计算机研究与发展
Journal of Computer Research and Development
https://crad.ict.ac.cn
63(1): 66−89, 2026

多示例多标记学习综述

王梅霞[1] 王慧颖[2] 黄苗苗[1] 赵宇海[1]
[1](东北大学计算机科学与工程学院 沈阳 110819)
[2](国网辽宁省电力有限公司信息通信分公司 沈阳 110065)
(sonic_wmx@163.com)

Survey on Multi-Instance Multi-Label Learning

Wang Meixia[1], Wang Huiying[2], Huang Miaomiao[1], and Zhao Yuhai[1]
[1](*School of Computer Science and Engineering*, *Northeastern University*, *Shenyang* 110819)
[2](*State Grid Liaoning Information and Communication Company*, *Shenyang* 110065)

Abstract Multi-instance multi-label learning (MIML) is an important machine learning framework designed to address complex tasks where each object is represented as a bag of multiple instances and can be associated with multiple labels simultaneously. It has been widely applied in domains such as drug activity prediction and medical image analysis. Different from traditional multi-instance learning and multi-label learning, MIML simultaneously captures the diversity in both the input space (multiple instances) and the output space (multiple labels), enabling more comprehensive modeling and understanding of real-world complex objects. However, this expressiveness also brings significant challenges, including high computational complexity, difficulties in model optimization, and limited generalization capabilities. Therefore, MIML has attracted increasing attention from the research community. Despite this growing interest, a comprehensive survey of MIML is still lacking. We first introduce the formal problem definition and notations related to MIML. Then, from the perspective of data complexity, we categorize MIML problems into three main types: standard MIML, multi-source MIML, and MIML with imprecise labeling. These are further refined into nine subcategories based on solution strategies, instance sources, and label completeness, which are systematically discussed and analyzed. We also summarize commonly used datasets and benchmark comparisons for MIML methods. Finally, we present five representative application scenarios of MIML in real-world tasks, outline four promising future research directions, and conclude the paper.

Key words multi-instance multi-label learning (MIML); multi-view; imprecise labeling; degradation methods; direct solution methods; active learning

摘 要 多示例多标记学习(multi-instance multi-label learning, MIML)是一种重要的机器学习框架,用于处理每个对象被表示为多个示例构成的多示例包,并且是一种能够同时与多个标记相关联的复杂学习任务,被广泛应用于药物活性预测、医学图像分析等领域。与传统的多示例学习和多标记学习不同,MIML同时考虑了输入空间(多示例)和输出空间(多标记)的多样性,使得模型能更全面地描述和理解现实世界的复杂对象,但同样也面临着计算复杂性高、模型难以优化和泛化能力受限等挑战。因此,MIML问题受到了研究者们的广泛关注。然而,目前尚缺乏对MIML研究的完整综述。首先给出与MIML相关的问题和符号定义;其次从数据复杂性的角度将MIML划分为标准MIML、多元MIML和非精确标记的

收稿日期: 2024-08-01; **修回日期**: 2025-07-11
基金项目: 国家自然科学基金项目(62432003, 92267206, 62032013)
This work was supported by the National Natural Science Foundation of China (62432003, 92267206, 62032013).
通信作者: 赵宇海(zhaoyuhai@mail.neu.edu.cn)

MIML 三类，并且分别从问题求解策略、示例来源和标记完整性的角度，将 3 类 MIML 细化为 9 个小类进行介绍和分析；然后给出了 MIML 方法的常用数据集和实验对比结果；最后介绍了 5 种常见的 MIML 实际应用场景，展望了 MIML 领域的 4 个未来研究方向并对全文进行总结。

关键词 多示例多标记学习；多视图；非精确标记；退化方法；直接求解方法；主动学习

中图法分类号 TP181

DOI: 10.7544/issn1000-1239.202440643 **CSTR:** 32373.14.issn1000-1239.202440643

随着信息技术的迅速发展，机器学习受到越来越多的关注。机器学习旨在建立输入空间和输出空间之间的映射关系，这种映射关系可能是简单的 1 对 1 关系，也可能是复杂的多对 1、1 对多和多对多关系。具体来说，单示例单标记学习（single-instance single-label learning, SISL）旨在建模单个特征示例和单个类别标记之间的 1 对 1 映射关系。多示例学习[1-2]用 1 组示例表示单个对象，旨在建模 1 组特征示例和单个类别标记之间的多对 1 映射关系，适用于处理输入空间多样性的数据。多标记学习[3-4]允许每个对象同时具有多个标记，旨在建模单个特征示例和多个类别标记之间的 1 对多映射关系，适用于处理输出空间多样性的数据。多示例多标记学习（multi-instance multi-label learning, MIML）结合了多示例学习和多标记学习的特点，将每个对象表示为多个示例构成的多示例包[5]，并且能够同时与多个标记相关联。MIML 旨在建模 1 组特征示例和多个类别标记之间的多对多映射关系，能够处理同时具有输入空间和输出空间多样性的数据，能更确切地描述现实世界的复杂对象，提升机器学习模型的性能[6]。例如，在计算机视觉领域，每幅图像通常包含多个物体，并且同时具有多个标记，通过将每幅图像表示为 1 个多示例包以及图像中各个区域表示为包中的 1 个示例来预测图像所具有的多个类别标记[7]。尽管 MIML 结合了多示例学习和多标记学习的优势，但是多示例学习方法和多标记学习方法及其组合往往无法同时考虑多个特征示例和多个类别标记，以及它们之间的复杂关系和依赖性。同时，多个特征示例和多个类别标记的组合使得模型参数空间庞大，导致 MIML 通常面临着计算复杂性高、模型难以优化和泛化能力受限等挑战。因此，MIML 问题受到了诸多学者的广泛关注，具有重要的研究意义。

目前已有一些有关多示例学习或多标记学习的综述，如多示例学习算法综述[8-9]、多示例学习问题研究进展综述[10]、数字病理学研究的多示例学习综述[11]、自动代码审查的深度多示例学习综述[12]、医学图像分析中的多示例学习综述[13]、多标签分类综述[14]、基于监督和半监督学习的多标记分类综述[15]以及多标记深度学习综述[16]等。但是这些综述倾向于研究多示例学习或多标记学习及其在特定应用场景中的研究和发展现状，还没有关于 MIML 研究方法和发展现状及其未来研究方向的综述。为了填补在 MIML 综述方面的研究空白，本文梳理总结了 MIML 的各类研究方法和这些方法之间的区别与联系，并总结了 MIML 的常用数据集、典型应用场景及其未来研究方向。

本文首先根据数据复杂性，将 MIML 任务划分为 3 类：1）标准 MIML 任务。该类任务是最基本的 MIML 问题形式，即每个样本被表示为多个示例构成的包，并且每个包均与一组预定义的标记相关联，其目标是预测每个包的正确标记集合。这类任务通常假设多示例包和包标记都是明确且一致的，即包和标记之间的关系是固定的，不存在额外的复杂性或不确定性。2）多元 MIML 任务。该类任务在标准 MIML 任务的基础上考虑了特征空间的复杂性，即允许每个样本包含来自多个不同源头或具有不同表示形式的示例。3）非精确标记的 MIML 任务。该类任务在标准 MIML 任务的基础上考虑了标记空间的不确定性，即专注于在标记信息不完整、不确定或存在异常时，解决更为复杂的 MIML 问题。

其次，针对标准 MIML 任务，根据不同的问题求解策略将其进一步划分为 3 类：1）基于退化的 MIML 方法。该类方法旨在识别传统监督学习问题的等价物，通过分解示例和标记以多示例学习或多标记学习为中间问题，将 MIML 问题退化为传统的监督学习问题进行求解。根据其退化过程，该类方法又可细分为基于多示例学习的退化方法和基于多标记学习的退化方法 2 类，它们分别以多示例学习和多标记学习为中间问题进行退化。这类方法比较直观，可以简化问题的求解难度和复杂性，但是容易导致示例和标记层面的信息损失。2）基于直接求解的 MIML 方法。该类方法旨在以示例特征或包特征为

基础,通过学习特征和标记之间的关系,直接建模多示例包和多个类别标记之间的映射关系。根据其特征的表达形式,该类方法又可细分为基于包的方法和基于示例的方法2类,它们分别以包特征和示例特征为基础进行求解。这类方法的核心思想是直接求解复杂的MIML问题,能够有效地考虑示例和标记层面的相关性,不易造成信息损失。3)基于主动学习的MIML方法。该类方法不再是被动地接受所有的训练样本进行学习,而是根据一定的策略,主动选择那些对模型性能提升影响最大的样本进行标记和学习,能更有效地发现数据中的规律和模式,提升分类性能和降低标注成本。但是这种方法的学习效果很大程度上取决于初始模型性能。

然后,针对多元MIML任务,根据不同的示例来源进一步将其划分为2类:1)多视图MIML。该类任务在标准MIML任务的基础上引入了多视图信息,能够从多个角度捕捉样本特征,并利用不同视图之间的互补信息来捕捉更丰富的语义。2)多类型对象的多视图MIML。该类任务在多视图MIML的基础上进一步引入了多类型对象信息。这些对象通过关系矩阵相互连接,形成异构多示例网络,使模型不仅能够捕捉单一类型对象内部的特征和关系,还能够利用与目标对象相关联的其他类型对象信息来辅助学习。通过整合多类型对象信息,模型可以更全面地捕获数据中丰富的语义。

最后,针对非精确标记的MIML任务,根据标记的完整性将其进一步划分为2类:1)新标记学习。该类任务在标准MIML任务的基础上引入了新标记识别的需求,不仅能够处理训练阶段使用过的已知标记,还能够处理训练阶段未出现但在实际应用中可能出现的新标记。通过利用标记之间的语义关联和模型的泛化能力,模型能够适应动态变化的标记集合,同时对已知和未知的标记进行识别和预测。2)缺标记学习。该类任务在标准MIML任务的基础上进一步引入了缺失标记处理需求,这要求模型能够应对数据集中标记信息不准确、部分缺失或带有噪声等情况。通过利用标记之间的关系和上下文信息,模型能够适应未被精确标注的标记集合,补全标记信息并提高标记预测的完整性和准确性。

虽然围绕MIML展开了很多研究,但是仍然存在很多问题和挑战。从数据处理、问题求解和未来发展的角度可以分为4类:1)数据构建。现有的MIML方法通常依赖于手动构建MIML数据,难以从原始的图像、文本、视频、音频等数据中自动构建MIML格式的数据进行端到端的学习。2)数据融合。现有的MIML方法通常通过最大池化和注意力机制来融合多个示例的信息,但这些方法往往依赖于特定的假设,难以在更广泛的场景中有效应用。3)标记预测。MIML模型在训练过程中只能使用包标记而不能使用示例标记,导致次优的包级和示例级多标记分类性能。4)社区发展。现有的MIML方法在实验标准化、透明化和可复现性方面仍存在不足之处。

1 相关问题与符号定义

为了更好地描述本文介绍的内容,首先用图1来展示MIML与单示例学习、多标记学习和多示例学习之间的区别与联系。

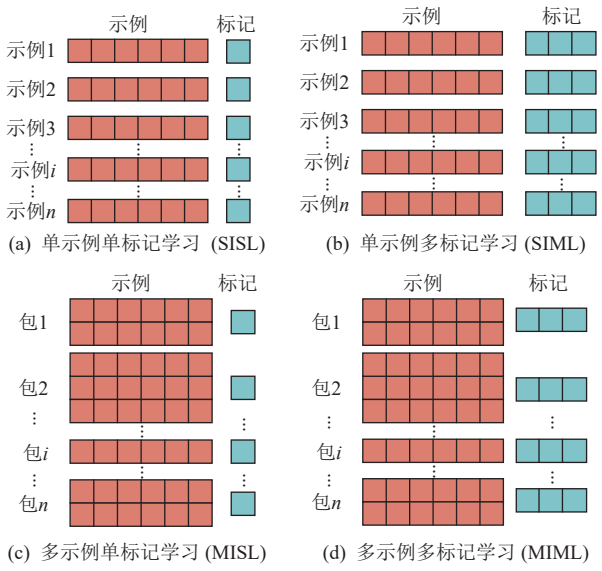


Fig. 1 Comparison of different learning frameworks

图1 不同学习框架对比

1)SISL如图1(a)所示。具体来说,给定一组由N个样本构成的数据集$\{(\boldsymbol{x}_1,y_1),(\boldsymbol{x}_2,y_2),\cdots,(\boldsymbol{x}_N,y_N)\}$。其中$\boldsymbol{x}_i\in\mathcal{X}$, $y_i\in\mathcal{Y}$, $\mathcal{X}$表示特征空间,$\mathcal{Y}$表示标记空间。SISL任务旨在利用示例$\boldsymbol{x}_i$和标记y_i通过机器学习方法获得示例和对应单个类别标记之间的映射关系$f:\mathcal{X}\to\mathcal{Y}$,并在通过训练得到映射关系模型后,用其预测未知的示例标记。

2)单示例多标记学习(single-instance multi-label learning, SIML)具有输出空间多样性,如图1(b)所示。具体来说,给定一组由N个样本构成的数据集$\{(\boldsymbol{x}_1,\boldsymbol{Y}_1),(\boldsymbol{x}_2,\boldsymbol{Y}_2),\cdots,(\boldsymbol{x}_N,\boldsymbol{Y}_N)\}$。其中$\boldsymbol{x}_i\in\mathcal{X}$, $\boldsymbol{Y}_i\in\mathcal{Y}$, $\boldsymbol{Y}_i=(y_{i1},y_{i2},\cdots,y_{iL})^{\mathrm{T}}$, $\mathcal{X}$表示特征空间,$\mathcal{Y}$表示标记空间,L表示数据集中的标记数,$\sum_{j=1}^{L}y_{ij}\geqslant 1$。如果示例$\boldsymbol{x}_i$具有

第j个标记，则$y_{ij}=1$，否则$y_{ij}=0$。SIML任务旨在利用示例$\boldsymbol{x}_i$和标记$\boldsymbol{Y}_i$通过机器学习方法获得单个示例和对应多个类别标记之间的映射关系$f:\mathcal{X}\to 2^{\mathcal{Y}}$，并在通过训练得到映射关系模型后，用其预测未知的示例标记。

3）多示例单标记学习（multi-instance single-label，MISL）具有输入空间多样性，如图1（c）所示。具体来说，给定一组由N个样本构成的数据集$\{(\boldsymbol{X}_1,y_1),(\boldsymbol{X}_2,y_2),\cdots,(\boldsymbol{X}_N,y_N)\}$。其中$\boldsymbol{X}_i\in\mathcal{X}$，$y_i\in\mathcal{Y}$，$\mathcal{X}$表示特征空间，$\mathcal{Y}$表示标记空间。多示例包$\boldsymbol{X}_i=(\boldsymbol{x}_{i1},\boldsymbol{x}_{i2},\cdots,\boldsymbol{x}_{in_i})^{\mathrm{T}}$，由$n_i$个示例构成。MISL任务旨在利用多示例包$\boldsymbol{X}_i$和包标记$y_i$通过机器学习方法获得多示例包和对应单个类别标记之间的映射关系$f:2^{\mathcal{X}}\to\mathcal{Y}$，并在通过训练得到映射关系模型后，用其预测未知的包标记和示例标记。

4）MIML结合了多示例学习和多标记学习的特点，同时具有输入空间和输出空间多样性，如图1（d）所示。具体来说，给定一组由N个样本构成的数据集$\{(\boldsymbol{X}_1,\boldsymbol{Y}_1),(\boldsymbol{X}_2,\boldsymbol{Y}_2),\cdots,(\boldsymbol{X}_N,\boldsymbol{Y}_N)\}$。其中$\boldsymbol{X}_i\in\mathcal{X}$，$\boldsymbol{Y}_i\in\mathcal{Y}$，$\mathcal{X}$表示特征空间，$\mathcal{Y}$表示标记空间。多示例包$\boldsymbol{X}_i=(\boldsymbol{x}_{i1},\boldsymbol{x}_{i2},\cdots,\boldsymbol{x}_{in_i})^{\mathrm{T}}$由$n_i$个示例构成，相应包标记$\boldsymbol{Y}_i=(y_{i1},y_{i2},\cdots,y_{iL})^{\mathrm{T}}$，$\sum_{j=1}^{L}y_{ij}\geqslant 1$，$L$表示数据集中的标记数。如果包$\boldsymbol{X}_i$有第$j$个标记，$y_{ij}=1$，否则$y_{ij}=0$。MIML任务旨在利用多示例包$\boldsymbol{X}_i$和包标记$\boldsymbol{Y}_i$，通过机器学习的方式获得多示例包和对应多个类别标记之间的映射关系$f:2^{\mathcal{X}}\to 2^{\mathcal{Y}}$，并在通过训练得到映射关系模型后，用其预测未知的包标记和示例标记。

2 标准MIML

如前文所述，MIML任务根据数据复杂性可以分为3类，本节主要介绍标准MIML任务。标准MIML任务假设输入多示例包仅来自单个视图，并且输出标记是完整的和准确的。其目标是根据训练数据中的多示例包和相应的多个类别标记训练特征分类器，能够有效地根据输入多示例包的不同判断它们所属的多个类别。根据不同的问题求解策略可以将标准MIML任务的求解方法进一步划分为基于退化的MIML方法、基于直接求解的MIML方法和基于主动学习的MIML方法3类分别展开介绍。

2.1 基于退化的MIML方法

基于退化的MIML方法旨在识别传统监督学习（SISL）的等价物，即将复杂的MIML任务退化为经典的监督学习任务进行求解，如图2所示。根据其退化过程又可细分为基于多示例学习的退化方法和基于多标记学习的退化方法2类。表1中总结了经典的基于退化的MIML方法及其所采用的机器学习技术。

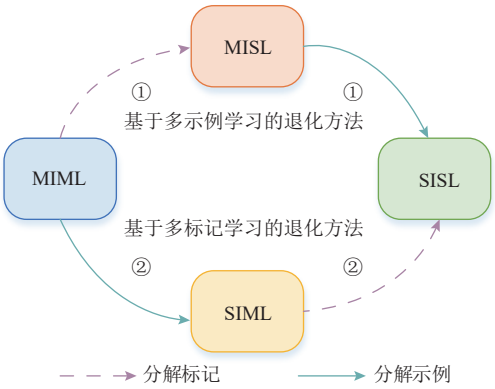


Fig. 2 Degradation based MIML

图2 基于退化的MIML

Table 1 Degradation Based MIML Methods

表1 基于退化的MIML方法

分类	方法	技术
基于多示例学习的退化方法	MIMLBOOST[5]	集成方法
	文献[17]	其他方法
基于多标记学习的退化方法	MIMLSVM[5]	核方法
	SISL-MIML[18]	核方法
	MIMLSVM+[19]	核方法
	E-MIMLSVM+[19]	核方法
	MIMLDML[20]	核方法
	En-MIMLSVM[21]	核方法
	CPNMIML[22]	其他方法
	MIML-ELM[23]	其他方法

2.1.1 基于多示例学习的退化方法

基于多示例学习的退化方法以多示例学习为基础将MIML问题退化为SISL问题进行求解，如图2中①所示。该方法首先为多示例包分配单个标记，将MIML问题$f_{\mathrm{MIML}}:2^{\mathcal{X}}\to 2^{\mathcal{Y}}$退化为多示例学习问题，学习多示例包到单个包标记的映射函数$f_{\mathrm{MISL}}:2^{\mathcal{X}}\times\mathcal{Y}\to\{1,0\}$。其中对于任意的$y\in\mathcal{Y}$，$y\in Y_i^+$时$f_{\mathrm{MISL}}(\boldsymbol{X}_i,y)=1$，否则$f_{\mathrm{MISL}}(\boldsymbol{X}_i,y)=0$，$Y_i^+$是$\boldsymbol{X}_i$的正标记集。然后，将多示例学习进一步蜕变为传统的监督学习任务进行求解，学习单个特征示例到单个类别标记的映射函数$f_{\mathrm{SISL}}:\mathcal{X}\times\mathcal{Y}\to\{1,0\}$，即在某种约束下将目标函数$f_{\mathrm{MISL}}(\boldsymbol{X}_i,y)$转化为$f_{\mathrm{SISL}}(\boldsymbol{x}_{ij},y)$，其中$j\in\{1,2,\cdots,n_i\}$，并且对于任意的$y\in\mathcal{Y}$，如果$y\in Y_i^+$，那

么$f_{\mathrm{SISL}}(\boldsymbol{x}_{ij},y)=1$, 否则$f_{\mathrm{SISL}}(\boldsymbol{x}_{ij},y)=0$, $\boldsymbol{x}_{ij}$可以由$\boldsymbol{X}_i$通过分解和筛选等方式获得。

Zhou 等人[5]提出 MIML 问题, 并提出一种基于多示例学习的退化方法 MIMLBOOST。该方法首先将 MIML 问题$\{(\boldsymbol{X}_1,\boldsymbol{Y}_1),(\boldsymbol{X}_2,\boldsymbol{Y}_2),\cdots,(\boldsymbol{X}_N,\boldsymbol{Y}_N)\}$退化为 MISL 问题$\{[(\boldsymbol{X}_1,y_1),\psi(\boldsymbol{X}_1,y_1)],[(\boldsymbol{X}_1,y_2),\psi(\boldsymbol{X}_1,y_2)],\cdots,[(\boldsymbol{X}_N,y_L),\psi(\boldsymbol{X}_N,y_L)]\}$。其中$(\boldsymbol{X}_i,y_j)$表示包含$n_i$个示例的多示例包$\{(\boldsymbol{x}_{i1},y_j),(\boldsymbol{x}_{i2},y_j),\cdots,(\boldsymbol{x}_{in_i},y_j)\}$, $\psi(\boldsymbol{X}_i,y_j)\in\{1,0\}$表示$(\boldsymbol{X}_i,y_j)$的标记。然后, 采用 MIBOOSTING[24]算法通过将包标记$\psi(\boldsymbol{X}_i,y_j)$分配给包中的每个示例, 将 MISL 问题进一步蜕变为 SISL 问题进行求解。MIMLBOOST 算法通过多示例学习将复杂的 MIML 问题分解为简单的 SISL 问题, 算法实现简单。但是, 退化过程中容易损失标记之间、示例之间以及示例与标记之间的依赖关系, 导致分类效果不理想, 并且为每个示例分配单个标记的方式会产生较高的计算成本, 使其在处理大规模 MIML 问题时存在明显不足。此外, Wang 等人[17]从理论层面对 MIML 的可学习性进行了探索, 通过证明 MISL 的泛化边界将 MIML 视为多个具有标记间相关性的 MISL 任务, 并证明了由多示例单标记假设类构建的 MIML 假设类是 PAC(probably approximately correct)可学习的。

总体而言, 基于多示例学习的退化方法可以逐步简化问题的求解难度并利用多示例学习的技术和理论基础。但是退化过程可能会丢失关键信息, 尤其是标记之间的关联性。更适用于标记间关系较弱且计算资源有限的场景。

2.1.2 基于多标记学习的退化方法

基于多标记学习的退化方法以多标记学习为基础将 MIML 问题退化为 SISL 问题进行求解, 如图 2 中②所示。该方法首先通过合适的问题转换函数ϕ, 将 MIML 问题$f_{\mathrm{MIML}}:2^{\mathcal{X}}\to 2^{\mathcal{Y}}$退化为 SIML 问题$f_{\mathrm{SIML}}:\mathcal{Z}\to 2^{\mathcal{Y}}$, 即学习一个有效的转换函数$\phi:2^{\mathcal{X}}\to\mathcal{Z}$, 使得对于任意的$\boldsymbol{z}_i\in\mathcal{Z}$存在$f_{\mathrm{SIML}}(\boldsymbol{z}_i)=f_{\mathrm{MIML}}(\boldsymbol{X}_i)$。然后, 进一步将多标记学习问题蜕变为经典的机器学习问题$f_{\mathrm{SISL}}:\mathcal{Z}\times\mathcal{Y}\to\{1,0\}$进行求解。其中对于任意的$y\in\mathcal{Y}$, 如果$y\in Y_i^+$, $f_{\mathrm{SISL}}(\boldsymbol{z}_i,y)=1$, 否则$f_{\mathrm{SISL}}(\boldsymbol{z}_i,y)=0$。根据退化后所采用的经典机器学习技术的不同, 将基于多标记学习的退化方法分为 2 类分别进行介绍。

1)核方法

核方法采用核技术来解决退化后的 SISL 问题。Zhou 等人[5]提出基于多标记学习的退化方法 MIML-SVM, 该方法借助核方法解决退化后的 MIML 问题。首先对 MIML 数据$\{(\boldsymbol{X}_1,\boldsymbol{Y}_1),(\boldsymbol{X}_2,\boldsymbol{Y}_2),\cdots,(\boldsymbol{X}_N,\boldsymbol{Y}_N)\}$进行 k-medoids 聚类, 得到 k 个聚类中心$\{\boldsymbol{Q}_1,\boldsymbol{Q}_2,\cdots,\boldsymbol{Q}_k\}$。然后, 利用这些聚类中心, 将原始的多示例包$\boldsymbol{X}_i$转化为 k 维的数值向量$\boldsymbol{z}_i$。其中$\boldsymbol{z}_i$的第 j个值 ($j\in\{1,2,\cdots,k\}$)由$\boldsymbol{X}_i$到聚类中心$\boldsymbol{Q}_j$的距离构成, $\boldsymbol{z}_i$用于编码包$\boldsymbol{X}_i$与第 j个簇之间的关系。这样原始的 MIML 问题可以退化为多标记学习问题$\{(\boldsymbol{z}_1,\boldsymbol{Y}_1),(\boldsymbol{z}_2,\boldsymbol{Y}_2),\cdots,(\boldsymbol{z}_N,\boldsymbol{Y}_N)\}$。最后, 采用 MLSVM 方法[25]将多标记学习问题分解为多个独立的二分类问题, 采用支持向量机(support vector machine, SVM)技术进行求解。MIMLSVM 算法通过多标记学习将复杂的 MIML 问题分解为简单的 SISL 问题, 并采用 SVM 方法进行求解, 算法实现简单、时间效率高。但是, 将多标记学习问题分解为多个独立的二分类问题容易损失标记之间的相关性, 并且增加了额外的计算成本, 不适用于处理大规模的 MIML 数据。Nguyen[18]基于 SVM 提出一种新的 MIML 方法 SISL-MIML。该方法借助二次规划和整数规划, 通过同时为多示例包中的所有示例在标记集中寻找最匹配的单个标记, 显式地利用了示例与标记之间的联系, 并将 MIML 问题直接退化为 SISL 问题, 然后采用 SVM 方法进行求解。虽然 SISL-MIML 能够在一定程度上考虑示例与标记之间的关联性, 但却忽略了标记之间的潜在相关性, 限制了其在处理更复杂 MIML 场景时的性能。

随着研究的进一步深入, 学者们基于 MIMLSVM, 设计了一系列 MIML 方法。其中 Li 等人[19]提出 2 种基于 SVM 的 MIML 方法。①MIMLSVM+。该方法基于 MIMLSVM 的思想, 简单地采用了一种退化策略, 将多标记学习问题分解为一系列的二分类问题, 并为每个项构造 1 个 SVM 分类器, 但是这种方法忽略了标记之间的相关性。②E-MIMLSVM+。该方法在 MIMLSVM+的基础上, 通过基于核的多任务学习技术结合了标记之间的相关性, 扩展了 MIMLSVM+算法。但是这种方法由于同时学习多个任务, 导致更多的包参与优化过程, 消耗了大量的训练时间和存储空间, 分类效率不高。Xu 等人[20]提出一种用于 MIML 的距离度量框架 MIMLDML。该方法用马氏距离代替 MIMLSVM 中的欧氏距离, 同时考虑了标记数据和未标记数据之间的不平衡性, 但是这种方法在退化过程中容易损失示例和标记之间的相关性。Xu 等人[21]提出 En-MIMLSVM。该方法首先从多数类中独立地抽样几个子集, 然后使用这些子集和少数类来训练多个分类器, 将所构造的分类器组合起来得到最终的决策。En-MIMLSVM 在 MIMLSVM 的基础上融入了集成学习的思想, 旨在解决类不平衡

问题，但在应对数据退化过程中产生的信息损失时，其效果仍显不足。

2）其他方法

神经网络和极限学习机（extreme learning machine, ELM）方法也常用于解决退化后的 MIML 问题。Yan 等人[22]提出一种将概率潜在语义分析与神经网络相结合的方法 CPNMIML，利用概率潜在语义分析模型将 MIML 问题转化为 SIML 问题，然后采用神经网络进行求解。该方法所采用的概率潜在语义分析模型可以学习到潜在的语义相关性，降低噪声的影响，但是学习过程中忽略了标记之间的相关性。Yin 等人[23]基于 ELM 提出一种新的 MIML 方法 MIMLELM。该方法首先利用 Akaike 信息量准则（Akaike information criterion, AIC）自动确定集群数量。其次，通过 *k*-medoids 聚类将 MIML 问题转化为多标记学习问题。然后，通过在训练过程中多次使用每个标记数据，将多标记学习问题退化为经典的 SISL 问题。最后，使用 ELM 分类器进行求解。MIMLELM 方法能够借助 AIC 自动确定集群数量，并使用 ELM 分类器进一步提升分类性能，但是退化过程仍然容易引起示例和标记级的信息损失。

总体而言，基于多标记学习的退化方法能够简化问题的求解难度和复杂性并利用多标记学习的技术和理论基础。但是退化过程中容易损失信息，尤其是示例之间的结构信息。更适用于示例间结构较简单且计算资源有限的场景。

综上所述，基于退化的 MIML 方法，是一种直观的解决方法，分别以多示例学习和多标记学习为基础，通过问题退化来识别传统监督学习问题的等价物，并将复杂的 MIML 问题蜕变为传统的监督学习问题进行求解。这类方法的优点在于简单直观，但是退化过程往往无法有效考虑示例间、标记间以及示例与标记间的依赖关系，容易造成信息损失。更适用于数据特征独立、标记关系较弱且计算资源有限的任务。

2.2 基于直接求解的 MIML 方法

2.1 节介绍的基于退化的 MIML 方法以多示例学习或者多标记学习为基础，将 MIML 问题退化为经典的监督学习问题，然后采用经典的机器学习方法进行求解。尽管这种方法可以简化问题的求解难度，并且能够有效地利用现有的机器学习技术，但是在退化过程中容易产生信息损失。理想情况下，应该考虑到示例、包和标记之间的各种依赖关系。随着机器学习的快速发展，基于直接求解的 MIML 方法不断涌现。它们不需要显示地借助多示例学习或者多标记学习将 MIML 问题退化为传统的监督学习问题，而是以示例特征或包特征为单位，通过学习特征和标记之间的关系，直接建模多示例包和多个类别标记之间的映射关系，解决复杂的 MIML 问题。本节重点关注基于直接求解的 MIML 方法，并基于特征的表达形式将其划分为基包的方法和基于示例的方法 2 类分别展开介绍。

2.2.1 基于包的方法

基于包的方法以包特征为基本单位进行求解，而无需将 MIML 问题退化为 SISL 问题，如图 3 中①所示。该类方法首先通过示例信息学习包特征表示，然后通过学习包特征和包标记之间的关系来建模多示例包和多个类别标记之间的映射关系。根据其求解过程中所采用的机器学习技术大致可以分为 4 类，如表 2 所示。

1）最近邻方法

基于最近邻的方法借助示例特征计算不同包之

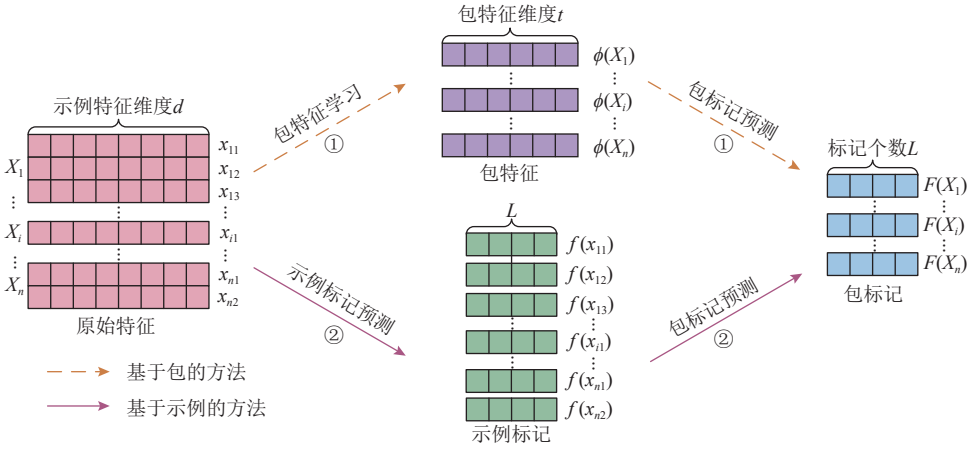


Fig. 3 MIML methods based on direct solution

图 3 基于直接求解的 MIML 方法

Table 2 MIML Methods Based on Bag

表 2 基于包的 MIML 方法

类别	方法	思路
最近邻方法	MIML-*k*NN[26]	基于平均霍夫距离确定包的邻居集合并利用 *k*NN 方法进行求解
	Markov-MIML[27]	基于马尔可夫概率转移矩阵计算类概率，并利用近邻信息学习正确标记
	文献 [28]	将 3 种霍夫距离和 5 种基于 *k*NN 的多标记学习算法进行组合来解决 MIML 问题
径向基神经网络方法	MIMLRBF[29]	利用 RBF 神经网络在第 1 层和第 2 层分别进行 *k*-medoids 聚类和 SVD 的权值优化
	IMIMLRBF[30]	利用数据密度优化 MIMLRBF 网络第 1 层的质心，以应对样本不平衡问题
	MIMLRBF-SD[31]	在 MIMLRBF 网络第 2 层采用“SVD+SD”的方法，避免数据噪声的影响
	MIMLRBF-GBMO[32]	利用 GBMO 和基于梯度的快速收敛参数估计方法来提高 RBF 网络的收敛性和精度
	文献 [33]	在第 2 层将随机梯度下降和动量方法相结合以避免 SVD 方法所造成的误差
	文献 [34]	将 RBF 神经网络中的 RBF 设为一种广泛应用于插值的多次函数进行求解
深度学习方法	文献 [35]	借助稀疏字典和深度神经网络进行 MIML 学习，以有效考虑示例与标记之间的关系
	MIML-LLMC[36]	借助并行多注意力机制和局部标记流形相关性求解 MIML 问题
其他方法	MK_MIML[37]	根据不同尺度的参数和图构造多个核函数，并进行多核融合，然后利用 SVM 分类
	KISAR[38]	利用特征向量映射和稀疏编码器，能够发现哪个示例触发了哪个标记

间的距离来得到包特征表示，然后利用 k 最近邻（k-nearest neighbor, kNN）算法进行求解。Zhang[26] 提出一种基于 kNN 的 MIML 方法 MIML-kNN。该方法首先基于平均霍夫距离，通过度量示例之间的欧氏距离来衡量不同包之间的距离。其次，基于不同包之间的距离确定针对多示例包 $\boldsymbol{X}_i$ 的 r 近邻集合和将多示例包 $\boldsymbol{X}_i$ 视为其 c 近邻的包集。其中将多示例包 $\boldsymbol{X}_i$ 视为其邻居的那些实例被称为“引用者”。然后，根据包 $\boldsymbol{X}_i$ 的邻居和引用者的标记信息构建一个标记计数向量作为包 $\boldsymbol{X}_i$ 的特征表示。最后，将包 $\boldsymbol{X}_i$ 的特征表示经过线性分类器进行预测。该方法的优势在于借助 kNN 算法充分利用了示例与标记之间的相关性，但是这种方法独立地处理每个标记，并未考虑标记之间的依赖关系。

针对 MIML-kNN 算法，诸多研究者进行了深入的改进工作。例如，Wu 等人[27] 为了避免 MIML 数据学习中所需的极高计算成本，并保持良好的分类性能，对 MIML-kNN 算法进行了优化，提出一种基于马尔可夫链的 MIML 算法 Markov-MIML。该方法首先在示例之间构造亲和矩阵来表达示例之间的亲和关系。然后，利用示例与对象之间的关系矩阵将示例之间的亲和信息传递到对象之间，得到包特征表示。最后，通过马尔可夫链从待学习对象的邻居中接收标记信息来更新自身的标记，直至收敛。近年来，Zafra 等人[28] 为了在 MIML-kNN 算法中考虑标记之间的关系并且有效提升算法的分类性能，提出一种新的基于 kNN 的 MIML 算法来直接处理 MIML 数据。该方法认为对象间距离的度量是影响 kNN 算法最大的因素之一。因此，采用最大霍夫距离、最小霍夫距离和平均霍夫距离来计算任意 2 个包之间的距离，并且分别采用 5 种多标记学习算法 BRkNN, MLkNN, DMLkNN, IBLR_ML, MLDGC 来解决 MIML 问题。这些基于 kNN 的多标记学习算法可以更好地保留标记之间的相关性。

2）径向基神经网络方法

基于径向基函数（radial basis function, RBF）神经网络的方法在隐藏层通过计算多示例包与基函数中心点之间的距离来得到包特征表示，然后在输出层学习包与包标记之间的映射关系。Zhang 等人[29] 提出一种基于 RBF 神经网络的 MIML 算法 MIMLRBF。该算法由 1 个 2 层的神经网络构成，第 1 层为聚类阶段，利用示例特征并基于平均霍夫距离和 k-medoids 聚类来学习聚类中心，通过计算多示例包与基函数中心点之间的距离来学习包特征表示。第 2 层为优化阶段，利用奇异值分解（singular value decomposition, SVD）方法对权值进行优化。MIMLRBF 算法的神经网络结构如图 4 所示。其中输入是多示例包 $\boldsymbol{X}_i=(\boldsymbol{x}_{i1},\boldsymbol{x}_{i2},\cdots,\boldsymbol{x}_{in_i})^{\mathrm{T}}$，每个示例 $\boldsymbol{x}_{ij}$ 有 1 个 d 维的特征向量 $(x_{ij1},x_{ij2},\cdots,x_{ijd})^{\mathrm{T}}$。该神经网络的第 1 层有 L 组包 $\bigcup_{l=1}^{L}\{C_1^l,C_2^l,\cdots,C_{M_l}^l\}$。其中 M_l 表示第 l 个类所保留的包个数，第 1 层的包总数为 $M=\sum_{l=1}^{L}M_l$。该神经网络的每个输出单元都与一个可能的类标记相关，第 1 层和第 2 层之间的权值 $\boldsymbol{W}=(w_{jl})_{(M+1)\times L}$ 由基函数到输出单元之间

的线表示。同时偏置项被表示为来自额外基函数ϕ_0的权重，它的输出被固定为1。该方法采用RBF神经网络来直接解决MIML问题并且可以显式地利用示例和标记之间的关系。

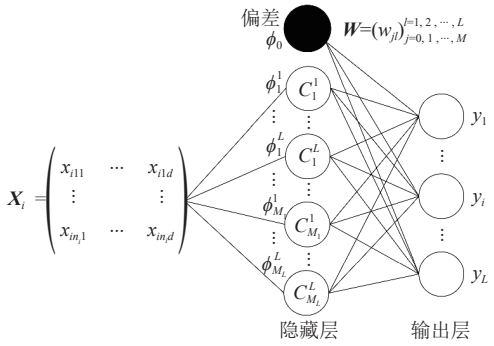


Fig. 4 Neural network structure of MIMLRBF [29]

图4 MIMLRBF神经网结构[29]

针对MIMLRBF方法，诸多学者探索了多样化的改进策略。其中Li等人[30]认为当某些类的样本数量比其他类小得多时就会出现不平衡问题，影响算法的分类性能。为此提出IMIMLRBF方法，通过平衡不同类别的隐单元数来缓解这一问题。具体来说，该方法在第1个聚类阶段根据数据密度计算每个类别的初始质心，并用*k*-medoids方法来优化质心。然后，利用这些调整好的单元，在第2阶段采用SVD方法优化第1层和第2层之间的权值。改进后的方法可用于处理不平衡样本问题。同时，Li等人[31]发现当训练数据有噪声或不容易识别时，用于计算输出层权值的SVD方法会导致网络性能的整体误差增强。为此提出一种MIMLRBF-SD方法。该方法在第2层权重优化阶段采用SVD方法初始化权重之后，采用最速下降法(steepest descent, SD)对权值进行优化，以避免在训练数据有噪声或不容易识别时，用于计算输出层权重的SVD方法会导致网络性能的整体误差增强。

Abdechiri等人[32]认为RBF神经网络虽然可以建模示例与标记之间的复杂关系，但是参数估计比较困难。为了克服这一难题，提高RBF神经网络的收敛性和预测精度，提出一种由气体布朗运动优化(gases Brownian motion optimization, GBMO)算法和基于梯度的快速收敛参数估计方法组成的新混合算法MIMLRBF-GBMO，来同步优化RBF神经网络的关键参数。该算法充分利用了GBMO算法的鲁棒搜索能力，在广阔的参数空间中高效搜索最优解，同时结合梯度下降法的效率优势，对参数进行精细化调整。具体来说，该方法首先利用GBMO算法来确定网络节点的宽度和中心。然后，针对每个输入量固定、隐藏节点数固定的分子，采用结构化非线性参数优化方法对参数进行优化，以确保参数调整的精确性和效率。近年来，Bao等人[33]针对MIMLRBF中所存在的关键问题进行了深入探索与优化。具体来说，为了避免直接使用SVD方法可能引入的误差，在参数求解过程中，将随机梯度下降和动量方法相结合用于解决隐藏层和输出层的权重优化问题。并进一步认识到RBF神经网络中隐藏层的激活函数是一个高斯函数，其性能取决于中心位置、自由基函数的宽度和权值。为了克服传统MIMLRBF方法在这一方面的局限性，他们创新性地将RBF拓展为多次函数，这种函数形式在插值领域有着广泛的应用基础，从而为RBF神经网络的求解提供了更为灵活和强大的工具[34]。

3)深度学习方法

基于包的方法中，部分方法基于深度学习技术通过学习包特征表示进行包标记预测。例如，Zhang等人[35]提出一种基于稀疏编码和深度神经网络的MIML算法。该方法首先基于示例特征学习一个稀疏字典，并基于稀疏字典对每个示例完成稀疏编码。然后，将每个示例的残差拼接起来构成包特征表示。最后，使用深度前馈神经网络进行包标记预测。然而，该方法认为不同示例之间不存在任何关系，即满足独立同分布的假设，这在一定程度上限制了其处理复杂依赖关系的能力。Yang等人[36]提出一种基于并行注意力和局部标记流形相关性的算法MIML-LLMC。该算法首先基于并行多注意力机制将每个包转换为多个融合向量，每个标记对应1个向量，注意力机制的使用有助于识别包中会触发各自的标记的示例，以提高方法的可解释性。然后，基于标记流行的生成和聚类构建了局部标记流行相关向量，以提供标记向量的高水平表示。最后，将每个融合向量连接到局部标记流行相关向量上作为包特征表示，以独立地预测每个包标记。通过对注意力和局部标记流行相关向量的利用，MIML-LLMC能够在发现示例-标记关系的同时利用局部标记相关性。

4)其他方法

除了上述3种方法之外，也有许多其他方法以包特征为单位来直接求解标准MIML问题。例如，Chen等人[37]提出一种基于多核融合的MIML方法MK_MIML。该方法首先基于高斯距离和阈值为每个包推导出一个亲和矩阵构成图来描述包内示例之间的相关性。其次，通过高斯RBF核将所构造的图映射为高维空间的特征向量作为包特征表示，这样多示

例包就可以被转化为单示例。然后，考虑到不同标记的预测对应不同尺度的图，MK_MIML 引入多核融合方法，根据不同尺度的参数和图构造多个核函数，并在融合步骤中考虑了核的凸组合。最后，利用 SVM 方法进行分类。采用的核函数和多核融合方法如式(1)和式(2)所示。该方法可以有效地考虑包内示例之间的相关性。

$$t_g\left(\boldsymbol{X}_i,\boldsymbol{X}_j\right)=\frac{\sum_{a=1}^{n_i}\sum_{b=1}^{n_j}W_{ia}W_{jb}\,t\left(\boldsymbol{x}_{ia},\boldsymbol{x}_{jb}\right)}{\sum_{a=1}^{n_i}W_{ia}\sum_{b=1}^{n_j}W_{jb}},\tag{1}$$

$$T(\boldsymbol{x},\boldsymbol{x}')=\sum_{o=1}^{O}d_oT_o(\boldsymbol{x},\boldsymbol{x}'),\tag{2}$$

$$d_o\geqslant 0,\ \sum_{o=1}^{O}d_o=1\ 。$$

其中 $t(\boldsymbol{x}_{ia},\boldsymbol{x}_{jb})=\exp(-\gamma\|\boldsymbol{x}_{ia}-\boldsymbol{x}_{jb}\|^2)$，$O$ 为核数量。

Li 等人[38]认为包标记与多示例包相关联的原因是该包中有一些示例而不是所有示例会触发这个标记，并将这些示例称为标记的关键示例，还考虑到高度相关的标记通常会共享一些关键示例，因此开发了一种 2 阶段方法。首先，研究从 1 个示例包到 1 个特征向量的映射，其中每个特征值度量该包与 1 组相似示例之间的关联程度。然后，使用稀疏预测器作为标记来选择示例的关键组，从而选择关键示例。最后，将任务表述为一个凸优化问题，并给出了一种交替优化解。文献 [38] 所提出的方法 KISAR 的关键优势在于能够发现包和标记之间的合理关系。但是该方法是一种 2 阶段方法，容易造成信息损失。

总体而言，基于包的方法直接对包进行建模，能够从全局视角捕捉包的特征，避免处理复杂的示例标记关系，降低模型复杂度。但是无法捕捉示例级的细粒度信息，限制了细粒度标记预测能力，适用于处理示例间关系较弱以及不需要细粒度的示例级标记预测且全局特征更为重要的场景。

2.2.2 基于示例的方法

基于示例的方法以示例特征为单位进行求解，而无需将 MIML 问题退化为 SISL 问题，如图 3 中②所示。该类方法首先学习示例级标记信息。然后，基于示例级标记获得包标记，以建模多示例包和多个类别标记之间的映射关系。根据其求解过程中所采用的机器学习技术，该类方法大致可以分为 3 类，如表 3 所示。

1)概率方法

基于概率的方法借助贝叶斯和隐藏条件随机场等概率模型，通过学习示例标记进行包标记预测。Yang 等人[39]基于潜在狄利克雷分布提出一种概率生成模型 DBA。该方法假设数据具有树状结构，即包中的每个示例均具有单个标记，类似于潜在狄利克雷分配。通过采样类标记独立地生成每个示例，即

Table 3 MIML Methods Based on Instance

表 3 基于示例的 MIML 方法

类别	方法	思路
概率方法	DBA[39]	基于潜在狄利克雷分布提出一种解决 MIML 问题的概率生成模型
	文献 [40]	基于隐藏条件随机场，在单一的框架中捕获示例与标记之间以及不同标记之间的关系
	MIMLGP[41]	通过假设一个具有高斯过程先验的潜在函数，利用贝叶斯方法解决 MIML 问题
	MIML-AI[42]	基于辅助信息，并且通过在标记和辅助信息之间混合监督，提出一种判别概率模型
	MIML-CRFs[43]	利用条件随机场捕获示例的概率分布以及示例间的上下文和示例与标记之间的相关性
神经网络方法	MIMLNN[44]	利用 MLP 神经网络求解 MIML 问题
	DeepMIML[45]	使用卷积神经网络从原始图像数据中自动生成多示例包进行端到端学习
	MIML-FCN+[46]	使用双流全卷积神经网络借助包级特权信息和新的特权损失进行求解
	MIML-GAN[47]	使用 GAN 通过生成器和判别器之间的博弈进行包级和示例级的标记预测
	MetaMIML[48]	借助异构网络和元学习表达不同包之间的依赖关系、上下文信息、属性和结构信息，并且快速适应使用少量标记包的新任务
	AB-MIML[49]	利用注意力机制聚合示例信息并通过溯因推理校正伪标记以改进 MIML 方法
其他方法	M3MIML[50]	基于最大间隔方法，将任务表述为二次规划问题，能够利用示例和标记之间的关系
	文献 [51]	提出一种用于 MIML 距离度量学习的迭代算法，利用马氏距离来衡量包之间的距离
	文献 [52]	通过利用 L_1 正则化计算包与类之间的距离进行 MIML 距离度量学习
	MIMLfast[53]	利用特征向量映射、标记子概念和近似排名损失进行求解，可用于处理大规模数据

学习示例级标记信息，并在所采样标记的条件下，根据伯努利分布生成包标记。Zha 等人[40]提出一种基于隐藏条件随机场的综合 MIML 方法，能够在一个单一的框架中捕获标记与示例之间的关系以及不同标记之间的相关性。该方法通过为示例$\boldsymbol{x}$引入中间隐藏变量$\boldsymbol{h}$，利用隐藏条件随机场来建模标记与示例之间、不同示例标记之间、示例与包标记之间以及不同包标记之间的依赖关系，并通过极大似然准则借助期望最大化(expectation maximization, EM)算法进行求解，预测包标记和示例标记。

He 等人[41]提出一种基于高斯过程的 MIML 方法 MIMLGP。该方法通过在每个标记的示例空间中假设一个具有高斯过程先验的潜在函数，并利用贝叶斯方法对这些函数进行积分，用于推断示例在不同标记上的预测概率，同时通过示例预测概率得到包预测概率，而标记之间的相关性可以通过潜在函数的协方差矩阵来确定。该方法可以有效地捕获示例与标记之间以及不同标记之间的相关性。Nguyen 等人[42]提出一种混合监督的 MIML 算法 MIML-AI。该方法利用部分带标记的多示例包和部分带有辅助标记信息的多示例包进行训练，通过预测示例标记结合辅助信息进行包标记预测。Li 等人[43]提出一种基于条件随机场的 MIML 算法 MIML-CRFs。该方法通过高斯混合模型来学习所有示例的概率潜在分布，同时考虑了相邻示例之间的空间上下文以及示例和标记之间的相关性，以预测示例标记，并用包中所有示例标记作为包标记。基于概率的方法通过显式地建模示例与标记之间的映射关系来预测包标记，可以有效地保留示例与标记之间的相关性。

2)神经网络方法

基于示例的方法中，部分方法基于多层感知机(multi-layer perceptron, MLP)等浅层神经网络以及卷积神经网络等深度神经网络模型通过学习示例标记信息，进行包标记预测。浅层神经网络方法中，Chen 等人[44]基于 MLP 提出 MIMLNN 算法。该方法由 2 个阶段的 MLP 组成，第 1 个阶段的 MLP 以示例特征作为输入，学习示例标记，捕获示例与标记之间的关系。第 2 个阶段的 MLP 将第 1 个阶段 MLP 的输出合并作为输入，学习包标记，捕获不同标记之间的关系。MIMLNN 验证了 MIML 任务在神经网络结构中的适用性，为后续采用深度神经网络解决 MIML 问题奠定了坚实的基础。

深度神经网络方法中，Feng 等人[45]提出一种 DeepMIML 方法，能够利用深度神经网络生成 MIML 的示例表示，而不是通过传统方法进行手工设计。该方法首先定义一个能够建模单个示例和标记之间关系的 2D 子概念层。其次，采用按列池化操作实现标记预测，如图 5 所示。然后，通过堆叠多个 2D 子概念层并假设每个包中具有相同数量的示例(具有不同数量的示例用 0 填充)，将 2D 子概念层扩展到 3D 张量层，其中张量的每个切片都是每个示例的 2D 子概念层。最后，分别采用垂直和水平池化操作得到示例-标记得分层和包-标记得分层以探索示例和标记之间的关系，如图 6 所示。DeepMIML 的优势在于能够自动生成多示例，并且子概念学习组件能够捕获示例与标记之间的依赖关系，识别关键示例。

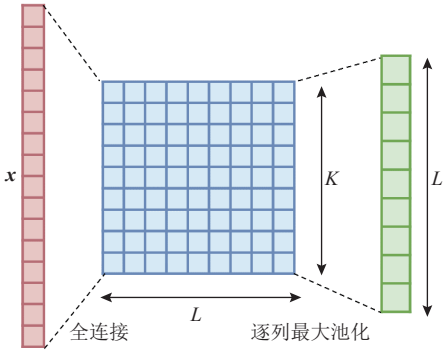


Fig. 5 2D sub-concept layer network structure[45]

图 5 2D 子概念层网络结构[45]

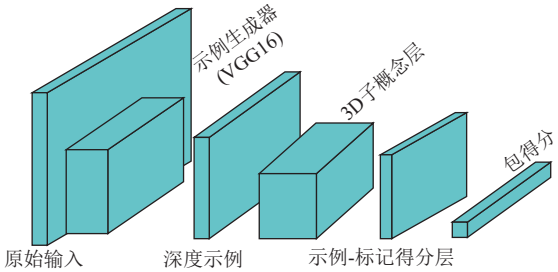


Fig. 6 Network structure of DeepMIML [45]

图 6 DeepMIML 网络结构[45]

Yang 等人[46]提出一种基于双流全卷积网络的方法 MIML-FCN+。该方法使用 2 组深度神经网络框架来分别处理训练包和特权包，并使用最近邻方法学习示例之间的相关性。其中训练包结合深度神经网络结构 VGG16 进行示例级标记预测，如式(3)所示。

$$f(\boldsymbol{x})=\varphi(\boldsymbol{x})\boldsymbol{W}, \tag{3}$$

其中$\varphi(\boldsymbol{x})$是所采用的深度神经网络结构，$\boldsymbol{W}$是要学习的网络参数，用于将示例从特征空间映射到标记空间，$f(\boldsymbol{x})$表示示例$\boldsymbol{x}$的预测结果。然后，采用最大池化方法聚合示例预测得到包预测结果，如式(4)所示。

$$F(\boldsymbol{X})=\max_{\boldsymbol{X}=(\boldsymbol{x})}\varphi(\boldsymbol{x})\boldsymbol{W}, \tag{4}$$

其中$F(\boldsymbol{X})$表示包$\boldsymbol{X}$的预测结果。此外，特权包中使用包$\boldsymbol{X}$的边框和标题作为特权信息通过神经网络结构进行特征提取，并通过新的特权损失将两者统一起来进行学习。因此，MIML-FCN+可以有效地利用包级特权信息和示例之间的相关性，并且很容易与其他深度学习模型进行集成。

此外，Pan 等人[47]基于生成对抗网络(generative adversarial network, GAN)提出一种用于重叠信号波形识别的 MIML 方法 MIML-GAN。该方法在生成器与判别器之间进行博弈以寻找动态平衡，其中判别器负责进行示例级标记预测，并通过最大池化操作聚合为包标记。同时，模型通过最小化预测包标记与真实包标记之间的交叉熵损失来实现优化。该方法也可被用于解决其他 MIML 任务，并且可以同时进行包级和示例级的标记预测。Qiu 等人[48]提出一种基于异构网络嵌入和元学习的 MetaMIML 方法，可以有效解决具有多类型对象的标准 MIML 任务。该方法首先利用具备网络嵌入能力的上下文学习器，提取包级结构信息并生成上下文表示。然后，通过任务学习器挖掘可迁移的元知识，以便在训练样本有限的情境下快速适应新任务。最后，融合结构信息与属性信息，联合预测包级和示例级标记。Wu 等人[49]提出一种溯因 MIML 方法(AB-MIML)。该方法首先将图像划分为多个图像块(即示例)，并利用神经网络提取其特征。然后，通过注意力机制对示例特征进行加权聚合，生成初始的示例级与包级伪标记。为了提高标记质量，这些伪标记被送入融合领域知识的逻辑知识库中，通过溯因推理进行校正，使其更接近真实标记，进一步提升模型的分类精度与可解释性。

3)其他方法

除了上述几种技术之外，基于示例的 MIML 方法中部分学者采用其他机器学习技术进行求解。例如，Zhang 等人[50]基于最大间隔方法提出 M3MIML 方法。该方法假设每个类都有 1 个线性模型，学习任务被表述为二次规划问题，并以对偶形式出现。其中每个类的输出被设置为所有 MIML 包中示例相对于相应线性模型的最大预测。然后，将所有可能类的输出组合起来，以定义分类系统上 MIML 实例的边界。这样，可以通过在每个示例上定义 1 个特定的间隔来直接考虑示例与标记之间的关系。这种方法的优势在于直接利用了示例与标记之间的关系，不容易遗漏信息，但是优化过程太多，导致算法效率不高，尤其是在训练数据量较多的情况下。Jin 等人[51]和 Verma 等人[52]基于距离度量策略，通过最小化示例与标记之间的距离来最小化包与包标记之间的距离，同时最大化不同标记之间的距离进行求解。

Huang 等人[53]为了有效地处理大数据集并且识别特定于标记的关键示例，提出一种快速 MIML 方法 MIMLfast。该方法首先从原始特征中学习了所有标记的共享空间。其次从共享空间中训练了特定标记的线性模型。为了识别用于表示特定标记的关键示例，在示例级别训练分类模型，如式(5)所示。然后选择具有最大预测的示例作为关键示例如式(6)所示。

$$f_l(\boldsymbol{x})=\boldsymbol{w}_l^{\mathrm{T}}\boldsymbol{W}_0\boldsymbol{x}, \tag{5}$$

$$f_l(\boldsymbol{X})=\max_{\boldsymbol{X}=(\boldsymbol{x})}f_l(\boldsymbol{x}), \tag{6}$$

其中$f_l(\boldsymbol{x})$表示示例$\boldsymbol{x}$在标记l上的预测得分，矩阵$\boldsymbol{W}_0$用于将原始特征向量映射到共享空间，$\boldsymbol{w}_l$是标记l上的权重向量。最后，在每次迭代中随机抽样 1 个三元组，包括 1 个包、1 个相关的包标记和 1 个不相关的标记，并优化模型以在违反此顺序时使相关标记排在不相关标记之前，损失函数如式(7)所示。

$$\mathcal{L}\left(\boldsymbol{X},Y^{+},Y^{-}\right)=\eta\left(\boldsymbol{X},Y^{+}\right)\left[1+f_{Y^{-}}(\boldsymbol{X})-f_{Y^{+}}(\boldsymbol{X})\right], \tag{7}$$

其中Y^{+}为包$\boldsymbol{X}$的相关标记，Y^{-}为包$\boldsymbol{X}$的不相关标记，$\eta(\boldsymbol{X},Y^{+})=\sum_{i=1}^{R(\boldsymbol{X},l)}1/i$，$R(\boldsymbol{X},l)=\sum_{j\in Y^{-}}I[f_j(\boldsymbol{X})>f_l(\boldsymbol{X})]$。

MIMLfast 的显著优势在于算法效率高，能够有效地扩展至大规模的 MIML 数据集，并且能够识别针对特定标记的关键示例。

总体而言，基于示例的方法能够充分捕捉细粒度的示例级特征信息，并进行示例级标记预测。然而，这种方法需要对每个示例单独建模，导致较高的计算成本，并且对包内示例的噪声更加敏感。适用于包内示例间关系复杂、示例标记预测对任务至关重要且需要示例级解释的场景。

综上所述，基于直接求解的 MIML 方法直接通过学习包特征或示例特征与多个类别标记之间的映射关系来解决标准 MIML 问题，而不需要显式地将 MIML 问题退化为传统的监督学习问题进行求解。这类方法可以有效地保留示例之间、标记之间以及示例与标记之间的依赖关系，不易造成信息损失。但是计算复杂度较高、开发难度较大，适用于对模型性能要求较高且计算资源充足的场景。

2.3 基于主动学习的 MIML 方法

相较于上述通过被动学习技术来解决 MIML 问

题的退化方法和直接求解方法，基于主动学习的MIML方法通过主动学习技术来解决标准MIML问题。主动学习算法在学习过程中具有选择新样本的能力，以提高模型性能并减少数据标注成本。具体来说，在这类方法中通常有一小部分的数据$D_s=\{(\boldsymbol{X}_1,\boldsymbol{Y}_1),(\boldsymbol{X}_2,\boldsymbol{Y}_2),\cdots,(\boldsymbol{X}_s,\boldsymbol{Y}_s)\}$在初始时被标记，而大量的数据在初始时是未被标记的$D_u=\{(\boldsymbol{X}_{s+1},\boldsymbol{Y}_{s+1}),(\boldsymbol{X}_{s+2},\boldsymbol{Y}_{s+2}),\cdots,(\boldsymbol{X}_u,\boldsymbol{Y}_u)\}$。其目的是利用少量标记数据$D_s$来逐步地学习大量未标记数据$D_u$的未知标记。

Retz等人[54]将主动学习应用于MIML领域并提出一种MIML主动学习方法。该方法使用MIMLSVM作为基分类器，以解决包评分的难题。然后，引入了2个选择标准，从未标记的数据集中选择信息最丰富的数据。最后，采用单示例多标记主动学习策略来解决退化后的MIML问题。该方法将主动学习应用于MIML问题，但是退化过程容易产生信息损失。Huang等人[55]也提出一种基于主动学习的MIML方法。该方法以MIMLfast算法作为基分类器，通过输入和输出空间中的多样性和不确定性，主动查询最有价值的信息，解决MIML主动学习问题，并且不需要对问题进行退化。Ruiz等人[56]认为主动学习算法的效率不仅依赖于查询策略的设计，还依赖于基分类器的选择。因此，在基于池的场景中引入了一种主动的MIML方法，并使用MIML-kNN作为基分类器。该方法首先计算了包距离矩阵$\boldsymbol{P}$，然后从这个矩阵中提取出标记包的距离子矩阵$\boldsymbol{P}_s$，并将其用于MIML-kNN模型。对于包$\boldsymbol{X}_i$的分类任务，需要向训练过的MIML-kNN模型提供相应的矩阵。最后，引入了8种查询策略来进一步优化算法的分类性能。该方法在多种基准数据集上均获得了显著的改善结果，但仍倾向于在MIML算法和主动学习查询策略之间进行有机组合。

Wu等人[57]受MIMLfast方法的分类策略和AUDI算法中使用的主动查询方案的启发，以MIMLfast为基分类器，提出一种多示例多标记主动学习算法MIMLAL-R，用于细菌基因组的蛋白质功能注释。Su等人[58]提出了一种具有成本效益的多示例多标记主动学习方法CM2AL。该方法利用包和示例信息来选择信息丰富的包-标记对，避免对负包-标记对的过度查询，并且通过包和示例之间的协作帮助量化包-标记对的不确定性，有效降低查询成本。类似地，Su等人[59]提出了一种新的多示例多标记主动学习框架，首先对包进行特征筛选，以有效地消除无用特征对注释的干扰。然后，量化特征和包之间的相关性，作为选择包-标记对的基础。最后，查询所选包-标记对中最有可能的正子样本标记对。该方法可以有效地利用包与特征之间的关系来降低标注成本。

综上所述，基于主动学习的MIML方法能够有效降低数据标注成本，并在有限的标注数据条件下提升模型性能。但是这类方法高度依赖于查询策略的选择，且对未标记数据的评估可能带来额外的计算开销，适用于数据稀缺和数据标注成本较高的场景。

3 多元MIML

多元MIML任务在标准MIML任务的基础上，从输入空间的角度出发，考虑了多示例包的多样性和复杂性。其目标是根据多元的多示例包和相应的多个类别标记训练一个特征分类器，能够有效地根据输入多示例包的不同来判断它们所属的类别。根据多示例包的来源或类型可以将多元MIML任务进一步划分为多视图MIML和多类型对象多视图MIML两类分别展开介绍。

3.1 多视图MIML

标准MIML任务通常认为输入数据来自单个特征视图，而不是来自其他特征视图的信息。多视图MIML(multi-view multi-instance multi-label learning, M3L)任务是指在标准MIML任务的基础上，考虑具有多视图性质的输入数据。N个样本构成的M3L数据集可以表示为$D=\{([\boldsymbol{X}_1^1,\boldsymbol{X}_1^2,\cdots,\boldsymbol{X}_1^V],\boldsymbol{Y}_1),([\boldsymbol{X}_2^1,\boldsymbol{X}_2^2,\cdots,\boldsymbol{X}_2^V],\boldsymbol{Y}_2),\cdots,([\boldsymbol{X}_N^1,\boldsymbol{X}_N^2,\cdots,\boldsymbol{X}_N^V],\boldsymbol{Y}_N)\}$，其中$N$表示数据集中的样本数(包个数)，$V$表示数据集中的视图数，$\boldsymbol{X}_i^j$表示第$i$个多示例包的第$j$个视图信息，$\boldsymbol{X}_i^j=(\boldsymbol{x}_{i1}^j,\boldsymbol{x}_{i2}^j,\cdots,\boldsymbol{x}_{in_i}^j)^{\mathrm{T}}$，$\boldsymbol{x}_{ik}^j$表示第$i$个多示例包在第$j$个视图下的第$k$个示例。M3L的目标是利用多视图的样本特征进行多标记预测。

Nguyen等人[60]认为通过同时利用视觉和文本信息，可以获得更好的性能，并提出了多视图多示例多标记的潜在狄利克雷分配方法M3LDA。该方法由视觉标记部分、文本标记部分和标记主题部分组成，其中视觉标记部分和文本标记部分专门用于从视觉和文本空间到标记空间的映射，标记主题部分用于融合标记关系。其基本思想是，由视觉信息决定的主题和由文本信息决定的主题应该是一致的，从而导致正确的标记分配。M3LDA方法将标准的MIML框架扩展到多视图设置，同时利用视觉和文本信息，并且能够被扩展到更多的视图，但是该方法只适用于离散视图，并且没有考虑标记间的相关性。此后，

Nguyen 等人[61]受贝叶斯网络的启发，又提出一种更通用、更有效的基于层次贝叶斯网络的多视图多示例多标记混合方法 MIMLmix，并推导出一种有效的基于变分推理的学习算法。该方法较少出现标记不平衡问题并且允许来自离散和连续视图的示例，比 M3LDA 方法更通用、更有效，特别是在具有大量标记的数据集中，但是这种方法忽略了标记间的相关性。

Yang 等人[62]认为上述 M3L 方法很少考虑标记之间的相关性。为了同时学习标记预测和探索标记相关性，提出一种新的多视图多示例多标记深度网络方法 M3DN。该方法为每个视图构建了独立的深度网络，并在包级施加视图一致性约束，要求不同视图上基于同一包的预测能够产生相似的标记相关性，从而实现了标记预测与标记相关性探索的同步进行。M3DN 不仅将 M3L 问题推广到了深度学习领域，还成功地将标记间相关性纳入考虑范畴，但是这种方法对每个视图使用单独的网络，忽略了视图之间的一致性。为了充分利用未标记数据中的丰富信息，Yang 等人[63]进一步提出一种半监督的 M3DN 方法 M3DNS。Xing 等人[64]认为现有的 M3L 解决方案只是部分地探索包、示例和标记的内在联系以及它们之间的关系，容易产生性能损失。因此，提出一种基于协同矩阵分解的解决方案 M3Lcmf。该方法首先构建由包、示例和标记为节点组成的异构网络，通过多个关系数据矩阵对不同类型的关系进行编码。然后，将子网络的关联矩阵协同分解为低秩矩阵，得到节点的低秩表示和它们之间的潜在关系，并且有选择性地整合多个特征视图的包和示例。最后，在分解目标中引入一个聚合项，不仅可以将示例标记关联聚合到包级还可以反向指导这些关联的预测。该方法借助异构网络和协同矩阵分解有效地保留了包、示例和标记之间的各种关系。

Xing 等人[65]指出现有的 M3L 方法通常基于包与视图间完全对应及具有完整训练注释的理想化假设。然而，在现实场景中包之间的对应关系及包的注释信息往往都不完整。为此，引入弱监督 M3L 问题，并提出一种基于多视图字典学习的弱监督方法 WSM3L。该方法首先引入一个统一的目标函数来寻找不同视图包之间的匹配，并预测包标记。它利用包的异构特征来学习多视图协调的字典和每个视图包的个体编码向量。然后，利用编码向量的标记相似度和特征相似度来寻找视图之间的匹配。最后，使用邻近包标记联合补充包标记，将包标记分发到示例上，并将示例标记反向聚合到它们的原始包上。这样，WSM3L 可以以连贯的方式预测包标记和示例标记。该方法的优势在于不仅可以处理弱配对的包，还可以有效处理仅有部分标注的训练包。Hang 等人[66]认为在 M3L 中有效利用标记之间的相关性非常重要，并且提出一种基于图卷积神经网络的方法 M3GN。该方法首先使用神经网络学习每个示例的特征表示，然后基于注意力机制聚合示例特征得到包特征表示，最后在所有标记上构建一个图，通过图神经网络来学习标记之间的相关性，并将这些标记嵌入映射到相互相关的标记分类器中。该方法可以有效捕获标记相关性、过滤数据噪声和识别关键示例。

综上所述，M3L 通过整合来自不同视图的数据充分利用各个视图的互补信息，提高模型的泛化能力、准确性和鲁棒性。但是其计算复杂度较高、视图间关系建模较为困难，且标注成本较高。适用于处理具有多视图和多模态数据的场景。

3.2 多类型对象多视图 MIML

经典的 M3L 方法通常仅针对同构网络或单一类型对象的集合进行建模，却忽略了多类型对象之间复杂而多样的相互关联。在实际情况中，复杂对象的标记不仅取决于其自身的属性，还深受与其他类型对象之间关系的影响，而 M3L 缺乏同时建模多类型对象的能力。多类型对象多视图 MIML(multi-typed objects multi-view multi-instance multi-label learning, M4L)在 M3L 的基础上，能够处理同时具有多种类型对象的 MIML 问题。M4L 的数据实例如图 7 所示，该实例具有 3 种类型的对象，如作者、会议和论文。其中论文由图像和文本视图进行编码，每个视图中的论文对象(包)进一步由不同的示例(例如，a_2中的文本段落$\boldsymbol{x}_{21}^1$和图像区域$\boldsymbol{x}_{22}^2$)组成，这些对象

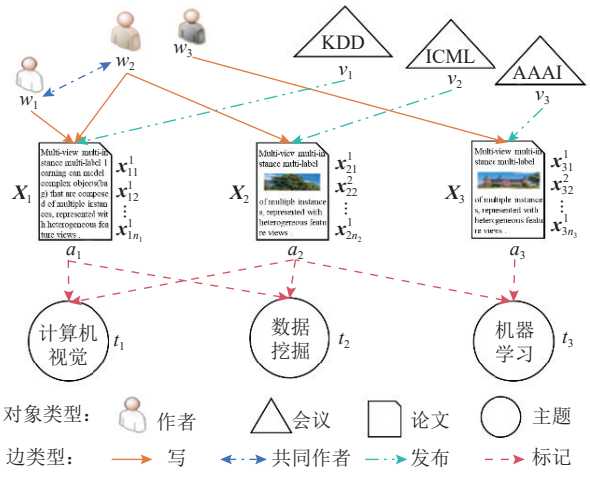


Fig. 7 Example of M4L data[67]

图 7 M4L 数据实例[67]

同时被标注为多个语义标记(例如,数据挖掘、计算机视觉、机器学习等)。值得注意的是,M3L 仅考虑论文本身的特征,而 M4L 不仅考虑论文本身的特征,还能整合与论文相关的其他类型对象,如作者和会议的信息,从而更全面地建模复杂数据中的多层次关系。M4L 的目标是在考虑不同类型对象之间相互关系的同时,学习具有多类型对象和多视图信息的多示例包与多个类别标记之间的映射关系。

Yang 等人[68]提出 M4L 范式,并提出一种基于联合矩阵分解的解决方案 M4L-JMF。该方法首先构建一个以对象、示例和标记作为节点的异构网络,对多类型对象之间的各种关系进行编码。其次,将异构网络的块关联数据矩阵与各类型节点的属性数据矩阵联合分解为低秩矩阵来进一步探索包、示例和语义标记的潜在表示。然后,利用这些潜在表示来预测包和标记之间的关系。最后,为了考虑包和示例之间的关系,进一步引入调度和聚合术语,将包标记分配到单个示例中,并以一种连贯的方式将示例标记聚合到其相应的包中。然而,M4L-JMF 只能学习这些对象的线性表示,忽略了复杂的非线性关系,表达能力有限,并且由于矩阵分解的高度复杂性,不能扩展到大规模的数据集中。此后,Yang 等人[67]又提出一种基于异构网络嵌入的 DeepM4L 方法。该方法首先构建一个属性异构网络来编码多类型对象之间的关系。其次,利用示例嵌入网络来学习示例的特征表示和异构网络嵌入中的包嵌入空间信息。然后,为了增强模型在视图内和视图间的标记一致性,使用多示例池化层为每个包生成跨视图的标记得分张量。最后,利用多视图学习技术融合多视图包的标记得分张量,并同时进行包级和示例级的标记预测。

综上所述,M4L 能够同时处理多视图数据和不同类型对象之间的复杂关系,捕捉更丰富的语义和结构信息,提升模型的表达能力。但是模型复杂度高,数据标注和获取的成本高。M4L 适用于需要同时建模多个实体及其关系,并且具有多视图或多模态数据的任务。

4 非精确标记的 MIML

非精确标记的 MIML 任务在标准 MIML 任务的基础上,从输出空间的角度考虑标记的不确定性。其目标是根据多示例包和不精确的包标记训练特征分类器,能够有效地根据输入多示例包的不同判断它们精确的多个包标记。根据标记的完整性和准确性可以将非精确标记的 MIML 任务进一步划分为新标记学习和缺标记学习 2 类分别展开介绍。

4.1 新标记学习

标准的 MIML 任务通常假设所有的包都属于一个固定的标记集,然而这一假设经常被违反,一些新的标记可能存在,但隐藏在训练数据中。这意味着即使因为某种原因部分标记不存在于初始的目标标记集中,但是仍然希望能够使用所训练的模型对这部分未知的标记进行预测。具有这种现象的学习问题被称为新标记学习问题。给定一组由N个样本构成的示例包$\mathcal{X}=\{\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_N\}$及包标记集$\mathcal{Y}=\{\boldsymbol{Y}_1,\boldsymbol{Y}_2,\cdots,\boldsymbol{Y}_N\}$。其中每个包被表示为$\boldsymbol{X}_i=(\boldsymbol{x}_{i1},\boldsymbol{x}_{i2},\cdots,\boldsymbol{x}_{in_i})^{\mathrm{T}}$,由$n_i$个示例构成,每个包标记被表示为$\boldsymbol{Y}_i=(y_{i1},y_{i2}\cdots,y_{iq},y_{iq+1},\cdots,y_{iL})^{\mathrm{T}}$。$L$表示标记集中的真实标记数,$q$表示训练过程中所知道的标记数。训练过程中包$\boldsymbol{X}_i$的包标记被表示为$\hat{\boldsymbol{Y}}_i=(y_{i1},y_{i2},\cdots,y_{iq})^{\mathrm{T}}$, $y_{ij}\in\{0,1\}$, 其中$y_{ij}=1$表示包$\boldsymbol{X}_i$具有标记j, $y_{ij}=0$表示包$\boldsymbol{X}_i$不具有标记j。MIML 的新标记学习任务是指利用包和已知的包标记通过机器学习的方式获得包和对应 L 个类别标记之间的映射关系。

Lou 等人[69]打破了包中所有示例均属于已知标记集的假设,专注于研究包中可能包含的新类示例(包中存在具有未知新标记的示例)的场景。其目标在于确定包中的示例属于已知类还是新类,并设计了一个检测策略,通过为每个示例引入一个针对标记的得分函数进行判断。具体来说,他们认为如果示例$\boldsymbol{x}_i$在已知标记集上的最大得分大于阈值ϵ,那么示例$\boldsymbol{x}_i$属于已知标记,否则示例$\boldsymbol{x}_i$属于新标记。Pham 等人[70]基于上述思想,为原始数据引入 1 个新标记,通过示例特征进行示例级标记预测,并聚合示例标记得到包标记。朱越等人[71]针对 MIML 中的新标记学习任务提出一种端到端的 M3L 方法 EM3NL,在 M3L 条件下解决新标记学习问题。该方法基于卷积神经网络产生多示例包,并通过最小化包上观察标记的错分损失和对新标记预测值排序损失的惩罚以及对多视图预测不一致的惩罚,同时学习图像、文本 2 个视图的特征表示以及已知标记和新标记的预测函数。然而,上述 3 种方法只能识别新标记的存在,而无法同时处理多个新标记。

Zhu 等人[72]借助正则化和高效的增广拉格朗日优化方法,提出一种能够发现 MIML 问题中多个新标记的方法 DMNL。该方法首先利用包中所有的示例,根据已知的包标记,设计了一个基于平方误差的包依赖损失项,并引入一种重新调整策略,以考虑所

有示例对包标记的贡献,使得每个包标记被视为同等重要。其次,利用示例之间的结构来正则化假设空间,并且基于独立于包的示例在具有相同标记时彼此接近,具有不同标记时彼此远离的聚类假设,设计了一个包独立的聚类正则化项。然后,在非负正交约束条件下,将损失项和正则化项结合起来。最后,采用交替优化策略,并遵循增广的拉格朗日优化框架进行求解。该方法能够在示例级和包级同时进行多个新标记的预测,并且随着包标记数量的增加,计算成本呈线性增长,提升了算法计算效率。

MIML 中的新标记学习允许模型在训练阶段未见过的标记上进行预测,从而显著提升模型对未知数据的泛化能力,使模型能够适应动态变化的标记空间并充分利用数据中的隐藏信息以提升整体性能。但是这类算法设计复杂、标注难度大、性能不稳定,适用于标记集合可能会随着时间推移不断变化的任务。

4.2 缺标记学习

标准 MIML 问题通常假设训练集中所有的包标记都是已知的和准确的,然而这一假设经常被违反,一些包的标记可能是未知的或缺失的。这意味着即使因为某种原因部分包标记不准确或者未知,但是仍然希望能够进行准确的标记预测。具有这种现象的学习问题被称为缺标记学习问题。给定一组由N个样本构成的多示例包$\mathcal{X}=\{\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_N\}$及其标记集$\mathcal{Y}=\{\boldsymbol{Y}_1,\boldsymbol{Y}_2,\cdots,\boldsymbol{Y}_N\}$,包$\boldsymbol{X}_i$被表示为$\boldsymbol{X}_i=(\boldsymbol{x}_{i1},\boldsymbol{x}_{i2},\cdots,\boldsymbol{x}_{in_i})^{\mathrm{T}}$,由$n_i$个示例构成,每个包标记被表示为$\boldsymbol{Y}_i=(y_{i1},y_{i2},\cdots,y_{iL})^{\mathrm{T}}$,$y_{ij}\in\{0,1,-1\}^L$,其中$L$表示标记集中的真实标记数,$y_{ij}=1$表示包$\boldsymbol{X}_i$具有标记$j$,$y_{ij}=0$表示包$\boldsymbol{X}_i$不具有标记$j$,$y_{ij}=-1$表示包$\boldsymbol{X}_i$的标记$j$未知。MIML 的缺标记学习任务是指利用多示例包和已知的包标记通过机器学习的方式获得包和对应L个类别标记之间的映射关系,并进行准确的标记预测。

Yang 等人[73]提出弱标记的 MIML 任务,并基于最大间隔技术提出新的解决方法 MIMLwel。该方法首先将每个多示例包映射为一个单一的特征向量,将 MIML 问题转化为 SIML 问题。其中每个元素用于度量包与一组相似示例的关联程度。然后,使用稀疏预测器学习包标记,使得每个包标记的类均值最大化。最后,在一个通用框架中,通过一种有效的块坐标下降方法进行求解。然而,该方法假设只有正标记可以缺失,负标记不会缺失。

Nguyen 等人[74]认为随着数据量的增加,获取用于训练给定学习任务数据的瓶颈在于数据中手动标记示例的成本。为了缓解这种问题,考虑在多示例学习下引入一种新的缺少包标记的 MIML 任务,即被分配给一组示例的多个包标记仅部分可用,使得每个示例的标记比例相当低,从而显著地减少标记需求。具体来说,为了解决不完全标记(缺标记)的 MIML 问题,首先提出一种新的具有潜在示例标记的判别图模型,通过假设每个包中的示例和包标记在包之间是独立的,以及潜在多类示例标记条件独立于包,能够同时进行包级和示例级标记预测。然后,为基于最大似然的推理推导出一种计算高效的替代方法,提出一种 EM 方法,通过引入一个新的重构模型实现有效的 E 步计算。最后,为进一步提高模型参数推断步骤的有效性,考虑一种通过最大化边际似然而不是联合对数似然的近似推断方法,以进一步降低训练过程中可用标记函数的复杂性。该方法的显著优势在于能够处理任意种类的缺失标记,而不是假设缺失的是正标记。

Zhang 等人[75]认为,在 MIML 中,由于标记分布的固有倾斜性和标记间的依赖关系,标记不平衡问题是一个关键但很少被研究的挑战。为了解决这一问题,提出了一种基于纠错编码集成和自适应阈值策略的新方法 IMIMLC。该方法首先通过 Fisher 向量提取每个对象的结构信息,从而消除不精确监督的影响。接着,为了减轻由标记不平衡带来的干扰,IMIMLC 通过随机选择子任务的纠错编码构建了一个新颖的集成模型。该模型在小规模数据块上训练二元基分类器以增强分类器的多样性,并提高模型对不平衡问题的鲁棒性。最后,IMIMLC 通过最大化分类间隔,自适应地为每个标记学习阈值,从而有效避免了多标记语义差异和不平衡比率导致的不准确预测。该方法的显著优势在于纠错编码和自适应阈值策略能够有效缓解标记分布不平衡对模型性能的影响。

MIML 中的缺标记学习能够处理数据标记不完整的 MIML 问题,提升模型的鲁棒性,减少数据标注成本,充分利用未标注数据并灵活应对真实应用场景。但是其计算复杂性较高,依赖于一定的标记分布假设,标记难度较大,模型性能不稳定。适用于处理可能存在标记信息不准确、缺失或带有噪声的任务。

5 MIML 方法评价

为了便于对 MIML 方法的效果进行对比分析,本节首先介绍 MIML 领域常用的数据集,然后对一些代表性 MIML 方法的实验结果进行对比和分析。

5.1 常用数据集

目前，MIML 领域常用的代表性数据集涵盖了音频、图像和文本等多个重要领域。相关数据集的统计信息如表 4 所示。

Table 4 Common Datasets for MIML

表 4 MIML 常用数据集

领域	数据集	示例数	包数	标记数
音频	Birds[76]	10 232	548	13
图像	MSRC v2[77]	1 758	591	23
	Scene[5]	18 000	2 000	5
	MS COCO[78]	37 880	10 000	80
文本	Letter Carroll[76]	717	166	26
	Letter Frost[76]	565	144	26
	Reuters[50]	7 119	2 000	7

Birds 数据集是一个音频信号处理的 MIML 数据集，旨在通过自动识别音频中的鸟类鸣叫声来生成鸟类种群的声学调查。每段 10 s 的音频录音被视为 1 个包，而频谱图中的每个片段被视为 1 个示例，其标记对应于录音中出现的鸟类种类。MSRC v2 数据集是 MSRC 图像数据集的一个子集。每幅图像被视为 1 个包，图像的真实分割中每个连续区域被视为 1 个示例，其标记对应于图像类别。Scene 数据集是一个自然场景图像数据集。每幅图像被视为 1 个包，通过对图像进行平滑和下采样处理生成 9 个示例，这些示例描述了每个颜色块及其邻近块之间的颜色差异，其标记对应于图像所属的类别。MS COCO 数据集是一个大型图像数据集，每幅图像被视为 1 个包，图像的子区域被视为示例，其标记对应于图像所属的类别。Letter Carroll 和 Letter Frost 数据集是基于 Letter Recognition 数据集构成的 2 个合成的文本 MIML 数据集，由 2 首诗中提取的单词构成。每个单词被视为 1 个包，每个字母对应 1 个示例，其标记集是英文字母表。Reuters 数据集来自广泛研究的 Reuters-21578 文本数据集，数据集中的文档通过滑动窗口技术表示为多个示例。每篇文档是 1 个包，每个示例是 1 个文本片段，并通过“词袋”模型进行特征提取。

5.2 效果对比

本节以汉明损失作为评价指标，对经典及最新的 MIML 方法在 6 个通用公共数据集上的性能进行对比与分析。汉明损失是衡量多标记分类性能的常用评价指标，用于计算样本中预测错误的标记数占总标记数的比例，其值越小分类效果越好。在本次实验中，选取的 MIML 方法涵盖了多种类型，包括基于退化的 MIMLSVM 和 MIMLBOOST 方法，基于直接求解的 MIMLfast, MIML-LLMC, DMIML*k*NN, MIMLDGC 方法，基于主动学习的 CM2AL 方法和非精确标记的 MIMLwel 方法。所有实验均采用 3 折交叉验证取平均值和标准差，如表 5 和表 6 所示。

从表 5 可以看出，经过严格的 3 折交叉验证，标准 MIML 算法中，MIML-LLMC 和 DMIML*k*NN 在多标记分类性能上表现优异。同时，相较于基于退化

Table 5 Hamming Loss of the Standard MIML Methods

表 5 标准 MIML 方法的汉明损失

数据集	MIMLSVM	MIMLBOOST	MIMLfast	MIML-LLMC	DMIML*k*NN	MIMLDGC	CM2AL
Birds	0.187±0.004	0.165±0.006	0.055±0.012	0.158±0.012	**0.048±0.005**	0.051±0.007	0.109±0.011
MSRC v2	0.086±0.003	0.117±0.002	0.091±0.006	**0.072±0.005**	0.076±0.001	0.075±0.002	0.081±0.001
Scene	0.331±0.006	0.247±0.005	0.198±0.008	**0.174±0.005**	0.184±0.004	0.198±0.005	0.182±0.009
Letter Carroll	0.155±0.005	0.151±0.008	0.126±0.005	**0.121±0.016**	0.127±0.002	0.128±0.007	0.135±0.002
Letter Frost	0.148±0.007	0.140±0.009	0.132±0.002	0.127±0.004	**0.120±0.009**	0.121±0.006	0.124±0.006
Reuters	0.175±0.006	0.159±0.003	0.076±0.009	**0.030±0.002**	0.063±0.001	0.062±0.002	0.038±0.003

注：黑体数值表示该行最好的结果。“±”前表示平均值，“±”后表示标准差。

Table 6 Hamming Loss of MIMLwel Method Under Different Datasets and Weak Label Ratios

表 6 MIMLwel 方法在不同数据集及弱标记比例下的汉明损失

弱标记比例/%	Birds	MSRC v2	Scene	Letter Carroll	Letter Frost	Reuters
10	0.163±0.007	0.109±0.004	0.247±0.005	0.151±0.010	0.139±0.006	0.157±0.002
20	0.151±0.008	0.103±0.004	0.228±0.004	0.150±0.008	0.138±0.009	0.142±0.002
30	0.137±0.016	**0.100±0.003**	0.220±0.005	0.146±0.007	**0.128±0.009**	0.125±0.008
40	**0.119±0.010**	**0.100±0.003**	**0.213±0.002**	**0.141±0.018**	0.129±0.013	**0.107±0.005**

注：黑体数值表示该列最好的结果。“±”前表示平均值，“±”后表示标准差。

的 MIML 方法，基于直接求解的 MIML 方法展现出更优的性能。然而，由于 CM2AL 算法减少了对包标记的使用，其性能稍逊色于部分标准 MIML 方法。实验结果表明，在标准 MIML 任务中，基于包的 DMIMLkNN 和 MIML-LLMC 方法在包级多标记分类任务中具有显著的优势。从表 6 可以看出，弱标记的 MIMLwel 算法在不同弱标记比例下的多标记分类性能有所变化。总体而言，随着弱标记比例的增加，汉明损失逐渐下降，分类效果也逐步提升。这是因为随着弱标记比例的增加，算法可以获取更多标记信息，从而更好地学习标记之间的相关性和特征表示，最终提高分类准确性。总体而言，所有实验结果均可复现，并且与原论文的实验结果一致。

6 应用场景及未来研究方向

MIML 在自然语言处理、计算机视觉、生物信息学、医学和信号处理等领域都有着非常重要的应用和广阔的前景。本节给出 MIML 的现有应用场景和一些未来可能的研究方向。

6.1 应用场景

1）自然语言处理

MIML 应用于自然语言处理的第 1 步就是多示例包的构建。常见的方法是通过文本划分技术将文本划分为多个片段，整个文本样本被表示为 1 个多示例包，由被表示为示例的多个文本片段构成。或者是从文本中提取多个关键词，将文本表示为多示例包，由多个被表示为示例的关键词构成。与传统的单示例表示方法相比，多示例表示方法可以更好地表示文本中多个语义片段的细粒度特征信息。MIML 在自然语言处理领域已经被广泛地应用于文本分类、方面级情感分析等问题。例如，Ren 等人[79]提出一种基于特征缩减的 MIML 方法进行在线生物医学出版物分类，解决学习过程中训练实例远少于特征数量和需要向文档分配多个非排他标记的问题。Sahlin 等人[80]提出一种基于规则的 MIML 方法进行人工短消息服务活动评估。Mylonas 等人[81]提出一种基于 MIML 的多标记弱监督学习方法用于解决新兴的医学主题标题词汇表的不断演变问题。为了有效学习同一上下文中其他跨度标记之间的依赖关系，Chen 等人[82]将宣传技术分类作为一个 MIML 问题，并提出一个简单的基于 RoBERTa 的模型，用于同时对每篇文章中的所有跨度进行分类。为进一步提高我国专利分类的准确性，Liu 等人[83]提出一种层次 MIML 方法将中文专利表示为多示例多标记的形式，以每项专利为多示例包，使用序列号作为分割标记，将权利要求划分为多个独立文本来构成包中的多个示例，并且与作为标记的多个 IPC 号相关联，解决中文专利多 IPC 号分类问题。此外，在方面级情感分析方面 Li 等人[84]提出一种用于方面类别情绪分析的 MIML 网络，将句子视为包、单词视为示例、表示方面类别的单词作为方面类别的关键示例，实现给定一个句子和句子中提到的方面类别来预测句子的方面类别情绪。

2）计算机视觉

在计算机视觉领域，每幅图像可以被划分为多个感兴趣区域，每幅图像被表示为 1 个多示例包，每个感兴趣区域被表示为包中的示例，可以同时与 1 个或者多个标记相关联。与传统的单示例表示方法相比，多示例表示方法可以更好地表示图像中多个感兴趣区域的细粒度信息，而不需要使用难以获得的细粒度图像标记。MIML 在计算机视觉领域已被广泛应用于处理图像分类[85-86]、图像检索、视频字幕生成和抗干扰识别等方面[87]。例如，Song 等人[88]提出一种用于多示例多标记图像分类的深度多模态卷积神经网络，通过将卷积神经网络与 MIML 相结合，将每幅图像表示为多示例包，并继承了卷积神经网络和 MIML 的优点。利用卷积神经网络的架构自动生成 MIML 的示例表示，利用标记相关性分组并结合标记组的文本上下文生成多模态示例，能够有效地区分属于不同组的视觉上相似的对象。Li 等人[43]提出一种基于集成上下文信息的图像分类方法，通过多示例条件随机场框架联合建模示例的潜在概率分布、相邻示例之间的空间上下文以及示例与标记之间的相关性。Chen 等人[89]将多标记图像检索的哈希学习作为多示例排序学习问题，提出一种基于多示例排序的端到端深度哈希框架，解决检索过程中哈希码次优的问题。Ji 等人[90]提出一种基于编码器-解码器重构结构的多示例多标记对偶学习方法，通过标题生成模块和视频重建模块生成视频字幕。Yao 等人[91]提出一种基于 MIML 的干扰识别对策，促进现代电子设备在复杂电磁环境中具有识别和对抗干扰的能力。Julia 等人[92]利用改进版的 MobileNetV2 模型作为特征提取器，结合 DeepMIML 和 LSTM 模型作为多标记分类器，将 MIML 技术应用于游戏录像中的事件分类问题。

3）生物信息学

多示例学习的研究最早起源于 Dietterich 等人[93]

在药物活性预测问题上的应用，其任务是预测药物分子能否与相应的受体蛋白质相结合。其中每个药物分子的活性取决于一些特定但未知的低能构象。因此，每个药物分子被表示为多个示例组成的包，而不是标准监督学习中单一的特征向量。随着多示例学习的快速发展，MIML 在生物信息学领域受到许多关注，被广泛应用于药物预测[94]、基因功能注释和检测、蛋白质结构检测[95-96]等领域。例如，蛋白质注释问题中，每个结构域包含多个蛋白质，适合用多个示例描述，且与多个不同功能相关联，可能属于多个类别。为此，Meng 等人[97]提出一种基于 SVM 的 MIML 框架 PDAMIML，结合 SVM 和自交叉协方差，有效地利用位置特定评分矩阵的特征，解决蛋白质结构域注释问题。Liu 等人[98]通过学习多核融合函数神经网络，提出 MK-EnMIMLNN 算法进行水稻蛋白质功能预测。Mei 等人[99]提出一种多示例多标记迁移学习方法来同时重建 27 条人类信号通路，并建模它们的交叉对话，以进行合理的蛋白质预测。在基因功能注释和检测方面，Xu 等人[100]提出一种 MIML 框架，使用多个示例进行鲁棒性多标记分类，对肝细胞癌进行基因突变预测。Li 等人[101]改进了多示例层次聚类，利用基因本体层次来注释基因功能，将基因本体层次与 MIML 框架结合，使用多标记 SVM 和 *k*NN 算法来预测基因功能。Li 等人[102]提出一种层次的 MIML 框架 HMIML，分别在图像级和基因级使用卷积神经网络，识别不同基因果蝇所对应的图像。在临床药物预测方面，Shang 等人[103]将病人的多项疾病诊断作为多个示例输入，将多个药物作为输出标记，提出一种知识引导的 MIML 网络 KG-MIML-Net，将 2 种类型的附加知识合并到 1 个编码器-解码器模型中为临床患者进行药物预测。Zhang 等人[104]利用深度学习和注意力机制，将 MIML 应用于免疫荧光图像中蛋白质亚细胞定位的预测，显著提升了蛋白质多定位的识别准确性。

4）医学

在医学领域 MIML 常被应用于医学图像分析方面的研究。例如，Loukas 等人[105]将手术图像表示为多示例包，将图像中不同感兴趣区域表示为包中的多个示例，提出一种基于变分贝叶斯高斯混合模型的 MIML 方法用于腹腔镜手术操作视频帧的学习和分类。Feng 等人[106]提出一种基于注意力多示例分类机制和 1D ResNet 网络的端到端深度神经网络模型 MIC-ResNet，用于 12 导联心电图分类任务。Lai 等人[107]提出一种新的单阶段广义 MIML 框架 BMIML 用于医学图像分类任务。与此同时，MIML 在乳腺癌的组织病理学研究方面也有许多应用。例如，Gecer 等人[7]将全幻灯片数字乳腺癌组织病理学图像表示为多示例包，包含多个感兴趣的幻灯片区域，采用 MIML 方法对幻灯片进行多标记分类，识别可能的多个乳腺癌类别。Mercan 等人[108]通过在弱监督学习场景中利用病理学家的查看记录及其幻灯片级注释，使用 4 种不同的 MIML 算法对整个幻灯片乳腺癌组织病理学图像进行幻灯片级和感兴趣区域级的乳腺癌类别诊断和预测。此外，MIML 在中医药领域也具有广泛的应用前景。例如，Yan 等人[109]通过患者在 5 个元音样本上的语音数据，利用 MIML 实现对传统中医的客观听诊任务。Li 等人[110]提出一种基于 MIML 的相似草药加权抽样方法，解决传统中药配方评价问题。

5）信号处理

在信号处理领域，MIML 也常被用于信号识别、视频和音频识别等方面。例如，Pan 等人[47]提出一种基于 GAN 的 MIML 算法用于重叠信号波形识别，通过弱监督学习精确地拟合由重叠信号变换得到的时频图像。Dong 等人[111]提出一种用于视频社会关系识别的 2 阶段 MIML 模型，在第 1 阶段获取有效的社会关系特征，并通过生成多个示例来挖掘哪些时刻的信息能够更好地反映任务之间的社会关系。在第 2 阶段重点发现视频中存在哪些社会关系，探索社会关系之间的潜在相关性，以促进对视频中社会关系的识别。Zhang 等人[112]提出一种基于时空预裁剪的 MIML 建模网络，以识别未裁剪视频中的动作并定位相应的帧。Zeng 等人[113]将视频拍摄视为示例，从视频级的标记中学习镜头标记，降低标记成本。并且提出一种多标记、多示例的主题模型对视频进行标注，通过控制视频标记和镜头之间的关系来学习标记语义，解决粗粒度视频镜头注释问题。Pan 等人[114]提出一种基于深度卷积神经网络的 MIML 框架，以端到端的方式处理信号，可以自动识别仅由单类型信号训练的多重低截获率雷达信号。在音频信号处理方面，Wu 等人[115]将音乐情绪识别建模为一种新的多标记多层次多示例多视图学习问题。其中音乐被表述为一个层次的多示例结构，多个情绪标记至少对应于 1 个示例，并且每层有多个视图。然后，提出一种层次音乐情感识别模型，基于句子级音乐和歌词特征借助贝叶斯模型进行音乐情绪识别。Briggs 等人[116]采用 MIML 表示来处理多只声音重叠的鸟类，并提出使用了这种表示的新算法进行鸟类生物多样

性识别。此外, MIML 还可以被广泛应用于推荐系统[117]、跨模态检索[118]和行为分析[119-120]等领域[121]。

6.2 面临的挑战和未来研究方向

尽管 MIML 问题已有较长的研究历史, 并在近年来取得了显著进展, 但该领域仍面临许多挑战, 值得进一步深入探索。本节将从数据处理、问题求解和未来发展的角度, 提出当前 MIML 算法的不足, 并探讨潜在的研究方向。

1)数据构建

MIML 任务不同于传统的图像分类或文本分类任务, 它通常需要构建多对多的映射关系, 即将多个示例组合成 1 个包, 并为该包分配多个标记。然而, 现有的 MIML 数据大多依赖手动构建, 且缺乏端到端的自动构建方法。这使得 MIML 任务往往被划分为数据构建和模型分类 2 个独立的阶段。在此过程中, 数据构建的质量直接影响到最终的分类性能, 而由于缺乏通用的自动构建方法, MIML 的应用场景也受到了一定限制。虽然一些研究已尝试自动化数据构建[45], 但这些方法通常只适用于特定类型的任务(如图像数据), 并未广泛适用于复杂或多模态数据。因此, 如何设计出能够适应多种任务特征的自动化数据构建方法, 仍是未来亟待解决的关键问题。

2)数据融合

MIML 的核心问题之一是如何有效地融合多个示例的信息, 以准确地反映样本的整体特征。每个示例可能包含不同的特征和信息, 并对样本的标记产生不同程度的影响。如何平衡这些示例的作用、避免某些示例的噪声或异常值对整体分类结果产生误导, 是数据融合中的关键问题[66]。此外, 数据融合还需要考虑示例之间的相关性及其对不同标记的贡献。尽管已有一些方法尝试通过最大池化和注意力机制[53]等技术来解决这一问题, 但在计算复杂度和模型泛化能力方面, 仍然面临着诸多挑战。

3)标记预测

MIML 的显著优势是能够在没有示例级标记的情况下, 同时进行包级和示例级标记预测, 这使其在许多实际应用中, 尤其是在医学图像分析、文本段落标记和视频片段标记等任务中, 具有广泛的应用前景。然而, 这一双粒度预测的优势也带来了挑战。由于缺乏示例级标记的监督信息, MIML 中的包级和示例级预测精度通常都会受到较大影响, 尤其是示例级标记预测。在没有精确示例级标记的情况下, 模型难以准确捕捉示例与标记之间的复杂关系, 从而导致示例级预测的准确性较低, 进而影响最终的分类结果。同时, 包级和示例级之间的关联性、标记之间的重叠性以及示例之间的相互依赖性, 也使得标记预测变得更加复杂。尽管一些研究通过秩损失函数[122]、逻辑回归模型[123-124]等优化策略尝试提高示例级预测的精度, 但这些方法大多局限于特定场景, 未能普遍适用[125]。因此, 如何在弱监督条件下提升示例级标记预测的精度, 并有效减少包级标记对示例级预测误差的影响, 仍是未来研究的关键方向。

4)社区发展

MIML 领域的实验可复现性问题一直是一个亟待解决的挑战。目前, MIML 相关研究中的实验设置通常缺乏统一的标准和详细描述, 这使得不同方法之间的比较变得困难, 并且许多实验结果难以复现。不同研究在数据划分、模型选择、评价指标等方面的设置差异使得横向比较和深入理解问题变得复杂。因此, 未来的 MIML 研究应更加注重实验过程的透明化和标准化, 明确列出实验设置、数据划分方法和评价方式等关键细节。只有确保实验设置的一致性和透明度, 才能促进不同研究成果之间的公平对比和深入讨论, 推动 MIML 学术社区的发展, 提升该领域研究的可信度和可复现性。

7 总结

MIML 不仅是机器学习领域的基础问题, 而且有着广泛的应用, 具有重要的研究价值。本文总结了 MIML 领域的进展和发展现状及其未来研究方向。MIML 为处理具有复杂结构和多语义依赖的现实世界对象提供了有效的建模框架。本文以数据复杂性为核心视角构建了系统化的 MIML 分类体系, 涵盖从标准任务到多源建模及非精确标记条件下的扩展场景。研究现状显示, 直接求解方法在性能上逐渐成为主流, 但计算效率仍是其主要瓶颈; 多视图与多类型对象建模强化了语义表达能力, 但也显著提高了模型复杂度; 同时, 对标记不确定性的鲁棒建模能力正在成为衡量模型实用性的关键指标。未来研究应聚焦于自动化数据构建、示例级标记预测以及评估体系标准化, 以推动 MIML 的规模化应用。

作者贡献声明: 王梅霞负责研究构思、结构设计、文献搜集与整理、主要内容撰写与修改; 王慧颖和黄苗苗负责文献搜集与整理、部分内容的撰写与审核; 赵宇海负责研究构思、监督指导、内容修改。

参 考 文 献

[1] Wei Xiushen, Xu Shulin, An Peng, et al. Multi-instance learning with incremental classes[J]. Journal of Computer Research and Development, 2022, 59(8): 1723−1731 (in Chinese)
(魏秀参, 徐书林, 安鹏, 等. 面向增量分类的多示例学习[J]. 计算机研究与发展, 2022, 59(8): 1723−1731)

[2] Ren Jie, Hou Bojian, Jiang Yuan. Deep forest for multiple instance learning[J]. Journal of Computer Research and Development, 2019, 56(8): 1670−1676 (in Chinese)
(任婕, 侯博建, 姜远. 多示例学习下的深度森林架构[J]. 计算机研究与发展, 2019, 56(8): 1670−1676)

[3] Li Zhongnian, Huangfu Zhiyu, Yang Kaijie, et al. Semi-supervised open vocabulary multi-label learning based on graph prompting[J]. Journal of Computer Research and Development, 2025, 62(2): 432−442 (in Chinese)
(李仲年, 皇甫志宇, 杨凯杰, 等. 基于图提示的半监督开放词汇多标记学习[J]. 计算机研究与发展, 2025, 62(2): 432−442)

[4] Duan Jie, Hu Qinghua, Zhang Lingjun, et al. Feature selection for multi-label classification based on neighborhood rough sets[J]. Journal of Computer Research and Development, 2015, 52(1): 56−65 (in Chinese)
(段洁, 胡清华, 张灵均, 等. 基于领域粗糙集的多标记分类特征选择算法[J]. 计算机研究与发展, 2015, 52(1): 56−65)

[5] Zhou Zhihua, Zhang Minling. Multi-instance multi-label learning with application to scene classification[C]//Proc of the 20th Annual Conf on Neural Information Processing Systems. Cambridge, MA: MIT, 2006: 1609−1616

[6] Zhou Zhihua, Zhang Minling, Huang Shengjun, et al. Multi-instance multi-label learning[J]. Artificial Intelligence, 2012, 176(1): 2291−2320

[7] Gecer B, Yalcinkaya O, Tasar O, et al. Evaluation of joint multi-instance multi-label learning for breast cancer diagnosis[J]. arXiv preprint, arXiv: 1510.02942, 2015

[8] Tian Yuan, Hao Wenning, Jin Dawei, et al. A review of latest multi-instance learning[C]//Proc of the 4th Int Conf on Computer Science and Artificial Intelligence. New York: ACM, 2020: 41−45

[9] Zhao Lu, Yuan Liming, Hao Kun. Review of multi-instance learning algorithms[J]. Computer Science, 2022, 49(S1): 93−99 (in Chinese)
(赵璐, 袁立明, 郝琨. 多示例学习算法综述[J]. 计算机科学, 2022, 49(S1): 93−99)

[10] Tian Yingjie, Xu Dongkuan, Zhang Chunhua. A review of multi-instance learning research[J]. Operations Research Transactions, 2018, 22(2): 1−17 (in Chinese)
(田英杰, 胥栋宽, 张春华. 多示例学习问题研究进展综述[J]. 运筹学学报, 2018, 22(2): 1−17)

[11] Gadermayr M, Tschuchnig M. Multiple instance learning for digital pathology: A review of the state-of-the-art, limitations & future potential[J]. Computerized Medical Imaging and Graphics, 2024, 112: 102337

[12] Li Hengyi, Shi Shuting, Thung F, et al. DeepReview: Automatic code review using deep multi-instance learning[C]//Proc of the 23rd Pacific-Asia Conf on Knowledge Discovery and Data Mining. Berlin: Springer, 2019: 318−330

[13] Cheplygina V, Bruijne M, Pluim J P W. Not-so-supervised: A survey of semi-supervised, multi-instance, and transfer learning in medical image analysis[J]. Medical Image Analysis, 2019, 54: 280−296

[14] Li Dongmei, Yang Yu, Meng Xianghao, et al. Review on multi-label classification[J]. Journal of Frontiers of Computer Science and Technology, 2023, 17(11): 2529−2542 (in Chinese)
(李冬梅, 杨宇, 孟湘皓, 等. 多标签分类综述[J]. 计算机科学与探索, 2023, 17(11): 2529−2542)

[15] Han Meng, Wu Hongxin, Chen Zhiqiang, et al. A survey of multi-label classification based on supervised and semi-supervised learning[J]. International Journal of Machine Learning and Cybernetics, 2023, 14(3): 697−724

[16] Tarekegn A N, Ullah M, Cheikh F A. Deep learning for multi-label learning: A comprehensive survey[J]. arXiv preprint, arXiv: 2401.16549, 2024

[17] Wang Wei, Zhou Zhihua. Learnability of multi-instance multi-label learning[J]. Chinese Science Bulletin, 2012, 57: 2488−2491

[18] Nguyen N. A new SVM approach to multi-instance multi-label learning[C]//Proc of the 10th IEEE Int Conf on Data Mining. Los Alamitos, CA: IEEE Computer Society, 2010: 384−392

[19] Li Yingxin, Ji Shuiwang, Kumar S, et al. Drosophila gene expression pattern annotation through multi-instance multi-label learning[J]. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2011, 9(1): 98−112

[20] Xu Yonghui, Min Huaqing, Song Hengjie, et al. Multi-instance multi-label distance metric learning for genome-wide protein function prediction[J]. Computational Biology and Chemistry, 2016, 63: 30−40

[21] Xu Xinshun, Xue Xiangyang, Zhou Zhihua. Ensemble multi-instance multi-label learning approach for video annotation task[C]//Proc of the 19th ACM Int Conf on Multimedia. New York: ACM, 2011: 1153−1156

[22] Yan Kaobi, Li Zhixin, Zhang Canlong. A new multi-instance multi-label learning approach for image and text classification[J]. Multimedia Tools and Applications, 2016, 75(13): 7875−7890

[23] Yin Ying, Zhao Yuhai, Li Chengguang, et al. Improving multi-instance multi-label learning by extreme learning machine[J]. Applied Sciences, 2016, 6(6): 160

[24] Xu Xin, Frank E. Logistic regression and boosting for labeled bags of instances[C]//Proc of the 8th Pacific-Asia Conf on Knowledge Discovery and Data Mining. Berlin: Springer, 2004: 272−281

[25] Boutell M R, Luo Jiebo, Shen Xipeng, et al. Learning multi-label scene classification[J]. Pattern Recognition, 2004, 37(9): 1757−1771

[26] Zhang Minling. A *k*-nearest neighbor based multi-instance multi-label learning algorithm[C]//Proc of the 22nd IEEE Int Conf on Tools with Artificial Intelligence. Los Alamitos, CA: IEEE Computer Society, 2010: 207−212

[27] Wu Qingyao, Ng M K, Ye Yunming. Markov-MIML: A Markov chain-based multi-instance multi-label learning algorithm[J]. Knowledge and Information Systems, 2013, 37(1): 83−104

[28] Zafra A, Gibaja E. Nearest neighbor-based approaches for multi-instance multi-label classification[J]. Expert Systems with Applications, 2023, 232: 120876

[29] Zhang Minling, Wang Zhijian. MIMLRBF: RBF neural networks for multi-instance multi-label learning[J]. Neurocomputing, 2009, 72(16/17/18): 3951−3956

[30] Li Cunhe, Shi Guoqiang. Improvement of learning algorithm for the multi-instance multi-label RBF neural networks trained with imbalanced samples[J]. Journal of Information Science and Engineering, 2013, 29(4): 765−776

[31] Li Cunhe, Shi Guoqiang. Weights optimization for multi-instance multi-label RBF neural networks using steepest descent method[J]. Neural Computing and Applications, 2013, 22(7/8): 1563−1569

[32] Abdechiri M, Faez K. Efficacy of utilizing a hybrid algorithmic method in enhancing the functionality of multi-instance multi-label radial basis function neural networks[J]. Applied Soft Computing, 2015, 34: 788−798

[33] Bao Xiang, Liu Guifeng, Wang Manrong. Text categorization by multi-instance multi-label and momentum stochastic gradient descent strategy[C]//Proc of the 3rd Int Conf on Algorithms, Computing and Artificial Intelligence. New York: ACM, 2020: 1−4

[34] Bao Xiang, Liu Guifeng, Yang Guoli, et al. Multi-instance multi-label text categorization algorithm based on multi-quadric function radial basis network model[C]//Proc of the 3rd Int Conf on Artificial Intelligence and Big Data. Piscataway, NJ: IEEE, 2020: 133−136

[35] Zhang Hu, Wu Wei, Wang Ding. Multi-instance multi-label learning of natural scene images: Via sparse coding and multi-layer neural network[J]. IET Computer Vision, 2018, 12(3): 305−311

[36] Yang Mei, Tang Wentao, Min Fan. Multi-instance multi-label learning based on parallel attention and local label manifold correlation[C]//Proc of the 9th IEEE Int Conf on Data Science and Advanced Analytics. Piscataway, NJ: IEEE, 2022: 1−10

[37] Chen Tongtong, Liu Chanjuan, Zou Hailin, et al. A multi-instance multi-label scene classification method based on multi-kernel fusion[C]//Proc of the 2015 SAI Intelligent Systems Conf. Piscataway, NJ: IEEE, 2015: 782−787

[38] Li Yufeng, Hu Juhua, Jiang Yuan, et al. Towards discovering what patterns trigger what labels[C]//Proc of the 26th AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2012: 1012−1018

[39] Yang Shuanghong, Zha Hongyuan, Hu Baogang. Dirichlet-Bernoulli alignment: A generative model for multi-class multi-label multi-instance corpora[C/OL]//Proc of the 23rd Neural Information Processing Systems. 2009: 2143−2150. [2025-03-26]. https://proceedings.neurips.cc/paper/2009/hash/6c8349cc7260ae62e3b1396831a8398f-Abstract.html

[40] Zha Zhengjun, Hua Xiansheng, Mei Tao, et al. Joint multi-label multi-instance learning for image classification[C]//Proc of the 2008 IEEE Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2008: 1−8

[41] He Jianjun, Gu Hong, Wang Zhelong. Bayesian multi-instance multi-label learning using Gaussian process prior[J]. Machine Learning, 2012, 88(1/2): 273−295

[42] Nguyen T, Raich R, Fern X Z, et al. MIML-AI: Mixed-supervision multi-instance multi-label learning with auxiliary information[C]// Proc of the 27th Int Workshop on Machine Learning for Signal Processing. Piscataway, NJ: IEEE, 2017: 1−6

[43] Li Xingyue, Wan Shouhong, Zou Chang, et al. Multi-instance multi-label learning for image categorization based on integrated contextual information [C]//Proc of the 9th Int Conf on Image and Graphics. Berlin: Springer, 2017: 639−650

[44] Chen Zenghai, Chi Zheru, Fu Hong, et al. Multi-instance multi-label image classification: A neural approach[J]. Neurocomputing, 2013, 99: 298−306

[45] Feng Ji, Zhou Zhihua. Deep MIML network[C]//Proc of the 31st AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2017: 1884−1890

[46] Yang Hao, Zhou T J, Cai Jianfei, et al. MIML-FCN+: Multi-instance multi-label learning via fully convolutional networks with privileged information [C]//Proc of the 2017 IEEE Conf on Computer Vision and Pattern Recognition. Los Alamitos, CA: IEEE Computer Society, 2017: 5996−6004

[47] Pan Zesi, Wang Bo, Zhang Ruibin, et al. MIML-GAN: A GAN-based algorithm for multi-instance multi-label learning on overlapping signal waveform recognition[J]. IEEE Transactions on Signal Processing, 2023, 71: 859−872

[48] Qiu Sichao, Wang Mengyi, Yang Yuanlin, et al. Meta multi-instance multi-label learning by heterogeneous network fusion[J]. Information Fusion, 2023, 94: 272−283

[49] Wu Ziyuan, Guo Wei, Zhou Wei, et al. Abductive multi-instance multi-label learning for periodontal disease classification with prior domain knowledge[J]. Medical Image Analysis, 2025, 101: 103452

[50] Zhang Minling, Zhou Zhihua. M3MIML: A maximum margin method for multi-instance multi-label learning[C]//Proc of the 8th IEEE Int Conf on Data Mining. Los Alamitos, CA: IEEE Computer Society, 2008: 688−697

[51] Jin Rong, Wang Shijun, Zhou Zhihua. Learning a distance metric from multi-instance multi-label data[C]//Proc of the 2009 IEEE Conf on Computer Vision and Pattern Recognition. Los Alamitos, CA: IEEE Computer Society, 2009: 896−902

[52] Verma Y, Jawahar C V. A robust distance with correlated metric learning for multi-instance multi-label data[C]//Proc of the 24th ACM Int Conf on Multimedia. New York: ACM, 2016: 441−445

[53] Huang Shengjun, Gao Wei, Zhou Zhihua. Fast multi-instance multi-label learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(11): 2614−2627

[54] Retz R, Schwenker F. Active multi-instance multi-label learning[C]// Proc of the 2nd European Conf on Data Analysis. Berlin: Springer, 2014: 91−101

[55] Huang Shengjun, Gao Nengneng, Chen Songcan. Multi-instance multi-label active learning[C]//Proc of the 26th Int Joint Conf on Artificial Intelligence. Palo Alto, CA: IJCAI, 2017: 1886−1892

[56] Ruiz A T, Thiam P, Schwenker F, et al. A *k*-nearest neighbor based algorithm for multi-instance multi-label active learning[C]//Proc of the 8th Int Association for Pattern Recognition Workshop on Artificial Neural Networks in Pattern Recognition. Berlin: Springer, 2018: 139−151

[57] Wu Jiansheng, Zhu Wenyong, Jiang Ye, et al. Predicting protein functions of bacteria genomes via multi-instance multi-label active learning[C]//Proc of the 3rd Int Conf on Integrated Circuits and Microsystems. Piscataway, NJ: IEEE, 2018: 302−307

[58] Su Cong, Yan Zhongmin, Yu Guoxian. Cost-effective multi-instance multilabel active learning[J]. International Journal of Intelligent Systems, 2021, 36(12): 7177−7203

[59] Su Guoliang, Wu Zhangquan, Ye Yujia, et al. Cost-efficient multi-instance multi-label active learning via correlation of features[C]//Proc of the 2023 IEEE Int Conf on Image Processing. Piscataway, NJ: IEEE, 2023: 410−414

[60] Nguyen C T, Zhan Dechuan, Zhou Zhihua. Multi-modal image annotation with multi-instance multi-label LDA[C]//Proc of the 23rd Int Joint Conf on Artificial Intelligence. Palo Alto, CA: IJCAI, 2013: 1558−1564

[61] Nguyen C T, Wang Xiaoliang, Liu Jing, et al. Labeling complicated objects: Multi-view multi-instance multi-label learning[C]//Proc of the 28th AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2014: 2013−2019

[62] Yang Yang, Wu Yifeng, Zhan Dechuan, et al. Complex object classification: A multi-modal multi-instance multi-label deep network with optimal transport[C]//Proc of the 24th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining. New York: ACM, 2018: 2594−2603

[63] Yang Yang, Fu Zhaoyang, Zhan Dechuan, et al. Semi-supervised multi-modal multi-instance multi-label deep network with optimal transport[J]. IEEE Transactions on Knowledge and Data Engineering, 2021, 33(2): 696−709

[64] Xing Yuying, Yu Guoxian, Domeniconi C, et al. Multi-view multi-instance multi-label learning based on collaborative matrix factorization[C]//Proc of the 33rd AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2019: 5508−5515

[65] Xing Yuying, Yu Guoxian, Wang Jun, et al. Weakly-supervised multi-view multi-instance multi-label learning[C]//Proc of the 29th Int Joint Conf on Artificial Intelligence. Palo Alto, CA: IJCAI, 2020: 3124−3130

[66] Hang Cheng, Wang Wei, Zhan Dechuan. Multi-modal multi-instance multi-label learning with graph convolutional network[C]//Proc of the 2021 Int Joint Conf on Neural Networks. Piscataway, NJ: IEEE, 2021: 1−8

[67] Yang Yuanlin, Yu Guoxian, Domeniconi C, et al. Deep multi-type objects multi-view multi-instance multi-label learning[C]//Proc of the 2021 SIAM Int Conf on Data Mining. Philadelphia, PA: SIAM, 2021: 486−494

[68] Yang Yuanlin, Yu Guoxian, Wang Jun, et al. Multi-typed objects multi-view multi-instance multi-label learning[C]//Proc of the 20th Int Conf on Data Mining. Piscataway, NJ: IEEE, 2020: 1370−1375

[69] Lou Qi, Raich R, Briggs F, et al. Novelty detection under multi-label multi-instance framework[C]//Proc of the 2013 IEEE Int Workshop on Machine Learning for Signal Processing. Piscataway, NJ: IEEE, 2013: 1−6

[70] Pham A T, Raich R, Fern X Z, et al. Multi-instance multi-label learning in the presence of novel class instances[C]//Proc of the 32nd Int Conf on Machine Learning. Brookline, MA: PMLR, 2015: 2427−2435

[71] Zhu Yue, Jiang Yuan, Zhou Zhihua. Multi-instance multi-label new label learning[J]. Science China Information Sciences, 2018, 48(12): 1670−1680 (in Chinese)
(朱越, 姜远, 周志华. 一种基于多示例多标记学习的新标记学习方法[J]. 中国科学: 信息科学, 2018, 48(12): 1670−1680)

[72] Zhu Yue, Ting Kaiming, Zhou Zhihua. Discover multiple novel labels in multi-instance multi-label learning[C]//Proc of the 31st AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2017, 2977−2984

[73] Yang Shujun, Jiang Yuan, Zhou Zhihua. Multi-instance multi-label learning with weak label[C]//Proc of the 23rd Int Joint Conf on Artificial Intelligence. Palo Alto, CA: IJCAI, 2013: 1862−1868

[74] Nguyen T, Raich R. Incomplete label multiple instance multiple label learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 44(3): 1320−1337

[75] Zhang Xinyue, Luo Tingjin, Liu Yueying, et al. Imbalanced multi-instance multi-label learning via coding ensemble and adaptive thresholds[C]//Proc of the 32nd ACM Int Conf on Multimedia. New York: ACM, 2024: 5413−5422

[76] Briggs F, Fern X Z, Raich R. Rank-loss support instance machines for MIML instance annotation[C]//Proc of the 18th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2012: 534−542

[77] Winn J, Criminisi A, Minka T. Object categorization by learned universal visual dictionary[C]//Proc of the 10th IEEE Int Conf on Computer Vision. Los Alamitos, CA: IEEE Computer Society, 2005: 1800−1807

[78] Lin Tsungyi, Maire M, Belongie S, et al. Microsoft COCO: Common objects in context[C]//Proc of the 13th European Conf on Computer Vision. Berlin, Springer, 2014: 740−755

[79] Ren Dong, Ma Long, Zhang Yanqing, et al. Online biomedical publication classification using multi-instance multi-label algorithms with feature reduction[C]//Proc of the 14th Int Conf on Cognitive Informatics & Cognitive Computing. Los Alamitos, CA: IEEE Computer Society, 2015: 234−241

[80] Sahlin J, Sundell H, Alm H, et al. Evaluating artificial short message service campaigns through rule based multi-instance multi-label classification [C]//Proc of the AAAI 2021 Spring Symp on Combining Machine Learning and Knowledge Engineering. Palo Alto, CA: AAAI, 2021: 1−12

[81] Mylonas N, Karlos S, Tsoumakas G. A multi-instance multi-label weakly supervised approach for dealing with emerging MeSH descriptors[C]//Proc of the 19th Int Conf on Artificial Intelligence in Medicine. Berlin: Springer, 2021: 397−407

[82] Chen Anni, Dhingra B. Hierarchical multi-instance multi-label learning for detecting propaganda techniques[C]//Proc of the 8th Workshop on Representation Learning for NLP. Stroudsburg, PA: ACL, 2023: 155−163

[83] Liu Yunduo, Xu Fang, Zhao Yushan, et al. Hierarchical multi-instance multi-label learning for Chinese patent text classification[J]. Connection Science, 2024, 36(1): 2295818

[84] Li Yuncong, Yin Cunxiang, Zhong Shenghua, et al. Multi-instance multi-label learning networks for aspect-category sentiment analysis[C]//Proc of the 2020 Conf on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2020: 3550−3560

[85] Xue Xiangyang, Zhang Wei, Zhang Jie, et al. Correlative multi-label multi-instance image annotation[C]//Proc of the 2011 Int Conf on Computer Vision. Los Alamitos, CA: IEEE Computer Society, 2011: 651−658

[86] Ding Xinmiao, Li Bing, Xiong Weihua, et al. Multi-instance multi-label learning combining hierarchical context and its application to image annotation[J]. IEEE Transactions on Multimedia, 2016, 18(8): 1616−1627

[87] Xu Xinshun, Jiang Yuan, Xue Xiangyang, et al. Semi-supervised multi-instance multi-label learning for video annotation task[C]// Proc of the 20th ACM Multimedia. New York: ACM, 2012: 737−740

[88] Song Lingyun, Liu Jun, Qian Buyue, et al. A deep multi-modal CNN for multi-instance multi-label image classification[J]. IEEE Transactions on Image Processing, 2018, 27(12): 6025−6038

[89] Chen Gang, Cheng Xiang, Su Sen, et al. Multiple-instance ranking based deep hashing for multi-label image retrieval[J]. Neurocomputing, 2020, 402: 89−99

[90] Ji Wanting, Wang Ruili. A multi-instance multi-label dual learning approach for video captioning[J]. ACM Transactions on Multimedia Computing Communications and Applications, 2021, 17(2s): 1−18

[91] Yao Di, Song Yuhang, Li Feng, et al. Anti-jamming method based on multi-instance multi-label learning[C]//Proc of the 2020 IET Int Radar Conf. 2020: 836−840

[92] Julia E, do Nascimento M Z, Faria M, et al. Deep learning-based models for performing multi-instance multi-label event classification in gameplay footage[C]// Proc of the 19th Int Joint Conf on Computer Vision, Imaging and Computer Graphics Theory and Applications. 2024: 110−121. DOI: 10.5220/0012358900003660

[93] Dietterich T G, Lathrop R H, Lozano-Pérez T. Solving the multiple instance problem with axis-parallel rectangles[J]. Artificial Intelligence, 1997, 89(1/2): 31−71

[94] Wu Jiansheng, Feng Qiaoyu, Yuan Jingzhou, et al. Predicting biological functions of G protein-coupled receptors based on fast multi-instance multi-label learning[J]. Journal of Computer Research and Development, 2018, 55(8): 1674−1682 (in Chinese)
(吴建盛, 冯巧遇, 袁京洲, 等. 基于快速多示例多标记学习的 G 蛋白偶联受体生物学功能预测[J]. 计算机研究与发展, 2018, 55(8): 1674−1682)

[95] Wu Feng, Liu Qiong, Hao Tianyong, et al. Online multi-instance multi-label learning for protein function prediction[C]//Proc of the 2016 IEEE Int Conf on Bioinformatics and Biomedicine. Los Alamitos, CA: IEEE Computer Society, 2016: 780−785

[96] Wu Jiansheng, Huang Shengjun, Zhou Zhihua. Genome-wide protein function prediction through multi-instance multi-label learning[J]. IEEE/ ACM Transactions on Computational Biology and Bioinformatics, 2014, 11(5): 891−902

[97] Meng Yang, Deng Lei, Chen Zhigang, et al. A multi-instance multi-label learning approach for protein domain annotation[C]//Proc of the 10th Int Conf Intelligent Computing in Bioinformatics. Berlin: Springer, 2014: 104−111

[98] Liu Jing, Tang Xinghua, Cui Shuanglong, et al. Predicting the function of rice proteins through multi-instance multi-label learning based on multiple features fusion[J]. Briefings in Bioinformatics, 2022, 23(3): bbac095

[99] Mei Suyu, Zhu Hao. Multi-label multi-instance transfer learning for simultaneous reconstruction and cross-talk modeling of multiple human signaling pathways[J]. BMC Bioinformatics, 2015, 16: 1−18

[100] Xu Kaixin, Zhao Ziyuan, Gu Jiapan, et al. Multi-instance multi-label learning for gene mutation prediction in hepatocellular carcinoma[C]//Proc of the 42nd Annual Int Conf of the IEEE Engineering in Medicine & Biology Society. Piscataway, NJ: IEEE, 2020: 6095−6098

[101] Li Zejun, Liao Bo, Li Yun, et al. Gene function prediction based on combining gene ontology hierarchy with multi-instance multi-label learning[J]. RSC Advances, 2018, 8(50): 28503−28509

[102] Li Tiange, Yang Yang, Shen Hongbin. HMIML: Hierarchical multi-instance multi-label learning of drosophila embryogenesis images using convolutional neural networks[C]//Proc of the 2018 IEEE Int Conf on Bioinformatics and Biomedicine. Los Alamitos, CA: IEEE Computer Society, 2018: 907−912

[103] Shang Junyuan, Hong Shenda, Zhou Yuxi, et al. Knowledge guided multi-instance multi-label learning via neural networks in medicines prediction [C/OL]// Proc of the 10th Asian Conf on Machine Learning.2018:831−846.[2025-03-26].https://proceedings.mlr.press/v95/shan g18c.html

[104] Zhang Zhihai, Wan Boyang, Zhu Sunwan, et al. MMLoc: A multi-instance multi-label learning approach for predicting protein subcellular localization from immunofluorescence images[C]//Proc of the 4th Int Conf on Neural Networks, Information and Communication Engineering. Piscataway, NJ: IEEE, 2024: 447−450

[105] Loukas C, Sgouros N P. Multi-instance multi-label learning for surgical image annotation[J]. The International Journal of Medical Robotics and Computer Assisted Surgery, 2020, 16(2): e2058

[106] Feng Yingjing, Vigmond E. Deep multi-label multi-instance classification on 12-lead ECG[C]//Proc of the 2020 Computing in Cardiology. Piscataway, NJ: IEEE, 2020: 1−4

[107] Lai Qi, Zhou Jianhang, Gan Yanfen, et al. Single-stage broad multi-instance multi-label learning (BMIML) with diverse inter-correlations and its application to medical image classification[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2023, 8(1): 828−839

[108] Mercan C, Aksoy S, Mercan E, et al. Multi-instance multi-label learning for multi-class classification of whole slide breast histopathology images[J]. IEEE Transactions on Medical Imaging, 2018, 37(1): 316−325

[109] Yan Jianjun, Shen Qingwei, Ren Jintao, et al. A multi-instance multi-label learning approach to objective auscultation analysis of traditional Chinese medicine[C]//Proc of the 4th Int Conf on Biomedical Engineering and Informatics. Piscataway, NJ: IEEE, 2011: 1626−1630

[110] Li Yongchun, Li Hong, Wang Qian, et al. Traditional Chinese

medicine formula evaluation using multi-instance multi-label framework[C]//Proc of the 2016 IEEE Int Conf on Bioinformatics and Biomedicine. Los Alamitos, CA: IEEE Computer Society, 2016: 484−488

[111] Dong Wenlong, Ma Zhongchen, Zhu Qing, et al. Two-stage multi-instance multi-label learning model for video social relationship recognition[C]//Proc of the 4th Int Conf on Intelligent Computing and Human-Computer Interaction. Piscataway, NJ: IEEE, 2023: 84−88

[112] Zhang Xiaoyu, Shi Haichao, Li Changsheng, et al. Multi-instance multi-label action recognition and localization based on spatio-temporal pre-trimming for untrimmed videos[C]//Proc of the 34th AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2020: 12886−12893

[113] Zeng Zehua, Gao Neng, Xue Cong, et al. Learning from audience interaction: Multi-instance multi-label topic model for video shots annotating[C]//Proc of the 24th IEEE Int Conf on Computer Supported Cooperative Work in Design. Piscataway, NJ: IEEE, 2021: 1075−1080

[114] Pan Zesi, Wang Shafei, Zhu Mengtao, et al. Automatic waveform recognition of overlapping LPI radar signals based on multi-instance multi-label learning[J]. IEEE Signal Processing Letters, 2020, 27: 1275−1279

[115] Wu Bin, Zhong Erheng, Horner A, et al. Music emotion recognition by multi-label multi-layer multi-instance multi-view learning[C]// Proc of the 22nd ACM Int Conf on Multimedia. New York: ACM, 2014: 117−126

[116] Briggs F, Fern X, Raich R, et al. Multi-instance multi-label class discovery: A computational approach for assessing bird biodiversity[C]//Proc of the 30th AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2016: 3807−3813

[117] Shen Chen, Jiao Jun, Yang Yahui, et al. Multi-instance multi-label learning for automatic tag recommendation[C]//Proc of the 2009 IEEE Int Conf on Systems, Man and Cybernetics. Piscataway, NJ: IEEE, 2009: 4910−4914

[118] Yang Yang, Cao Liyuan, Shi Haoyu, et al. Multi-instance multi-label learning for text-motion retrieval[C]//Proc of the 32nd ACM Int Conf on Multimedia. New York: ACM, 2024: 5829−5837

[119] Fan Wenping, Zhang Yao, Hao Qichen, et al. BAMBOO: A multi-instance multi-label approach towards VDI user logon behavior modeling[C/OL]// Proc of the 30th Int Joint Conf on Artificial Intelligence. 2021: 2367−2373. [2025-05-09]. https://www.ijcai.org/proceedings/2021/326

[120] Ma Yuling, Cui Chaoran, Yu Jun, et al. Multi-task MIML learning for pre-course student performance prediction[J]. Frontiers of Computer Science, 2020, 14(5): 145313

[121] Surdeanu M, Tibshirani J, Nallapati R, et al. Multi-instance multi-label learning for relation extraction[C]//Proc of the 2012 Joint Conf on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. Stroudsburg, PA: ACL, 2012: 455−465

[122] Briggs F, Fern X Z, Raich R, et al. Instance annotation for multi-instance multi-label learning[J]. ACM Transactions on Knowledge Discovery from Data, 2013, 7(3): 1−30

[123] Pham A T, Raich R. Kernel-based instance annotation in multi-instance multi-label learning[C]//Proc of the 2014 IEEE Int Workshop on Machine Learning for Signal Processing. Piscataway, NJ: IEEE, 2014: 1−6

[124] Pham A T, Raich R, Fern X Z. Simultaneous instance annotation and clustering in multi-instance multi-label learning[C]//Proc of the 25th Int Workshop on Machine Learning for Signal Processing. Piscataway, NJ: IEEE, 2015: 1−6

[125] Pham A T, Raich R, Fern X Z. Dynamic programming for instance annotation in multi-instance multi-label learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2381−2394



Wang Meixia, born in 1998. PhD candidate. Student member of CCF. Her main research interests include multi-instance learning, multi-label learning, and multi-modal sentiment analysis.

王梅霞, 1998 年生。博士研究生。CCF 学生会员。主要研究方向为多示例学习、多标记学习、多模态情感分析。



Wang Huiying, born in 1992. Master, engineer. Her main research interests include information system operation analysis, intelligent scheduling technology research, and data mining analysis and processing.

王慧颖, 1992 年生。硕士，工程师。主要研究方向为信息系统运行分析、智能调度技术研究、数据挖掘分析与处理。



Huang Miaomiao, born in 1999. PhD candidate. Her main research interests include machine learning and drug discovery.

黄苗苗, 1999 年生。博士研究生。主要研究方向为机器学习、药物发现。



Zhao Yuhai, born in 1975. PhD, professor. Senior member of CCF. His main research interests include intelligent computing and learning, databases, software engineering, and bioinformatics.

赵宇海, 1975 年生。博士，教授。CCF 高级会员。主要研究方向为智能计算与学习、数据库、软件工程、生物信息学。