Factify 2：多模态假新闻与讽刺新闻数据集

S Suryavardan*[1], Shreyash Mishra*[1], Parth Patwa[2], MeghaChakraborty[3], AnkuRani[3], Aishwarya Reganti[4], Aman Chadha†[5,6], Amitava Das[3], Amit Sheth[3], Manoj Chinnakotla[7], Asif Ekbal[8] 和 SrijanKumar[9]

[1]印度斯里城国际信息技术学院[2]
美国加州大学洛杉矶分校[3]美国
南卡罗来纳大学[4]美国卡内基梅
隆大学[5]美国斯坦福大学[6]美国亚
马逊人工智能研究院[7]美国微软[8]
印度帕特纳印度理工学院[9]美国
佐治亚理工学院

摘要

互联网为世界提供了一个开放平台，让人们表达观点、分享故事。尽管这极具价值，但也使假新闻成为当今社会最紧迫的问题之一。人工事实核查过程耗时，因此很难在误导性主张造成重大危害之前将其驳斥。这正是人们对自动事实或主张核验产生兴趣的原因。一些现有数据集旨在支持自动化事实核查技术的发展 [1, 2], 然而，其中大多数都以文本为基础。多模态事实核查受到的关注相对较少。本文提供了一个名为 FACTIFY 2 的多模态事实核查数据集，通过采用新的数据来源并加入讽刺作品文章，对 Factify 1 进行了改进。Factify 2 包含 50,000 条新的数据实例。与 FACTIFY 1.0 类似，我们设置了三个大类——支持、无证据和反驳，并根据视觉数据与文本数据之间的蕴含关系设置子类别。我们还提供了一个基于 BERT 和视觉 Transformer 的基线模型，其在测试集上的 F1 分数达到 65%。基线代码和数据集将在以下网址公开 https://github.com/surya1701/Factify-2.0。

关键词

假新闻，事实核查，多模态，数据集，机器学习，蕴含

1. 引言

随着社交媒体平台成为新闻传播的主要渠道，如何区分事实与假新闻已成为令人担忧的问题。虚假新闻文章通常表现为编造的故事

*共同贡献。

†本文工作与其在 Amazon 的职位无关。*De-Factify 2*：第二届多模态事实核查与仇恨言论检测研讨会，与 *AAAI 2023* 同期举行。*2023*年，美国华盛顿特区✉ suryavardan.s19@iiits.in (S. Suryavardan*);shreyash.m19@iiits.in (S. Mishra*); amitava@mailbox.sc.edu(A. Das)© 2022年本文作者享有本论文版权。允许根据知识共享署名 4.0 国际许可协议（CC BY 4.0）使用。CEUR Workshop Proceedings（CEUR-WS.org）



其中没有可核验的事实、来源或引言。有时，这些故事可能是旨在误导读者的宣传，也可能是为了经济利益而编写的“标题党”内容。在线复制、粘贴、点击和分享内容在技术上轻而易举，这助长了错误信息和虚假信息的传播。这在新冠肺炎-19 [3, 4, 5],选举 [6]等事件中引发了诸多挑战。在某些情况下，故事经过精心设计，旨在激起读者的情绪反应，并被发布在特定网站上，以诱使读者广泛分享。另一些情况下，“假新闻”文章可能由“机器人”生成并传播——这些计算机算法被设计成像人一样分享信息，但能够快速自动地完成这一过程 [7]。尽管已有一些大规模的假新闻识别工作，例如 FEVER [1] 和 LIAR [2], ，但这些数据集并未考虑现实世界中假新闻的演变。社交媒体平台上的假新闻检测还面临另一项阻碍：线上信息非常多样，涵盖大量主题，这增加了此项任务的复杂性。很多时候，仅靠计算机难以核验陈述的真实性和意图，因此相关工作必须依赖人类与技术协作，也就是采用人在回路中的模式 [8]。此外，有助于支持文本主张的视觉线索，可以让系统更有把握地检测虚假内容。上一版 FACTIFY 1 便针对这些问题推出了一个用于多模态事实核查的多模态事实核查数据集。该数据集包含图像、文本主张以及作为参考的文本文档/图像。它提出了一项多模态蕴含任务，使用三个类别（即支持、无证据和反驳）将这些主张与经过核验的文档/图像进行标注；下一节将解释每个类别。前两个类别还进一步细分为文本和多模态组成部分。因此，所有数据样本总共被标注为五种选项之一。数据取自美国和印度这两个大国中热门新闻频道的 Twitter 账号。Factify 2 是 Factify 的最新版本，我们发布了包含 5 万条实例的新数据，其中包括讽刺文章；这些文章采用了不同的假新闻呈现方式。

本文组织如下：第 2 节介绍相关工作。第 3 节介绍所提出的任务。第 4 节阐述数据收集和数据分布，第 5 节展示基线模型。第 6 节展示我们的基线模型结果。最后，第 7 节总结我们的任务，并讨论其进一步发展方向和开放性研究建议。

2. 相关工作

基于文本的数据集：近年来，已有许多用于事实核查和事实核验的文本数据集发布。LIAR [2] 数据集包含来自 Politifact[9] 的 1.3 万条陈述，并标注为 6 个细粒度类别。FEVER 提供了 18.5 万条经人工更新的维基百科主张及其相关支持文档，分为 `Support`、 `Refute`或 `NotEnoughInfo`。Patwa 等人 [10] 发布了一个包含 1 万条关于新冠肺炎的推文/文章的数据集，并将其标注为真或假。 [11]提出了一个用于证据抽取、文档检索、立场检测和主张验证的数据集。 [12] 创建了一个用于区分假新闻与讽刺作品的数据集。PUBHEALTH [13] 数据集包含 1.2 万条公共卫生声明，以及记者撰写的解释，以支持事实核查标签。其他数据集包括 [14, 15, 16, 17]。用于检测基于文本的假新闻的常见方法包括

CNN [18], 、RNN [19, 20], 、BERT [21, 22, 23], 等。

多模态数据集：在社交媒体时代，仅包含文本的数据库已不足以满足需求。要检测假新闻，关键是拓展范围，纳入图像和视频等其他模态。fakeddit [24] 数据集包含一百万个文本+图像样本，并标注为 6 个细粒度类别，用于假新闻检测。FakeNewsNet 提供时空和视觉数据，以及新闻和社交背景信息，以供分析和检测假新闻。它包含来自 Politifact 和 GossipCop 的约 20,000 篇多模态文章所对应的 Twitter 用户数据，例如位置、回复、转推、时间戳等。名为 MOCHEG [25] 的多模态事实核查数据集包含 21,184 条断言，每条断言都附有真实性标签（support、 refute和 not enough information）及解释说明。 [26]提供了一个视频数据集，其中包含 180 个经核实的视频和 200 个已辟谣的视频。其他一些数据集包括 [27, 28, 29]。

针对这项任务的建模方法多种多样且各具特色，涉及分类器、对抗训练、注意力机制等。SpotFake [30] 分别从 BERT 和 VGG 中提取文本与视觉表示，再将二者拼接后进行分类。EANN [31] 通过加入事件判别器，以对抗方式训练假新闻分类器，确保输入数据具有事件不变性，从而也能核验新出现的事件。CARM-N [32] 提出了一种多通道非卷积神经网络，通过同时从原始数据和融合后的文本信息中提取文本特征表示，减轻跨模态注意力融合可能产生的噪声信息的影响。其他方法还包括基于 BERT 的 CapsNet [33], 、跨模态相似度 [34] 和变分自编码器[35] 等 [36, 37, 38, 39]。

Factify 1： FACTIFY [40], 是规模最大的多模态事实核验公开数据集之一，包含 5 万个数据点，涵盖来自印度和美国的新闻。图像、文本和参考文本均属于 FACTIFY 的组成部分。它们分为三大类： Support、Insufficient和 Refute；根据是否包含视觉和文本数据，还可细分出其他类别。FACTIFY 2 沿用类似模式，并新增发布了 5 万个样本，其中包含讽刺文章及来自新数据源的数据。

对于 Factify1，研究人员使用了 BERT [41], RoBERTa [42], 和 BigBird [43] 提取文本特征，并使用 ResNet [43], DeiT [44], EfficientNet [45], 和 VGG [42] 提取视觉特征。所有方法的详情请参阅 [46]。

3. Factify任务

事实核查是一项难以完全自动化的任务，尤其是在多模态数据的情况下。由于需要结合视觉和文本模态进行整体评估，以确定主张的真实性，因此面临着固有的挑战。为此，我们将假新闻核查建模为一项多模态蕴含任务，以核验文本和图像的真实性。

该任务的表述方式与前文介绍的 Factify 1 类似。每个样本都包含一个需要进行核验或事实核查的主张。每个主张都附有一份支持文档，用于通过比较或基于蕴含的方法确定其真实性。主张和文档都是多模态的，即包含文本和视觉数据，从而支持用于事实核查的多模态蕴含。每个样本都包含文本、图像和光学字符识别的配对数据。

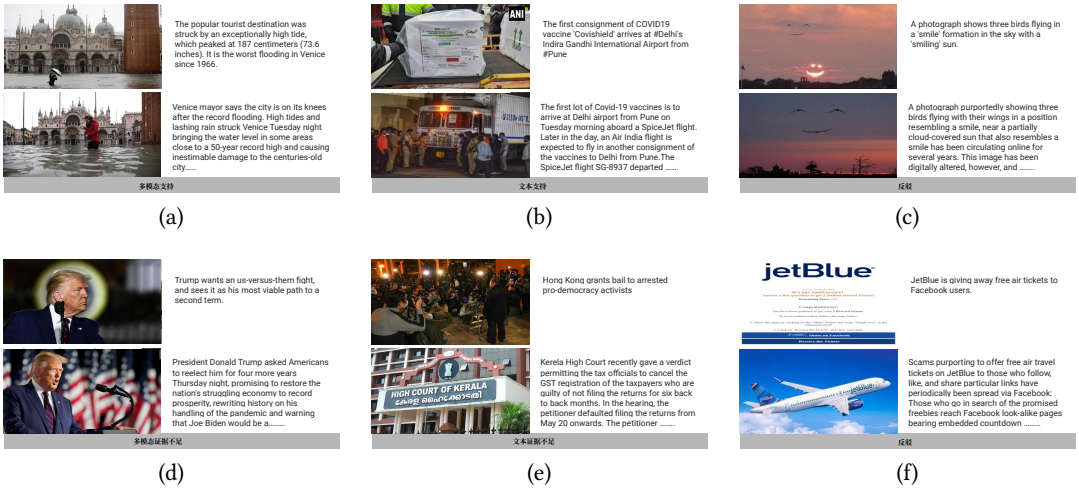


图1： 以下是涵盖全部5个类别的示例。在图像(a)和(b)中，文档文本支持图像中的主张文本；在图像(d)和(e)中，证据不足；而在图像(c)和f)中，文档文本反驳了主张。在图像(a)和(d)中，主张图像与文档图像相互蕴含；在图像(b)和(e)中则不相互蕴含。

我们定义了以下五个类别，用于描述主张与文档之间的蕴含关系：`Support_Text`、`Support_Multimodal`、`Insufficient_Text`、`Insufficient_Multimodal`和`Refute`。这些类别的具体描述如下：

- `Support_Text`：主张和文档的文本数据相互蕴含，但它们的图像不相互蕴含。
- `Support_Multimodal`：文本数据相互蕴含，并且主张和文档的图像也相似。
- `Insufficient_Text`：文本数据不相互蕴含，但主张和文档可能包含若干相同的词语，且图像也不相互蕴含。
- `Insufficient_Multimodal`：主张和文档文本不相互蕴含，但它们可能包含相同的词语，并且在这种情况下，图像也相互蕴含。
- `Refute`：文档中的文本和图像均与主张中的文本和图像相矛盾或对其加以反驳，因此表明给定的主张是错误的。

数据集中的一些示例如图1所示。

4. 数据

本节将介绍数据收集和数据分析。

4.1. 数据收集

数据收集过程包括两个独立的流程：(i) 收集用于支持和无证据类别的真实新闻文章，以及(ii) 收集用于无证据和反驳类别的虚假新闻文章

类别。最终目标是整理出一个数据集，其中包含主张及其对应支持文档的文本和图像。

数据集收集的第一部分与 FACTIFY 1 类似。我们按日期从知名新闻 Twitter 账号收集推文，具体包括《印度斯坦时报》、ANI，以及分别面向印度和美国的 ABC、CNN。这些账号的性质和格式，以及它们发布的推文，有助于我们实现收集真实新闻主张和文章的目标。为提高数据集的多样性和功能性，我们对比了各新闻账号发布的推文，以找出报道相同或相似新闻的推文。为此，我们采取了与 FACTIFY 1 类似的步骤，使用 Sentence BERT 对推文文本进行比较 [47] 并通过设定阈值来判断它们报道的是否为同一新闻。具体而言，我们使用预训练的 paraphrase-MiniLM-L6-v2[48] Sentence BERT（SBERT）变体 [49]，而非 BERT 或 RoBERTa 等替代方案，因为其生成的丰富句子嵌入带来了更优的性能 [47],，同时效率也高得多。如果报道的不是同一新闻，我们便使用 NLTK 库比较共同词语 [50]，将推文归类为相似或不相似。这有助于分别定义支持和无证据类别，如第 3 节所述。我们还利用对比推文对中的图像相似度，根据视觉蕴含关系进一步对数据进行分类。我们根据两项指标设定图像相似度阈值，将图像归类为存在蕴含或不存在蕴含：ResNet50 嵌入之间的余弦相似度和直方图相似度。对于收集到的数据，我们将一个账号发布的推文视为主张，并将另一账号发布的、与推文相关的新闻文章视为支持文档。

第二部分是从多个不同网站收集数据。反驳类别中的部分数据来自事实核查网站，与 FACTIFY 1 类似。我们从 Snopes [51], Factly [52] 和 Boom [53] 抓取了数据。这些网站提供了定义明确的主张，以及能够驳斥该主张的文档。在本次任务迭代中，我们新增了一个数据来源：收集了一些本质上是虚假的讽刺文章，但这些文章的写法让读者觉得内容真实。虽然我们抓取文章的网站（即 Fauxy [54] 和 EmpireNews[55], ）会注明文章内容不实，但我们仍将这些文章归入支持类别。这是因为如前所述，尽管主张本质上是虚假的，这些文章仍支持其主张。为了使主张具有多模态特征，我们通过搜索文章标题抓取图像。对于搜索这些标题得到的结果中，我们还手动标注了部分文章；若文章涉及这些讽刺性主张，则将其加入无证据和反驳类别。

4.2. 数据统计与分析

FACTIFY 的第二个版本与 FACTIFY1 具有相同的类别，共包含 50,000 个数据样本。这些样本平均分布在五个类别中，并分别按 70:15:15 的比例划分为训练集、验证集和测试集。

关键词对于识别或预测给定主张的真实性至关重要。通过分析主张及其文档，我们在图 2中找到了出现频率最高的词语。大多数词语都与政治相关，这表明新闻文章存在偏见。

图 3中的支持和无证据类别词云再次体现了该数据集的政治倾向。然而，在同一图像中，反驳类别的词语分布更加广泛，其中有多个与社交媒体相关的词语出现在反驳类别词云中

	训练集	验证集	Test	总计
Support_Multimodal	7000	1500	1500	10000
Support_Text	7000	1500	1500	10000
Insufficient_Multimodal	7000	1500	1500	10000
Insufficient_Text	7000	1500	1500	10000
Refute	7000	1500	1500	10000
总计	35000	7500	7500	**50000**

Tabl 1
eFACTIFY 2 数据集的分布统计。请注意，数据在各类别之间保持均衡。

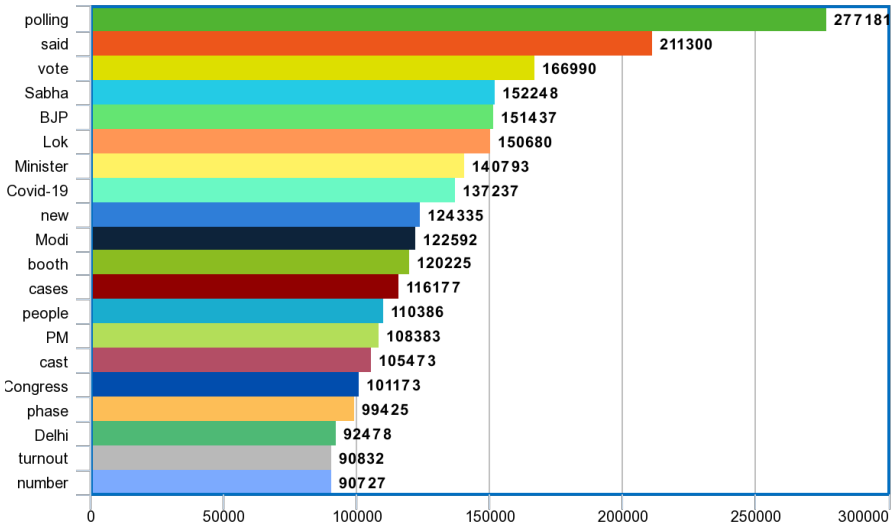


图 2： 从声明文档中提取的出现频率最高的 20 个词及其频率。许多词与政治有关。

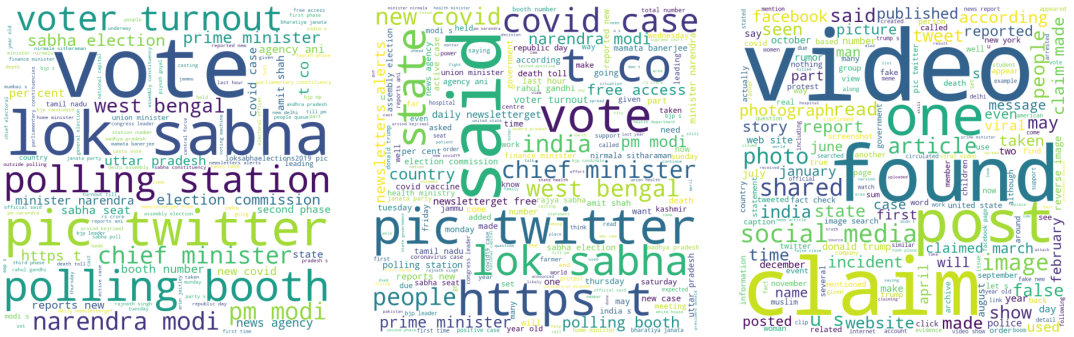
N 元语法	示例
一元语法	（投票），（选举）
二元组	（人民院，选举，投票站）
三元组	(总理,纳伦德拉,印度), (访问,新闻简报)
四元组	(总理,纳伦德拉,莫迪,印度), (阶段,人民院)

表2所有类别主张的 N 元语法示例。

。我们还在表 2 中展示了 FACTIFY 2 数据集独特的 N 元语法示例，以体现该数据集的词汇多样性。

5. 基线模型

在线信息交换中经常使用多种媒介。图片可能会歪曲主张并传播错误信息。我们必须同时考虑



(a) 支持 (b) 证据不足 (c) 反驳

图3： 词云展示了各类别中使用频率最高的词语。与政治和新冠肺炎相关的词语在支持类和证据不足类中占主导地位。

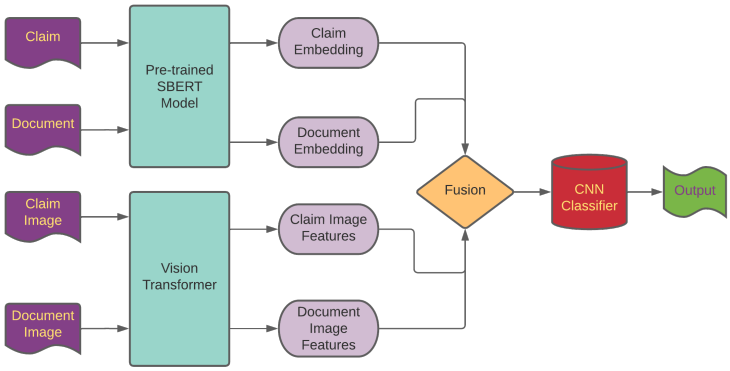


图4： 基线模型架构。从文档和主张中提取的文本和图像特征会被拼接起来，用于最终预测。

图像和文本，才能恰当地对主张进行分类。由于这是一种基于蕴含的技术，因此必须从主张与文档的图文配对中提取特征。视觉特征由预训练视觉Transformer模型（ViT）提取。 [56]得益于 ViT 对图像块进行的位置嵌入，ViT 模型在计算效率和准确性方面均优于传统卷积神经网络。模型使用预训练的 Sentence BERT 模型（具体而言，`stsb-mpnet-base-v2` 变体）生成主张和文档属性的句子嵌入。Sentence-BERT嵌入与 ViT 模型的池化输出拼接。拼接后的特征经过多层感知机处理后进行分类。调整多层感知机后，多模态特征将用于全部三个子任务。模型架构如图4所示。代码将在以下地址公开：https://github.com/surya1701/Factify-2.0。

6. 结果

表 3 中的基线结果展示了下文提及的一些多模态建模方法的宏平均 F1 分数。使用 ViT 提取视觉特征，并使用 Sentence-BERT 提取文本特征时，基线模型的得分为 0.6499。我们还将基线模型（ViT + SBERT- MPNet）与其他方法进行比较，例如 ViT + SBERT-RoBERTa，其中使用 SBERT-RoBERTa 模型代替 SBERT-RoBERTa 来生成文本嵌入。对于 Resnet50 + SBERT-RoBERTa 和 Resnet50 + SBERT-MPNet，使用简单的 ResNet50 模型提取视觉特征。与 ResNet 模型相比，使用视觉 Transformer 带来的性能提升，凸显了图像对该任务的重要性。

方法	宏平均 F1
Resnet50 + SBERT-RoBERTa	0.4504
Resnet50 + SBERT-MPNet	0.4727
ViT + SBERT-RoBERTa	0.6226
ViT + SBERT-MPNet	**0.6499**

表 3测试集上的基线得分。基于 ViT 的模型显著优于基于 ResNet 的模型。

7. 结论与未来工作

通过发布一个包含文本和图像两种模态输入的大型真实世界数据集，我们为本研究中构建多模态事实核查的机器学习方法迈出了重要一步。为凸显这一问题的难度和改进空间，我们进行了数据分析并发布了多模态基线。然而，仍有许多值得探索的研究方向，因为我们的工作仅触及了问题的表面。一种潜在的研究方向是为数据集补充推理内容，说明特定新闻为何是假的。另一种可能是使用符合总体数据分布的合成数据，从而增加反驳类别的复杂性。

参考文献

[1] J. Thorne, A. Vlachos, C. Christodoulopoulos, A. Mittal，FEVER：一个用于事实抽取和核验的大规模数据集，arXiv 预印本 arXiv:1803.05355（2018年）。[2] W. Y. Wang，“骗子，骗子，裤子着火了”：一个用于假新闻检测的新基准数据集，arXiv 预印本 arXiv:1705.00648（2017年）。[3]N. Karimi, J. Gambrell，伊朗数百人因中毒死亡，假新闻声称甲醇可治疗病毒，2020年。网址：https://www.timesofisrael.com/hundreds-die-of-poisoning-in-iran-as-fake-news-suggests-methanol-cure-for-virus/.[4] J. Bae, D. Gandhi, J. Kothari, S. Shankar, J. Bae, P. Patwa, R. Sukumaran, A. Chharia, S. Adhikesaven, S. Rathod, I. Nandutu, S. TV, V. Yu, K. Misra, S. Murali, A. Saxena, K. Jakimowicz, V. Sharma, R. Iyer, A. Mehra, A. Radunsky, P. Katiyar, A. James, J. Dalal, S. Anand, S. Ad- vani, J. Dhaliwal, R. Raskar，新冠肺炎大流行期间公平的疫苗分配所面临的挑战，2022年。 arXiv:2012.12263.[5] M. Morales, R. Barbar, D. Gandhi, S. Landage, J. Bae, A. Vats, J. Kothari, S. Shankar, R. Sukumaran, H. Mathur, K. Misra, A. Saxena, P. Patwa, S. T. V., M. Arseni, S. Ad- vani, K. Jakimowicz, S. Anand, P. Katiyar, A. Mehra, R. Iyer, S. Murali, A. Mahindra, M. Dmitrienko, S. Srivastava, A. Gangavarapu, S. Penrod, V. Sharma, A. Singh, R. Raskar，新冠肺炎检测失控：隐私、有效性、管理不善与误解，2021年。 arXiv:2101.01693.[6] S. Muhammed T, S. K. Mathew，错误信息之祸：社交媒体研究综述，《International Journal of Data Science and Analytics》13（2022年）271–285。[7] S. Desai, H. Mooney, J. A. Oehrli，假新闻、谎言与宣传：如何辨别事实与虚构，主题：新闻与时事。密歇根大学图书馆（2021年）。[8] X. Zhang, A. A. Ghorbani，网络假新闻概述：特征描述、检测与讨论，《Information Processing & Management》57（2020年）102025。[9] Politifact，https://www.politifact.com，2007年。访问日期：2022年。[10] P. Patwa, S. Sharma, S. Pykl, V. Guptha, G. Kumari, M. S. Akhtar, A. Ekbal, A. Das, T. Chakraborty，抗击信息疫情：COVID-19假新闻数据集，载于：紧急情况下打击区域语言网络敌意帖子（CONSTRAINT）2021，Springer，2021年，第21–29页。网址：http://dx.doi.org/10.1007/978-3-030-73696-5_3。doi:10.1007/978-3-030-73696-5_3.[11] A. Hanselowski, C. Stab, C. Schulz, Z. Li, I. Gurevych，用于自动事实核查不同任务的丰富标注语料库，载于：第23届计算自然语言学习会议（CoNLL）论文集，计算语言学协会，中国香港，2019年，第493–503页。网址：https://aclanthology.org/K19-1046。doi:10.18653/v1/K19-1046.[12] J. Golbeck, M. Mauriello, B. Auxier, K. H. Bhanushali, C. Bonk, M. A. Bouzaghrane, C. Bun- tain, R. Chanduka, P. Cheakalos, J. B. Everett, W. Falak, C. Gieringer, J. Graney, K. M. Hoffman, L. Huth, Z. Ma, M. Jha, M. Khan, V. Kori, E. Lewis, G. Mirano, W. T. Mohn IV, S. Mussenden, T. M. Nelson, S. Mcwillie, A. Pant, P. Shetye, R. Shrestha, A. Steinheimer, A. Subramanian, G. Visnansky，假新闻与讽刺作品：数据集与分析，载于：第10届ACM网络科学会议论文集，WebSci ’18，计算机协会，纽约，美国，2018年，第17–21页。网址：https://doi.org/10.1145/3201064.3201100。

doi:10.1145/3201064.3201100.[13] N. Kotonya, F. Toni, 《面向公共卫生声明的可解释自动事实核查》，载于：2020年自然语言处理实证方法会议论文集（EMNLP），计算语言学协会，线上，2020年，第7740–7754页。URL：https://aclanthology.org/2020.emnlp-main.623。doi:10.18653/v1/2020.emnlp-main.623R. Mihalcea, C. Strapparava, 《测谎仪：欺骗性语言自动识别探索》，载于：ACL-IJCNLP 2009会议短论文集，ACLShort '09，计算语言学协会，美国，2009年，第309–312页。[15] A. Kazemi, K. Garimella, D. Gaffney, S. A. Hale, 《超越英语的声明匹配，以扩大全球事实核查规模》，2021年。 arXiv:2106.00853.[16] T. Mitra, E. Gilbert, 《Credbank：附带可信度标注的大规模社交媒体语料库》，载于：ICWSM，2015年。[17] I. Augenstein, C. Lioma, D. Wang, L. Chaves Lima, C. Hansen, C. Hansen, J. Grue Simonsen, 《MultiFC：用于对声明进行基于证据的事实核查的真实多领域数据集》，载于：EMNLP，计算语言学协会，2019年。[18] H. Saleh, A. Alharbi, S. H. Alsamhi, 《OPCNN-Fake：用于假新闻检测的优化卷积神经网络》，IEEE Access 9（2021年）129471–129489。doi:10.1109/ACCESS.2021.3112806。[19] O. Ajao, D. Bhowmik, S. Zargari, 《使用混合CNN和RNN模型在Twitter上识别假新闻》，载于：第9届社交媒体与社会国际会议论文集，SMSociety '18，计算机协会，纽约，美国，2018年，第226–230页。URL：https://doi.org/10.1145/3217804.3217917。doi:10.1145/3217804.3217917.[20] J. A. Nasir, O. S. Khan, I. Varlamis, 《假新闻检测：一种基于混合CNN-RNN的深度学习方法》，《国际信息管理杂志：数据洞察》1（2021年）100007。URL：https://www.sciencedirect.com/science/article/pii/S2667096820300070。doi:https://doi.org/10.1016/j.jjimei.2020.100007.[21] R. K. Kaliyar, A. Goswami, P. Narang, 《FakeBERT：使用基于BERT的深度学习方法检测社交媒体上的假新闻》，Multimedia Tools and Applications 80（2021年）11765–11788。[22] P. Patwa, M. Bhardwaj, V. Guptha, G. Kumari, S. Sharma, S. PYKL, A. Das, A. Ekbal, S. Akhtar, T. Chakraborty, 《CONSTRAINT 2021共享任务概述：检测英语COVID-19假新闻和印地语敌意帖子》，载于：第一届紧急情况下打击区域语言网络敌意帖子的研讨会论文集（CONSTRAINT），Springer，2021年。[23] A. Glazkova, M. Glazkov, T. Trifonov, 《g2tmn在constraint@AAAI2021中的应用：利用CT-BERT和集成学习检测COVID-19假新闻》，载于：紧急情况下打击区域语言网络敌意帖子，施普林格国际出版，2021年，第116–127页。doi:10.1007/978-3-030-73696-5_12.[24] K. Nakamura, S. Levy, W. Y. Wang, 《r/fakeddit：用于细粒度假新闻检测的新型多模态基准数据集》，arXiv预印本 arXiv:1911.03854（2019年）。[25] B. M. Yao, A. Shah, L. Sun, J.-H. Cho, L. Huang, 《端到端多模态事实核查与解释生成：一项具有挑战性的数据集及模型》，arXiv预印本 arXiv:2205.12487（2022年）。

[26] O. Papadopoulou, M. Zampoglou, S. Papadopoulos, Y. Kompatsiaris, 《已辟谣和已核实的用户生成视频语料库》, Online Information Review 43（2019年）72–88。[27] J. C. S. Reis, P. de Freitas Melo, K. Garimella, J. M. Almeida, D. Eckles, F. Benevenuto, 《巴西和印度选举期间在WhatsApp上分享的经事实核查图像数据集》, 2020年。 arXiv:2005.02443.[28] S. Jindal, R. Sood, R. Singh, M. Vatsa, T. Chakraborty, 《Newsbag：用于假新闻检测的多模态基准数据集》, 载于：CEUR Workshop Proc.，第2560卷，2020年，第138–145页。[29] D. Zlatkova, P. Nakov, I. Koychev, 《事实核查遇上摆拍新闻摄影：核实关于图像的声明》, 载于：2019 年自然语言处理实证方法会议和第9届国际自然语言处理联合会议论文集（EMNLP-IJCNLP），计算语言学协会，香港，中国，2019年，第2099–2108页。URL：https://aclanthology.org/D19-1216。doi:10.18653/v1/D19-1216.[30] S. Singhal, R. R. Shah, T. Chakraborty, P. Kumaraguru, S. Satoh, 《Spotfake：用于假新闻检测的多模态框架》, 载于：2019年IEEE第五届多媒体大数据国际会议（BigMM），2019年，第39–47页。doi:10.1109/BigMM.2019.00-44.[31] Y. Wang, F. Ma, Z. Jin, Y. Yuan, G. Xun, K. Jha, L. Su, J. Gao, 《EANN：用于多模态假新闻检测的事件对抗神经网络》, 2018年，第849–857页。doi:10.1145/3219819.3219903。[32] C. Song, N. Ning, Y. Zhang, B. Wu, 《基于跨模态注意力残差和多通道卷积神经网络的多模态假新闻检测模型》, Information Processing & Management 58（2020年）。doi:10.1016/j.ipm.2020.102437.[33] B. Palani, S. Elango, V. Viswanathan K, 《CB-Fake：一种结合胶囊神经网络和BERT的自动假新闻检测多模态深度学习框架》, Multimedia Tools Appl. 81（2022年）5587–5620。URL：https://doi.org/10.1007/s11042-021-11782-3。doi:10.1007/s11042-021-11782-3.[34] X. Zhou, J. Wu, R. Zafarani, 《SAFE：相似性感知多模态虚假新闻检测》, 载于：知识发现与数据挖掘进展：第24届亚太会议，PAKDD 2020，新加坡，2020年5月11–14日，论文集，第二部分，Springer，2020年，第354–367页。[35] D. Khattar, J. Singh, M. Gupta, V. Varma, 《MVAE：用于假新闻检测的多模态变分自编码器》, 2019年，第2915–2921页。doi:10.1145/3308558.3313552.[36] V. Pérez-Rosas, B. Kleinberg, A. Lefevre, R. Mihalcea, 《假新闻自动检测》, 2017年。arXiv:1708.07104.[37] J. Ma, W. Gao, K.-F. Wong, 《使用树结构递归神经网络检测Twitter上的谣言》, 载于：第56届计算语言学协会年会论文集（第1卷：长篇论文），计算语言学协会，墨尔本，澳大利亚，2018年，第1980–1989页。URL：https://aclanthology.org/P18-1184。doi:10.18653/v1/P18-1184.[38] S. R. Sahoo, B. B. Gupta, 《基于多种特征和深度学习的社交网络假新闻自动检测方法》, Applied Soft Computing 100（2021年）106983。[39] Z. Guo, M. Schlichtkrull, A. Vlachos, 《自动事实核查综述》, 计算语言学协会会刊 10（2022年）178–206。[40] S. Mishra, S. Suryavardan, A. Bhaskar, P. Chopra, A. Reganti, P. Patwa, A. Das,

T. Chakraborty, A. Sheth, A. Ekbal 等，《Factify：多模态事实核查数据集》，载于：第一届多模态事实核查与仇恨言论检测研讨会论文集（DE-FACTIFY），2022年。[41] A. Dhankar, O. Zaiane, F. Bolduc，《Uofa-truth 团队参与 Factify 2022：一种多模态事实核查的简易方法》（2022年）。[42] Y. Zhuang, Y. Zhang，《Yet 团队参与 Factify 2022：用于事实核查的单模态与双模态基于 RoBERTa 的模型》，载于：《De-Factify：多模态事实核查与仇恨言论检测研讨会论文集》，CEUR，2022年。[43] J. Gao, H.-F. Hoffmann, S. Oikonomou, D. Kiskovski, A. Bandhakavi，《Logically 团队参与 Factify 2022：多模态事实核查》（2022年）。[44] W.-Y. Wang, W.-C. Peng，《Team yao 团队参与 Factify 2022：利用预训练模型和协同注意力网络进行多模态事实核查》（2022年）。[45] N. Hulke, B. R. Siva, A. Raj, A. A. Saifee，《Tyche 团队参与 Factify 2022：用于多模态事实核查的融合网络》（2021年）。[46] P. Patwa, S. Mishra, S. Suryavardan, A. Bhaskar, P. Chopra, A. Reganti, A. Das, T. Chakraborty, A. Sheth, A. Ekbal 等，《事实核查中的多模态蕴涵基准测试》，载于：《De-Factify：多模态事实核查与仇恨言论检测研讨会论文集》，CEUR，2022年。[47] N. Reimers, I. Gurevych，《Sentence-BERT：使用孪生 BERT 网络生成句子嵌入》，载于：2019年自然语言处理实证方法会议论文集，计算语言学协会，2019年。网址：https: //arxiv.org/abs/1908.10084。[48]paraphrase-minilm-l6-v2, https://huggingface.co/sentence-transformers/ paraphrase-MiniLM-L6-v2, 2019年。访问时间：2022年。[49] N. Reimers, I. Gurevych，《Sentence-BERT：使用孪生 BERT 网络生成句子嵌入》，arXiv 预印本 arXiv:1908.10084（2019年）。[50] S. Bird, E. Klein, E. Loper，《使用 Python 进行自然语言处理：利用自然语言工具包分析文本》，O'Reilly Media, Inc.，2009年。[51] Snopes, https://www.snopes.com/，1994年。访问时间：2022年。[52] Factly, https://factly.in/category/english/，2016年。访问时间：2022年。[53] Boomlive, https://www.boomlive.in/fact-check，2014年。访问时间：2022年。[54] Fauxy, https://thefauxy.com/，2018年。访问时间：2022年。[55] Empirenews, https://empirenews.net/，2014年。访问时间：2022年。[56] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. De-hghani, M. Minderer, G. Heigold, S. Gelly 等，《一张图像胜过 16×16 个词：用于大规模图像识别的 Transformer》，arXiv 预印本 arXiv:2010.11929（2020年）。