计算机科学
COMPUTER SCIENCE

融入标签相关性先验的多视图多标签学习

盛思柔, 欧阳宵, 陶红, 侯臣平

引用本文

盛思柔, 欧阳宵, 陶红, 侯臣平. 融入标签相关性先验的多视图多标签学习[J]. 计算机科学, 2025, 52(2): 58-66.

SHENG Sirou, OUYANG Xiao, TAO Hong, HOU Chenping. Multi-view Multi-label Learning with Label Correlation Priors [J]. Computer Science, 2025, 52(2): 58-66.

相似文章推荐（请使用火狐或 IE 浏览器查看文章）

Similar articles recommended (Please use Firefox or IE to view the article)

基于申威众核处理器的Office口令恢复向量化研究

Study on Office Password Recovery Vectorization Technology Based on Sunway Many-core Processor

计算机科学, 2022, 49(11A): 210900176-5. https://doi.org/10.11896/jsjkx.210900176

基于边际概率分布匹配的主动标记分布学习

Active Label Distribution Learning Based on Marginal Probability Distribution Matching

计算机科学, 2020, 47(9): 190-197. https://doi.org/10.11896/jsjkx.200700077

基于深度学习的多标签生成研究进展

Research and Development of Multi-label Generation Based on Deep Learning

计算机科学, 2020, 47(3): 192-199. https://doi.org/10.11896/jsjkx.190300137

一种基于先验信息的混合数据聚类个数确定算法

Algorithm to Determine Number of Clusters for Mixed Data Based on Prior Information

计算机科学, 2016, 43(2): 101-104. https://doi.org/10.11896/j.issn.1002-137X.2016.02.023

一种基于Bayes估计的WSN节点信任度计算模型

Trust Computation Model of Nodes Based on Bayes Estimation in Wireless Sensor Networks

计算机科学, 2013, 40(10): 61-64.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.240600030

融入标签相关性先验的多视图多标签学习

盛思柔[1,2]　欧阳宵[1]　陶　红[1]　侯臣平[1]
1 国防科技大学理学院　长沙 410073
2 湘潭大学数学与计算科学学院　湖南 湘潭 411105
(202221511110@smail.xtu.edu.cn)

摘　要　多视图多标签学习作为一种广泛应用于图像分类、文本挖掘和生物信息学等多个领域的机器学习和数据挖掘技术,正受到研究者们的广泛关注。在此框架下,样本通常由多个视图进行表征,并且可以同时关联到多个标签。尽管已有大量方法被提出,但许多方法未能充分整合先验信息来提升学习性能,这往往导致预测性能不尽如人意。针对这一问题,文中提出了一种新的多视图多标签学习方法,称为融入标签相关性先验的多视图多标签学习(Multi-view multi-label Learning with Label Correlation Priors,MERIT)。该方法在无标签的训练数据的情况下,仅利用标签相关性先验作为弱监督信息来获取多标签预测模型,从而减少对大量标注数据的依赖。它不仅能自适应地调整不同视图的权重,还能最小化样本相似性与标签相似性之间的差异,从而更准确地描述同一组样本间的相似性。在8个多视图多标签数据集上的实验结果表明,与同类方法相比,MERIT展现出了更优越的性能。

关键词:多视图多标签;标签相关性;样本相似性;先验信息;自加权策略
中图分类号　TP399

Multi-view Multi-label Learning with Label Correlation Priors

SHENG Sirou[1,2],OUYANG Xiao[1],TAO Hong[1] and HOU Chenping[1]
1 School of Science,National University of Defense Technology,Changsha 410073,China
2 School of Mathematics and Computational Science,Xiangtan University,Xiangtan,Hunan 411105,China

Abstract　Multi-view multi-label learning,as a widely used machine learning and data mining technique in fields such as image classification,text mining,and bioinformatics,is receiving extensive attention from researchers. In this framework,samples are typically represented by multiple views and can be associated with multiple labels simultaneously. Although many methods have been proposed,many of them fail to fully integrate prior information to enhance learning performance,which often leads to unsatisfactory prediction performance. Aiming at this issue,this paper proposes a new multi-view multi-label learning method called multi-view multi-label learning with label correlation priors(MERIT). In the absence of labeled training data,this method acquires a multi-label prediction model by using only the prior of label correlations as weak supervision,thereby reducing the dependence on a large amount of annotated data. It not only adaptively adjusts the weights of different views but also accurately characterizes the similarity among samples of the same group by minimizing the discrepancy between sample similarity and label similarity. Experimental results on 8 multi-view multi-label datasets show that MERIT exhibits superior performance compared to similar methods.

Keywords　Multi-view multi-label,Label correlation,Sample similarity,Prior information,Auto-weighted strategy

1　引言

多视图多标签数据集[1]指每个样本具有多个视图,并且在每个视图中一个样本可以有多个标签。相比传统的单视图学习,多视图数据提供了不同角度的数据,数据来源更加丰富,不同的视图之间能够提供互补信息。若采用合适的方法将不同视图的数据进行整合,则能够提高分类器性能,给出较为准确的预测结果。多标签数据区别于以往的多类别数据,

到稿日期:2024-06-03　返修日期:2024-08-21
基金项目:国家重点研发计划(2022ZD0114803);国家自然科学基金重点项目(62136005);国家自然科学基金(61922087,62006238);湖南省自然科学基金(2023JJ20052)
This work was supported by the National Key Research and Development Program of China(2022ZD0114803),Key Program of the National Natural Science Foundation of China(62136005),National Natural Science Foundation of China(61922087,62006238) and Natural Science Foundation of Hunan Province,China(2023JJ20052).
通信作者:侯臣平(houchenping@nudt.edu.cn)

多类别数据指每个样本属于多个类别中的某一类，数据的标签通常用序号 1，2，3…或者 one-hot 向量来表示；而多标签数据则指每个样本可以同时属于多个类别，因此数据标签经常用 0，1 来表示样本是否属于该类别。

近年来，多视图多标签问题在各领域引起了诸多关注。比如，在图像分析中，可以使用多标签图像分类器对新上传的图像进行标注，以方便基于文本的搜索引擎进行检索[2]；在信号处理中，信号同时存在于时域和频域两个不同的视图中，通过对信号进行多标签分解，可以更全面地描述其特性，从而加深对信号的理解[3]；在文本分类中，一篇文章同时属于多个主题，因此可以为文章赋予多个标签[4]。

然而，由于多视图多标签数据的复杂性，其分析面临诸多挑战。目前，标签获取难和标签含噪的问题普遍存在，因此数据对应的标签数据获取较难。WSMLLC[5] 方法提出将标签相关性作为先验信息，为单视图数据预测标签，但该方法只能处理单视图数据，未能扩展到多视图。因此，本文提出仅利用标签相关性作为先验信息，为多视图数据构建一种不依赖于有标记数据的多标签预测模型。

针对多视图多标签数据处理中的问题，本文提出以下解决方案：1）将历史信息与当前信息进行结合，不仅能挖掘两者之间的联系，还能充分利用这些先验信息对当前信息进行处理，从而提高预测的准确性和效率；2）由于标签获取难度较大，为多视图数据分配多标签的任务变得复杂，因此若能使用历史场景中提供的标签相关性，则能减少对大量标注数据的依赖；3）若平等地分配每个视图的权重，可能会忽略不同视图的实际影响，相反，如果为每个视图分配不同的权重，则能更好地评估它们对最终结果的贡献。

为此，本文创造性地提出了一个融入标签相关性先验的多视图多标签学习（MERIT）方法，与先前提出的多视图多标签学习方法相比，该方法具有两个优势。一方面，MERIT 使用历史的标签相关性信息作为先验知识，为多视图数据集预测对应的标签，在一定程度上缓解了对有标签数据的依赖；另一方面，MERIT 在描述样本相似度时分别从视图角度和标签角度进行衡量，这种方法综合了样本特征和标签信息，能更全面地评估样本间的相似性。

综上所述，本文提出的 MERIT 方法能挖掘出数据中的潜在规律和结构，具有广泛的应用前景。本文的主要贡献如下：

1）提出了一种融入标签相关性先验的多视图多标签学习框架，其利用易于获取的标签相关性作为先验知识来指导学习过程，能够实现不依赖于有标记数据的多视图多标签预测。

2）分别从视图角度和标签角度描述样本间的相似性，不仅能够更全面地理解数据，而且当这两种描述的相似性一致时，可以增强模型对数据关系的解释能力。

3）MERIT 通过为不同的视图分配动态权重，能够更加精确地衡量各个视图对预测结果的贡献。在优化过程中，该方法能够自适应地调整这些权重，使其在迭代过程中持续优化权重分配，从而实现预测性能的最优化。

4）设计了多组实验，将本文方法与当前具有代表性的 6 种多标签方法进行对比。从多个评估模型性能的指标结果来看，本文方法显著优于其他同类方法。

2 相关工作

2.1 多标签学习

多标签学习的目的是预测一个样本所属的多个类别。与单标签学习不同，每个样本可以有多个标签，在多标签学习中，通常使用一个向量来表示标签集合，其中每个元素表示样本属于该类别的概率或得分。多标签学习的目标是最小化预测误差和提高标签之间的关联性。根据利用标签相关性的方式，现有的多标签学习方法大致可分为 3 类[6]：一阶策略、二阶策略以及高阶策略。

一阶策略假设标签之间是相互独立的，简单地将多标签问题转换为多个单标签分类问题，如文献[7]提出的二元相关性（Binary Relevance，BR）算法，该算法为每个标签独立地训练一个分类器，又如 LIFT[8] 首先对每个标签的正例和负例进行聚类分析，构造每个标签特有的特征，然后通过查询聚类结果进行训练和测试，但一阶策略在模型构建时没有考虑到标签之间的相关性。二阶策略考虑标签之间的成对关系来处理多标签学习任务，LPLC[9] 利用局部正负成对标签相关性，试图为所有训练样本的真实标签找到与其正负相关的类标签；JFSC[10] 考虑成对标签相关性来同时学习共享特征和标签特异性特征，并在学习到的低维数据表示上同时构建多标签分类器。高阶策略则利用所有类标签或者类标签子集之间的关系[11]，考虑了所有其他标签对每个标签的影响，如 GCC[12] 用标签空间扩充原始特征空间并将其聚类成组，然后分别学习每组中的标签依赖图，并在每组特定的标签依赖图上构建分类器链。

本文方法将先验信息融入学习过程，采用的是二阶策略。与一阶策略相比，该方法不仅考虑了单个标签的信息，还利用了标签之间的相关性。具体来说，它通过成对地分析标签之间的关系，从而在考虑标签相互依赖的同时，避免了高阶策略可能带来过高计算复杂度的问题。这种方法旨在平衡性能提升与计算成本之间的关系，使得在多标签学习任务中能够更高效地处理标签间的相互作用。

2.2 多视图多标签学习

在多视图多标签学习中，每个样本可以用多个视图来表示，并且该样本可同时属于多个类别标签。多视图多标签学习应用范围广泛，成为近年研究的热点。当前的多视图多标签方法分为 3 类，即前期融合、后期融合以及中期融合。

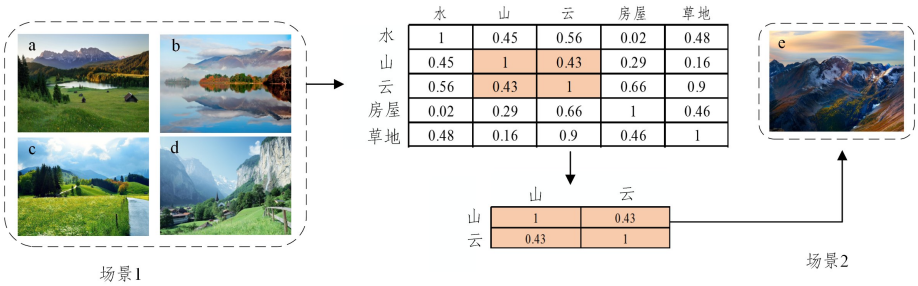
前期融合指在特征层面将来自不同视图的特征直接合并，然后应用分类器进行学习，例如 lrMMC[13]，ICM2L[14]，iMVWL[15] 中均使用了该技巧。lrMMC 首先通过矩阵分解寻找所有视图的低维共同表示，然后通过矩阵补全的方法预测未知样本的类标签；ICM2L 是一种基于个性和共性的方法，采用集成策略作出预测。然而，不同的视图可能具有不同的数据分布，这些方法将多个视图投影到统一的子空间中，却难以保证不同视图间信息的充分利用；iMVWL 则同时从具有弱标签、局部标签相关性、预测器的不完全视图的子空间中，学习一个共享子空间。

后期融合则是在决策层面进行融合，即每个视图单独学习一个分类器，然后将这些分类器的决策进行融合，例如

MLCT[16]，MLSO[17]，MMP[18]在视图上分别采用不同的分类器得到最终的结果。MLCT 利用共同训练方法，通过标签共现信息来处理多标签类别不平衡问题；MLSO 为每个视图构建一个支持向量机分类器；MMP 同时探讨视图异质性和标签异质性。但由于这些方法为每个视图都构造一个单独的分类器，因此方法的复杂性相应增加。

中期融合在模型的训练过程中进行，通常在不同的学习阶段或者不同的模型层上进行，较常见于深度方法。例如 Radhika 等[19]提出在卷积层上进行中期融合可以取得更好的效果；CRBSP[20]运用深度学习方法，结合中期融合策略，进行更为灵活的融合位置选择，将预测任务转化为二分类问题。

本文采用的是中期融合策略，该方法既没有一开始就将所有视图直接进行拼接，也没有首先单独为每个视图构建分类器再将分类器融合，而是先区分了不同视图的影响，并且与构建分类器的过程同时进行。因此，与前期融合方法不同，后者仅基于一个共享的表征来构建分类器，忽略了视图之间的差异性；同样，与后期融合方法不同，其仅为每个视图单独构建分类器，未能有效地整合视图之间的联合信息。图 1 给出了本文利用标签相关性先验信息的示例。



注：图中，(a)为山＋水＋房屋＋草地；(b)为山＋水＋云；(c)为山＋云＋草地；(d)为山＋房屋＋云＋草地；(e)为山＋云；场景 1 表示历史场景中的多标签图像数据集；场景 2 表示当前场景的多标签数据集。

图 1 利用标签相关性先验信息示例

Fig. 1 Example of using label correlation priors

3 融入标签相关性先验的多视图多标签学习

3.1 符号与定义

一般来说，对于矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$，它的第 i 行第 j 列元素记为 a_{ij}，第 i 行元素记为 $a_{i\cdot}$，第 j 列元素表示为 $a_{\cdot j}$。给定数据集 $\boldsymbol{X}=\{X^v\}_{v=1}^m$，其表示具有 n 个样本及 m 个视图的多视图数据集。而 $X^v\in\mathbb{R}^{n\times d_v}$ 为第 v 个视图的特征空间，其中 d_v 表示第 v 个视图的维数，d 表示所有视图的维数之和。

本文的目标就是在给定数据集 $\boldsymbol{X}\in\mathbb{R}^{n\times d}$ 以及历史标签相关性矩阵 $\boldsymbol{B}\in\mathbb{R}^{c\times c}$ 的前提下，且当数据的标签矩阵 $\boldsymbol{Y}\in\mathbb{R}^{n\times c}$ 未知的情况下，通过 MERIT 模型得出样本的标签指示矩阵 $\boldsymbol{T}\in\mathbb{R}^{n\times c}$，其中的元素 $t_{ik}\geqslant 0$ 表示将第 i 个样本分配给第 k 个类别的标签置信度，并且记 $\boldsymbol{t}_i=[t_{i1},t_{i2},\cdots,t_{ic}]^{\mathrm{T}}$ 为第 i 个样本在每个类别上的标签置信度。表 1 列出了本文使用的符号及对应的含义。

表 1 符号说明

Table 1 Symbol notations

符号	含义
n	样本数
v	视图索引
m	视图数
d_v	第 v 个视图的维数
d	多视图数据的总维数
c	标签数
$\boldsymbol{X}$	数据矩阵
X^v	第 v 个视图的数据矩阵
$\boldsymbol{\alpha}$	视图权重向量
$\boldsymbol{S}$	样本相关性矩阵
$\boldsymbol{B}$	标签相关性矩阵
$\boldsymbol{T}$	标签指示矩阵
$\tilde{\boldsymbol{T}}$	规范化的标签指示矩阵

3.2 模型与方法

3.2.1 自适应调整视图权重

在多视图场景中，一个样本可以由多个不同的视图来表示，如果这些视图能合理地组合到一起，那么最终结果的准确性将会得到提升。

对于多视图数据集 $\boldsymbol{X}$，将不同视图的样本相似度结合起来组成一个样本相关性矩阵 $\boldsymbol{S}\in[0,1]^{n\times n}$，它能与所有数据点建立起联系，因此样本间相似的概率可以用 $\boldsymbol{S}$ 表示。一般来说，样本间的距离用 $\|\mathrm{x}_i^{(v)}-\mathrm{x}_j^{(v)}\|_2^2$ 表示，而距离接近的样本点应当有更大的相似性，因此 s_{ij} 与样本 x_i 和 x_j 间的距离应当具有负相关关系。如式(1)所示，其中 λ_1 用来控制权重分布，第 2 项是为了防止得出最接近 x_i 的样本点概率为 1 而其他样本点概率为 0 的平凡解[21]。其中 $\boldsymbol{\alpha}=[\alpha^{(1)},\alpha^{(2)},\cdots,\alpha^{(m)}]^{\mathrm{T}}$ 是视图权重向量，λ_1 是调整所有视图权重分布的参数。

$$\begin{cases}\min\limits_{\boldsymbol{S},\boldsymbol{\alpha}}\sum\limits_{v=1}^{m}\sum\limits_{i,j=1}^{n}(\alpha^{(v)})^{\lambda_1}\|x_i^{(v)}-x_j^{(v)}\|_2^2 s_{ij}+\beta\|\boldsymbol{s}\|_F^2\\ \text{s.t.}\quad\forall i,j,s_i^{\mathrm{T}}\mathbf{1}=1,s_{ij}\geqslant 0\end{cases}\tag{1}$$

3.2.2 先验信息

计算类别相似性通常用 $\boldsymbol{t}_i^{\mathrm{T}}\boldsymbol{t}_j$ 点乘的形式，但这种方法将每个标签视为独立的，没有考虑到标签之间的相关性[22]。对于没有共同标签的样本，点乘形式赋予类别的相似性为 0，但如果它们所分配的类别之间有密切关系，而两者没有共同标签，它们仍然可以具有强相关关系。

因此，对于历史信息中多视图数据集对应的多标签矩阵，本文引入先验的标签相关性矩阵 $\boldsymbol{B}=(b_{kl})_{c\times c}$ 来表示标签之间的相似性，采用余弦相似度计算，元素 $b_{kl}\geqslant 0$ 代表两类标签之间的相似性。不再以简单的点乘形式来表示标签相似性，而将标签相关性以加权的形式表示，即用 $\boldsymbol{t}_i^{\mathrm{T}}\boldsymbol{B}\boldsymbol{t}_j$ 来表示。对于

样本x_i与x_j，我们期望从多视图数据中获取的样本相似性与从标签角度来衡量的相关性尽可能接近，即$s_{ij}\approx \boldsymbol{t}_i^{\mathrm{T}}\boldsymbol{B}\boldsymbol{t}_j$，即求解如下优化问题：

$$\begin{cases}\arg\min\limits_{T}\sum\limits_{i,j=1}^{n}(s_{ij}-\sum\limits_{k,l=1}^{c}t_{ik}b_{kl}t_{lj})^2\\ \text{s.t.}\ \ t_{ik}\geqslant 0;i,j=1,\cdots,n;k,l=1,\cdots,c\end{cases}\tag{2}$$

将两项结合，得到的目标函数如式(3)所示：

$$\begin{cases}\min\limits_{\boldsymbol{S},\boldsymbol{T},\boldsymbol{\alpha}}\sum\limits_{v=1}^{m}\sum\limits_{i,j=1}^{n}(\alpha^{(v)})^{\lambda_1}\parallel x_i^{(v)}-x_j^{(v)}\parallel_2^2 s_{ij}+\lambda_2\sum\limits_{i,j=1}^{n}(s_{ij}-\\ \qquad\sum\limits_{k,l=1}^{c}t_{ik}b_{kl}t_{lj})^2\\ \text{s.t.}\ \ \forall i,j,s_{ij},t_{ik},\boldsymbol{\alpha}\geqslant 0,\boldsymbol{s}_i^{\mathrm{T}}\mathbf{1}=1,\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{1}=1\\ \qquad i,j=1,\cdots,n;k,l=1,\cdots,c;v=1,\cdots,m\end{cases}\tag{3}$$

3.2.3 模型描述

本文提出利用标签相关性的多视图多标签模型，即先通过多视图数据集 X，将每个视图上的样本相似度矩阵组合得到一个共有的样本相似性矩阵 $\boldsymbol{S}$，同时考虑样本间距离与相似度之间的负相关关系，并且为不同视图赋予权重向量 $\boldsymbol{\alpha}=[\alpha^{(1)},\alpha^{(2)},\cdots,\alpha^{(m)}]^{\mathrm{T}}$，在优化过程中自适应调整视图权重以区分不同视图对分类效果的影响；接着把从历史信息中获取的标签相关性作为先验信息输入；最后利用从多视图数据集上描述的样本相关性与从标签角度衡量的样本相似度的一致性，建立一个多视图多标签模型。图 2 给出了 MERIT 算法的流程。

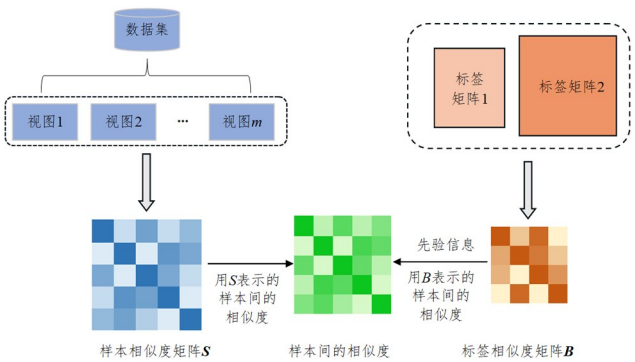


图 2 MERIT 算法的描述

Fig. 2 Description of MERIT algorithm

3.3 模型求解

本文采用交替优化方法来优化目标函数，由目标函数可知，需要优化的 3 个变量分别是 $\boldsymbol{S}$，$\boldsymbol{\alpha}$ 和 $\boldsymbol{T}$。在优化过程中，采取交替迭代的方式进行更新。

3.3.1 固定 $\boldsymbol{\alpha}$，$\boldsymbol{T}$，求解 $\boldsymbol{S}$

简化目标函数式(3)，即求解式(4)：

$$\min_{\boldsymbol{S}\in[0,1]^{n\times n},\boldsymbol{s}_i^{\mathrm{T}}\mathbf{1}=1}\sum_{v=1}^{m}\sum_{i,j=1}^{n}(\alpha^{(v)})^{\lambda_1}\parallel x_i^{(v)}-x_j^{(v)}\parallel_2^2 s_{ij}+\lambda_2\sum_{i,j=1}^{n}(s_{ij}-\sum_{k,l=1}^{c}t_{ik}b_{kl}t_{jl})^2\tag{4}$$

为了将计算过程简化，将向量$\boldsymbol{d}_i$的第 j 个元素记为 $d_{ij}^x=\sum\limits_{v=1}^{m}(\alpha^{(v)})^{\lambda_1}\parallel x_i^{(v)}-x_j^{(v)}\parallel_2^2$，此时的优化问题转变为求解式(5)：

$$\min_{\boldsymbol{S}}\sum_{i,j=1}^{n}\left(s_{ij}+\frac{d_{ij}^x}{2\lambda_2}-[\boldsymbol{T}\boldsymbol{B}\boldsymbol{T}^{\mathrm{T}}]_{ij}\right)^2\tag{5}$$

其中，$[\boldsymbol{T}\boldsymbol{B}\boldsymbol{T}^{\mathrm{T}}]_{ij}$表示第 i 个样本与第 j 个样本之间的相似性，是从标签角度来衡量的，并且记 $h_{ij}=\dfrac{d_{ij}^x}{2\lambda_2}-[\boldsymbol{T}\boldsymbol{B}\boldsymbol{T}^{\mathrm{T}}]_{ij}$。最终将问题简化为：

$$\begin{cases}\min\limits_{\boldsymbol{S}}\sum\limits_{i,j=1}^{n}(s_{ij}+h_{ij})^2\\ \text{s.t.}\ \ \boldsymbol{s}_i^{\mathrm{T}}\mathbf{1}=1;s_{ij}\geqslant 0\end{cases}\tag{6}$$

该函数可以通过拉格朗日乘数法以及 KKT 条件来求解，其中 k 为邻居数，设置为 5。

$$s_{ij}=\left(\frac{1}{k}(1+\sum_{j=1}^{k}h_{ij})-h_{ij}\right)_+\tag{7}$$

由于第 i 个样本与第 j 个样本之间的相似度s_{ij}与s_{ji}是一致的，因此需要保证 $\boldsymbol{S}$ 为对称矩阵，故对 $\boldsymbol{S}$ 进行对称化，采用 $\boldsymbol{S}=(\boldsymbol{S}+\boldsymbol{S}^{\mathrm{T}})/2$。

3.3.2 固定 $\boldsymbol{S}$，$\boldsymbol{T}$，求解 $\boldsymbol{\alpha}$

将与 $\boldsymbol{\alpha}$ 无关的项剔除，此时的目标函数变为：

$$\min_{\boldsymbol{\alpha}\geqslant 0,\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{1}=1}\sum_{v=1}^{m}\sum_{i,j=1}^{n}(\alpha^{(v)})^{\lambda_1}\parallel x_i^{v}-x_j^{v}\parallel_2^2 s_{ij}\tag{8}$$

式(8)的拉格朗日函数为：

$$L(\boldsymbol{\alpha},\eta)=\sum_{v=1}^{m}\sum_{i,j=1}^{n}(\alpha^{(v)})^{\lambda_1}\parallel x_i^{(v)}-x_j^{(v)}\parallel_2^2 s_{ij}-\eta(\sum_{v=1}^{m}\alpha^{(v)}-1)\tag{9}$$

对$\alpha^{(v)}$求偏导并令其等于 0，得：

$$\alpha^{(v)}=\left(\frac{\eta}{\lambda_1\sum\limits_{i,j=1}^{n}\parallel x_i^{(v)}-x_j^{(v)}\parallel_2^2 s_{ij}}\right)^{\frac{1}{\lambda_1-1}}\tag{10}$$

由于对 $\boldsymbol{\alpha}$ 有约束条件$\sum\limits_{v=1}^{m}\alpha^{(v)}=1$，因此求解出的 $\boldsymbol{\alpha}$ 为：

$$\alpha^{(v)}=\frac{(\sum\limits_{i,j=1}^{n}\parallel x_i^{(v)}-x_j^{(v)}\parallel_2^2 s_{ij})^{\frac{1}{\lambda_1-1}}}{(\sum\limits_{v=1}^{m}\sum\limits_{i,j=1}^{n}\parallel x_i^{(v)}-x_j^{(v)}\parallel_2^2 s_{ij})^{\frac{1}{\lambda_1-1}}}\tag{11}$$

3.3.3 固定 $\boldsymbol{S}$，$\boldsymbol{\alpha}$，求解 $\boldsymbol{T}$

将与 $\boldsymbol{T}$ 无关的项剔除，即求解目标函数第 2 项的上界。在式(12)中，$\bar{\boldsymbol{T}}$ 表示 $\boldsymbol{T}$ 在上一次迭代中获得的值。

$$\begin{aligned}&\sum_{i,j=1}^{n}(s_{ij}-\sum_{k,l=1}^{c}t_{ik}b_{kl}t_{jl})^2\\ &\leqslant\sum_{i,j=1}^{n}\sum_{k,l=1}^{c}\frac{\bar{t}_{ik}b_{kl}\bar{t}_{jl}}{[\bar{\boldsymbol{T}}\boldsymbol{B}\bar{\boldsymbol{T}}^{\mathrm{T}}]_{ij}}\left(s_{ij}-[\bar{\boldsymbol{T}}\boldsymbol{B}\bar{\boldsymbol{T}}^{\mathrm{T}}]_{ij}\frac{t_{ik}b_{kl}t_{jl}}{[\bar{\boldsymbol{T}}\boldsymbol{B}\bar{\boldsymbol{T}}^{\mathrm{T}}]_{ij}}\right)^2\\ &\leqslant\sum_{i,j=1}^{n}(s_{ij}^2+\sum_{l=1}^{c}[\bar{\boldsymbol{T}}\boldsymbol{B}\bar{\boldsymbol{T}}^{\mathrm{T}}]_{ij}[\bar{\boldsymbol{T}}\boldsymbol{B}]_{il}\frac{t_{jl}^4}{\bar{t}_{jl}^{\,3}}-4\sum_{l=1}^{c}s_{ij}[\bar{\boldsymbol{T}}\boldsymbol{B}]_{jl}\bar{t}_{jl}\\ &\qquad\lg t_{jl}-2s_{ij}[\bar{\boldsymbol{T}}\boldsymbol{B}\bar{\boldsymbol{T}}^{\mathrm{T}}]_{ij}+4\sum_{k=1}^{c}s_{ij}\bar{t}_{ik}[\boldsymbol{B}\bar{\boldsymbol{T}}^{\mathrm{T}}]_{kj}\lg\bar{t}_{ik})\end{aligned}\tag{12}$$

对边界函数关于t_{jl}求导并令其等于 0，得到的结果如式(13)所示：

$$t_{jl}=\left[\frac{\sum\limits_{i=1}^{n}s_{ij}[\bar{\boldsymbol{T}}\boldsymbol{B}]_{il}\bar{\boldsymbol{T}}_{jl}^4}{\sum\limits_{i=1}^{n}[\bar{\boldsymbol{T}}\boldsymbol{B}\bar{\boldsymbol{T}}^{\mathrm{T}}]_{ij}[\bar{\boldsymbol{T}}\boldsymbol{B}]_{il}}\right]^{\frac{1}{4}}\tag{13}$$

最后，由于求出的 $\boldsymbol{T}$ 矩阵中的值是未经规范化的，因此需要对其进行归一化，以便用阈值对其作出判断。即：

$$\tilde{t}_{jl}=\frac{t_{jl}}{\sum_{l=1}^{c} t_{jl}}, j=1,\cdots,n; l=1,\cdots,c \tag{14}$$

3.4 算法描述

MERIT 的算法描述如算法 1 所示。

算法 1 MERIT

输入：多视图数据集 **X**；历史标签相关性矩阵 **B**；阈值 thresh=1/c

输出：图权重向量 **α**；通过式(14)归一化得到标签指示矩阵 $\tilde{\mathbf{T}}$

初始化：

1. 初始化样本相关性矩阵 **S**
2. 初始化视图权重向量 $\alpha^{(v)}=1/m$
3. 随机初始化标签指示矩阵 **T**

过程：

重复

1. 固定 **α**，**T**，通过式(7)更新 **S**
2. 固定 **S**，**T**，通过式(11)更新 **α**
3. 固定 **α**，**S**，通过式(13)更新 **T**

直至收敛

判断

若 $\tilde{t}_{ij}>$thresh，则预测的标签为 1，否则为 0

3.5 收敛性证明

证明：令

$$F(\boldsymbol{S},\boldsymbol{\alpha},\boldsymbol{T})=\min_{\boldsymbol{S},\boldsymbol{\alpha},\boldsymbol{T}}\sum_{v=1}^{m}\sum_{i,j=1}^{n}(\alpha^{(v)})^{\lambda_1}\|x_i^{(v)}-x_j^{(v)}\|_2^2 s_{ij}+\lambda_2\sum_{i,j=1}^{n}(s_{ij}-\sum_{k,l=1}^{c}t_{ik}b_{kl}t_{lj})^2 \tag{15}$$

本文用 $\bar{\boldsymbol{S}},\bar{\boldsymbol{T}},\bar{\boldsymbol{\alpha}}$ 表示上一步迭代得到的值，用 $\boldsymbol{S}^*,\boldsymbol{T}^*,\boldsymbol{\alpha}^*$ 表示当前步骤获得的最优解，此时：

$$\boldsymbol{S}^*=\min\sum_{v=1}^{m}\sum_{i,j=1}^{n}(\alpha^{(v)})^{\lambda_1}\|x_i^{(v)}-x_j^{(v)}\|_2^2 s_{ij}+\lambda_2\sum_{i,j=1}^{n}(s_{ij}-\sum_{k,l=1}^{c}t_{ik}b_{kl}t_{lj})^2 \tag{16}$$

因此有：

$$F(\boldsymbol{S}^*,\bar{\boldsymbol{\alpha}},\bar{\boldsymbol{T}})\leqslant F(\bar{\boldsymbol{S}},\bar{\boldsymbol{\alpha}},\bar{\boldsymbol{T}}) \tag{17}$$

$$F(\boldsymbol{S}^*,\boldsymbol{\alpha}^*,\bar{\boldsymbol{T}})\leqslant F(\boldsymbol{S}^*,\bar{\boldsymbol{\alpha}},\bar{\boldsymbol{T}}) \tag{18}$$

$$F(\boldsymbol{S}^*,\boldsymbol{\alpha}^*,\bar{\boldsymbol{T}})\leqslant F(\bar{\boldsymbol{S}},\bar{\boldsymbol{\alpha}},\bar{\boldsymbol{T}}) \tag{19}$$

最后证明 **T** 的更新不会增加目标函数的值，记为：

$$f(\boldsymbol{T})=\lambda_2\sum_{i,j=1}^{n}(s_{ij}-\sum_{k,l=1}^{c}t_{ik}b_{kl}t_{lj})^2 \tag{20}$$

当前迭代中原函数放大后的项记为：

$$g(\boldsymbol{T},\bar{\boldsymbol{T}})=\sum_{i,j=1}^{n}(s_{ij}^2+\sum_{l=1}^{c}[\bar{\boldsymbol{T}}\boldsymbol{B}\bar{\boldsymbol{T}}^{\mathrm{T}}]_{ij}[\bar{\boldsymbol{T}}\boldsymbol{B}]_{il}\frac{t_{jl}^4}{\bar{t}_{jl}^3}-4\sum_{l=1}^{c}s_{ij}[\bar{\boldsymbol{T}}\boldsymbol{B}]_{jl}\bar{t}_{jl}\lg t_{jl}-2s_{ij}[\bar{\boldsymbol{T}}\boldsymbol{B}\bar{\boldsymbol{T}}^{\mathrm{T}}]_{ij}+4\sum_{k=1}^{c}s_{ij}\bar{t}_{ik}[\boldsymbol{B}\bar{\boldsymbol{T}}^{\mathrm{T}}]_{ij}\lg\bar{t}_{ik}) \tag{21}$$

且 $\sum_{l=1}^{c}[\bar{\boldsymbol{T}}\boldsymbol{B}\bar{\boldsymbol{T}}^{\mathrm{T}}]_{ij}[\bar{\boldsymbol{T}}\boldsymbol{B}]_{il}\frac{t_{jl}^4}{\bar{t}_{jl}^3}=(\sum_{k,l=1}^{c}\bar{t}_{ik}b_{kl}\bar{t}_{jl})^2$，取 $\boldsymbol{T}=\bar{\boldsymbol{T}}$，即有：

$$g(\bar{\boldsymbol{T}},\bar{\boldsymbol{T}})=f(\bar{\boldsymbol{T}}) \tag{22}$$

由于 $\boldsymbol{T}^*=\arg\min_{T} g(\boldsymbol{T},\bar{\boldsymbol{T}})$，因此

$$g(\boldsymbol{T}^*,\bar{\boldsymbol{T}})\leqslant g(\bar{\boldsymbol{T}},\bar{\boldsymbol{T}})=f(\bar{\boldsymbol{T}}) \tag{23}$$

$$f(\boldsymbol{T}^*)\leqslant g(\boldsymbol{T}^*,\bar{\boldsymbol{T}}) \tag{24}$$

$$f(\boldsymbol{T}^*)\leqslant f(\bar{\boldsymbol{T}}) \tag{25}$$

即：

$$F(\boldsymbol{S}^*,\boldsymbol{\alpha}^*,\boldsymbol{T}^*)\leqslant F(\bar{\boldsymbol{S}},\bar{\boldsymbol{\alpha}},\bar{\boldsymbol{T}}) \tag{26}$$

因此，随着迭代次数的增加，目标函数是收敛的。

3.6 复杂度分析

1)空间复杂度分析。算法的主要空间开销为样本相似度矩阵 $\boldsymbol{S}$、视图权重向量 $\boldsymbol{\alpha}$、标签指示矩阵 $\boldsymbol{T}$，其空间复杂度分别为 $O(n^2)$，$O(m)$，$O(nc)$，因此 MERIT 的空间复杂度为 $O(n^2+m+nc)$。由于 $m,c\ll n$，因此算法的空间复杂度为 $O(n^2)$。

2)时间复杂度分析。首先更新 $\boldsymbol{S}$ 的最主要计算为 h_{ij} 及 $\bar{\boldsymbol{T}}\boldsymbol{B}\bar{\boldsymbol{T}}^{\mathrm{T}}$，时间复杂度分别为 $O(nd)$ 及 $O(nc^2+n^2c)$，因此其计算复杂度为 $O(nd+nc^2+n^2c)$；更新 $\boldsymbol{\alpha}$ 主要在于其分母的更新，因此其计算复杂度为 $O(nd)$；更新 $\boldsymbol{T}$ 的主要计算在于矩阵 $\bar{\boldsymbol{T}}\boldsymbol{B}\bar{\boldsymbol{T}}^{\mathrm{T}}$ 的计算，因此复杂度为 $O(nc^2+n^2c)$。MERIT 的时间复杂度为 $O(n^2c)$。

4 实验分析

4.1 数据集及实验设置

4.1.1 数据集描述

本文在 cal500[1)]，emotions[1)]，languagelog[2)]，enron[1)]，scene[1)]，pascal07[3)]，nus_wide_scene[4)]，mirflickr[5)] 共 8 个数据集上进行了实验，这些数据集覆盖了多个领域。

其中，cal500，emotions 为音乐数据集，其在每首歌曲上有多个特征视图，标签为音频表现的情绪；languagelog，enron 为文本数据集，它们来源于不同的文本内容，languagelog 数据集来源于语言论坛讨论，而 enron 数据集则与电子邮件通信相关；scene，pascal07，nus_wide_scene，mirflickr 则是图像数据集，一般包含颜色、距离等视图。其中 scene 数据集由自然场景图像组成，其中每张图像都与一组标签进行人工关联。对于 pascal07 数据集，选择了其中具有代表性的 DenseHue，Gist，RGB 这 3 个视图，这些视图分别捕捉了图像的颜色、全局结构和原始像素信息。nus_wide_scene 选取全部视图，共有 5 个，即 CH，CM，CORR，EDH，WH，其标签为图像中出现的物品类别。mirflickr 数据集则选择了具有代表性的 5 个视图，即 DenseHue，DenseSift，Gist，HarrisSift，Hsv，这些视图综合了颜色、纹理、形状和结构等多种图像特征。数据集的描述如表 2 所列。

1) https://mulan.sourceforge.net/

2) http://languagelog.ldc.upenn.edu/nll/

3) http://lear.inrialpes.fr/people/guillaumin/data.php

4) https://lms.comp.nus.edu.sg/wpcontent/uploads/2019/research/nuswide/NUS-WIDE.html

5) http://www.flickr.com/

表 2 实验数据集

Table 2 Experimental datasets

数据集	样本数	视图数	标签数	维度数
cal500	87	2	174	[34,34]
emotions	297	2	6	[64,8]
languagelog	730	2	75	[500,504]
enron	851	2	53	[300,701]
scene	1 204	2	6	[49,245]
pascal07	4 982	3	20	[100,512,4 096]
nus_wide_scene	3 340	5	31	[63,224,143,72,127]
mirflickr	12 500	5	38	[100,1 000,512,1 000,4 096]

4.1.2 评价指标

本文采用 MacroF1 与 MicroF1 两个评价指标，两者的值越大，分类效果越好。

1）MacroF1

首先定义精确率与召回率两个概念。精确率表示在所有模型预测为正类的样本中，实际上为正类的比例。而召回率则表示在所有实际为正类的样本中，被模型正确预测为正类的比例。两者的计算式分别为：

$$Precision=\frac{TP}{TP+FP} \tag{27}$$

$$Recall=\frac{TP}{TP+FN} \tag{28}$$

其中，TP 为真正例，FP 为假负例，FN 为假负例。

F1 分数依赖于精确率和召回率，其计算式为：

$$F1=\frac{2\times Precision\times Recall}{Precision+Recall} \tag{29}$$

因此宏观 F1 分数（MacroF1）的计算方式是先分别计算每个类别的 F1 分数，然后对这些 F1 分数取算术平均值，其中 C 表示类别数量，$F1_i$ 是第 i 个类别的 $F1$ 分数，具体计算式如下：

$$MacroF1=\frac{1}{C}\sum_{i=1}^{C}F1_i \tag{30}$$

2）MicroF1

TP_i，FP_i，FN_i 分别为第 i 个类别的真正例、假正例、假反例分数量，则其对应的精确率、召回率分别表示为：

$$Precision=\frac{\sum_{i=1}^{C}TP_i}{\sum_{i=1}^{C}TP_i+\sum_{i=1}^{C}FP_i} \tag{31}$$

$$Recall=\frac{\sum_{i=1}^{C}TP_i}{\sum_{i=1}^{C}TP_i+\sum_{i=1}^{C}FN_i} \tag{32}$$

微观 F1 分数（MicroF1）的计算方式是将所有类的 TP，FP，FN 汇总起来计算精确率和召回率，再用这些值来计算 F1 分数：

$$MicroF1=\frac{2\times Precision\times Recall}{Precision+Recall} \tag{33}$$

4.1.3 实验设置

本文将数据集分为两部分，一部分用于测试，比例为 50%；另一部分则作为监督方法的训练集或为多标签学习方法提供标签相关性矩阵，在每个数据集上实验 5 次，并记录下 5 次重复实验的均值和方差。对比实验设置如下：

1）WSMLLC[5] 是一种弱监督多标签学习方法，该方法与本文类似，第一种将标签相关性作为先验信息，但只在单视图上做实验，因此将所有视图中效果最佳的记为 WSMLLC-B；另一种将所有视图拼成一个视图作为输入，将该实验的结果记为 WSMLLC-C。

2）ICM2L[14] 是一种多标签多视图方法，该方法利用多视图多标签数据的共性和特性，并将其纳入到联合模型中，从而改善分类性能。

3）McWL[23] 是基于矩阵完善的多视图弱标签学习方法，旨在通过利用多视图的数据和弱标签信息来提高学习性能。

4）iMVWL[15] 是不完全多视图弱标签学习方法，该方法在处理缺失标签方面表现出了卓越的性能。

5）有监督分类方法 LIFT[8] 是单视图多标签方法，因此本文将该方法分成两种。第一种将视图分离，对每个视图单独做实验，随后选取所有视图中效果最佳的作为实验结果，记为 LIFT-B；另一种将所有视图拼成一个视图作为输入，将该实验的结果记为 LIFT-C。

6）CNMF[22] 是采用非负矩阵分解的半监督多标签学习方法，运用了标签相关性矩阵。该方法为单视图方法，因此将所有视图拼接的结果记为 CNMF-C。

LIFT，ICM2L，McWL 均是监督方法，因此将有监督分类方法的训练集比例设置为 5%；WSMLLC 利用了标签相关性先验信息，因此将用于计算标签相关性的历史标签比例设置为 50%。

本文的参数 λ_1 主要候选集为 $\{10^{-3},10^{-2},10^{-1},1,10,10^2,10^3\}$，$\lambda_2$ 则在 $\{1.4,1.6,1.8,2,3\}$ 中调节，类别划分的阈值 z 设定为 $1/C$。对于其他对比实验，本文采用给定的默认参数。

4.2 对比实验结果

为了验证本文方法的性能，本小节在 8 个多视图多标签数据集上与 WSMLLC-B，WSMLLC-C，McWL，ICM2L，iMVWL 进行了实验。在多标签学习的评价指标方面，本文采用了 2 种评价标准：宏观平均 F1 分数（MacroF1）和微观平均 F1 分数（MicroF1）。这 2 个指标的数值越高，表明本文方法的性能越优。不同数据表示下的 MacroF1 和 MicroF1 如表 3 和表 4 所列。

表 3 不同数据表示下的 MacroF1

Table 3 MacroF1 with different data representations

数据集	WSMLLC-B	WSMLLC-C	ICM2L	McWL	iMVWL	MERIT
cal500	0.1507	0.0437	0.0536	0.0541	0.1165	**0.2320**
	0.0421	0.0119	0.0010	0.0016	0.0079	**0.0022**
emotions	0.0845	0.2528	0.2984	0.1510	0.1922	**0.3761**
	0.0646	0.0428	0.0933	0.0012	0.0600	**0.0078**
languagelog	0.2058	0.2179	0.0675	0.0825	0.1992	**0.2788**
	0.0036	0.0032	0.0005	0.0338	0.0138	**0.0021**
enron	0.0485	0.0509	0.0596	0.0634	0.0596	**0.1101**
	0.0039	0.0012	0.0001	0.0084	0.0000	**0.0016**
scene	0.2234	0.2144	0.2125	0.0935	0.2277	**0.2623**
	0.0123	0.0154	0.0893	0.0005	0.0431	**0.0092**
pascal07	0.1056	0.1006	0.0093	0.0209	0.0736	**0.1185**
	0.0007	0.0013	0.0001	0.0159	0.0034	**0.0015**
nus_wide_scene	0.0855	0.0873	0.0186	0.0186	0.0886	**0.1642**
	0.0036	0.0026	0.0001	0.0002	0.0083	**0.0039**
mirflickr	0.1477	0.1389	0.0093	0.0611	0.0093	**0.1741**
	0.0008	0.0007	0.0000	0.0473	0.0001	**0.0011**

注：第一行是均值；第二行是方差；加粗数据表示最佳结果；带下划线的数据表示次优结果。

表 4 不同数据表示下的 MicroF1

Table 4 MicroF1 with different data representations

数据集	WSMLLC-B	WSMLLC-C	ICM2L	McWL	iMVWL	MERIT
cal500	0.1730	0.1539	0.1834	0.1845	**0.3763**	0.2338
	0.0573	0.0214	0.0020	0.0034	**0.0164**	0.0022
emotions	0.2143	0.2682	**0.4527**	0.3103	0.3272	0.3829
	0.1428	0.0441	**0.0830**	0.0037	0.0351	0.0089
languagelog	0.2238	0.2390	0.1857	0.2429	**0.4495**	0.3076
	0.0029	0.0030	0.0011	0.1291	**0.0148**	0.0024
enron	0.0541	0.0572	0.0227	0.0227	0.0224	**0.1123**
	0.0037	0.0016	0.0007	0.0001	0.0002	**0.0018**
scene	0.2241	0.2165	**0.2997**	0.2138	0.2465	0.2679
	0.0122	0.0156	**0.0318**	0.0013	0.0327	0.0082
pascal07	0.1145	0.1063	0.0553	0.1023	**0.2164**	0.1307
	0.0013	0.0017	0.0004	0.1053	**0.0166**	0.0015
nus_wide_scene	0.1038	0.1026	0.1091	0.1092	**0.4466**	0.1695
	0.0050	0.0047	0.0006	0.0013	**0.0155**	0.0041
mirflickr	0.1664	0.1498	0.0395	**0.2245**	0.0394	0.2027
	0.0014	0.0007	0.0002	**0.1688**	0.0002	0.0028

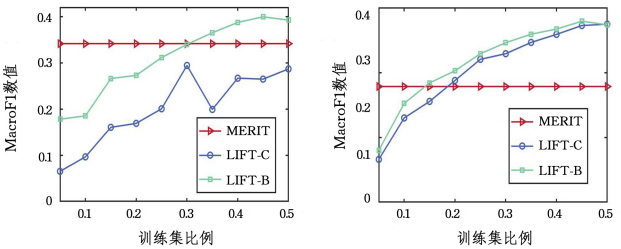
注：第一行是均值；第二行是方差；加粗数据表示最佳结果；带下划线的数据表示次优结果。

与单视图方法 WSMLLC 相比，本文引入的自适应调整视图权重，能更好地结合不同视图的信息，从而提升学习性能；与有监督的多视图方法 ICM2L，McWL，iMVWL 相比，本文采用的先验机制显著改善了学习效果，进一步说明了采用先验信息的合理性，且本文方法仅利用先验的标签相关性，就达到了其他利用有标记数据的监督方法的效果。

4.3 训练集比例对实验结果的影响

为了研究训练集比例对实验结果的影响，本文在 emotions，languagelog 数据集上与单视图的 LIFT-C，LIFT-B 两种方法进行了对比实验。在实验中，将训练集比例设定在 5%～50%之间，以 5%为步长。

对比实验结果如图 3 所示，从图中可以看出，随着训练集比例的提高，LIFT 方法的表现也随之增强。这一结果验证了在监督学习方法中，训练集比例的增加通常会带来更好的训练效果。在 emotions 数据集的 MacroF1 指标上，LIFT-B 方法与 MERIT 相交的点约为 30%，而在 languagelog 数据集上相交的点约为 15%，即在该点处本文提出的方法与 LIFT-B 相当，这一点显示出了 MERIT 在利用先验信息进行多标签分类上的优势。



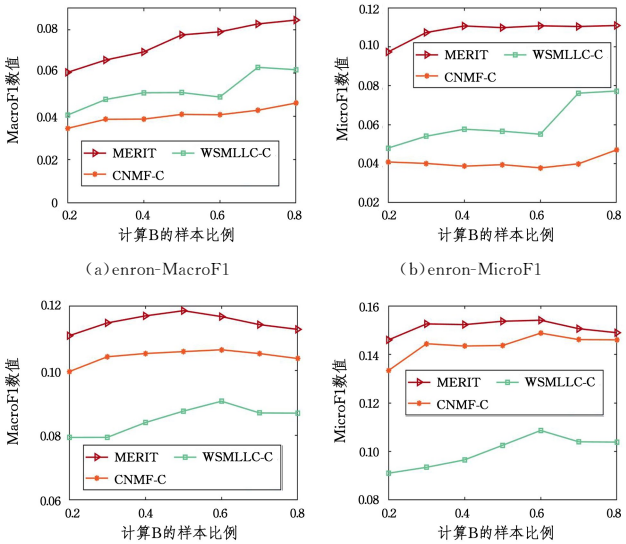
(a) emotions-MacroF1 (b) languagelog-MacroF1

图 3 训练集比例变化结果

Fig. 3 Comparative experimental results with different training set proportion

4.4 标签比例对实验结果的影响

为了研究计算标签相关性矩阵 $\boldsymbol{B}$ 的样本比例对实验结果的影响，本节在 enron 和 nus_wide_scene 数据集上进行实验，将 MERIT 与同样采用了标签比例的单视图方法 WSMLLC 和 CNMF 进行了比较，并将所有视图按顺序组合而成进行实验的结果记为 WSMLLC-C 与 CNMF-C，如图 4 所示。



(a) enron-MacroF1 (b) enron-MicroF1

(c) nus_wide_scene-MacroF1 (d) nus_wide_scene-MicroF1

图 4 标签比例变化比较结果

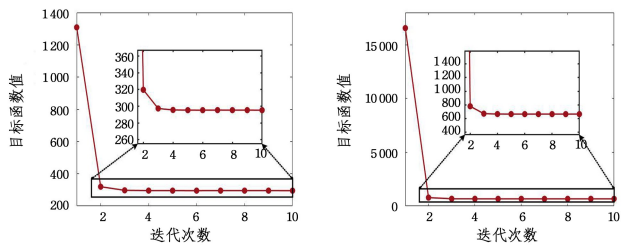
Fig. 4 Comparative experimental results with different label proportion

在 enron 和 nus_wide_scene 数据集的 MacroF1，MicroF1 数值上，调节计算标签比例的样本比例会对数据集产生一定的积极影响，这是因为计算 $\boldsymbol{B}$ 所使用的数据量有所增加，使其更可能接近于训练数据中真实的标签相关性矩阵；此外，MERIT 比 WSMLLC-C 线条更为平稳，与其相比，在计算 $\boldsymbol{B}$ 的样本比例上展现出了更强的鲁棒性。

4.5 迭代次数对目标函数的影响

本文在第 2 章中已经证明了目标函数的收敛性。为了进一步研究迭代次数对目标函数收敛性的具体影响，本节进行了相应的实验验证，即当连续两次迭代的目标函数差值小于前一次迭代的目标函数差值时，就认为该算法已经收敛。

在 emotions 和 nus_wide_scene 数据集上，将迭代次数设置为 1～10，并将每次迭代的目标函数值绘制在折线图上。从图 5 中可以观察到，大约在第 10 次迭代时，目标函数的值开始趋于稳定，表明目标函数在大约第 10 次迭代后基本收敛。这一实验结果与理论证明的收敛性相吻合，验证了算法在实际应用中的有效性。此外，这也提供了在实际操作中选择合适的迭代次数的依据。



(a) emotions (b) nus_wide_scene

图 5 MERIT 收敛速度

Fig. 5 Convergence speed of MERIT

4.6 阈值变化对实验结果的影响

为了研究不同阈值对模型性能的影响，我们在 emotions，nus_scene 和 pascal07 数据集上进行了实验。实验中设定了多个阈值选项，包括 $\left\{\frac{2}{C},\frac{1}{C},\frac{1}{2C},\frac{1}{3C},\frac{1}{4C},\frac{1}{5C}\right\}$，其中 C 代表标签数。实验结果如表 5 所列，在阈值设为 $\frac{2}{C}$ 时，模型性能的指标值降到最小。且随着阈值的减小，指标值呈现上升趋势，这表明模型性能随着阈值的调整而有所改善。

这些结果表明，通过适当调整阈值，我们可以进一步提高模型的性能。因此，选择合适的阈值对于优化模型至关重要。在实际应用中，应根据具体情况和需求来确定最佳的阈值。

表 5 阈值变化对实验结果的影响

Table 5 Influence of different thresholds on experimental outcomes

数据集	阈值	MacroF1	MicroF1
emotions	2/C	0.0818±0.0153	0.0851±0.0166
	1/C	0.3577±0.0151	0.3670±0.0155
	1/2C	0.4087±0.0120	0.4156±0.0112
	1/3C	0.4207±0.0154	0.4276±0.0160
	1/4C	0.4216±0.0048	0.4285±0.0055
	1/5C	0.4267±0.0091	0.4333±0.0093
nus_wide_scene	2/C	0.0362±0.0060	0.0437±0.0063
	1/C	0.1065±0.0019	0.1401±0.0020
	1/2C	0.1197±0.0017	0.1523±0.0022
	1/3C	0.1230±0.0006	0.1570±0.0015
	1/4C	0.1235±0.0010	0.1573±0.0013
	1/5C	0.1251±0.0014	0.1608±0.0017
pascal07	2/C	0.0573±0.0026	0.0597±0.0025
	1/C	0.1178±0.0017	0.1296±0.0019
	1/2C	0.1260±0.0011	0.1387±0.0014
	1/3C	0.1273±0.0013	0.1407±0.0013
	1/4C	0.1275±0.0013	0.1421±0.0012
	1/5C	0.1281±0.0008	0.1424±0.0008

4.7 运行时间对比

表 6 列出了各方法在不同数据集上的运行时间，单位为秒(s)。从表中可以观察到，其中 WSMLLC-B 方法在所有数据集上运行较快，其采用所有视图中效果最佳的视图作为结果，因此只是所有视图中某一视图的结果，并未将所有视图的结果纳入，虽运行时间较短，但效果却并未达到最优；与 ICM2L，McWL，iMVWL 的运行时间相比，本文方法比其他多视图方法更快，且效果大部分比以上方法好。此外，MERIT 在处理大量数据时展现出了更快的运行速度，这一优势主要归功于 MERIT 算法复杂度较低，且能以较快速度收敛。

表 6 运行时间比较

Table 6 Running time comparison

数据集	WSMLLC-B	WSMLLC-C	ICM2L	McWL	iMVWL	MERIT
cal_500	4.14	**3.62**	90.01	93.21	554.52	5.94
emotions	**0.30**	0.65	4.57	0.73	12.23	0.58
languagelog	**17.29**	22.65	106.26	90.53	331.96	17.95
enron	29.14	61.37	78.34	81.43	173.92	**17.78**
scene	**4.00**	8.06	14.02	13.50	45.00	17.41
pascal07	**57.83**	714.32	538.27	182.17	523.35	153.95
nus_wide_scene	61.74	**44.52**	185.12	177.39	452.13	223.28
mirflickr	**467.08**	4918.73	1728.99	1063.66	2402.23	573.60

注：加粗数据表示算法的最快运行时间，带下划线的数据表示次优运行时间。

结束语 本文提出了一种新的多视图多标签学习方法，该方法将标签相关性作为一种先验信息，用于求解模型，并引入了一种迭代算法，其收敛性得到了验证。通过自适应调整每个视图的权重系数，MERIT 能够自动平衡不同视图的贡献。在 8 个真实数据集上的评估结果表明，该方法在多数情况下具有优良性能。

在未来的研究中，我们将从以下几个方向进行探索。首先，考虑与其他有监督分类方法进行结合，以提升小样本学习场景下的学习性能；其次，将探索如何将先验信息应用于其他方法，以充分利用这部分信息，提升学习效果；最后，将进一步扩展模型的应用范围，探究其在更广泛场景下的有效性和鲁棒性。

参 考 文 献

[1] CHENG Y，LI Q，WANG Y，et al. Multi-view Multi-label Learning with View Feature Attention Allocation[J]. Neurocomputing，2022，501：857-874.

[2] ZHANG J，WEI G，SUN F. Synthetic Multi-view Clustering with Missing Relationships and Instances[J]. Image and Vision Computing，2023，134：104669.

[3] TANG Q，LIANG J，ZHU F. A Comparative Review on Multimodal Sensors Fusion Based on Deep Learning[J]. Signal Processing，2023，213：109165.

[4] AL-SALEMI B，AYOB M，KENDALL G，et al. Multi-label Arabic Text Categorization：A Bench-mark and Baseline Comparison of Multi-label Learning Algorithms[J]. Information Processing & Management，2019，56(1)：212-227.

[5] OUYANG X，TAO H，FAN R D，et al. Weakly Supervised Multi label Learning Using Prior Label Correlation Information[J]. Journal of Software，2023，34(4)：1732-1748.

[6] ZHANG M，ZHOU Z. A Review on Multi-Label Learning Algorithms[J]. IEEE Transactions on Knowledge and Data Engineering，2014，26(8)：1819-1837.

[7] BOUTELL M，LUO J，SHEN X，et al. Learning Multi-label Scene Classification[J]. Pattern Recognition，2004，37(9)：1757-1771.

[8] ZHANG M，WU L. Lift：Multi-Label Learning with Label-Specific Features[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2015，37(1)：107-120.

[9] HUANG J，LI G，WANG S，et al. Multi-label Classification by Exploiting Local Positive and Negative Pairwise Label Correlation[J]. Neurocomputing，2017，257：164-174.

[10] HUANG J，LI G，HUANG Q，et al. Joint Feature Selection and Classification for Multilabel Learning[J]. IEEE Transactions on Cybernetics，2018，48(3)：876-889.

[11] MELKI G，CANO A，KECMAN V，et al. Multi-target Support Vector Regression via Cor-relation Regressor Chains[J]. Information Sciences，2017，415-416：53-69.

[12] HUANG J，LI G，WANG S，et al. Group Sensitive Classifier Chains for Multi-label Classifi-cation[C]//2015 IEEE International Conference on Multimedia and Expo(ICME). Turin，Italy：IEEE，2015：1-6.

[13] LIU M，LUO Y，TAO D，et al. Low-Rank Multi-View Learning

in Matrix Completion for Multi-Label Image Classification[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2015, 29(1): 2778-2784.

[14] TAN Q, YU G, WANG J, et al. Individuality- and Commonality-Based Multiview Multilabel Learning[J]. IEEE Transactions on Cyber-netics, 2021, 51(3): 1716-1727.

[15] TAN Q, YU G, DOMENICONI C, et al. Incomplete Multi-View Weak-Label Learning[C]// Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18. International Joint Conferences on Artificial Intelligence Organization, 2018: 2703-2709.

[16] XING Y, YU G, DOMENICONI C, et al. Multi-Label Co-Training[C]// Proceedings of the Twenty-Seventh International Joint Con-ference on Artificial Intelligence. Stockholm, Sweden: International Joint Conferences on Artificial Intelligence Organization, 2018: 2882-2888.

[17] ZHANG J, LI C, SUN Z, et al. Towards a Unified Multi-source-based Optimization Framework for Multi-label Learning[J]. Applied Soft Computing, 2019, 76: 425-435.

[18] HE Z, CHEN C, BU J, et al. Multi-view Based Multi-label Propagation for Image Anno-tation[J]. Neurocomputing, 2015, 168: 853-860.

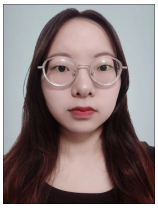
[19] RADHIKA K, ORUGANTI V R M. Deep Multimodal Fusion for Subject-Independent Stress Detection[C]// 2021 11th International Conference on Cloud Computing, Data Science & Engineering(Confluence). Noida, India: IEEE, 2021: 105-109.

[20] LIU N, ZHANG Z, WU Y, et al. CRBSP: Prediction of Circ-RNA-RBP Binding Sites Based on Multimodal Intermediate Fusion[J]. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2023, 20(5): 2898-2906.

[21] NIE F, CAI G, LI X. Multi-View Clustering and Semi-Supervised Classification with Adaptive Neighbours[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2017.

[22] LIU Y, JIN R, YANG L. Semi-supervised Multi-label Learning by Constrained Non-negative Matrix Factorization [C]// AAAI. 2006: 421-426.

[23] TAN Q, YU G, DOMENICONI C, et al. Multi-view Weak-label Learning Based on Matrix Completion[M]// Proceedings of the 2018 SIAM International Conference on Data Mining (SDM). Society for Industrial and Applied Mathematics, 2018: 450-458.



SHENG Sirou, born in 2001, postgraduate. Her main research interests include machine learning and pattern recognition.

HOU Chenping, born in 1982, Ph. D, professor, is a senior member of CCF (No. 21237S). His main research interests include pattern recognition, machine learning, data mining and computer vision.

(责任编辑:喻藜)