近, 不相似样本则更远离. 在多模态虚假新闻检测中, 这种机制十分有效. 因为借助对比学习可以更精确地对齐两种模态的表示, 通过缩小模态间的差距增强模型对多模态信息的整合能力. 我们首先利用联合视觉增强的文本特征向量 F_i^{Vb} 和社交图特征向量 G_{ie} 进行增强特征对齐. 设集合 $F^{Vb} = \{F_1^{Vb}, F_2^{Vb}, F_3^{Vb}, \dots, F_N^{Vb}\}$ 、 $G_e = \{G_{1e}, G_{2e}, G_{3e}, \dots, G_{Ne}\}$ 为联合视觉增强的文本特征集合和社交图特征集合, N 是帖子总数. 我们将每个对应的文本-社交图对视为 N 个正样本对, 其余 $N^2 - N$ 个未匹配的视为负样本对. 然后利用特征向量的点积并结合温度参数 τ 进行缩放, 以此计算相似度分数. 通过这种方式, 能够获取文本到社交图的相似度 $p_{b \rightarrow G}$ 和社交图到文本的相似度 $p_{G \rightarrow b}$. 公式如式 (8):

$$\begin{cases} p_{b \rightarrow G} = \frac{\exp(F_i^{Vb} \cdot G_{je} / \tau)}{\sum_{k=1}^N \exp(F_i^{Vb} \cdot G_{ke} / \tau)} \\ p_{G \rightarrow b} = \frac{\exp(G_{ie} \cdot F_j^{Vb} / \tau)}{\sum_{k=1}^N \exp(G_{ie} \cdot F_k^{Vb} / \tau)} \end{cases} \quad (8)$$

对比损失用于评估预测相似度与真实标签间的差异. 针对文本到社交图和社交图到文本两个方向, 我们均使用交叉熵损失进行计算. 具体而言, 文本到社交图的交叉熵损失与社交图到文本的交叉熵损失的公式如式 (9) 所示:

$$\begin{cases} L_{b \rightarrow G} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N y_{ij}^{b \rightarrow G} \log p_{ij}^{b \rightarrow G} \\ L_{G \rightarrow b} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N y_{ij}^{G \rightarrow b} \log p_{ij}^{G \rightarrow b} \end{cases} \quad (9)$$

其中, $y_{ij}^{b \rightarrow G}$ 和 $y_{ij}^{G \rightarrow b}$ 分别代表第 i 和第 j 对是否为正样本对的二值标签: 如果相等则为 1; 如果不相等则为 0. 整体对比损失 L_{con} 取两部分损失之和的平均值用于共同优化文本与社交图的对齐效果, 公式如式 (10):

$$L_{\text{con}} = \frac{L_{b \rightarrow G} + L_{G \rightarrow b}}{2} \quad (10)$$

经过对比学习机制对齐后得到对齐的文本特征 $\hat{b}_i \in \mathbb{R}^{512}$ 和对齐的社交图特征 $\hat{G}_i \in \mathbb{R}^{512}$, 并结合文本增强的联合视觉特征 $F_i^{bV} \in \mathbb{R}^{512}$, 三者两两之间通过跨模态交叉注意力得到联合视觉-对齐文本融合特征 $\hat{F}_i^{Vb} \in \mathbb{R}^{512}$ 、联合视觉-对齐社交图融合特征 $\hat{F}_i^{VG} \in \mathbb{R}^{512}$ 、对

齐文本-联合视觉特征 $\hat{F}_i^{bV} \in \mathbb{R}^{512}$ 、对齐文本-对齐社交图特征 $\hat{F}_i^{bG} \in \mathbb{R}^{512}$ 、对齐社交图-对齐文本特征 $\hat{F}_i^{Gv} \in \mathbb{R}^{512}$ 、对齐社交图-联合视觉特征 $\hat{F}_i^{GV} \in \mathbb{R}^{512}$ 。它们从多角度融合了各模态的特征。将6个增强后的多模态特征向量拼接,作为新闻帖子 x_i 的跨模态全局特征 $F_i \in \mathbb{R}^{3072}$,公式如式(11):

$$F_i = \text{Concat}(\hat{F}_i^{Vb}, \hat{F}_i^{VG}, \hat{F}_i^{bv}, \hat{F}_i^{bG}, \hat{F}_i^{Gv}, \hat{F}_i^{GV}) \quad (11)$$

2.4 结合标签优化器的分类

直接使用跨模态全局特征经过 *Softmax* 归一化得到的预测标签进行预测,忽略了新闻帖子在社交图中与邻居节点之间的细微关系。例如,社交图中相似的帖子节点之间可能存在复杂的正负关联,单纯依赖 *Softmax* 得到的预测标签难以捕捉这些关系,从而影响整体预测的准确性。基于这些不足,我们通过全局标签优化器(global label optimization)全局性地捕捉社交图中相似帖子节点之间的正负关系从而更新新闻帖子的预测标签。具体来说,首先把新闻帖子 x_i 的跨模态全局特征 F_i 输入到 MLP 中映射为二维向量后通过 *Softmax* 函数得到初步类别概率标签 \hat{y}_i 。而全局标签优化器由双向标签注意力组成。通过双向标签注意力捕捉相似帖子节点之间的正负相关性。社交图特征通过揭示新闻帖子在社交网络中的传播模式和相似度关系,能够提供更为精准的全局上下文信息,从而优化标签预测。相比之下,跨模态全局特征虽然融合了多模态信息,但未能深入捕捉新闻帖子之间的文本正负关联。导致模型难以捕捉社交图中特定节点间的复杂关系和传播模式。因此我们通过相似帖子节点之间对齐的社交图特征 \hat{G}_i 和 \hat{G}_j 计算它们之间的双向标签注意力权重,记为 $q_{ij} \in [-1, 1]$,如式(12)所示:

$$q_{ij} = \tanh(B[H\hat{G}_i, H\hat{G}_j]) \quad (12)$$

其中, $B \in \mathbb{R}^{1 \times 2d}$ 和 $H \in \mathbb{R}^{d \times d}$ 为可学习的线性变换矩阵, $[\cdot, \cdot]$ 表示拼接操作。通过变换社交图特征,使其能与相似帖子的社交图特征进行对比,然后通过 *tanh* 函数得到最终的注意力权重。

最后基于双向标签注意力权重计算出最终的标签概率分布公式为:

$$\hat{y}_i^F = \text{Softmax}\left(\hat{y}_i + \sum_{j \in N_i} q_{ij} \times \hat{y}_j\right) \quad (13)$$

其中, \times 代表逐元素相乘, \hat{y}_i^F 代表新闻帖子 x_i 的最终预

测标签。双向注意力能够有效捕捉相似帖子之间的正负相关性,进而优化标签优化。例如,当两条相似帖子 x_i 与 x_j 内容相反时,帖子 x_j 会对 x_i 的真实标签概率产生抑制作用。

本文选用交叉熵损失函数作为 MFND-SGLO 的分类损失函数,通过梯度下降法减少分类损失并指导 MFND-SGLO 的参数更新。公式如式(14):

$$L_{\text{classify}} = -y \log(\hat{y}_i^F) - (1-y) \log(1-\hat{y}_i^F) \quad (14)$$

其中, $y \in \{0, 1\}$ 代表样本的真实标签: 0 代表虚假, 1 代表真实。最终的多模态对比交叉熵损失 L_{total} 如式(15)所示:

$$L_{\text{total}} = L_{\text{con}} + \mu L_{\text{classify}} \quad (15)$$

其中, μ 是平衡超参数。

3 实验分析

3.1 实验数据

PHEME 数据集^[25]收集到的新闻数据来自 Twitter, 涵盖 5 个突发新闻事件, 每个事件都包含一组新闻, 其中包含大量的图像和文本, 并且每一条新闻都有对应的事件和真假标签。Weibo 数据集由 Jin 等人^[26]收集, 源自中国社交媒体平台新浪微博。真实新闻采编自权威的中国新闻机构, 包括但不限于新华社。虚假新闻来源于微博用户发布的内容, 并经过微博官方辟谣系统进行验证证实为虚假新闻。时间跨度为 2012 年 5 月-2016 年 1 月。本文把训练集、测试集和验证集按 7:2:1 的比例进行划分。PHEME 和 Weibo 数据集的具体信息如表 2 所示。

表 2 PHEME 与 Weibo 数据集统计信息(条)

统计信息	PHEME	Weibo
真实新闻	590	1058
虚假新闻	1428	4206
新闻总数	2018	5264
用户节点数	894	985
评论节点数	7388	6963

3.2 实验设置

使用 PyTorch 实现整体模型, GPU 的型号为 RTX 3090, 显存 24 GB, Python 版本为 3.8。模型的相关参数如表 3 所示。

3.3 对比实验

为验证模型的有效性, 选择了 10 种基线方法在

PHEME 和 Weibo 数据集上进行对比. 基线方法分别是 MVAE^[12]、SAFE^[11]、MCAN^[9]、MCNN^[14]、HMCAN^[10]、CAFE^[13]、MFAN^[17]、MGIN-AG^[18]、MRAN^[15]、MDF-FND^[16].

表 3 实验相关参数

模型参数	参数设置
θ	0.6
优化器	AdamW
学习率	0.002
训练次数	80
丢弃率	0.5
μ	2
τ	0.07
batch_size	64

对整体模型训练 5 次, 每次共训练 80 轮, 5 次训练得到的平均值作为最终实验结果. 在训练过程中, 本文采用早停法, 一旦损失函数在连续 10 轮的训练进程中未出现下降趋势, 就立即停止训练. 实验结果如表 4、表 5 所示, 本文表格中加粗均表示最优结果.

表 4 PHEME 数据集实验结果对比 (%)

方法	准确率	精确率	召回率	F1分数
MVAE	77.83	73.82	73.45	72.21
SAFE	81.25	79.92	79.11	79.69
MCAN	80.74	79.21	79.64	80.15
CAFE	77.40	71.80	72.86	72.07
MFAN	87.90	85.66	83.96	85.02
MGIN-AG	87.60	86.80	88.90	87.20
MRAN	87.00	88.90	80.80	87.00
Ours	90.91	90.84	90.90	90.87

表 5 Weibo 数据集实验结果对比

方法	准确率	精确率	召回率	F1分数
MVAE	71.86	70.32	70.32	70.53
SAFE	84.91	83.81	82.19	83.01
MCNN	84.60	80.90	85.70	83.20
HMCAN	85.61	89.07	84.57	86.97
CAFE	84.83	84.00	86.83	84.93
MFAN	87.52	87.31	86.45	86.80
MDF-FND	82.00	85.00	89.00	87.00
Ours	89.15	89.49	89.15	89.46

从表 4 和表 5 可以看出, 本文方法相比基线方法在 PHEME 数据集上准确率提升了 3.01%, F1 分数提升了 3.67%; 在 Weibo 数据集上准确率提升了 1.63%, F1 分数提升了 2.46%. 这表明结合图像描述特征加强视觉特征的表达, 通过对比学习机制实现模态间的有效对齐, 并借助全局动态加权机制强化社交图特征的表达和捕捉相似新闻帖子的正负关系并据此更新预测

标签对模型性能的提升起到了积极作用.

3.4 消融实验

为了证明 MFND-SGLO 中各模块的重要性, 我们进行了 7 种消融实验并评估了不同设置. MFND-SGLO-V (i) 去除视觉模态和图像描述特征, MFND-SGLO-S (ii) 去除社交图特征与跨模态对齐, MFND-SGLO-L (iii) 去除全局标签优化器, MFND-SGLO-T (iv) 去除图像描述特征, MFND-SGLO-D (v) 不使用对比损失函数进行模态对齐, MFND-SGLO-Q (vi) 使用跨模态全局特征计算相似帖子节点之间的正负关系. MFND-SGLO-a (vii) 表示不考虑相似新闻帖子之间的正负关系, 把 q_{ij} 的取值修改为 [0, 1] 之间. viii 表示本文方法. 消融实验的结果如表 6 所示.

表 6 不同数据集消融实验结果 (%)

实验序号	PHEME		Weibo	
	准确率	F1分数	准确率	F1分数
i	83.38	81.30	83.22	82.20
ii	85.97	86.30	85.20	85.40
iii	88.40	88.50	88.02	87.80
iv	89.87	89.71	88.14	88.14
v	89.56	89.20	88.06	88.10
vi	88.73	88.16	88.55	88.33
vii	89.60	89.60	88.57	88.57
viii	90.91	90.87	89.15	89.46

从表 6 可见, 比对 i 与 ii 显示, 去除社交图或视觉模态都导致模型性能下降, 说明社交图和视觉特征的融入增强了模型的多模态检测能力. 比对 iv 与 viii 显示, 图像描述的融入增强了视觉特征的表达, 提升了模型性能. 通过对比 v 与 viii, 不使用对比学习进行模态间对齐使得模型性能出现了下降. 比对 iii 与 viii 发现, 去除标签优化器后在 PHEME 数据集准确率和 F1 分数分别下降了 2.51% 和 2.37%, 在 Weibo 数据集分别下降了 1.01% 和 1.32%, 表明标签优化有效减少了过拟合, 尤其在 PHEME 数据集上效果更显著. 比对 vi、vii 和 viii, 无论是使用相似帖子节点之间跨模态全局特征计算它们之间的双向标签注意力权重, 还是把 q_{ij} 的取值修改为 [0, 1] 之间, 模型的性能都出现了下降. 这说明考虑相似新闻帖子节点之间正负关系的同时使用相似帖子节点之间对齐的社交图特征计算彼此的双向标签注意力权重对模型性能的提升起到了积极作用.

3.5 参数敏感性实验

为了探究模型对学习率和丢弃率的敏感性, 我们在 PHEME 和 Weibo 数据集上进行实验. 首先, 固定其

他参数不变,逐步调整丢弃率进行测试.然后,保持其他参数固定,调整学习率.模型在两个数据集上的准确率和 $F1$ 分数的变化如表 7 和表 8 所示.

表 7 不同数据集丢弃率实验结果 (%)

丢弃率	PHEME		Weibo	
	准确率	$F1$ 分数	准确率	$F1$ 分数
0.1	84.42	84.87	83.15	83.43
0.2	86.40	86.78	85.26	84.88
0.3	87.23	87.41	86.57	85.96
0.4	89.23	89.25	88.15	87.87
0.5	90.91	90.87	89.15	89.46
0.6	89.00	88.00	87.85	88.51
0.7	86.49	86.78	85.66	86.62
0.8	84.57	85.58	83.09	84.29
0.9	80.65	82.80	81.75	80.70

表 8 不同数据集学习率实验结果 (%)

学习率	PHEME		Weibo	
	准确率	$F1$ 分数	准确率	$F1$ 分数
0.0005	83.44	83.46	84.80	84.81
0.001	85.71	84.70	85.76	85.57
0.002	90.91	90.87	89.15	89.46
0.003	88.31	88.26	86.10	85.48
0.004	85.97	85.01	82.71	81.43
0.005	87.79	87.70	83.70	83.70

从表 7 和表 8 中可以看出,丢弃率和学习率显著影响模型性能.丢弃率小于 0.5 时,模型表现较差,可能导致过拟合;增加丢弃率可提升性能,0.5 时最佳,但进一步增加会导致性能下降.学习率过低时,模型学习不充分;0.002 时性能最佳,但过高的学习率会导致震荡,影响收敛和最终性能.

为了探究损失函数中参数 μ 对准确率和 $F1$ 分数的影响,固定其他超参数不变,调整该参数的值,并在 PHEME 和 Weibo 数据集上进行实验.比对结果如表 9 所示.

表 9 不同 μ 值下的准确率与 $F1$ 分数比较 (%)

μ	PHEME		Weibo	
	准确率	$F1$ 分数	准确率	$F1$ 分数
0.5	84.68	85.01	83.59	83.56
1.0	86.75	87.00	85.78	86.61
1.5	90.13	90.20	88.14	88.08
2.0	90.91	90.87	89.15	89.46
2.5	82.33	81.13	85.08	85.23
3.0	88.31	87.79	87.46	87.38
3.5	88.05	87.63	87.05	87.35
4.0	87.83	87.33	86.83	86.33
4.5	86.05	87.03	86.05	86.03
5.0	85.05	86.00	84.05	84.00

由表 9 可见,随着 μ 值增加,模型的准确率和 $F1$ 分数提升,且在 μ 值为 2 时表现最佳.进一步增大 μ 值导致性能下降,说明分类损失函数过度主导会导致过拟合.因此,本文将该参数值设置为 2.

为了探究模型对缓解初始社交图稀疏性的阈值 θ 的敏感性,本文在 PHEME 和 Weibo 数据集上进行了实验.固定其他参数不变,调整阈值,并通过实验选择最佳超参数.比对结果如表 10 所示.

表 10 不同 θ 值下的准确率与 $F1$ 分数比较 (%)

θ	PHEME		Weibo	
	准确率	$F1$ 分数	准确率	$F1$ 分数
0.4	85.97	86.38	80.68	80.87
0.5	88.83	88.75	89.00	88.95
0.6	90.91	90.87	89.15	89.46
0.7	87.27	87.51	87.50	86.50
0.8	89.09	89.14	87.46	87.46
0.9	84.02	84.38	82.71	82.33

从表 10 中可以看出,在控制其他变量不变的情况下,参数 θ 的调整会显著影响模型在数据集上的检测性能.实验表明,当节点相似度阈值设定为 0.6 时执行构建潜在连接任务时,模型能够达到最佳性能.

为了探究训练批次大小对模型性能的影响,本文在 PHEME 和 Weibo 数据集上进行了实验,固定其他参数不变,调整批次大小.模型在两个数据集上的准确率和 $F1$ 分数变化如表 11 所示.

表 11 不同批次大小下的准确率与 $F1$ 分数比较 (%)

批次大小	PHEME		Weibo	
	准确率	$F1$ 分数	准确率	$F1$ 分数
16	83.08	80.91	81.70	80.75
32	85.45	84.60	88.14	88.10
48	90.56	90.69	88.95	88.33
64	90.91	90.87	89.15	89.46
80	88.57	88.94	87.12	86.98
96	85.19	83.83	85.08	84.95
112	85.97	86.36	84.71	84.55
128	85.71	84.48	84.41	83.35

从表 11 中可见,较小的批次大小导致更新噪声大,模型难以收敛;而过大的批次大小则降低更新频率,减慢收敛速度.批次大小为 64 时,模型在充分利用数据的同时,保持了训练效率,达到了理想平衡.

3.6 算法收敛性讨论

为了从理论角度验证 MFND-SGLO 模型的稳定性与可靠性,本节对其算法的收敛性进行分析与讨论.

首先,从损失函数角度看, MFND-SGLO 采用多模

态对比交叉熵损失, 该损失函数由分类损失 L_{classify} 与对比学习损失 L_{con} 两部分组成. 分类损失 L_{classify} 为标准交叉熵损失, 其梯度有界且满足 Lipschitz 连续条件. 对比学习损失 L_{con} 通过对文本与社交图特征进行对齐, 本质上是在特征空间上施加正则化约束, 使得正样本在嵌入空间中更接近, 负样本更远离. 由于 L_{con} 中包含 *Softmax* 归一化的相似度计算, 梯度同样有界. 因此, 整体损失函数在参数空间中是光滑可微的, 且在使用 AdamW 优化器时, 能够保证梯度更新的稳定性.

其次, 从标签优化器的内部机制看, 全局标签优化器通过双向标签注意力机制 (式 (12)) 计算相似帖子节点之间的注意力权重 q_{ij} , 并基于该权重更新预测标签 (式 (13)). 该过程可视为一种迭代更新机制, 类似于消息传播算法. 具体而言, 注意力权重 q_{ij} 经过 *Softmax* 归一化, 即权重归一化约束. 这保证了标签更新过程中不会出现梯度爆炸或梯度消失问题. 进一步地, 由于注意力权重依赖社交图特征的余弦相似度 (用于构建潜在连接), 而余弦相似度的取值范围是有界的, 因此注意力权重的计算是有界的. 在迭代更新过程中, 标签的更新如式 (13) 所示, 这一过程等价于在加权邻居标签上进行平均, 因此具有收缩性. 在实际训练中, 标签优化器在有限步内完成更新, 因此不会出现震荡或发散问题.

第三, 从图注意力网络 GAT-SE 的角度看, GAT-SE 通过注意力机制聚合邻居节点信息, 并通过全局动态加权机制对特征通道进行自适应调整. 注意力权重的计算同样经过归一化, 保证了权重分配的稳定性. 此外, GAT-SE 中的模块通过全连接层学习通道权重, 该过程中激活函数 (*LeakyReLU* 与 *Sigmoid*) 均为标准且常用的函数, 配合权重归一化约束, 共同保证了梯度回传的稳定性, 避免了梯度爆炸. 在训练过程中, GAT-SE 的梯度回传路径经过多个全连接层, 但由于注意力权重的归一化约束, 梯度在反向传播时不会累积放大, 从而保证了训练的稳定性.

最后, 从实验验证角度来看, 本文在 PHEME 与 Weibo 数据集上进行了 80 轮训练, 并采用早停法防止过拟合. 训练过程中, 损失曲线呈现出总体下降并最终收敛的趋势, 且在验证集上未出现震荡或发散现象, 验证了模型的收敛性. 此外, 参数敏感性实验表明, 在合理的学习率 (0.002) 与批次大小 (64) 下, 模型能够稳定

收敛, 进一步证明了算法的鲁棒性.

4 结论与展望

我们提出了一种基于社交图与标签优化的多模态虚假新闻检测模型. 首先, 利用预训练的模型提取文本、视觉与社交图的特征, 通过图像描述增强视觉特征. 为了更有效地学习社交图特征, 我们对其进行全局动态加权. 并通过对比学习机制加强文本特征与社交图特征的跨模态对齐. 随后结合标签优化机制捕捉相似新闻帖子的正负相关性并进行预测标签更新. 最后, 使用多模态对比交叉熵损失训练整个模型. 实验结果表明, 该方法在大多数场景优于基线模型. 然而, 模型也有局限. 首先, 对于外部知识的融合还有所欠缺. 其次, 该模型在应对复杂多模态信息时计算开销较大, 这或许会限制其应用的广泛性. 未来的工作将考虑结合音频、视频等更具创新性的新闻内容形式来检测虚假新闻.

参考文献

- 1 第 55 次《中国互联网络发展状况统计报告》发布. 传媒论坛, 2025, 8(2): 121.
- 2 Melchior C, Oliveira M. A systematic literature review of the motivations to share fake news on social media platforms and how to fight them. *New Media & Society*, 2023, 26(2): 1127–1150.
- 3 Zaheer H, Bashir M. Detecting fake news for COVID-19 using deep learning: A review. *Multimedia Tools and Applications*, 2024, 83(30): 74469–74502. [doi: 10.1007/s11042-024-18564-7]
- 4 王炎, 应龙. 基于 DeepLIFT 算法的可解释多模态假新闻检测. *计算机系统应用*, 2025, 34(3): 161–170. [doi: 10.15888/j.cnki.csa.009823]
- 5 王振琦, 陈涛, 张宝宇, 等. 基于交叉特征感知融合的多领域虚假新闻检测. *计算机系统应用*, 2024, 33(3): 264–272. [doi: 10.15888/j.cnki.csa.009439]
- 6 李铭伟, 陈浩鹏, 李凤环, 等. 基于自监督异质子图注意力网络的虚假新闻检测. *计算机系统应用*, 2025, 34(2): 237–245. [doi: 10.15888/j.cnki.csa.009766]
- 7 Agarwal I, Rana D, Panwala K, *et al.* Analysis of contextual features' granularity for fake news detection. *Multimedia Tools and Applications*, 2024, 83(17): 51835–51851.
- 8 Luvembe AM, Li WM, Li SH, *et al.* Dual emotion based fake news detection: A deep attention-weight update approach. *Information Processing & Management*, 2023,

- 60(4): 103354.
- 9 Wu Y, Zhan PW, Zhang YJ, *et al.* Multimodal fusion with co-attention networks for fake news detection. Proceedings of the 2021 Findings of ACL: ACL-IJCNLP 2021. ACL, 2021. 2560–2569.
- 10 Qian SS, Wang JG, Hu J, *et al.* Hierarchical multi-modal contextual attention network for fake news detection. Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2021. 153–162.
- 11 Zhou XY, Wu JD, Zafarani R. SAFE: Similarity-aware multi-modal fake news detection. Proceedings of the 24th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining. Singapore: Springer, 2020. 354–367.
- 12 Khattar D, Goud JS, Gupta M, *et al.* MVAE: Multimodal variational autoencoder for fake news detection. Proceedings of the 2019 World Wide Web Conference. San Francisco: ACM, 2019. 2915–2921.
- 13 Chen YX, Li DS, Zhang P, *et al.* Cross-modal ambiguity learning for multimodal fake news detection. Proceedings of the 2022 ACM Web Conference. Lyon: ACM, 2022. 2897–2905.
- 14 Xue JX, Wang YB, Tian YC, *et al.* Detecting fake news by exploring the consistency of multimodal data. Information Processing & Management, 2021, 58(5): 102610.
- 15 Yang HY, Zhang JJ, Zhang L, *et al.* MRAN: Multimodal relationship-aware attention network for fake news detection. Computer Standards & Interfaces, 2024, 89: 103822.
- 16 Lv HZ, Yang WZ, Yin YB, *et al.* MDF-FND: A dynamic fusion model for multimodal fake news detection. Knowledge-based Systems, 2025, 317: 113417. [doi: [10.1016/j.knosys.2025.113417](https://doi.org/10.1016/j.knosys.2025.113417)]
- 17 Zheng JQ, Zhang X, Guo SC, *et al.* MFAN: Multi-modal feature-enhanced attention networks for rumor detection. Proceedings of the 31st International Joint Conference on Artificial Intelligence. Vienna: IJCAI, 2022. 2413–2419.
- 18 Sun TN, Qian Z, Li PF, *et al.* Graph interactive network with adaptive gradient for multi-modal rumor detection. Proceedings of the 2023 ACM International Conference on Multimedia Retrieval. Thessaloniki: ACM, 2023. 316–324.
- 19 Radford A, Kim JW, Hallacy C, *et al.* Learning transferable visual models from natural language supervision. Proceedings of the 38th International Conference on Machine Learning. PMLR, 2021. 8748–8763.
- 20 Gardazi NM, Daud A, Malik MK, *et al.* BERT applications in natural language processing: A review. Artificial Intelligence Review, 2025, 58(6): 166. [doi: [10.1007/s10462-025-11162-5](https://doi.org/10.1007/s10462-025-11162-5)]
- 21 Li JN, Li DX, Xiong CM, *et al.* BLIP: Bootstrapping language-image pre-training for unified vision-language understanding and generation. Proceedings of the 39th International Conference on Machine Learning. Baltimore: PMLR, 2022. 12888–12900.
- 22 Vrahatis AG, Lazaros K, Kotsiantis S. Graph attention networks: A comprehensive review of methods and applications. Future Internet, 2024, 16(9): 318. [doi: [10.3390/fi16090318](https://doi.org/10.3390/fi16090318)]
- 23 Zhao JH, Ren RX, Wu Y, *et al.* SEAttention-residual based channel estimation for mmWave massive MIMO systems in IoV scenarios. Digital Communications and Networks, 2025, 11(3): 778–786. [doi: [10.1016/j.dcan.2024.04.005](https://doi.org/10.1016/j.dcan.2024.04.005)]
- 24 Wang LZ, Zhang C, Xu HB, *et al.* Cross-modal contrastive learning for multimodal fake news detection. Proceedings of the 31st ACM International Conference on Multimedia. Ottawa: ACM, 2023. 5696–5704.
- 25 Zubiaga A, Liakata M, Procter R. Exploiting context for rumour detection in social media. Proceedings of the 9th International Conference on Social Informatics. Oxford: Springer, 2017. 109–123.
- 26 Jin ZW, Cao J, Guo H, *et al.* Multimodal fusion with recurrent neural networks for rumor detection on microblogs. Proceedings of the 25th ACM International Conference on Multimedia. Mountain View: ACM, 2017. 795–816.

(校对责编: 李慧鑫)