基于混合粒度全局图的多标签文本分类方法*

王　哲，温秀梅

（河北建筑工程学院　信息工程学院，河北　张家口　075000）

摘　要：多标签文本分类旨在为每个文本实例分配多个标签。传统多标签文本分类方法通常依赖于粗粒度的特征表示，忽视了文本中多层次、多尺度的语义信息。为了解决该问题，提出一种基于混合粒度全局图的多标签文本分类方法，通过 MHA 提取细粒度的文本特征，捕捉词与标签之间的交互信息，同时使用 Bi-LSTM 提取粗粒度的文本特征。随后，通过门控融合机制将两种特征融合得到具有多层次语义的混合粒度特征。将混合粒度词表示、文本和标签作为节点构建全局图，并通过图卷积网络处理全局图以进行分类。在 AAPD、RCV1-V2 两个数据集上进行实验，实验结果表明，所提出方法能有效提升模型性能。

关键词：多标签文本分类；多头注意力机制；双向长短期记忆网络；门控融合机制；图卷积网络

中图分类号：TP393　　**文献标识码：**A　　**DOI：**10. 19358/j. issn. 2097 - 1788. 2025. 06. 006

引用格式：王哲，温秀梅. 基于混合粒度全局图的多标签文本分类方法［J］. 网络安全与数据治理，2025，44(6)：42 - 48.

A multi-label text classification method based on a mixed-granularity global graph

Wang Zhe，Wen Xiumei

（College of Information Engineering，Hebei University of Architecture，Zhangjiakou 075000，China）

Abstract：Multi-label text classification is designed to assign multiple labels to each instance of text. Traditional multi-label text classification methods usually rely on coarse-grained feature representations，ignoring the multi-level and multi-scale semantic information in the text. In order to solve this problem，this paper proposes a multi-label text classification method based on mixed granularity global graph，which extracts fine-grained text features through MHA to capture the interaction information between words and labels，and uses Bi-LSTM to extract coarse-grained text features. Subsequently，the two features are fused through the gated fusion mechanism to obtain mixed granular features with multi-level semantics. The fused mixed granular word representations，texts，and labels are used together to construct a global graph，and the global graph is processed through a graph convolutional network for classification. Experiments are carried out on two datasets，AAPD and RCV1-V2，and the experimental results show that the proposed method can effectively improve the performance of the model.

Key words：multi-label text classification；multi-head attention mechanism；bidirectional long short-term memory network；gated fusion mechanism；graph convolutional networks

0　引言

多标签文本分类是一项基本的文本挖掘任务。在许多应用程序中，文本可能对应于多个相互排斥的标签[1]。多标签文本分类可以有效地降低人工成本，具有广泛的应用前景[2]。与传统的单标签分类任务不同，多标签文本分类旨在为每个文本实例同时分配多个标签。多标签文本分类任务可以广泛应用于情感分析[3]、档案管理[4]、期刊分类[5]、新闻过滤[6]等领域，尤其在处理复杂、信息密集的文本时展现出独特的优势。例如，在情感分析任务中，一个文本可能同时表达多个情感，而在新闻处理中，一个事件可能涉及多个主题，因此需要高效的多标签分类方法来捕捉文本中的多层次语义信息。

尽管深度学习方法在自然语言处理任务中取得了显著进展，但现有的多标签文本分类方法依然面临许多挑战。大多数方法仍然依赖于词袋（Bag of Words，BoW）[7]和 n 元语法（n-gram）[8]模型或简单的深度学习架构，如

*基金项目：河北建筑工程学院研究生创新基金（XY2025029）

卷积神经网络（Convolutional Neural Network，CNN）[9]和长短期记忆网络（Long Short-Term Memory Network，LSTM）[10]。这些方法主要关注从文本中提取低级特征，忽视了文本中的高层次语义关系与标签之间的潜在关联性。在多标签分类任务中，标签之间的相关性非常复杂，不同标签之间往往存在一定的依赖关系和上下游信息，这使得现有方法在处理长文本或包含多重语义关系的复杂任务时，表现出信息丢失和语义理解不足的问题。

一方面，传统的基于特征的浅层模型无法有效捕捉文本中的上下文信息和语义层次，因此其分类效果在多标签任务中往往较差。另一方面，虽然现代的深度学习模型如 BERT[11-12]、GPT[13]等在单标签分类任务中取得了很大成功，但由于其关注点过于集中在文本本身的语言表达上，仍未能充分考虑标签之间的关联性。此外，许多现有方法采用的是逐标签独立学习的策略，即每个标签的预测都是基于输入文本的独立决策，这种做法未能充分挖掘标签间的共现和相互影响。

为了应对这些挑战，本文提出了一种基于混合粒度全局图（Hybrid Granularity Global Graph，HGG）的多标签文本分类方法。本文的核心创新在于引入了一个混合粒度特征提取机制，该机制结合了细粒度和粗粒度两种层次的文本特征，从而能够更全面地理解文本中的复杂语义。具体来说，细粒度特征提取通过捕捉单词、短语以及句子层次的语义信息，帮助模型更好地理解细节和局部语境；而粗粒度特征提取则通过全局信息建模，使得模型能够从宏观层面理解文本的主要内容和潜在意图。这种结合不同粒度特征的方式，有助于更好地平衡局部与全局信息，从而提高多标签文本分类的精度。

近年来，图神经网络（Graph Neural Network，GNN）在推荐系统[14]、计算机视觉[15]等领域得到广泛应用，许多改进的模型被相继提出，例如图卷积网络（Graph Convolutional Network，GCN）[16]、图注意力网络（Graph Attention Network，GAT）[17]等，它们在不同的应用场景中展现了出色的表现。在多标签文本分类任务中，GNN 能够通过构建标签之间的关联图，捕捉标签间的共现关系和依赖性，进而提升标签预测的准确性和全面性。特别是在多标签任务中，通过 GNN 的图结构传播机制，可以有效地增强标签之间的相互影响，从而提升多标签分类模型的性能。此外，GNN 的可扩展性和灵活性使得它能够适应多样化的数据结构，成为处理复杂数据和任务的一种理想工具。

因此，本文还提出了一种全局图结构，并利用 GCN 来捕捉文本和混合粒度词表示之间的全局关系。通过构建全局图，能够捕捉到标签间的潜在关联，并通过 GCN 的传播机制，强化节点之间的相互影响，从而提升模型的分类效果。

综上所述，本文的主要贡献为以下几点：

（1）针对现有模型通常只考虑文本粗粒度的特征，本文在文本特征提取阶段，采用多头注意力机制（Multi-Head Attention，MHA）提取包含词与标签交互信息的细粒度文本特征。

（2）设计一种门控融合机制，将粗粒度文本特征与细粒度文本特征融合得到混合粒度的文本特征。

（3）将词的混合粒度表示、文本和标签作为节点构建全局图，并通过 GCN 处理全局图，以捕捉文本中词语、标签与文本之间的潜在语义关联和全局结构信息。

（4）在 AAPD 和 RCV1-V2 两个公开数据集上进行多标签文本分类任务验证本文方法的有效性。

1　相关工作

1.1　基于深度学习的多标签文本分类

多标签文本分类是数据挖掘和自然语言处理中一个实际而有意义的任务，它意味着一个实例可以同时与多个标签相关联[18]。近年来，许多研究者通过深度学习方法提升了多标签文本分类的性能。深度学习算法对传统机器学习方法中的人工特征提取做出了巨大贡献[19]。

早期的多标签文本分类方法多依赖于传统的特征工程和机器学习模型，如支持向量机（Support Vector Machine，SVM）[20]和决策树（Decision Tree，DT）[21]等。这些方法通常忽略了标签之间的关系，且对文本的语义理解较为粗糙。随着神经网络的发展，CNN 和循环神经网络（Recurrent Neural Network，RNN）[22]被广泛应用于文本分类任务。Brownlee[23]提出了一种基于 CNN 的多标签文本分类方法，利用卷积核提取局部特征，虽然在某些任务中表现良好，但其方法难以捕捉文本中的长距离依赖性。相比之下，LSTM 由于能够处理长序列的依赖关系，成为多标签文本分类的常用模型，Yang 等人[5]提出的序列生成模型（Sequence Generation Model，SGM）基于 Bi-LSTM（Bidirectional Long Short-Term Memory Network）编码器，通过双向信息对文本上下文进行建模，提高了分类精度。然而，LSTM 类模型仍存在着处理复杂关系能力不足的问题。Yao 等人[24]提出的 TextGCN 首次通过构建图结构来捕捉文本中的复杂关系，利用文档节点和词节点构建图结构，将文本分类任务转化为节点分类任务。图结构的引入使得模型能够更好地建模文本中的语义和标签之间的依赖关系。GCN 和其他 GNN 架构逐渐成为多标签文本分类的研究热点，这些方法能够利用图结构在节点之间传递信息，从而捕捉文本中复杂的关

系和标签间的交互信息。

1.2 标签语义信息

在多标签文本分类中，标签不仅仅是分类的结果，通常还包含着重要的语义信息。标签语义信息的利用可以从多个方面入手。首先，标签的定义本身就包含了丰富的语义信息，这些信息可以通过预训练的词嵌入或知识图谱等方式进行表示和提取。例如，可以使用 Word2Vec[25]、BERT 等模型将标签转换为向量，这些向量能够捕捉到标签之间的语义相似性和差异性。也可以将这些语义特征作为额外的输入信息，整合到多标签文本分类模型中，从而提高模型的分类精度。其次，标签之间的关联也是标签语义信息的重要组成部分。在实际应用中，标签之间往往存在着复杂的层次结构或相关性。通过构建标签之间的语义关系图，并利用 GNN 等模型进行建模，可以进一步挖掘和利用这些关联信息，从而增强模型对文本的多标签预测能力。

Huang 等人[26]提出了如何通过点击数据来学习标签之间的关系，并应用于 Web 搜索领域，其提出的标签嵌入思想为后续的多标签学习和标签嵌入技术奠定了基础，但其无法捕捉标签间的关联性。Lee 等人[27]提出用图结构对标签关系建模，通过构造结构化知识图确定标签之间的拓扑和语义关系，以提高算法性能。

2 HGG 模型构建

本文提出的模型主要由五部分组成：粗粒度文本特征提取、细粒度文本特征提取、特征融合、构建全局图、GCN 分类。模型整体架构如图 1 所示。

本文设计并预训练一个 Bi-LSTM 编码器，来提取包含文本上下文语义信息的粗粒度文本特征；采用多头注意力机制提取包含词与标签交互信息的细粒度文本特征；通过门控融合机制将两种粒度的文本特征融合得到混合粒度的文本特征；将词的混合粒度特征表示与文本作为节点构建全局图；采用 GCN 处理全局图并实现分类任务。

2.1 粗粒度文本特征提取

相比于 CNN，LSTM 在处理文本数据时具有独特优势。LSTM 能够有效捕捉文本中的长距离依赖关系，这对于需要考虑上下文信息的任务尤为重要。作为一种递归神经网络，LSTM 通过内部记忆单元逐步处理每个输入词，并关联前后信息，因此在处理文本时能更好地保留上下文的语义。虽然 CNN 能够通过卷积操作有效提取局部特征，但对于需要全局语义理解和长距离依赖的任务，其能力相对有限。此外，LSTM 的门控机制使其能够在学习过程中自动调整对过去信息的记忆，从而更灵活地提取文本中的关键特征，尤其在处理较长文本时，能够更好地捕捉其中的重要信息。然而，传统 LSTM 容易受到梯度爆炸的影响。因此，本文采用 Bi-LSTM 来提取粗粒度的文本特征，不仅解决了 CNN 在特征提取上的局限性，还有效缓解了 LSTM 的梯度爆炸问题。

将文本 d 输入到 Bi-LSTM 中，得到第 i 个单词前向和后向隐藏状态如下：

$$\vec{h}_i = \text{LSTM}\ (\vec{h}_{i-1},\ w_i) \tag{1}$$

$$\overleftarrow{h}_i = \text{LSTM}\ (\overleftarrow{h}_{i-1},\ w_i) \tag{2}$$

$$\boldsymbol{H} = \{h_1,\ h_2,\ \cdots,\ h_m\} \tag{3}$$

通过将 i 时刻的前向与后向两个隐藏状态拼接，得到第 i 个单词的最终表示 $h_i = \{\vec{h}_i,\ \overleftarrow{h}_i\}$，$\boldsymbol{H}$ 为粗粒度的文本表示。

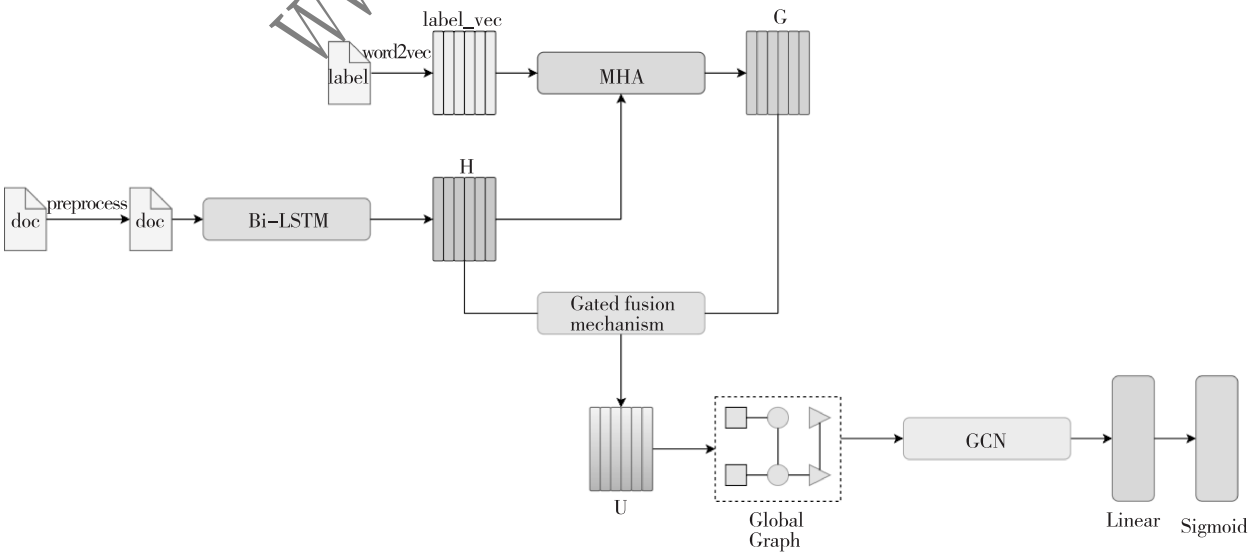

图 1 HGG 模型架构

2.2 细粒度文本特征提取

多头注意力机制能够通过并行计算多个注意力头来捕捉不同层次的语义关联。在多标签分类任务中，文本和标签之间的关系往往复杂且多样，传统的单一注意力头只能从一个角度理解这种关系。在本文中，将粗粒度的文本表示 $\boldsymbol{H}$ 作为查询向量，标签语义向量 $\boldsymbol{Y}$ 作为键向量和值向量，通过将查询向量与多个键向量进行匹配，从不同的子空间中计算多个注意力权重，从多个角度捕捉文本和标签之间不同的语义联系。这种并行化的计算方式不仅增强了模型在捕捉复杂关系时的能力，还能够在每个注意力头中提取到细粒度的特征信息。例如，一些注意力头可能会集中在某些标签与文本中的特定词汇之间的紧密联系，而其他头则可能关注到标签和文本中更为隐性或间接的关联。通过对多个头的输出进行拼接或加权，模型能够综合不同角度的信息，得到一个更加丰富和准确的文本表示，这个表示不仅包含了文本的语义信息，还融入了标签的语义信息，从而为最终的标签预测提供了更为细致和全面的特征表示。它不仅仅是对文本信息的直接编码，还在多个语义空间中对文本和标签进行交互式建模，具体实现如式（4）和式（5）所示：

$$\text{head}_i = \text{SoftMax}\left(\frac{[\boldsymbol{H}\boldsymbol{W}_i^{\text{Q}}][\boldsymbol{Y}\boldsymbol{W}_i^{\text{K}}]^{\text{T}}}{\sqrt{k/r}}\right)[\boldsymbol{Y}\boldsymbol{W}_i^{\text{V}}] \tag{4}$$

$$\boldsymbol{G} = \text{concat}(\text{head}_1, \text{head}_2, \cdots, \text{head}_r)\ \boldsymbol{W} \tag{5}$$

其中，head_i 为第 i 个注意力头的计算结果；$\{\boldsymbol{W}_i^{\text{Q}}、\boldsymbol{W}_i^{\text{K}}、\boldsymbol{W}_i^{\text{V}}\} \in \mathbf{R}^{d\times d/r}$ 分别为第 i 个注意力头中查询向量、键向量、值向量的参数矩阵，r 为注意力头的个数，d 为文本特征向量和标签语义向量维度；$\boldsymbol{G}$ 为词与标签细粒度的交互线索。

2.3 门控融合机制

Bi-LSTM 编码器输出的是粗粒度的文本表示，多头注意力机制输出的是细粒度的文本表示，为了有效融合这两种不同层次的表示，本文采用了门控融合机制，$\boldsymbol{U}$ 即为融合后的混合粒度文本表示。其中融合系数 μ 的取值通过训练自动调整。融合过程如式（6）所示：

$$\boldsymbol{U} = \mu \odot \boldsymbol{H} + (1-\mu)\ \odot \boldsymbol{G} \tag{6}$$

2.4 构建全局图

通过 Bi-LSTM、多头注意力机制和门控融合机制三个模块提取出了混合粒度的词表示，将其与文本节点和标签节点一起构建全局图。虽然在单词节点中已经融入了标签信息，但由于标签之间存在某种内在的关系，这种关系无法通过单词的语义嵌入来完全捕捉。通过将标签单独作为节点并建立标签之间的边，可以让模型更好地捕捉标签间的相互关系，明确标签结构，增强模型对标签语义的理解。其次将标签节点与词节点建边能够更直观地表现出某些单词与特定标签相关，增加了图的可解释性。邻接矩阵的构建和权值计算过程如式（7）所示：

$$A_{ij} = \begin{cases} \text{PMI}(i,j), & i,j = \text{word/label} \\ \text{IDF}(i,j), & \text{word}_j \in \text{document}_i \\ 1, & i = j \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

IDF、PMI 计算过程如式（8）~式（11）所示：

$$\text{IDF}(w_j) = \log\left(\frac{1+N}{1+\text{df}(w_j)}\right) \tag{8}$$

$$\text{PMI}(i,j) = \log\frac{p(i,j)}{p(i)\ p(j)} \tag{9}$$

$$p(i,j) = \frac{W(i,j)}{W} \tag{10}$$

$$p(i) = \frac{W(i)}{W} \tag{11}$$

其中，N 表示文本总数，$\text{df}(w_j)$ 表示包含单词 w_j 的文本数量，$p(i,j)$ 表示单词 i 和单词 j 在一个文本中同时出现的概率，$p(i)$ 表示单词 i 出现的频率，$p(j)$ 表示单词 j 出现的频率，$W(i,j)$ 表示包含单词 i 和单词 j 的滑动窗口数量，$W(i)$ 表示包含单词 i 的滑动窗口数量，W 表示滑动窗口总数。

2.5 GCN 分类

相比于其他 GNN，GCN 通过图卷积操作高效地聚合局部和全局信息，捕捉文本的复杂语义关系，且更适用于语料库级别的图。因此本文使用 GCN 来处理全局图并实现分类。设 $\boldsymbol{H} \in \mathbf{R}^{n\times m}$ 为 n 个节点每个节点维度为 m 的特征矩阵，$\boldsymbol{A}$ 为邻接矩阵，将 $\boldsymbol{H}$ 和 $\boldsymbol{A}$ 输入到 GCN 中，通过每一层的图卷积操作来更新节点的表示，使节点的表示不仅包含自身的信息，还结合了邻居节点的信息。具体的更新过程如式（12）所示：

$$\boldsymbol{H}^{(l+1)} = \text{ReLU}(\hat{\boldsymbol{A}}\boldsymbol{H}^{(l)}\boldsymbol{W}^{(l)} + \boldsymbol{b}^{(l)}) \tag{12}$$

其中，$\hat{\boldsymbol{A}}$ 是正则化的对称邻接矩阵，$\boldsymbol{W}$ 是参数矩阵，$\boldsymbol{b}$ 是偏置向量，ReLU 是激活函数。每一层的隐藏表示 $\boldsymbol{H}^{(l+1)}$ 由上一层的隐藏表示 $\boldsymbol{H}^{(l)}$ 得到。

将 GCN 最后一层处理后的节点表示 $\boldsymbol{H}^{(\text{last})}$ 通过一个全连接层映射到结果空间，每个节点的输出 $\boldsymbol{Z}$ 由式（13）得到：

$$\boldsymbol{Z} = \text{sigmoid}(f(\boldsymbol{H}^{(\text{last})})) \tag{13}$$

其中，f 是全连接层，sigmoid 是常用的激活函数。

3 实验

3.1 数据集

本文实验使用 AAPD、RCV1-V2 两个公开数据集，这

两个数据集在多标签文本分类任务中被广泛应用，数据集具体介绍如表 1 所示，其中 N 表示样本数，L 表示标签总数，L_{avg} 表示平均每个样本对应的标签数。

表 1　AAPD 数据集和 RCV1-V2 数据集

	AAPD	RCV1-V2
N	55 840	804 414
L	54	103
L_{avg}	2. 4	3. 2

AAPD 数据集是一个用于多标签文本分类任务的学术数据集，包含了 55 840 篇学术论文的摘要及其对应的多个主题标签。每篇论文可以被分配多个主题，数据集中定义了 54 个不同的标签，涵盖了多个学术领域。多标签文本分类任务中，其目标是根据每篇论文的摘要内容，预测与之相关的主题标签。

RCV1-V2 是路透社新闻数据集，该数据集由804 414 篇新闻报道文章和 103 个主题构成。

3. 2　评价指标

Micro-$F1$ 是一个综合性的评价指标，适用于多类别、不平衡数据和多标签分类任务，能够综合考虑所有类别的表现，并且能够避免某些类别样本较少时评价指标失真的问题。具体的计算过程如下：

$$\text{Micro-}F_1 = \frac{2PR}{P+R} \tag{14}$$

$$P = \frac{\sum_{t\in Y} \text{TP}_t}{\sum_{t\in Y} \text{TP}_t + \text{FP}_t} \tag{15}$$

$$R = \frac{\sum_{t\in Y} \text{TP}_t}{\sum_{t\in Y} \text{TP}_t + \text{FN}_t} \tag{16}$$

其中，TP_t 表示标签集合 Y 中第 t 个标签的真阳性，FP_t 表示标签集合 Y 中第 t 个标签的假阳性，FN_t 表示标签集合 Y 中第 t 个标签的假阴性，P 和 R 分别是精确度和召回率。

3. 3　基线模型

为了验证本文所提模型的有效性，采用如下基线模型进行对比实验。

TextCNN：采用多种不同大小的卷积核提取不同尺度的文本特征，并通过池化操作提取出重要的特征，从而有效降低计算复杂度，同时保持模型性能。

SGM[5]：基于 Bi-LSTM 编码器，通过双向信息对文本上下文进行建模。

BBN：同时考虑了表示学习和分类器学习，并通过双分支结构使这两部分协同优化。

HTTN：通过将头部标签中的元知识迁移到尾部标签，解决了尾部标签数据稀缺的问题。这种方法依赖于层次结构的设计，结合了多层次建模和知识迁移机制，从而提升了尾部标签的分类效果。

TextGCN：首次将文本分类任务构建为全局图并转化为节点分类。

SGC：简化了图卷积的计算，通过消除非线性变换和将多层卷积合并为一次邻接矩阵的幂操作，减少了计算和内存开销，使得它在处理大规模图数据时具有显著的优势。

Induct-GCN：通过高效的图卷积操作、节点特征、邻接矩阵分解、层次化信息传播以及不依赖全图训练的优化方法实现，从而大大提高了模型在动态和大规模图数据中的适应性、计算效率和泛化能力。

3. 4　实验结果及分析

本文在 AAPD 和 RCV1-V2 两个公开数据集上对 HGG 模型进行实验，并与 7 个基线模型进行对比，采用 Micro-$F1$ 作为主要评价指标，以评估各个模型的分类性能。

3. 4. 1　AAPD 数据集实验结果

在 AAPD 数据集上进行实验，结果如表 2 所示。

表 2　不同模型在 AAPD 数据集上的实验结果

Model	P	R	Micro-$F1$
TextCNN	**0. 847**	0. 549	0. 667
SGM	0. 748	0. 659	0. 701
BBN	0. 724	0. 638	0. 678
HTTN	0. 751	0. 647	0. 695
TextGCN	0. 762	0. 671	0. 714
SGC	0. 758	0. 631	0. 689
Induct-GCN	0. 772	0. 681	0. 724
HGG	0. 793	**0. 686**	**0. 736**

3. 4. 2　RCV1-V2 数据集实验结果

在 RCV1-V2 数据集上进行实验，结果如表 3 所示。

表 3　不同模型在 RCV1-V2 数据集上的实验结果

Model	P	R	Micro-$F1$
TextCNN	**0. 924**	0. 796	0. 855
SGM	0. 887	0. 849	0. 868
BBN	0. 878	0. 831	0. 854
HTTN	0. 882	0. 843	0. 862
TextGCN	0. 892	0. 871	0. 881
SGC	0. 884	0. 839	0. 861
Induct-GCN	0. 894	0. 883	0. 889
HGG	0. 897	**0. 886**	**0. 892**

3. 4. 3　结果分析

根据表 2 和表 3，对 HGG 模型与 7 个基线模型进行

结果分析。

TextCNN：CNN 模型在准确率指标上表现出色，主要得益于其层次化特征学习、字符级建模等特性，但其难以捕捉包含上下文语义信息的文本特征。

SGM：基于 Bi-LSTM 的 SGM 模型能够较好地对文本的上下文信息进行建模，但其难以处理标签之间的关系。

BBN：BBN 通过双分支结构提高了鲁棒性，在多个实验中表现较好，但其在信息融合和全局建模方面仍有不足，无法与基于图卷积的模型相比。

HTTN：HTTN 适用于复杂任务，尤其在长尾标签分类和多标签分类中表现突出，但其较高的计算复杂度和调参难度使得它在实际应用中需要平衡性能和资源消耗。

TextGCN：TextGCN 通过图卷积网络捕捉文本间的关系，但在处理多标签任务时，缺乏对语义理解的深度，导致其在性能上稍逊。

SGC：由于其简化的图卷积操作，SGC 在计算效率上占优，但在复杂任务中，未能充分建模文本的全局依赖和语义结构。

Induct-GCN：虽然 Induct-GCN 具有较高的计算效率和较好的泛化能力，但其在信息融合和特征表示方面的能力稍显不足。

HGG：HGG 模型在处理大规模数据集和多标签任务时，能够更好地融合文本和标签间的关系，因此在对比实验中取得了最优的性能，证明了其在多标签文本分类任务中的优势。

3.5 消融实验

为了验证本文提出模型的有效性，将粗粒度文本表示的提取模块 Bi-LSTM 和细粒度文本表示的提取模块 MHA 作为实验的消融变量进行有效性验证。其中“w/o Bi-LSTM”表示的是去掉粗粒度文本表示后的细粒度文本表示，“w/o MHA”表示的是去掉细粒度文本表示后的粗粒度文本表示。表 4 为在 AAPD 数据集上的消融实验结果，表 5 为在 RCV1-V2 数据集上的消融实验结果。

表 4　AAPD 数据集上的消融实验

Model	*P*	*R*	Micro-*F*1
w/o Bi-LSTM	0.751	0.663	0.704
w/o MHA	0.748	0.659	0.701
HGG	**0.793**	**0.686**	**0.736**

表 5　RCV1-V2 数据集上的消融实验

Model	*P*	*R*	Micro-*F*1
w/o Bi-LSTM	0.883	0.842	0.862
w/o MHA	0.887	0.849	0.868
HGG	**0.897**	**0.886**	**0.892**

实验结果表明，在 AAPD 和 RCV1-V2 数据集上，HGG 模型在准确率、召回率和 Micro-*F*1 值等评估指标上均优于去除任意一个模块的模型。这证明去除 Bi-LSTM 或 MHA 模块会导致模型性能的下降，尤其是在长文本和复杂文档分类任务中。这也证明了 Bi-LSTM 和 MHA 模块在捕获文本信息的不同粒度特征方面的协同作用。

4　结论

本文针对现有多标签文本分类模型通常依赖于粗粒度的特征表示，忽视了文本中多层次、多尺度的语义信息这一问题提出了一种新的模型——HGG 模型，该模型利用 Bi-LSTM 提取文本本身粗粒度级别的特征，然后通过 MHA 提取包含标签语义信息的细粒度文本特征，接着采用门控融合机制将两种特征融合得到混合粒度的文本表示。随后将词的混合粒度表示作为节点与文本节点和标签节点构建全局图，并使用 GCN 处理全局图。实验结果表明，本文模型在两个公开多标签文本分类数据集上取得了良好的效果。

虽然本文所提出的模型在两个公开数据集上均取得了不错的效果，但是该模型在文本本身语义特征提取层面还有待提升，未来将在该方面进行进一步的探究。

参考文献

[1] ZENG D L, ZHA E Z, KUANG J Y, et al. Multi-label text classification based on semantic-sensitive graph convolutional network [J]. Knowledge-Based Systems, 2024. DOI: 10.1016/j.knosys.2023.111303.

[2] WANG J Y, CHEN Z J, QIN Y, et al. Multi-aspect co-attentional collaborative filtering for extreme multi-label text classification [J]. Knowledge-Based Systems, 2023. DOI: 10.1016/j.knosys.2022.110110.

[3] AL QABLAN T A, MOHD NOOR M H, AL BETAR M A, et al. A survey on sentiment analysis and its applications [J]. Neural Computing and Applications, 2023, 35 (29): 21567 - 21601.

[4] JIANG X. Intelligent classification method of archive data based on multigranular semantics [J]. Computational Intelligence and Neuroscience, 2022, 2022 (1): 7559523.

[5] YANG P, SUN X, LI W, et al. SGM: sequence generation model for multi-label classification [J]. arXiv preprint arXiv: 1806.04822, 2018.

[6] RUBIN V L, CHEN Y, CONROY N K. Deception detection for news: three types of fakes [C]//Proceedings of the Association for Information Science and Technology, 2015, 52 (1): 1 - 4.

[7] PU W, LIU N, YAN S, et al. Local word bag model for text categorization [C]//Seventh IEEE International Conference on Data Mining (ICDM 2007), 2007: 625 - 630.

[8] ROBERTSON A M, WILLETT P. Applications of n-grams in tex-

tual information systems [J]. Journal of Documentation, 1998, 54 (1): 48 – 67.

[9] PRAKHYA S, VENKATARAM V, KALITA J. Open set text classification using CNNs [C]// Proceedings of the 14th International Conference on Natural Language Processing (ICON-2017), 2017: 466 – 475.

[10] LIU G, GUO J. Bidirectional LSTM with attention mechanism and convolutional layer for text classification [J]. Neurocomputing, 2019, 337: 325 – 338.

[11] SUN C, QIU X, XU Y, et al. How to fine-tune bert for text classification? [C]//Chinese Computational Linguistics: 18th China National Conference, 2019: 194 – 206.

[12] XIONG Y, FENG Y, WU H, et al. Fusing label embedding into BERT: an efficient improvement for text classification [C]//Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, 2021: 1743 – 1750.

[13] RADFORD A. Improving language understanding by generative pre-training [R]. 2018.

[14] GAO C, ZHENG Y, LI N, et al. A survey of graph neural networks for recommender systems: challenges, methods, and directions [J]. ACM Transactions on Recommender Systems, 2023, 1 (1): 1 – 51.

[15] GIRALDO J H, JAVED S, WERGHI N, et al. Graph CNN for moving object detection in complex environmentsfrom unseen videos [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 225 – 233.

[16] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks [J]. arXiv preprint arXiv: 1609. 02907, 2016.

[17] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks [J]. arXiv preprint arXiv: 1710. 10903, 2017.

[18] TIAN X, QIN Y, HUANG R, et al. Alabel information aware model for multi-label text classification [J]. Neural Processing Letters, 2024, 56 (5): 242.

[19] XIN Y, ZHANG Z. An improved Chinese text multi-label classification method based on CNN [J]. Journal of Physics Conference Series, 2020, 1619: 012017.

[20] CORTES C. VAPNIK V. Support-vector networks [J]. Machine Learning, 1995, 20 (3): 273 – 297.

[21] LOH W Y. Classification and regression trees [J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2011. DOI: 10. 1002/widm. 8.

[22] ELMAN J L. Distributed representations, simple recurrent networks, and grammatical structure [J]. Machine Learning, 1991, 7: 195 – 225.

[23] BROWNLEE J. Multi-label classification with deep learning [Z]. Deep Learning, 2020.

[24] YAO L, MAO C, LUO Y. Graph convolutional networks for textclassification [C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33: 7370 – 7377.

[25] MIKOLOV T, SUTSKEVER I, CHEN K, et al. Distributed representations of words and phrases and their compositionality [C]// Advances in Neural Information Processing Systems, 2013: 3136 – 3144.

[26] HUANG P S, HE X, GAO J, et al. Learning deep structured semantic models for web search using clickthrough data [C]// Proceedings of the 22nd ACM International Conference on Information & Knowledge Management, 2013: 2333 – 2338.

[27] LEE C W, FANG W, YEH C K, et al. Multi-label zero-shot learning with structured knowledge graphs [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 1576 – 1585.

（收稿日期：2025 – 03 – 04）

作者简介：

王哲（2001 – ），男，硕士研究生，主要研究方向：机器学习、自然语言处理。

温秀梅（1972 – ），通信作者，女，硕士，教授，主要研究方向：图神经网络、数据挖掘。E-mail：xiumeiwen@ 163. com。

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部