·研究论文·

科技文献的多层次结构功能识别

刘昊坦[1,2]　刘家伟[1,2]　张　帆[1,2]　陆　伟[1,2]

(1.武汉大学信息管理学院,武汉,430072;　2.武汉大学信息检索与知识挖掘研究所,武汉,430072)

[摘　要]　实现科技文献结构功能的自动识别有助于提升细粒度信息检索、关键词抽取、引文分析等任务的效率。针对当前结构功能识别研究面临的文本内部依赖关系表达能力较弱、模型泛化迁移能力不足等问题,本研究利用图卷积神经网络捕捉单词节点间存在的固有依赖信息和拓扑结构,提升模型对科技文本建模表达能力,同时,还引入对抗学习思想,提升结构功能识别模型的泛化能力。选取 ScienceDirect 数据集,考察多种模型方法对章节标题、章节内容、章节段落三个不同层次的结构功能的识别效果,并在 PubMed-20k 的医学摘要结构功能数据集上进一步测试多种模型的跨领域迁移能力。研究结果表明,在章节标题层次,BERT+GCN 的识别效果最佳,$F1$ 值达到了 88%,比基线模型提升 3%;在章节内容层次,BERT+GAN 的识别效果最佳,$F1$ 值达到了 76%,比基线模型提升了 3%;在章节段落层次,$F1$ 值达到了 68%。BERT+GCN 的跨领域迁移能力相比其他模型更优,在跨领域数据上取得了 90%的 $F1$ 值。

[关键词]　结构功能　图卷积神经网络　对抗生成网络　科技文献　信息识别

[中图分类号]G350　[文献标识码]A　[文章编号]2095-2171(2024)03-0090-14

DOI:10.13365/j.jirm.2024.03.090

Multi-level Functional Structure Recognition of Scientific Literature

Liu Haotan[1,2]　Liu Jiawei[1,2]　Zhang Fan[1,2]　Lu Wei[1,2]

(1. School of Information Management, Wuhan University, Wuhan, 430072;　2. Information Retrieval and Knowledge Mining Laboratory, Wuhan University, Wuhan, 430072)

[**Abstract**]　The automatic recognition of structure function helps improve the efficiency of tasks such as fine-grained information retrieval, keyword extraction, and citation analysis. In response to the current challenges faced by structure function recognition research, including weak expression of internal textual dependencies and insufficient model generalization and transferability, this paper utilizes graph convolution neural networks to capture inherent dependency information and topological structures among word nodes, enhan-

[基金项目]　本文系国家自然科学基金重点项目"数智赋能的科技信息资源与知识管理理论变革"(72234005)和国家自然科学基金面上项目"基于机器阅读理解的科学命题文本论证逻辑识别"(72174157)的研究成果之一。(This is an outcome of the Key Project "Data and Intelligence Empowered Theoretic Change of Scientific Information Resource and Knowledge Management Theory"(72234005) and the project "Argumentation Logic Recognition of Scientific Proposition Text based on Machine Reading Comprehension"(72174157), both supported by National Natural Science Foundation of China.)

[作者简介]　刘昊坦,博士研究生,研究方向为人机智能交互与协同,信息检索与文本生成等;刘家伟,博士研究生,研究方向为人机智能交互与协同,信息检索等;张帆,副研究员,博士,研究方向为信息检索评价、用户行为分析等;陆伟(通讯作者),教授,博士,研究方向为信息检索,数据智能,创新评价等,Email:00007485@whu.edu.cn。(Liu Haotan, Ph.D. candidate, research on human-research interaction(HCI), information retrieval and text generation; Liu Jiawei, Ph.D. candidate, research on human-research interaction(HCI), information retrieval; Zhang Fan, associate professor, Ph.D., research on information retrieval evaluation, user behavior analysis; Lu Wei(corresponding author), professor, Ph.D. research on information retrieval, data intelligence, innovation evaluation, and so on.)

本文引用格式:刘昊坦,刘家伟,张帆,等.科技文献的多层次结构功能识别[J].信息资源管理学报,2024,14(3):90-103.

cing the modeling and representation capabilities of scientific publications. Additionally, adversarial learning is introduced to improve the generalization ability of the structure-function recognition model. The ScienceDirect dataset is selected to examine the recognition effectiveness of various model approaches for structure function at three different granularities: Header, Section, and Paragraph. Furthermore, we tested the transferability of multiple models across domains on PubMED-20k, a medical abstract structure function recognition dataset. Experimental results demonstrate that BERT+GCN get the best performance at the Header level, with an $F1$ value of 88%, which is a 3% improvement over baseline models. At the Section level, the combination of BERT and GAN achieves the best performance, which is also a 3% improvement over baseline models. At the section paragraph level, the $F1$ score reaches 68%. BERT+GCN exhibits superior cross-domain transferability compared to other models, achieving an $F1$ score of 90% on cross-domain data.

[**Keywords**] Functional Structure; Graph convolution network; Generative adversarial networks; Scientific literature; Information recognition

1 引言

随着科学研究的迅速发展，越来越多的研究者投身于科研工作，这也导致科技文献的数量呈爆炸式增长的趋势。相关数据显示，近几十年间，全球学术论文和专利的数量每年都在快速增加[1]，其中每年出版的论文数量已经超过 250 万篇。然而，随着科技文献数量爆炸式增长，科研人员获取、筛选合适的科技文献信息也变得更加困难，如何高效地获取、筛选科技文献信息已经成为备受科研人员关注的问题。在这样的背景下，科研人员在筛选和接收学术信息时，往往是任务导向，即有目的性地阅读文章的特定部分或章节，例如摘要、方法、结果等。不同的文章结构对不同的研究学者而言，其重要程度也有所差异[2]。

科技文献具有结构规范和格式规范的特点，同时具有逻辑性和层次性。科技文献的结构功能是对写作逻辑结构和章节在结构中所表达的功能的概括。文章内部的章节和段落的设置是作者精心设计和呈现的，对于作者充分展示论文内容、表达论文观点有着独特的功能。研究科技文献中的结构功能和组成要素，对于信息检索、语义理解、关键词功能、引文功能、创新性评价等科技文献的相关研究具有重要作用[3-4]。因此，从科学研究的角度而言，探索大规模高效准确的结构功能自动识别技术具有长远的研究意义和应用价值。结构功能根据学术文本的研究对象和研究粒度可以分为三个不同的研究层次，即章节标题（header）、章节内容（section）和章节段落（paragraph）。章节标题是使用科技文献中每章的小标题来进行结构功能的标签分类；章节内容从科技文献章节全部内容角度出发并进行识别，蕴含更全的文本特征信息；章节段落从每个章节内的不同段落进行结构功能分类。

然而，目前结构功能的识别方法存在两个方面的问题：一方面，文本内部依赖关系表达能力较弱，例如，在文本转化为向量时，忽视了文本内部间语义关系的建模；另一方面，相关数据类别分布不平衡，训练的模型存在过拟合现象，导致模型的泛化迁移能力有限。为了解决上述问题，本文在相关问题中将广泛使用的处理文本分类的方法[5-6]引入到科技文献的分类识别中，具体来说，即通过图卷积神经网络（Graph Convolutional Network，GCN）[7]对文字依赖关系和文本拓扑结构进行建模，提升模型对科技文本的建模表达能力，而对抗生成网络（Generative Adversarial Nets，GAN）[8]可以施加语义扰动，提升模型的泛化能力，从而对科技文献的结构功能进行更准确的识别。此外，考虑到科技文献结构功能识别存在不同层次，本文比较了多种模型对章节标题、章节内容、章节段落三个不同层次的结构功能的识别效果。在 ScienceDirect 数据集和 PubMed-20k 数据集上的实验结果表明，在章节标题层次，BERT+GCN 的识别效果最佳，$F1$ 值达到了 88%，与基线模型相比提升 3%，提升

效果显著；在章节内容层次，BERT＋GAN 的识别效果最佳，$F1$ 值达到了 76%，与基线模型相比提升了 3%；在章节段落层次，单 BERT 模型的识别效果最佳，$F1$ 值达到了 68%，BERT＋GCN 的鲁棒性相比其他模型更优，在跨领域数据上取得了 90%的 $F1$ 值。

2 相关研究

2.1 科技文献结构功能的分类

科技文献是学术科研创造成果的重要展示形式之一，是最主要的科学研究和知识传播的媒介。科技文献结构功能是从文本自身内容和结构角度出发，反映文本章节在文献内容结构中所表达的功能。对科技文献结构功能的识别和分析，不仅有助于读者快速了解一篇文章整体的结构框架，而且解释了结构框架在一篇科技文献中发挥的具体作用，能够方便读者根据自身需求进行快速检索和阅读。

科技文献结构功能的分类是科技文献结构功能识别的前提和基础。国内外对于科技文献结构功能对象分类的研究成果相当丰硕，如表 1 所示。结构功能的分类较早诞生于引文功能的相关研究，Hu 等[9]在引文分布的相关研究中将文章的结构统一划分为引言（introduction）、方法（methodology）、结果（result）、结论（finding and conclusion）四部分，并使用这四部分对科技文献中的引文分布进行了可视化分析。Ding 等[10]将科技文献结构划分为摘要（abstract）、引言、文献综述（literature review）、方法、结果、结论六种类别，并用这 6 种类别分别从参考文献和提及两种引文权重计算方法进行统计和分析。2014 年，陆伟等[11]首次提出了一种科技文献的结构功能研究框架，基于章节标题的识别将科技文献章节的结构功能分为五种，即引言、相关研究、方法、实验（experiment）、结论。2016 年，黄永等[12]在陆伟等[11]研究的基础上，从章节内容出发，将科技文献的结构功能识别研究转化为文本分类问题，引入词汇特征进行文本分类，也是将科技文献分为五类，即引言、相关研究（related research）、方法、实验、总结（outcome）。同年，黄永等[13]结合引文功能和语义理解引入段落位置和段落投票相关概念进行结构功能的标签分类。2019 年，王佳敏等[14]将章节标题、章节内容、章节段落三者结合对比，并采用投票法进行结构功能研究，通过将结构功能分为引言、相关研究（related work）、方法、实验、结论进行相关研究。2020 年，秦成磊等[15]参照国际背景下科技文献的组织往往遵循一种公用的标准模式，即 IMRAD[16]（introduction，materials and methods，results and discussion）的文本结构，将科技文献的结构功能分为引言、方法和材料（methods & materials）、结果和讨论（discussions）四部分。2021 年，刘忠宝等[17]将结构功能引入科技文献的摘要进行相关研究，将摘要文本依据结构功能划分为三类，即目的、方法、结果。本文采用的科技文献结构功能的分类主要是参照了陆伟等[11]、黄永等[12-13]和王佳敏等[14]的相关研究成果。

表 1 科技文献的结构功能的分类方法汇总

Table 1 Summary of Structure Functions of Academic Text

相关文献	标签分类	研究领域
Hu 等[9]	引言、方法、结果、结论	引文功能
Ding 等[10]	摘要、引言、文献综述、方法、结果、结论	引文功能
陆伟等[11]	引言、相关研究、方法、实验、结论	结构功能
黄永等[12]	引言、相关研究、方法、实验、总结	结构功能
黄永等[13]	引言、相关研究、方法、实验、结论	结构功能
王佳敏等[14]	引言、相关研究、方法、实验、结论	结构功能
秦成磊等[15]	引言、方法和材料、结果、讨论	结构功能
刘忠宝等[17]	目的、方法、结果	结构功能

2.2 基于文本分类的科技文献结构功能自动识别方法

科技文献的结构功能是可以概括性表达文章结构的实体，其反映的是文章部分与文章整体的关系，通过文章结构帮助作者进行快捷高效的检索。结构功能的特征提取和分类任务主要采用统计学方法和自然语言处理类的机器学习和深度学习方法，见表 2。1997 年，Yang 等[6]在文本分类的特征提取上评估了包含文档频率（Document Frequency，DF）、信息增益（Information Gain，IG）、指互信息（Mutual Information，MI）、卡方检验（χ^2-test，CHI）、术语强度（Term Strength，TS）五种方法，并使用了 kNN（k-Nearest

Neighbor）分类器进行文本分类和 LLSF（Linear Least Squares Fit）进行回归拟合，进而给出了不同文本分类特征提取方法的评价。2014 年，陆伟等[11]采用自定义词表和条件随机场模型训练使用到的特征，通过 CRF 学习和预测实验样本的多个特征，实现了结构功能的文本分类。2016 年，黄永等[12-13]使用 k-means 方法对相同类型的词汇进行聚类的同时，还计算不同类目下的词汇在不同章节内所占有的比例，并运用机器学习的方法（如支持向量机）进行结构功能的文本分类。2018 年，王东波等[18]将科技文献结构功能的识别研究转化为对以句子为单元的序列标注问题的研究，探索不同特征对结构功能的识别产生的影响，同时试图寻求最优模型。同年，Lu 等[19]提出了一种新的聚类算法用于结构功能的研究，即基于章节标题、章节内容、章节段落进行结构功能的分类识别，并通过实验证明，将计算机领域的科技文献转化为固定的结构模式进行研究，会更有利于提升科技文献的结构功能分类识别的效果。近 5 年，随着深度学习的进步和飞速发展，结构功能的相关研究也进入了新时期，新的研究成果不断涌现。2019 年，王佳敏等[14]通过融合标题功能、章节功能、段落功能三种不同层次的结构功能，结合卷积神经网络（Convolutional Neural Network，CNN）、长短时记忆网络（Long Short Term Memory，LSTM）、卷积神经网络＋长短时记忆网络（CNN＋LSTM，CLSTM）等深度学习方法进行科技文献结构功能的识别，采用了集成学习中的投票法融合不同层次和不同模型的识别结果，对比了新的深度学习模型同传统机器学习模型 SVM 进行文本分类的效果。但这些研究仍存在使用模型单一的问题，未能充分发挥深度学习模型在科技文献结构功能识别，尤其是章节功能识别中的优点。2020 年，秦成磊等[15]不仅对比分析在文本分类任务中表现较佳的深度学习模型-BERT 同传统机器学习模型在科技文献结构功能的识别结果来尝试得到最佳效果模型，而且引入注意力机制构建了能够捕获结构功能信息的多粒度的层次间注意力网络模型，并对其相关领域的适应性进行分析，展示了结构功能在科技文献中强大的表现力。2022 年，Ma 等[20]在结构功能任务中将上下文信息作为相对位置特征来优化深度学习模型，取得了很好的效果，极大地推动了结构功能研究进程。2023 年，毛进等[21]使用学术文本摘要中的序列信息，从摘要中的单句和文本出发，比较分析了多种主动学习模型的分类识别效果，将主动学习引入到结构功能的分类任务当中。在现有的规则学习和深度学习方法及研究基础上，本文对科技文献的结构功能识别和分类进行了更深层次的探索，通过引入 GCN 和 GAN 组成新的结构功能识别神经网络，并在不同层次的结构功能上进行分类效果统计分析，同时进行纠错实验和迁移实验，进一步探索文本分类效果提升的空间，证明模型的普适性。

表 2　科技文献结构功能的研究方法发展史

Table 2　Developing Research Methods for the Structure Functions of Academic Texts

发展阶段	代表文献	年份
统计学和机器学习	Yang 等[6]	1997
	陆伟等[11]	2014
	黄永等[12]	2016
	黄永等[13]	2016
	王东波等[18]	2018
	Lu 等[19]	2018
深度学习	王佳敏等[14]	2019
	秦成磊等[15]	2020
	Ma 等[20]	2022
主动学习	毛进等[21]	2023

3　科技文献的多层次结构功能识别任务与方法

3.1　任务定义

本文主要是针对科技文献正文内容（不包含标题、摘要、关键词、目录、参考文献等相关内容）来进行结构功能识别的相关研究，结合陆伟等[11,19]、黄永等[12-13]和王佳敏等[14]的相关研究成果将科技文献的正文内容依照结构功能划分为“引言（introduction）”“文献综述/相关研究（related work、literature review、background）”“方法（methods、methodology、

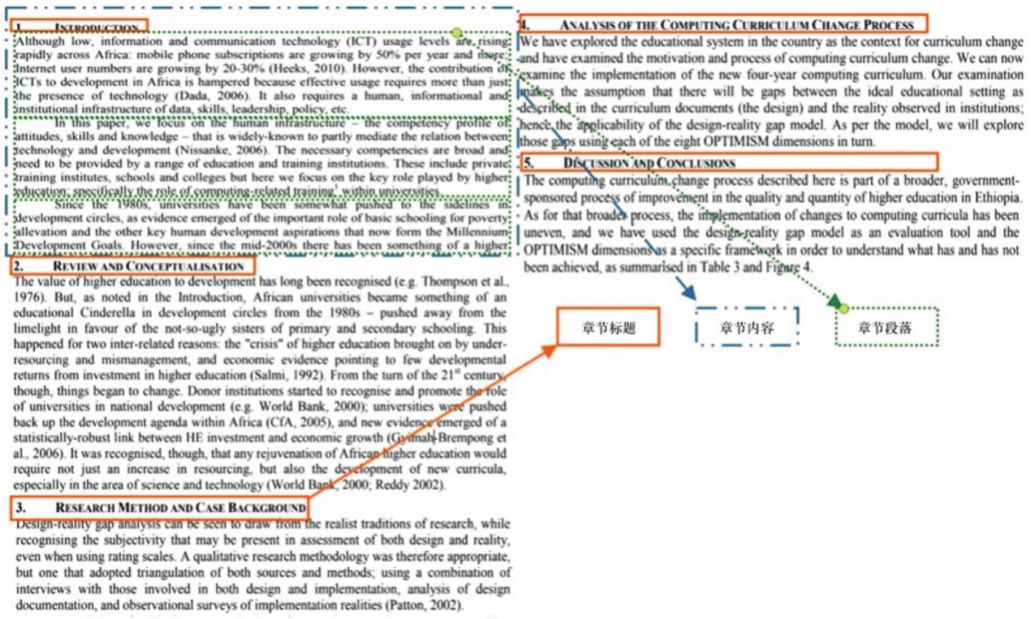

图 1　科技文献的多层次结构功能示例

Fig. 1　An Academic Texts Case of Structure Functions

model）”“实验（experiment、results、data and analysis）”“结论（conclusions、discussions、conclusions and discussions、outcomes）”五个部分。

科技文献的正文部分由多个章节组成，每一个章节都包含章节标题、章节内容以及章节内容中的不同章节段落，如图 1 所示。结构功能识别任务就是对科技文献中的章节按照所属功能进行标注，其中，研究可以根据不同实验对象和不同粒度文本进行科技文献的结构功能分类问题探索。 基于这一准则，结构功能的自动识别可划分为三个不同的层次：第一，基于章节标题的结构功能识别研究，即根据科技文献不同章节标题进行结构功能分类识别；第二，基于章节内容所属层次的结构功能识别研究，即根据科技文献不同章节全部所属内容进行结构功能的分类识别；第三，基于章节段落的结构功能识别研究，即对不同章节的所有不同段落进行结构功能的分类识别，其段落所属章节的标签定义即为该段落所拥有的结构功能。

本文以多种深度学习模型为实验基础，融合了科技文献章节标题、内容和段落三个层次对科技文献结构功能进行分类识别研究，具体的研究框架如图 2 所示，主要分为两大模块，前者是基于科技文献结构功能识别分类效果的研究，后者是对前者研究成果的纠错分析以及对文中提出的不同层级的最佳模型在生物学领域进行迁移实验的分析研

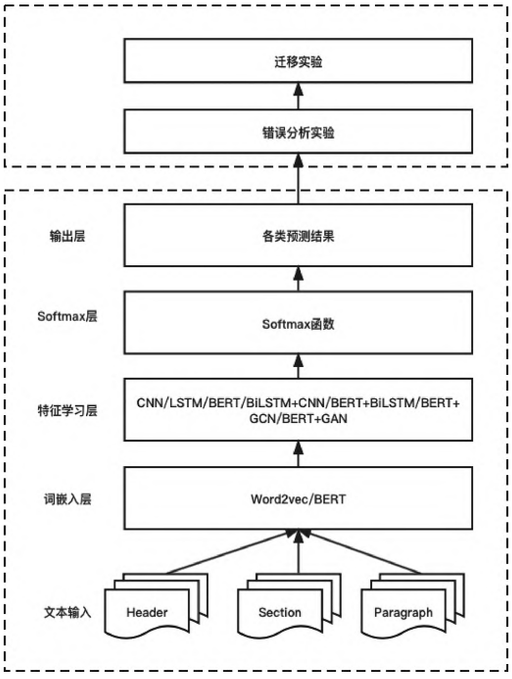

图 2　整体研究框架

Fig. 2　The Research Framework of Structure Functions

究，前后两者是递进和延展的关系。科技文献的结构功能研究多数是针对计算机类的科技文献，本文选择生物医学领域的学术文本作为迁移实验的对象，其论文格式大多采用国际统一的 IMRAD 结构，且秦成磊等[15]也通过实例论证了生物医学类论文用于结构功能任务研究的可行性。相比单一层次的识别研究，混合多层次研究不单单局限于某一学科或是某一模型，而是多层次多学科多模型，因此识别效果更好。

结构功能的分类模块共有五个部分：输入、词嵌入、特征学习、Softmax 层和最终输出。输入由章节标题、章节内容和章节段落构成的含标签的训练集和测试集部分，并统一将待分类文本分为同等长度。词嵌入将输入的内容转变成向量的形式表示，本研究采用 BERT 和 Word2vec 两种方式来进行词向量的生成。具体来说，CNN、LSTM、BiLSTM＋CNN 采用 Word2vec，BERT、BERT ＋ BiLSTM、BERT ＋ GCN、BERT ＋ GAN 采用 BERT。在 Word2vec 词嵌入中，每个词都用一个 K 维的实向量表示。它的优势是可以通过欧式距离、余弦相似度等方法进行多词间的相似度计算，而且较好地解决了稀疏性问题。相较于传统的语言模型，BERT 词嵌入的优势是使用了双向的 transformer 结构。作为一个预训练模型，它可以在预训练时，使用遮蔽语言模型（Masked Language Model，MLM），以及结合上下文内容，使用下一句预测（Next Sentence Prediction，NSP）进行两个任务的联合训练[22]。MLM 任务指的是一句话中随机遮盖 15％的词汇，让模型根据剩下的 85％词汇推测被抹掉的部分。NSP 指的是判断两个句子在实际预训练过程中是否是连续的，其本质是一个二分类的问题。采用 BERT 做词嵌入，可以使模型输出的词向量表示都尽可能准确、全面表达文本输入的整体信息。特征学习层是整个结构功能中最重要的部分，本文分别采用单模型和模型组合对比分析的方法，对已知类别的词向量表示进行特征学习，训练模型以进行结构功能的文本分类。激活函数选用 Softmax 函数，对特征学习层传递来的相关信息，计算出分类数据所属类别的概率。Softmax 函数是在深度学习文本分类任务中经常使用的一个归一化函数。不同类别的测试集的预测结果和每条输入数据的各个归属类别将在输出层进行输出统计。

3.2 核心模型构建

本文尝试将新的神经网络引入结构功能任务，提出 BERT＋GCN 和 BERT＋GAN 两个模型用于多层次结构功能识别任务。本节将分别进行介绍。

3.2.1 BERT＋GCN 模型

图神经网络（Graph Neural Network，GNN）是一种有效的直推式学习方法[23]，通过将原始文本的数据转换为图数据的模式，进而捕捉单词节点间存在的固有依赖信息和拓扑结构，帮助提升文本分类效果[24]。在一般的文本分类任务中，模型先构建了一个图形来模拟文本间各要素的关系。图形中的节点（nodes）代表了单词（word）和文本（documents）元素，而图形中边（edges）的构造是建立在这些节点之间的语义相似性上[25]。然后，将 GNN 应用到图中进行节点分类。GNN 等直推式学习方法的好处是：一方面，实例在训练集和测试集上的测试不仅仅依靠自己，还会通过邻居节点作出判断，增强了模型对异常数据值的免疫能力；另一方面，在训练时，由于模型通过图形的边将监督标签的影响传播到训练和测试上的实例中，因此，未标记的数据也有助于表征学习，从而提高模型整体的性能。现阶段，大多数 GNNs 都有着这样一种通用的架构，它们的滤波器参数共享在图中的任意位置，这样的 GNNs 模型称作 GCN。现有将 BERT 和 GCN 结合起来的研究工作，使用图来模拟单个文档样本中标记之间的关系，这属于归纳学习的范畴。与这些工作不同的是，本文参照 Yao 等[26]的方法搭建 BERT＋GCN 模型，使用图来模拟整个语料库中不同样本之间的关系，利用已标注和未标注文档之间的相似性，并使用 GCN 来学习它们的关系。

在 BERT＋GCN 模型中，本文使用 BERT 模型初始化文本图中文档节点的表示，这些

表示被用作 GCN 的输入。 然后，文档表示将使用 GCN 的图结构进行迭代，其输出被视为文档节点的最终表示，输入到 Softmax 层进行预测。 在构造文本图网络中，有两种节点类型，即文本和单词（这里指文本中不重复的单词）；有两种类型的构图边，即文本－单词和单词－单词，其中的各边都带有一定的权重 A_{ij}，并通过以下方式进行定义。

$$A_{ij}=\begin{cases}\mathrm{PMI} & (i,j)\ \text{其中}\ i,j\ \text{为单词，}\mathrm{PMI}(i,j)>0\\ \mathrm{TF-IDF}_{ij} & \text{其中}\ i\ \text{为文本，}j\ \text{为单词}\\ 1 & i=j\\ 0 & \text{其他}\end{cases} \tag{1}$$

点互信息（Point-wise Mutual Information，PMI）主要可以用来衡量两个词之间的关联度。 计算方法主要是通过公式 2、公式 3、公式 4 来计算单词间的权重。 其中，W_i 是所有滑动窗口中包含单词 i 的窗口个数；$W(i,j)$ 是指包含单词 i 与单词 j 的窗口个数，W 是总的滑动窗口个数。

$$\mathrm{PMI}(i,j)=\log\frac{p(i,j)}{p(i)p(j)} \tag{2}$$

$$p(i,j)=\frac{\#W(i,j)}{\#W} \tag{3}$$

$$p(i)=\frac{\#W(i)}{\#W} \tag{4}$$

在融合 BERT 与 GCN 训练时，由于 BERT 做词向量转化后输送到 GCN 里，进行联合训练时梯度同步回传，BERT 部分会产生梯度优化的问题。 针对这一问题，本文引入公式 5 和公式 6 采用插值更新。 融合分类概率是指 BERT 单独作用与文本得到的文本嵌入与 GCN 部分得到的文本嵌入相加，然后采用交叉熵损失函数进行一个分类预测，λ 控制着这两个目标之间的权衡。 λ ＝1 意味着使用完整的 BERT＋GCN 模型，λ ＝0 意味着只使用 BERT 模型。 当 λ ∈(0,1)时，能够平衡两个模型的预测，BERT＋GCN 模型可以得到更好的优化，本文 λ 值设为 0.7。

$$Z_{\mathrm{BERT}}=\mathrm{softmax}(WX) \tag{5}$$

$$Z=\lambda Z_{\mathrm{GCN}}+(1-\lambda)Z_{\mathrm{BERT}} \tag{6}$$

3.2.2 BERT＋GAN 模型

对抗生成网络主要由生成器（generator）和判别器（discriminator）两个部分组成（见图 3）。 通常情况下，这两部分模型使用神经网络来实现，但也可以采用映射的方法，将任意数据从一个空间映射到另一个空间的可区分系统来实现[27]。 生成器试图捕获实例的分布进而生成新的数据样本，而判别器本质是一个二进制的分类器，目的是尽可能准确地鉴别生成样本和真实样本。 GAN 的优化遵循极大极小全局优化的规则，参照公式 7，这种优化终止于一个鞍点，在这个鞍点上形成了一个相对于生成器的最小值和相对于判别器的最大值。 换言之，GAN 的优化目标是达到纳什均衡[28]。 因此，从这个角度看，生成器可以被认为是精准地捕获了真实样本的分布。 对抗生成网络在自然语言处理领域已经有广泛的研究与应用，例如信息检索、文本分类、文本生成[29]等任务。 在 BERT＋GAN 中，BERT 提供了基于上下文的词嵌入组成的句子，嵌入的同时，还捕捉了句子级的语义。BERT 可以对一个句子的单词、整个句子以及专用句组进行编码，可以解决文本分类、序列标记的问题。 通过引入对抗训练实现对 BERT 的微调（Fine-tune）来实现 BERT 扩展，进行更好的文本分类工作。 在 BERT 的基础上，增加 GAN 架构[29-30]，引入了一个具有对抗性作用的生成器 G 和一个用于判别实例样本的判定器 D。 G 和 D 均为多层感知器（Multi Layer Perceptron，MLP），D 的最后一层是一个 Softmax 激活层。 神经网络需要记忆一个原始数据分布和扰动数据分布，通过混合梯度的方法，增加了训练时间，试图拟合原始样本和有扰动的样本。 通过将 BERT 和对抗训练相结合，借鉴了强化学习的思路，通过最大化扰动的同时最小化对抗期望风险，提升模型自身的结构功能识别效果和泛化性。

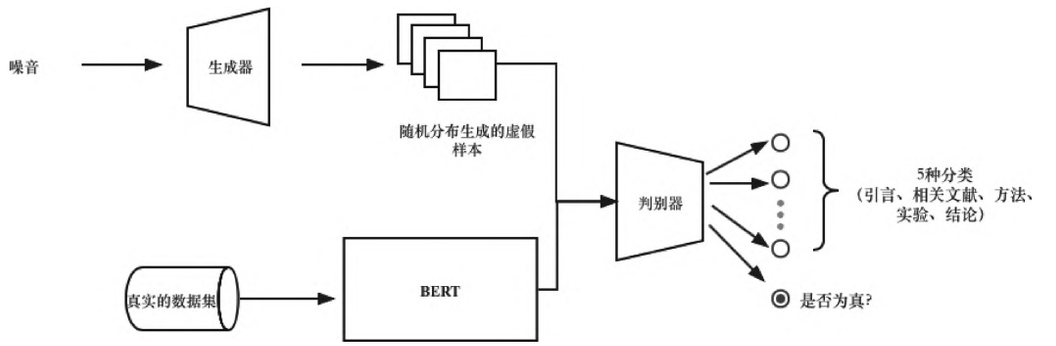

图 3　BERT＋GAN 神经网络结构图

Fig. 3　The Neural Network of BERT＋GAN

$$\arg_{\theta}\min E_{(x,y)\sim D}\left[\max_{\delta\in S}L(\theta,x+\delta,y)\right] \quad (7)$$

4　实验设置与结果分析

4.1　实验数据集

本文的实验数据选自 ScienceDirect 数据库中 2000—2022 年的计算机科学领域期刊文献，从中随机选取 4000 篇科技文献作为本次研究的数据集，包含不同种类的学术期刊 100 多种，按照 8∶1∶1 的比例随机抽取数据集中的 3200 篇作为科技文献的训练集，400 篇作为科技文献的测试集，400 篇作为科技文献的验证集。 每一篇科技文献均包含三个层次的文本：章节标题、章节内容、章节段落。其中包含章节标题 23275 条、章节内容 23275 条、章节段落 193344 条。

4.2　基线模型

本文分别采用独立的深度学习模型 CNN、LSTM、BERT 以及传统的具有代表性神经网络组合模型 BiLSTM＋CNN、BERT＋BiLSTM 作为基线模型，与提出的迁移图神经网络和对抗神经网络的模型组合 BERT＋GCN、BERT＋GAN 进行科技文献多层次结构功能识别任务对比，试图在不同层次找到最好的模型或模型组合。

（1）CNN

CNN 由输入、卷积、池化和输出四层结构构成。 输入层设置为特征矩阵 $S=n$（输入文本的长度）$\times d$（词向量纬度），其中 word embedding 的值设为 256，对于长度不足 n 的文本采用补零处理，采用 ReLU 函数作为每个神经元的激活函数，通过最大池化操作卷积层提取特征向量，试图找寻影响文本分类结果最大的因素。 同时，对于全连接层而言，池化层不仅固定了神经元的个数，而且控制了输出特征的长度。 最后，通过全联接的方式将所有得到的局部最优特征传输到输出层的输出节点，经由 Softmax 函数输出科技文献文本分类的判定结果。

（2）LSTM

LSTM 作为 RNN 的一种变式，其本质是使用了一个由记忆细胞、输入门、遗忘门和输出门组成的记忆单元替代了原本 RNN 中的隐藏单元。 记忆细胞通过状态参数记录信息，并通过相互交互的门单元控制记忆信息值的修改和传递，输入门和输出门分别对参数的输入和输出进行取舍，而遗忘门用来设置选择性遗忘的权重。 这样做的优势是可以进行选择性记忆和遗忘信息，从一定程度上避免了梯度消失的问题，同时也可以学习到长周期信息[31]。

（3）BiLSTM＋CNN[32]

BiLSTM 由前后双向的 LSTM 构成，在捕捉双向的语义依赖上表现较好。 BiLSTM＋CNN 的做法，本质上讲，是将 BiLSTM 的输出层和 CNN 的输入层相结合，将前者隐藏层的数值同后者池化结果较好地结合起来，最终通过全联接方式在输出层进行文本类别的输出工作。 CNN 模型结合 LSTM 模型具有较好地获取词汇间序列关系的优势，可以回避 LSTM 作为一个“有偏”模型，其词语次序和位置产生的负面影响。

（4）BERT＋BiLSTM[33]

BERT＋BiLSTM 模型的搭建，是在 BERT 后，结合 LSTM 层，进行文本信息的整合和句子顺序特征的提取，进而更细粒度地获取语义特征，使语义表达更加完整和准确，最后通过全连接层，经过 Softmax 函数进行分类输出。

4.3 评价指标

本文的评价指标选用准确率 P 值（$precision$）、召回率 R 值（$recall$）以及它们的调和平均值 $F1$ 对不同模型的分类识别结果进行评价，各指标的计算公式如下：

$$P = \text{科技文献中识别正确的结构功能数} / \text{科技文献中识别的结构功能总数} \quad (8)$$

$$R = \text{科技文献中识别正确的结构功能数} / \text{科技文献中实际的结构功能数} \quad (9)$$

$$F1 = 2 * P * R / (P + R) \quad (10)$$

整体准确率 P、召回率 R 和 $F1$ 值的加权算平均数值，作为衡量各模型整体性能的评价指标，其中 $F1$ 值作为选取同种分类任务中最佳模型的主要参考标准。

4.4 多层次结构功能识别效果

本研究分别采用 BERT、BiLSTM＋BERT、BERT＋GCN、BERT＋GAN 四种神经网络模型在开源的 Keras 计算框架上，以及采用 LSTM、CNN、BiLSTM＋CNN 三种神经网络模型在开源的 TensorFlow 计算框架上，对章节标题、章节内容、章节段落三个不同层次的科技文献数据进行结构功能的文本分类实验；并采用迭代学习的方式对各自的训练数据集学习 200 轮；采用验证集结果对模型进行调整，以寻求最优参数，通过不断调整相关超参数训练神经网络模型，直至得到最积极的实验结果。

通过训练好的模型分别对章节标题、章节内容、章节段落不同层次的测试集（test）进行分类考试，通过统计不同层次对应的不同类别和整体上的准确率、召回率和 $F1$ 值进行模型研究和分析，试图发现一定的规律，寻找到在各个任务中表现较佳的模型。不同层次的实验结果分别见表 3 至表 5。

从表 3 中可以看出，章节标题层次的普遍识别率均在 84％以上，其中 CNN、BiLSTM＋CNN、BERT＋BiLSTM、BERT＋GNN、BERT＋GAN 的识别效果均超过了 85％，其中效果最好的为 BERT＋GCN，其准确率为 91％、召回值为 86％、$F1$ 值达到了 88％。在不同的结构功能分类识别结果中，“introduction”的辨别效果最好，在七种不同的神经网络或神经网络组合中的 P 值、R 值和 $F1$ 值均取得了最好的成绩，在七种不同的神经网络或神经网络组合中的准确率、召回率和F1均取得了

表 3 章节标题层次在不同神经网络(组合)上的实验结果

Table 3 Results of Header on Different Neural Networks (Combinations)

章节标题 Header	CNN			LSTM			BERT			BiLSTM＋CNN			BERT＋BiLSTM			**BERT＋GCN**			BERT＋GAN		
	P	R	$F1$	P	R	$F1$	P	R	$F1$	P	R	$F1$	P	R	$F1$	P	R	$F1$	P	R	$F1$
introduction	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
literature review	0.90	0.63	0.74	0.84	0.62	0.71	0.69	0.66	0.67	0.87	0.61	0.72	0.96	0.62	0.76	0.97	0.70	0.81	0.78	0.65	0.71
methodology	0.73	0.86	0.79	0.69	0.88	0.77	0.72	0.73	0.73	0.70	0.84	0.76	0.77	0.93	0.81	0.79	0.90	0.83	0.74	0.82	0.78
result	0.74	0.78	0.79	0.83	0.72	0.77	0.75	0.79	0.77	0.78	0.75	0.77	0.86	0.83	0.85	0.87	0.79	0.83	0.81	0.81	0.81
conclusion	0.98	0.91	0.94	0.98	0.91	0.94	0.98	0.91	0.94	0.98	0.91	0.94	0.98	0.92	0.95	0.99	0.94	0.96	0.98	0.91	0.94
all(weighted)	0.86	0.85	0.85	0.85	0.84	0.84	0.86	0.83	0.84	0.85	0.84	0.84	0.87	0.87	0.87	**0.91**	**0.86**	**0.88**	0.85	0.85	0.85

表 4 章节内容层次在不同神经网络(组合)上的实验结果

Table 4 Results of Section on Different Neural Networks (Combinations)

章节内容 Section	CNN			LSTM			BERT			BiLSTM＋CNN			BERT＋BiLSTM			BERT＋GCN			**BERT＋GAN**		
	P	R	$F1$	P	R	$F1$	P	R	$F1$	P	R	$F1$	P	R	$F1$	P	R	$F1$	P	R	$F1$
introduction	0.93	0.87	0.90	0.89	0.65	0.75	0.77	0.86	0.81	0.93	0.82	0.87	0.90	0.90	0.85	0.90	0.87	0.88	0.78	0.92	0.84
literature review	0.49	0.54	0.51	0.16	0.13	0.14	0.56	0.46	0.41	0.43	0.42	0.42	0.46	0.40	0.43	0.65	0.34	0.45	0.66	0.56	0.60
methodology	0.65	0.68	0.67	0.46	0.04	0.07	0.58	0.76	0.66	0.65	0.55	0.60	0.63	0.70	0.66	0.65	0.74	0.69	0.64	0.66	0.65
result	0.81	0.68	0.74	0.46	0.78	0.58	0.75	0.77	0.72	0.70	0.73	0.71	0.81	0.75	0.75	0.80	0.72	0.76	0.73	0.74	0.73
conclusion	0.76	0.90	0.83	0.53	0.83	0.65	0.83	0.86	0.84	0.71	0.89	0.79	0.83	0.87	0.85	0.86	0.86	0.86	0.85	0.85	0.85
all(weighted)	0.75	0.74	0.73	0.53	0.52	0.46	0.74	0.72	0.71	0.71	0.71	0.70	0.74	0.75	0.75	0.75	0.74	0.74	**0.75**	**0.76**	**0.76**

表 5　章节段落层次在不同神经网络(组合)上的实验结果

Table 5　Results of Paragraph on Different Neural Networks (Combinations)

章节段落 Paragraph	CNN			LSTM			**BERT**			BiLSTM+CNN			BERT+BiLSTM			BERT+GCN			BERT+GAN		
	P	*R*	*F*1	*P*	*R*	*F*1	*P*	*R*	*F*1	*P*	*R*	*F*1	*P*	*R*	*F*1	*P*	*R*	*F*1	*P*	*R*	*F*1
introduction	0.75	0.72	0.73	0.72	0.75	0.73	0.49	0.75	0.59	0.77	0.69	0.73	0.77	0.69	0.73	0.79	0.78	0.79	0.81	0.72	0.77
literature review	0.44	0.19	0.27	0.43	0.19	0.26	0.56	0.02	0.03	0.22	0.09	0.12	0.22	0.09	0.12	0.52	0.21	0.30	0.36	0.36	0.36
methodology	0.59	0.61	0.60	0.59	0.68	0.63	0.56	0.74	0.64	0.51	0.67	0.58	0.51	0.67	0.58	0.52	0.70	0.59	0.61	0.62	0.62
result	0.62	0.74	0.68	0.68	0.70	0.69	0.68	0.55	0.61	0.67	0.64	0.66	0.67	0.64	0.66	0.69	0.52	0.62	0.66	0.76	0.71
conclusion	0.71	0.40	0.51	0.62	0.44	0.52	0.60	0.30	0.40	0.43	0.24	0.30	0.43	0.24	0.30	0.63	0.50	0.59	0.71	0.56	0.63
all(weighted)	0.62	0.62	0.61	0.63	0.64	0.62	**0.69**	**0.68**	**0.68**	0.57	0.59	0.57	0.57	0.59	0.57	0.68	0.62	0.66	0.67	0.66	0.66

100%的好成绩。“conclusion”的识别效果也表现较好，在七种神经网络或神经网络的组合中的准确率、召回率和 $F1$ 值也均超过了90%，其中 BERT+GCN 的 F1 值也取得最高值96%。“methodology”和“result”次之。“literature review”在七个模型中表现相对较差，但 $F1$ 值最高也达到了 81%，通过对相关语料进行内容分析发现，其部分章节标题的表述形式丰富多样，但因为语料规模自身的限制，增加了神经网络对特征识别的判别难度，导致基于章节标题的“literature review”功能识别表现较差，而其他结构功能的标题特征，尤其是“introduction”和“conclusion”表述较为集中和规范，具有明显的规律性特征，进而识别效果表现较佳。

从表 4 可以分析出，章节内容层次的普遍识别率均超过了 70%，有两组模型的 $F1$ 值达到了 75%及以上，其中表现最好的是 BERT+GAN，其准确率、召回率和 $F1$ 值分别达到了75%、76%、76%。相较于其他六组模型，单 LSTM 的表现最差，其 $F1$ 值仅有 46%，准确率和召回率也只有 53%和 52%。在单模型中，CNN 模型和 LSTM 模型在“introduction”类别上的准确率和 $F1$ 值最高，而 BERT 是在“conclusion”类别上表现突出，其准确率达到了 83%，远超出 CNN 和 LSTM 在“conclusion”模块的表现。在四组组合模型中，BiLSTM+CNN、BERT+BiLSTM、BERT+GCN 三组组合模型都在“introduction”类别上发挥了最好的作用，准确率均大于或等于 90%，$F1$ 值也大于等于 85%，但在“literature review”类别上的表现不如 BERT+GAN。BERT+GAN 在各类别模块中，表现最好的是“conclusion”类别，其准确率、召回率和 $F1$ 值均达到了85%。此外，BERT+GAN 组合模型不仅在整体上表现效果最好，超过其他六组模型，而且在“literature review”这个更容易被错分为其他类别的模块上表现最好，其 $F1$ 值达到了60%。“literature review”模块较其他模块的召回率普遍偏低，也证实了该模块更容易被错分为其他四种类别。

在章节标题层次，BERT+GCN 的表现优于 BERT+GAN，主要是因为章节标题数据集结构较为简短，多为单词或词组的形式，词节点关系明确。而 GCN 通过捕捉图结构中节点间的关系，并通过卷积操作在节点上进行信息的传播和特征提取，在章节标题级别数据上可以充分发挥其自身优势，识别准确率较高，文本分类效果较好。而在章节内容层次，输入文本长度远远大于标题层次，其中节点关系复杂且噪音较多，对 GCN 表达建模能力挑战较大。GAN 作为一种对抗生成模型，可以学习生成器网络生成的数据与原始数据之间的分布差异，通过对抗训练的方式提升分类的准确性和泛化性，从而减小噪声数据的影响，模型能够更好地学习到章节内容各个类别的特征和分布差异，所以在章节内容层次 BERT+GAN 表现优于 BERT+GCN。

从表 5 中可以发现，在章节段落层次中，七种神经网络或神经网络组合的整体表现相差不大，$F1$ 值均未超过 70%，其中准确率和召回率最高的均是单 BERT 模型，分别达到了69%和 68%。效果最好的也是单 BERT 模型，$F1$ 值达到了 68%。BiLSTM+CNN 组合模型在七种模型中表现较差，其准确率、召回率和 $F1$ 值分别为 57%、59%、57%。在五个

不同模块中，七个模型在“literature review”类别的召回率普遍偏低，这一点与章节内容层次识别结果类似，最低在单 BERT 模型上的召回率仅有 2%，但在 BERT＋GCN 和 BERT＋GAN 组合中召回率要远高于其他五种模型，R 值分别为 21%和 36%。七种模型均在“introduction”模块表现最好，在“methodology”模块表现得较为平均，$F1$ 值分布在 58%—64%，没有明显的差别。

对不同层次的识别结果进行整体分析，章节标题层次的科技文献结构功能分类效果最佳，章节内容的分类效果次之，而章节段落整体表现都较前两者差。溯其本源，和不同层次的数据构成有着密切的关系，章节标题层次的文本较短，并且多数按照一定的写作规律直接包含诸如“简介”“文献综述”“结论”等明显特征信息，规律分布更加明显，模型更容易学习到有用信息，因此神经网络模型表现最佳，其中因为 transformer 在短文本分类上又表现良好，所以 BERT 和 BERT 的组合模型发挥了强大的功能，在准确率和召回率上相对传统其他神经网络或神经网络组合有着明显的进步。而章节内容和章节段落层次所蕴含的信息较为丰富，文本较长，直接对这类文本进行分析，提升了模型获取有效信息的难度，干扰信息也比较多，因此神经网络模型表现较差。从七种模型分类效果来看，模型在章节标题和章节段落层次的识别相差较小，其中章节标题层次的识别相差最小，在章节内容层次的识别相差较大，其中 LSTM 同其他六种模型或模型组合有着明显的识别效果差距。从整体来看，BERT＋GCN 和 BERT＋GAN 在三个不同层次都有较好的表现，一定程度证明了新的神经网络组合在文本分类任务上可以发挥其强大的功能。传统的神经网络模型组合中，BiLSTM＋CNN 并不如 BERT＋BiLSTM 对不同层次的整体识别效果有显著的提升。从单一模型的效果角度来看，CNN 和 BERT 都表现较好，尤其是 BERT 发挥了在文本分类任务中的强大作用，LSTM 相较于前两者表现一般。

4.5 错误分析

为了进一步发现实验结果中的错误分类情况，本文以不同层次最好的模型或模型组合（章节标题选取 BERT＋GCN、章节内容选取 BERT＋GAN、章节段落选取 BERT）为分析对象，输出其分类结果，如表 6 至表 8 所示，其中行代表各个结构功能被划分为不同类别的比例，列代表不同结构功能被划分为这一类别结构功能的比例。

表 6　章节标题层次实验结果错分表

Table 6　The Misclassification Results at Header

章节标题(Header)	introduction	literature review	methodology	result	conclusion	all
introduction	0.99	0.00	0.01	0.00	0.00	1.00
literature review	0.00	0.48	0.44	0.08	0.00	1.00
methodology	0.00	0.03	0.86	0.11	0.00	1.00
result	0.00	0.01	0.21	0.77	0.01	1.00
conclusion	0.00	0.00	0.01	0.07	0.92	1.00
all	0.99	0.52	1.53	1.03	0.93	5.00

表 7　章节内容层次实验结果错分表

Table 7　The Misclassification Results at Section

章节内容(Section)	introduction	literature review	methodology	result	conclusion	all
introduction	0.83	0.08	0.07	0.01	0.01	1.00
literature review	0.10	0.43	0.39	0.08	0.00	1.00
methodology	0.03	0.08	0.64	0.23	0.02	1.00
result	0.02	0.01	0.17	0.74	0.06	1.00
conclusion	0.04	0.00	0.02	0.10	0.84	1.00
all	1.02	0.60	1.29	1.16	0.93	5.00

表 8　章节段落层次实验结果错分表

Table 8　The Misclassification Results at Paragraph

章节段落(Paragraph)	introduction	literature review	methodology	result	conclusion	all
introduction	0.77	0.08	0.11	0.03	0.01	1.00
literature review	0.14	0.33	0.44	0.08	0.01	1.00
methodology	0.04	0.05	0.63	0.27	0.01	1.00
result	0.02	0.01	0.20	0.73	0.04	1.00
conclusion	0.06	0.02	0.10	0.22	0.60	1.00
all	1.03	0.49	1.48	1.33	0.67	5.00

横向对比表 6 至表 8 可以发现，章节标题层次的结构功能分类效果优于章节内容和章节段落。以深度学习识别中表现较好的“introduction”为例，在章节标题层次中识别正确率高达 99%，在章节内容中识别正确率达到 83%，而在章节段落层次识别正确率仅有 77%，分类效果一般。在章节段落层次，“introduction”被错分为其余四种结构功能类别，这与文本段落特征不明显、缺少显著特征信息有关。不仅仅是“introduction”，其余四种结构功能在章节段落层次上，也不同程度地被错分为其余四种结构功能类别。

4.6　领域迁移实验分析

为进一步证实模型的泛化性，我们使用本文设计的模型进行文本分类任务，选用生物学领域中常用的 PubMed-20k 数据集进行跨领域摘要结构功能分类。表 9 总结了本文具有代表性的模型或模型组合的实验结果，可以观察到，BERT＋GCN 的表现优于单 BERT 和 BERT＋GAN，其准确率、召回率和 $F1$ 值均达到了 90%。表 10 中我们对比了本文模型和 SCIBERT 模型[34]的文本分类实验结果，SCIBERT 在 PubMed-20k 数据集上的分类结果 $F1$ 值为 87%，在单句单模型的文本分类任务中处于第一的位置。BERT＋GAN 的模型分类效果能与 SCIBERT 持平、而 BERT＋GCN 的组合模型效果超过了 SCIBERT。其结果体现了两组模型较好的适应性，不仅适用于某一领域单一特定数据集的结构功能分类，而是具有较好的鲁棒性，能在不同领域的数据集中进行结构功能分类时发挥强大的功能。从单句单模型文本的结构功能分类任务中也可以看出，GCN 在这里超过了 SCIBERT 的模型效果。

表 9　PubMed 20k RCT 数据集在 BERT、BERT＋GCN、BERT＋GAN 上的实验结果

Table 9　Results of PubMed 20k RCT dataset on BERT, BERT＋GCN and BERT＋GAN

PubMed 20k RCT	BERT			BERT＋GCN			BERT＋GAN		
	P	R	F1	P	R	F1	P	R	F1
background	0.66	0.81	0.73	0.73	0.85	0.78	0.68	0.77	0.72
conclusions	0.86	0.78	0.82	0.91	0.87	0.89	0.81	0.83	0.82
methods	0.92	0.95	0.93	0.95	0.96	0.96	0.92	0.95	0.93
objective	0.74	0.53	0.62	0.79	0.58	0.67	0.79	0.51	0.62
results	0.91	0.91	0.91	0.94	0.94	0.94	0.92	0.90	0.91
all(weighted)	0.86	0.86	0.86	**0.90**	**0.90**	**0.90**	0.87	0.87	0.87

表 10　本研究模型与生物医学数据集上的 SCIBERT 结果进行比较

Table 10　Comparison of Our Model with the SCIBERT Results on the Biomedical Dataset

Field	Dataset	SOTA	SCIBERT	BERT＋GAN	BERT＋GCN
CS	ScienceDirect-Header	0.88	0.84	0.85	**0.88**
	ScienceDirect-Section	0.76	0.71	**0.76**	0.74
	ScienceDirect-Paragraph	0.68	**0.68**	0.66	0.66
MED	PubMed 20k RCT	0.93	0.87	0.87	**0.90**

5 结论与展望

本文将图神经网络 GCN 和对抗神经网络 GAN 引入到科技文献的结构功能识别任务中，采用组合模型的方式，引入模型 BERT＋GCN 和 BERT＋GAN 对科技文献的三个不同层次——章节标题、章节内容、章节段落进行结构功能的识别研究。 就结构功能而言，从不同层次识别结果来看，章节标题层次表现最好，章节内容次之，章节段落相对表现欠佳。从模型角度而言，在章节标题层次中，BERT＋GCN 综合表现最佳，BERT＋GAN 次之，传统的 CNN 也表现较好。 在章节内容中，BERT＋GAN 组合效果最好，BERT＋BiLSTM 表现次之。 在章节段落上，这七种模型或模型组合表现差距不大，BERT＋GAN 虽然表现最好，但整体的 $F1$ 值仍未超过 70%，这也是下一步研究要重点关注的方向。 整体来看，本研究融合不同层次的结构功能文本类型，组合不同的神经网络模型来识别模型具有一定的合理性和可行性。 未来，可重点对章节段落层次进行相关研究，尝试引入新的特征来辅助结构功能的识别，例如，考虑不同章节内容的词汇功能分布和引用意图分布，还可以尝试加入词汇功能、引文句和引文功能等信息。

参考文献

[1] Johnson R，Watkinson A，Mabe M. The STM report：An overview of scientific and scholarly publishing[R/OL]. [2023-01-08]. https://www.stm-assoc.org/2018_10_04_STM_Report_2018.pdf.

[2] Alzahrani S，Palade V，Salim N，et al. Using structural information and citation evidence to detect significant plagiarism cases in scientific publications[J]. Journal of the American Society for Information Science and Technology，2012，63(2)：286-312.

[3] 马晓慧，赵文娟，刘忠宝. 基于深度学习的多学科多层次学术论文结构功能识别方法比较研究[J]. 情报科学，2021，39(8)：94-102.(Ma X H，Zhao W J，Liu Z B. Multi-Disciplinary and Multi-level Comparative Research on Methods of Academic Text Structure Function Recognition Based on Deep Learning[J]. Information Science，2021，39(8)：94-102.)

[4] 姜艺，黄永，夏义堃，等. 学术文本词汇功能识别——在关键词自动抽取中的应用[J]. 情报学报，2021，40(2)：152-162.(Jiang Y，Huang Y，Xia Y K，et al. Recognition of lexical functions in academic texts：Application in automatic keyword extraction[J]. Journal of the China Society for Scientific and Technical Information，2021，40(2)：152-162.)

[5] Gui J，Sun Z N，Wen Y G，et al. A review on generative adversarial networks：Algorithms，theory，and applications[J]. IEEE Transactions on Knowledge and Data Engineering，2023，35(4)：3313-3332.

[6] Yang Y M，Pedersen J O. A comparative study on feature selection in text categorization[C]//Proceedings of the Fourteenth International Conference on Machine Learning. San Francisco：Morgan Kaufmann Publishers，1997：412-420.

[7] Lu Z B，Du P，Nie J Y. VGCN-BERT：Augmenting BERT with graph embedding for text classification[C]//Proceedings of European Conference on Information Retrieval. Cham：Springer，2020：369-382.

[8] Goodfellow I，Pouget-Abadie J，Mirza M，et al. Generative adversarial networks[J]. Communications of the ACM，2020，63(11)：139-144.

[9] Hu Z G，Chen C M，Liu Z Y. Where are citations located in the body of scientific articles? A study of the distributions of citation locations[J]. Journal of Informetrics，2013，7(4)：887-896.

[10] Ding Y，Liu X Z，Guo C，et al. The distribution of references across texts：Some implications for citation analysis[J]. Journal of Informetrics，2013，7(3)：583-592.

[11] 陆伟，黄永，程齐凯. 学术文本的结构功能识别——功能框架及基于章节标题的识别[J]. 情报学报，2014，33(9)：979-985.(Lu W，Huang Y，Cheng Q K. The structure function of academic text and its classification[J]. Journal of the China Society for Scientific and Technical Information，2014，33(9)：979-985.)

[12] 黄永，陆伟，程齐凯. 学术文本的结构功能识别——基于章节内容的识别[J]. 情报学报，2016，35(3)：293-300.(Huang Y，Lu W，Cheng Q K. The structure function recognition of academic text chapter content based recognition[J]. Journal of the China Society for Scientific and Technical Information，2016，35(3)：293-300.)

[13] 黄永，陆伟，程齐凯，等. 学术文本的结构功能识别——基于段落的识别[J]. 情报学报，2016，35(5)：530-538.(Huang Y，Lu W，Cheng Q K，et al. The structure function recognition of academic text paragraph-based recognition[J]. Journal of the China Society for Scientific and Technical Information，2016，35(5)：530-538.)

[14] 王佳敏，陆伟，刘家伟，等. 多层次融合的学术文本结构功能识别研究[J]. 图书情报工作，2019，63(13)：95-104.(Wang J M，Lu W，Liu J W，et al. Research on structure function recognition of academic text based on multi-level

fusion[J]. Library And Information Service, 2019, 63(13): 95-104.)
[15] 秦成磊，章成志. 基于层次注意力网络模型的学术文本结构功能识别[J]. 数据分析与知识发现，2020，4(11): 26-42.(Qin C L, Zhang C Z. Recognizing structure functions of academic articles with hierarchica attention network[J]. Data Analysis and Knowledge Discovery, 2020, 4(11): 26-42.)
[16] Sollaci L B, Pereira M G. The introduction, methods, results, and discussion (IMRAD) structure: A fifty-year survey[J]. Journal of the Medical Library Association, 2004, 92(3): 364-367.
[17] 刘忠宝，王宇飞，张志剑. 基于深度学习模型的摘要结构功能识别方法研究[J]. 情报科学，2021，39(3): 107-112.(Liu Z B, Wang Y F, Zhang Z J. Research on the recognition method of abstract structure function based on deep learning model[J]. Information Science, 2021, 39(3): 107-112.)
[18] 王东波，高瑞卿，叶文豪，等. 不同特征下的学术文本结构功能自动识别研究[J]. 情报学报，2018，37(10): 997-1008.(Wang D B, Gao R Q, Ye W H, et al. Research on the structure recognition of academic texts under different characteristics[J]. Journal of the China Society for Scientific and Technical Information, 2018, 37(10): 997-1008.)
[19] Lu W, Huang Y, Bu Y, et al. Functional structure identification of scientific documents in computer science[J]. Scientometrics, 2018, 115(1): 463-486.
[20] Ma B W, Zhang C Z, Wang Y Z, et al. Enhancing identification of structure function of academic articles using contextual information[J]. Scientometrics, 2022, 127(2): 885-925.
[21] 毛进，陈子洋. 基于深度主动学习的科技文献摘要结构功能识别研究[J/OL]. 数据分析与知识发现. [2023-08-16]. http://kns.cnki.net/kcms/detail/10.1478.G2.20230815.1419.006.html.(Mao J, Chen Z Y. Structural function identification of scientific literature abstracts based on deep active learning[J/OL]. Data Analysis and Knowledge Discovery.[2023-08-16]. http://kns.cnki.net/kcms/detail/10.1478.G2.20230815.1419.006.html.)
[22] 唐晓波，彭映寒. 科技论文引用对象和引文功能的联合自动识别方法研究[J]. 现代情报，2022，42(6): 38-48.(Tang X B, Peng Y H. Research on joint automatic recognition method of citation objects and their relationships in scientifc papers[J]. Journal of Modern Information, 2022, 42(6): 38-48.)
[23] Liu X E, You X X, Zhang X A, et al. Tensor graph convolutional networks for text classification[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Washington, USA: AAAI Press, 2020: 8409-8416.
[24] 郑诚，倪显虎，张苏航，等. 结合 GNN 的信息融合用于归纳式文本分类[J]. 小型微型计算机系统，2023，44(6): 1170-1176.(Zheng C, Ni X H, Zhang S H, et al. Information fusion combined with GNN is used for inductive text classifcation[J]. Journal of Chinese Mini-Micro Computer Systems, 2023, 44(6): 1170-1176.)
[25] Lin Y X, Meng Y X, Sun X F, et al. BertGCN: Transductive text classification by combining GCN and BERT[EB/OL]. arXiv preprint, 2021. [2023-01-08]. https://arxiv.org/abs/2105.05727.
[26] Yao L A, Mao C S, Luo Y A. Graph convolutional networks for text classification[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Washington, USA: AAAI Press, 2019: 7370-7377.
[27] Li C L, Su Y X, Liu W J. Text-to-text generative adversarial networks[C]//2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil. Piscataway: IEEE, 2018: 1-7.
[28] Croce D, Castellucci G, Basili R. GAN-BERT: Generative adversarial learning for robust text classification with a bunch of labeled examples[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA, USA: Association for Computational Linguistics, 2020: 2114-2119.
[29] Aggarwal A, Mittal M, Battineni G. Generative adversarial network: An overview of theory and applications[J]. International Journal of Information Management Data Insights, 2021, 1(1): 100004.
[30] Chen T, Zhai X H, Ritter M, et al. Self-supervised GANs via auxiliary rotation loss[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA. Piscataway: IEEE, 2020: 12146-12155.
[31] Graves A, Graves A. Supervised sequence labelling[M]. Cham: Springer, 2012.
[32] Rhanoui M, Mikram M, Yousfi S, et al. A CNN-BiLSTM model for document-level sentiment analysis[J]. Machine Learning and Knowledge Extraction, 2019, 1(3): 832-847.
[33] Li W T, Gao S B, Zhou H, et al. The automatic text classification method based on BERT and feature union[C]//2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS). TianJin, China. Piscataway: IEEE, 2020: 774-777.
[34] Beltagy I, Lo K, Cohan A. SciBERT: A pretrained language model for scientific text[EB/OL]. arXiv preprint, 2019. [2023-01-08]. https://arxiv.org/abs/1903.10676.

（收稿日期：2023-04-19）