

## 基于深度多任务学习的社交图像标签和分组联合推荐

耿蕾蕾<sup>1</sup> 崔超然<sup>1</sup> 石成<sup>2</sup> 申朕<sup>2</sup> 尹义龙<sup>3</sup> 冯仕红<sup>1</sup>

1 山东财经大学计算机科学与技术学院 济南 250014

2 山东大学计算机科学与技术学院 济南 250014

3 山东大学软件学院 济南 250014

(leileigeng\_njust@163.com)

**摘要** 随着在线社交的多媒体共享网站的蓬勃发展,社交图像推荐逐渐成为研究热点。人们通常对社交图像进行标签化、分组化,使得图像数据更加易于管理。传统的图像标签或分组推荐方法往往只关注特定任务,忽略了标签推荐和分组推荐任务之间的隐含关系。多任务学习则可以充分挖掘不同任务对图像的共享或相互关联的隐含表示,融合多种任务抽取图像特征,对于提高单一任务的准确性具有积极意义。因此,文中提出了一种基于深度多任务学习的社交图像标签和分组联合推荐模型。该方法使用基于比较的偏序学习深度网络分别进行标签推荐和分组推荐,有效缓解了单任务中的数据稀疏性问题。此外,在处理社交图像视觉特征的卷积神经网络中,首先使用多任务学习将来自不同任务的中间层特征进行连接,然后通过卷积实现降维和特征的自动融合,使得不同任务的图像特征得到共享,同时降维后的融合特征能够满足下一层卷积神经网络的尺寸要求,使得单一任务的整体结构得以保持。从大量 Flickr 图片共享网站上爬取的真实数据集上的实验结果表明,与现有经典推荐算法相比,所提算法获得的准确率和召回率均有较大提升,证明了该方法的有效性和可行性。

**关键词:** 多任务学习; 分组推荐; 标签推荐; 偏序学习; 特征融合; 联合推荐

**中图分类号** TP391.3

## Social Image Tag and Group Joint Recommendation Based on Deep Multi-task Learning

GENG Lei-lei<sup>1</sup>, CUI Chao-ran<sup>1</sup>, SHI Cheng<sup>2</sup>, SHEN Zhen<sup>2</sup>, YIN Yi-long<sup>3</sup> and FENG Shi-hong<sup>1</sup>

1 School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan 250014, China

2 School of Computer Science and Technology, Shandong University, Jinan 250014, China

3 School of Software, Shandong University, Jinan 250014, China

**Abstract** With the rapid development of multimedia sharing websites, the social image recommendation has become a hot research topic recently. It is much easier to manage social images by tagging and grouping them. Traditional image recommendation methods tend to focus on a specific task just like tag or group recommendation, which ignore the correlation between the two tasks. By fusing image features extracted from multiple tasks, multi-task learning method exploit image sharing or correlation representation among different tasks to improve the accuracy of the single task. Therefore, this paper proposes a novel social image tag and group joint recommendation model based on deep multi-task learning. In the signal task, the tag and group recommendation are solved using comparison-based partial order deep learning method respectively for alleviating the data sparsity. Furthermore, the features extracted from their intermediate layer are saved for multi-task learning. In the convolutional neural network processing the visual features of social images, the features of two recommendations are connected. Then the dimension-reduction and automatic fusion of the features are realized by convolution. Hence image features extracted from different recommendations are shared. Moreover, the size of processed features is suitable for the next layer of the convolutional neural network so that the network architecture of single recommendation task can be maintained. The experimental results on a real Flickr dataset show that compared with the traditional methods, the accuracy and recall rate of the proposed algorithm are greatly improved, which

到稿日期:2019-10-22 返修日期:2020-03-22 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金(61701281,61876098,61573219);国家重点研发计划课题(2018YFC0830100,2018YFC0830102);山东省高等学校科技计划项目(KJ2018BAN047);山东省高等学校“青创科技计划”立项支持(2019KJN045);山东省高等学校优势学科人才团队培育计划

This work was supported by the National Natural Science Fund of China (61701281,61876098,61573219), National Key R&D Program of China (2018YFC0830100,2018YFC0830102), Shandong Province Higher Educational Science and Technology Program (KJ2018BAN047), Science and Technology Innovation Program for Distributed Young Talents of Shandong Province Higher Education Institutions (2019KJN045) and Fostering Project of Dominant Discipline and Talent Team of Shandong Province Higher Education Institutions.

通信作者:崔超然(crcui@sdufe.edu.cn)

proves the effectiveness and feasibility of the proposed method.

**Keywords** Multi-task learning, Group recommendation, Tag recommendation, Partial order learning, Feature fusion, Joint recommendation

## 1 引言

近年来,随着互联网技术的飞速发展,在线社交的多媒体共享网站得到了蓬勃发展,如微博、Facebook、WhatsApp、Flickr、Instagram等都给用户发布分享自己多媒体信息的功能。社交媒体中服务的重要特征是允许用户对多媒体信息进行点赞和评论、设置标签和分组信息、建立基于多媒体内容分享的关系网络等。这些交互行为改变了用户只能被动接受信息的情况,产生了以用户为中心的更加自主化、现代化的信息传播形式。在社会心理学中,人类的认知和行为相互影响,用户社交行为可视为人类对图像认知的反映。社交网络中的多媒体信息和用户的交互信息可提供更丰富的数据,通过对这些社交信息的科学理解,有助于避免海量数据中的信息过载,可使广大用户在社交媒体平台上获得更好的服务。因此,以社交图像为中心的标签推荐和分组推荐任务成为了推荐领域的热点问题<sup>[1-2]</sup>。

目前,大多数推荐工作通常尝试建立标签或分组与推荐图像视觉内容之间的联系,却忽略了标签推荐和分组推荐这两个任务之间的关联性。本文认为标签推荐和分组推荐两个任务之间存在某种关联性。例如,对“滑雪”“赛车”“足球”“体操”等标签点赞的用户更有可能对“运动”分组的图像点赞,因此对于一幅图像来说,标签信息和分组信息有较大的重复性。但是,通常情况下由用户设置的标签或分组信息并不完备,两个任务之间可以形成一定的互补性。另一方面,相对于标签信息,分组信息往往聚焦于更加宽泛的概念。标签和分组信息不同划分粒度,可以为综合利用标签和分组任务之间的关联信息以提升推荐效果提供可能。因此,本文提出了一种基于深度多任务学习的社交图像标签和分组联合推荐模型。针对单一任务,本文设计了基于偏序学习的深度网络,使得大量未标记数据得以利用,有效缓解了数据的稀疏性问题。另外,本文设计了深度多任务特征融合网络,以融合不同任务卷积层中的特征图,共享不同任务间的关联信息,从而使每个单一任务的性能都得以提高。

## 2 相关工作

近年来,图像标签和分组推荐问题受到了人们的广泛关注和研究。Qi<sup>[3]</sup>提出了一种基于子空间学习的图像标签推荐方法,该方法在图像视觉特征和标签内容语义共享的隐式子空间中计算视觉特征与标签的关联程度,进而进行推荐。Zhao等<sup>[4]</sup>将图像、群组、地理位置等元数据进行融合,提出了一种基于二分图的个性化图像标签推荐算法。然而,上述标签推荐算法仅考虑了标签频率信息或图像底层视觉特征,没有考虑图像视觉特征与标签内容语义之间的关系,从而导致在冷启动情况下标签推荐的效果不太理想。Gao等<sup>[1]</sup>和Nguye等<sup>[5]</sup>采用基于卷积神经网络的深度学习模型提取图像的特征,并通过监督学习的方式优化模型的标签标注,为图像

推荐个性化的图像标签。Li等<sup>[6]</sup>提出了一种基于卷积神经网络的多标签图像自动标注模型,该模型在 AlexNet 原本的结构上调整了池化间隔和学习速率,并设计了一种基于 Softmax 回归的多标签排名损失函数,取得了较好的多标签推荐效果。

图像分组推荐算法可分为基于文本的图像推荐方法和基于图像内容的图像推荐方法两类。早期的图像分组推荐算法主要为基于文本的方法,该类方法将用户输入的关键字信息与数据库中对应图像的标注信息进行相似度比对,然后将相似度较高的图像反馈给用户。由于包含标注信息的图像少,并且标注信息不准确,因此该方法的推荐效果有待提高<sup>[7]</sup>。目前,基于内容的图像分组推荐最为常用,即根据图像的特征(如颜色、形状、纹理等)进行推荐。Liu等<sup>[8]</sup>在稀疏表示的基础上联合考虑图像内容分析与用户偏好的问题,提出了一种新颖的主题推荐模型,较好地解决了冷启动问题。Zhang等<sup>[9]</sup>通过使用标签来表示用户的偏好,提出了一种基于用户标签模型的个性化社交图像推荐方法,该方法对图像标签的准确度要求很高,若要对海量图像加上精准的标签则需要巨大的工作量,因此该方法非常耗时,不具有实时性。同时,图像标签的添加很大程度上受到了用户因素的影响,并不能全面地代表用户的偏好。因此,由于内容特征易受到拍摄环境、拍摄设备和拍摄软件等外界因素的影响,基于内容的图像分组的推荐结果并不理想。

上述标签和分组推荐算法都只考虑了单一推荐任务,没有考虑标签推荐和分组推荐间的相互关系,但基于神经网络的多任务学习是综合利用相关任务监督信号的有力工具。近几年,深度多任务学习中的隐含层参数共享方式通常分为软共享和硬共享,其中硬共享方式最为常见。Long等<sup>[10]</sup>提出了深度关系网络,基于特定层中参数张量的正态先验挖掘多任务间的关系,并通过学习任务间的关系和可转移的特征一定程度地解决了特征层的负迁移和分类层的欠迁移问题。Lu等<sup>[11]</sup>提出了基于动态调整的多任务网络,使用贪心算法对初始的窄网络结构进行动态变更以获得局部最优解,但其只能学习简单任务间的交互。Misra等<sup>[12]</sup>提出了十字绣网络,使用两个参数结合网络,参数的不同设置决定了网络的结合方式,这种结构属于参数软共享。Lei等<sup>[13]</sup>将多源异构数据映射到同一隐空间,提出了可自动学习共享子空间和输入内容合理表示的比较深度学习网络。

本文以多任务学习为出发点,立足于综合考虑标签推荐和分组推荐间的相关性来提高推荐性能,提出了基于深度多任务学习的社交图像标签和分组联合推荐模型。本文方法将社交网络中的不同任务信息进行共享,缓解了单一任务中数据的稀疏性问题,同时利用关联信息提高了单一任务的性能。此外,本文方法不需要手动设计多任务特征的结合方式,可以根据任务间的关联关系自动进行特征的融合共享。

### 3 基于深度多任务学习的标签和分组推荐

#### 3.1 基于偏序学习深度网络的单任务推荐

社交网络中的数据具有稀疏性,即每张图像仅具有有限个标签或者分组信息,鉴于此,本文使用基于比较的偏序学习方法对单任务进行建模。由于标签推荐任务和分组推荐任务具有相似性,下文以标签推荐任务为例对本文构建的单任务模型进行介绍。

在标签推荐模型中, $\mathcal{S}$ 表示社交网络中的图像集合, $s$ 表示推荐模型的输入图像; $\mathcal{T}$ 表示任务集合, $\mathcal{T}_i$ 表示第*i*个标签的推荐任务; $t_i^+$ 表示第*i*个任务图像拥有的标签, $t_i^-$ 表示第*i*个任务图像未拥有的标签。第*i*个标签推荐任务的输入可以表示为三元组 $(s, t_i^+, t_i^-)$ 。图像映射函数 $\pi_i(s)$ 可将隐藏层的特征变换为任务相关的视觉特征。标签推荐的目标是学习能够将标签(任务属性)嵌入度量空间中的潜在映射 $\phi_i(t_i)$ 。在嵌入的度量空间中,图像与已拥有标签的距离相比其不具有标签的距离更近,如图1所示。

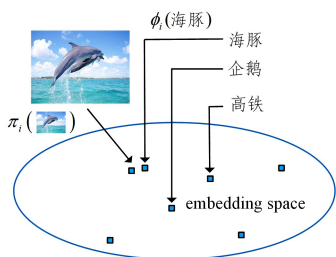


图1 通过优化偏序关系 $(s, t_i^+, t_i^-)$ 学习 $\phi_i(t_i)$ 和 $\pi_i(s)$

Fig. 1 Learn both  $\phi_i(t_i)$  and  $\pi_i(s)$  by optimizing partial order

为描述图像与标签间的距离,本文定义匹配度函数 $f_{s,t_i}$ 来描述嵌入空间中图像和任务属性之间的匹配程度,具体公式如下:

$$f_{s,t_i} = \text{sim}(\pi_i(s), \phi_i(t_i)) \quad (1)$$

其中, $\text{sim}(\cdot)$ 为相似性函数。事实上,图像拥有标签表示图像内容与标签在很大程度上是匹配的。然而,由于图像标注的不全面和不准确,部分图像不具有标签类型且标签解释存在模糊性。在这种情况下,相比未拥有的标签,图像与已拥有标签的相似性更大。本文的单任务推荐方法并非根据有限标签学习嵌入函数,而是基于图像标签之间的成对排名构建学习目标,即本文单任务推荐方法采用基于比较的偏序学习方式,其中偏序关系可表示为:

$$\mathcal{P}_i = \{(s, t_i^+, t_i^-) \mid t_i^+ \in \mathcal{T}_i^+ \wedge t_i^- \in \mathcal{T}_i^-\} \quad (2)$$

其中, $\mathcal{T}$ 表示标签集合, $\mathcal{T}_i^+$ 和 $\mathcal{T}_i^-$ 分别表示给定图像*s*所拥有和未拥有的标签集合。三元组 $(s, t_i^+, t_i^-)$ 表示图像*s*更符合标签 $t_i^+$ ,而不是 $t_i^-$ 。在独立假设下,单任务推荐方法学习的嵌入函数即为最大化所有偏序关系的似然概率,即:

$$\mathcal{L}_i = - \sum_{(s, t_i^+, t_i^-) \in \mathcal{P}_i} p(f_{s,t_i^+} > f_{s,t_i^-}) \quad (3)$$

其中,概率 $p(f_{s,t_i^+} > f_{s,t_i^-})$ 根据以式(4)进行近似估计:

$$\mathcal{L}_i = - \sum_{(s, t_i^+, t_i^-) \in \mathcal{P}_i} \ln p(f_{s,t_i^+} > f_{s,t_i^-}) \quad (4)$$

其中,匹配度函数 $f_{s,t_i^+}$ 和 $f_{s,t_i^-}$ 的差距越大,偏序关系成立的可能性就越大。基于此,已标注和未标注的标签均可作为随后

的学习过程做出贡献,从而有助于减小数据稀疏性带来的负面影响。同时,最大化所有偏序关系的似然概率可等同于最小化以下损失函数:

$$\mathcal{L}_i = - \sum_{(s, t_i^+, t_i^-) \in \mathcal{P}_i} \ln p(f_{s,t_i^+} > f_{s,t_i^-}) \quad (5)$$

因此,单一任务的整体网络结构如图2所示。第一层和第三层为完全参数共享的全连接网络,其输入是经过处理的标签信息;中间层为卷积神经网络,用于对图像的视觉信息进行特征提取;整体网络最后的全连接层的输出尺寸与第一层和第三层相同。分别计算图像与正例标签和负例标签的距离,将其差值作为最后激活层的输入。

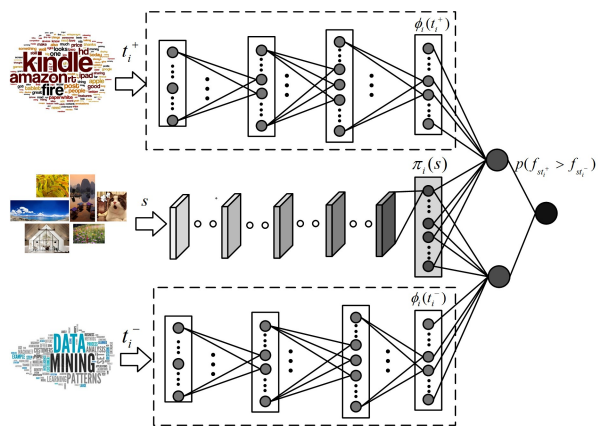


图2 单任务偏序学习深度网络

Fig. 2 Deep partial learning network for single task

#### 3.2 基于深度多任务学习的标签和分组联合推荐

本文使用相同的单任务网络分别对输入图像进行标签推荐和分组推荐,如第3.1节所述。基于深度多任务学习的标签和分组联合推荐通过构建一个不同任务间特征共享的网络,融合与特定任务相关的图像视觉特征,使得每个任务能够利用其他任务中的信息来提升当前单一任务的性能。因此,深度多任务学习网络的目标即为融合不同任务间的视觉特征,使得不同任务间相关的特征得以共享。

在单一任务中, $\phi_i(t_i)$ 和 $\pi_i(s)$ 分别表示任务属性(标签或分组)嵌入和图像特征嵌入的映射函数。本文定义 $\pi_i^l(s)$ 表示第*i*个任务的卷积神经网络(Convolutional Neural Networks, CNN)的中间层。那么, $\pi_1^l(s)$ 和 $\pi_2^l(s)$ 分别表示标签推荐和分组推荐的卷积神经网络第*l*层的特征图,如图3所示。虽然特征图 $\pi_1^l(s)$ 和 $\pi_2^l(s)$ 的输入为相同的图像,但其分别用于标签推荐和分组推荐两个不同的任务,并与其对应的推荐任务紧密相关。

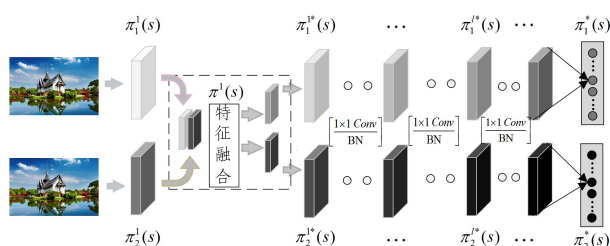


图3 多任务特征提取结构图

Fig. 3 Feature extraction network for multi-task

为了使不同任务的特征得以共享,本文的联合推荐方法在特征通道上对特征图进行堆叠。堆叠前单个任务的图像特征图的通道数为  $C$ ,则堆叠后图像特征图的通道数为  $2C$ 。此时,卷积神经网络特征图可用  $\pi^l(s)$  表示:

$$\pi^l(s) = [\pi_1^l(s), \pi_2^l(s)] \quad (6)$$

为保持原始多任务的网络结构,针对每个单任务,分别使用  $1 \times 1$  卷积对特征通道进行降维。因此,标签推荐和分组推荐的卷积神经网络的卷积层均包含  $C$  个卷积核,卷积核的步长和边缘填充大小分别为 1 和 0:

$$\begin{aligned} \pi_1^*(s) &= \pi^l(s) \cdot w_1 \\ \pi_2^*(s) &= \pi^l(s) \cdot w_2 \end{aligned} \quad (7)$$

其中,  $w_1$  和  $w_2$  分别是标签推荐和分组推荐任务的降维投影矩阵。通过这种方式,两个推荐任务的卷积层将重新输出新的图像特征图  $\pi_1^*(s)$  和  $\pi_2^*(s)$ 。相比特征图  $\pi_1^l(s)$  和  $\pi_2^l(s)$ ,新图像特征图  $\pi_1^*(s)$  和  $\pi_2^*(s)$  的大小保持不变,通道个数恢复为  $C$ 。然后,两个推荐任务的卷积层分别将新图像特征图  $\pi_1^*(s)$  和  $\pi_2^*(s)$  送入后续的全连接层。直观上看,跨分支连接层使两个网络分支之间以交换图像特征图的方式共享信息,并方便模型在训练过程中自动学习两个单任务分别需要

的信息。多任务之间的视觉特征由此得到共享,通过降维处理,融合的视觉特征也能够满足原始单任务的卷积神经网络结构的尺寸要求。于是,融合了多任务信息的单任务推荐目标和图像间的相似度可表示为:

$$f_{s,t_i}^* = \text{sim}(\pi_1^*(s), \phi_i(t_i)) \quad (8)$$

从而,深度多任务学习的总体损失函数可表示为:

$$L_{\text{total}} = \alpha \cdot L_{t_1}^* + \beta \cdot L_{t_2}^* \quad (9)$$

其中,  $L_{t_1}^*$  和  $L_{t_2}^*$  分别表示标签和分组推荐单任务的损失函数;  $\alpha$  和  $\beta$  分别为相应的均衡化参数,用于权衡不同任务损失在总体损失中的重要性,  $\alpha$  和  $\beta$  之和为 1。

综上所述,基于深度多任务学习的标签和分组联合推荐任务的整体网络结构如图 4 所示。其中, Task 1 和 Task 2 是 3.1 节所述的单一推荐任务模型,即分别是标签推荐和分组推荐。

两个推荐任务具有相同的图像输入,并针对该社交图像分别进行标签推荐任务和分组推荐任务。对于两个任务中图像处理的网络部分,使用本文提出的特征融合机制,然后结合两个任务的视觉特征分别计算各自的单任务损失,最后依据联合推荐任务的损失进行模型训练。

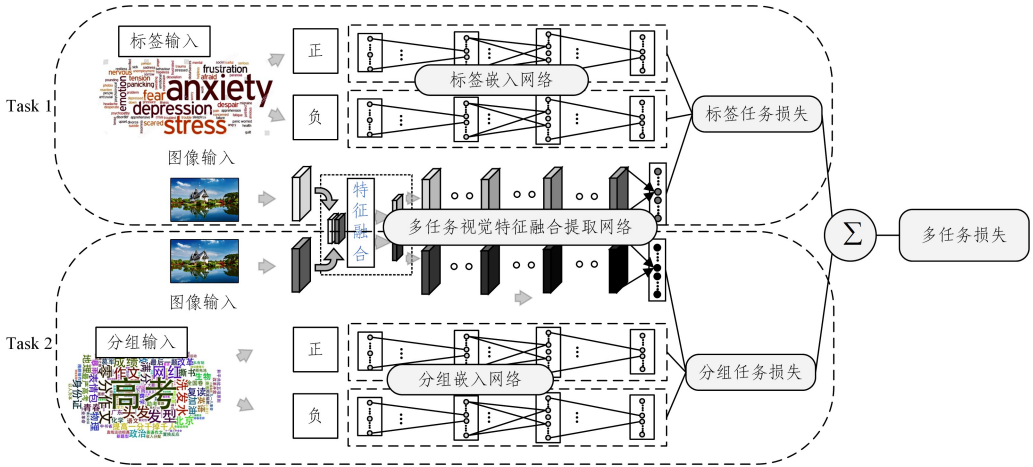


图 4 深度多任务学习的整体网络结构

Fig. 4 Network of deep multi-task learning

## 4 实验分析

本文中的实验均在工作站上运行,工作站配备了 12 核 3.50 GHz Intel Xeon 处理器、两个 Nvidia GTX 1080 GPU 和 128 GB RAM。

### 4.1 数据预处理

为了测评算法的性能,本文从 Flickr 图片共享网站上爬取了 22 万张社交图像进行实验。该图像数据集包含 97 778 个分组和 406 430 个标签,图像与标签的对应关系有 4 425 727 个,图像与分组的对应关系有 2 833 712 个。每张图像对应的标签和分组过多或过少,都会对模型训练造成不良影响,对应的标签和分组过多则会带来噪声,过少则模型无法训练。因此,为保证同一社交图像能够同时适用于标签推荐和分组推荐任务的训练和测试,首先对爬取的社交图像进行过滤,即限制每张图像的标签和分组数量在 6 至 200 之间。过滤后,满

足条件的图像为 7 812 张,标签个数为 1 855,分组个数为 872,图像与标签的对应关系有 334 125 个,图像与分组的对应关系有 128 029 个。此外,为了解决过拟合问题,从图像或其水平翻转图像中抽取  $224 \times 224$  的裁剪子图像,并使用子图像与图像数据集的均值像素所构建的模板图像做插值,将获得的插值图像作为深度多任务学习网络的输入图像。

### 4.2 评价方法

本文使用两种数据划分方式,并分别采用不同的评价方法,两种数据集的划分是在两次独立的训练过程中分别进行的。

采用第一种划分方式构造验证集合  $\mathcal{T}^*$  时,针对每张图像随机选择 1 个正例标签或分组  $t_i^+$  和 99 个负例标签或分组  $t_i^-$ ,并从  $\mathcal{T}^+$  中删除选择的  $t_i^+$ ,其余图像用于训练和测试。然后,使用训练后的模型对  $\mathcal{T}^*$  进行重排序。最后,采用平均位序<sup>[14]</sup> (Average Rank) 作为评估指标,其公式如下:

$$AR = \frac{1}{|\mathcal{T}^*|} \sum z_i^* \quad (10)$$

其中,  $z_i$  是测试标签或分组在排名列表中的位置,其按照在嵌入空间中与图像的相似程度产生。通俗来讲,如果一个标签或者分组属于某幅图像,那么其应该排在推荐结果列表中的最前面。显然,平均位序值越小,推荐效果就越好。

采用第二种划分方式构造验证集合  $\mathcal{T}^*$  时,针对标签和分组推荐两个任务,分别随机挑选 60% 的图像用于训练,20% 用于测试,剩余 20% 用于验证。与第一种划分方式不同,第二种数据划分方式中测试集中的图像均未在训练集中出现过。对于已训练好的模型来说,测试集中的图像都是全新的,仅包含视觉信息,不包含与推荐目标的任何交互信息,即为冷启动问题。测试集和验证集中的图像通过对其标签或分组进行随机抽样,抽取其拥有的 20 个标签(或分组)以及 80 个不具有的标签(或分组),最后将这 100 个数据混合用于测试。在这种数据划分方式下,采用准确率和召回率作为评估指标。

### 4.3 实验结果分析

为证明本文提出的融合多任务特征的深度多任务学习网络(Multi-task Deep Convolutional Networks, MTDCN)的有效性,将其与现有的经典推荐方法进行了比较,对比方法如下。1)基于张量分解的推荐算法(Modified Tensor Decomposition Model, MTD)<sup>[14]</sup>,该方法首先定义描述推荐图像的三维张量,然后建立融合谱聚类和惩罚机制的张量分解模型,从而在低维子空间中计算评分并推荐。2)基于图像元数据信息的推荐算法(Personalized Social Image Recommendation, PSIR)<sup>[15]</sup>,该方法首先根据图像内容对标签或分组信息进行重排序,然后计算用户偏好、图像、标签或分组之间的相关性,构建用户-图像-标签或用户-图像-分组三元图,实现了个性化的图像推荐系统,并有效缓解了数据稀疏性问题。3)基于二分图的图像推荐方法(Tag Recommendation Algorithm Based on Bipartite Graph Model, TG-BG)<sup>[4]</sup>,该方法利用二分图将初始标签和分组分值进行信息扩散,计算出最终的分值向量,并将该向量中分值较高的作为推荐结果。4)基于协调滤波卷积神经网络的算法(Collaborative Filtering by Convolution Deep Learning Model, CF-CDL)<sup>[16]</sup>,该方法采用深度学习方法获得图像的视觉特征,然后通过构建的加权最近邻计算相似度对社交图像进行标签或分组推荐。5)本文提出的用于解决单任务推荐的基于偏序学习的深度网络(Single Task Deep Convolutional Networks, STDCN)。

针对第一种数据划分方式,表 1 列出了不同方法的平均位序。实验结果表明,基于深度学习的算法的性能明显优于其他 3 种传统推荐算法,由此验证了端到端框架求解社交图像推荐任务的优势。然而,由于协调滤波卷积神经网络模型 CF-CDL 和基于偏序学习的深度模型 STDCN 没有挖掘标签推荐和分组推荐不同任务间的相关性,其性能落后于本文提出的基于深度多任务学习的标签和分组联合推荐算法。例如,与 STDCN 模型相比,MTDCN 算法获得的标签和分组推荐的准确率分别提高了 36.4% 和 38.5%。此外,由于其他推荐方法未利用图像的视觉信息和相关任务之间的信息,无法

进一步提升推荐的准确性,因此本文方法获得了最优的推荐结果。

在第二种划分的数据集中,将本文方法与 5 种推荐算法获得的准确率和召回率进行比较发现,本文基于多任务学习的联合推荐算法的性能明显高于单任务推荐方法,实验结果如图 5 和图 6 所示。

表 1 第一种划分方式的平均位序结果

Table 1 Average order based on first partition method

算法	标签推荐	分组推荐
MTD	31.1	29.4
TG-BG	25.6	24.4
PSIR	19.4	20.8
STDCN	15.1	16.9
CF-CDL	12.5	14.7
MTDCN	9.6	10.4

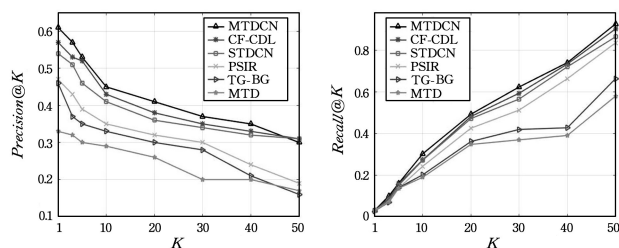


图 5 标签推荐任务的实验对比结果

Fig. 5 Comparison results in tag recommendation task

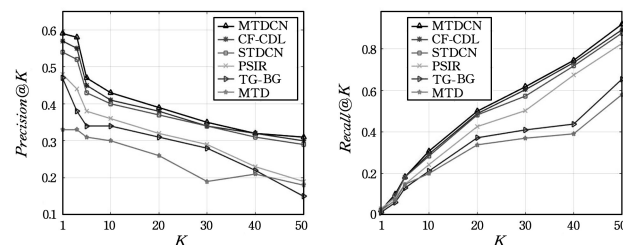


图 6 分组推荐任务的实验对比结果

Fig. 6 Comparison results in group recommendation task

本文中深度多任务学习的损失函数(见式(9))包含 2 个均衡化参数  $\alpha$  和  $\beta$ 。为探究其对联合推荐模型性能的影响,本文将均衡化参数  $\alpha$  和  $\beta$  在 0 到 1 范围内进行调整,并进行实验。图 7 给出了  $\alpha$  和  $\beta$  的不同取值下标签推荐和分组推荐的性能变化。图 7 中,横轴为  $\beta$ ,由于  $\alpha$  和  $\beta$  之和为 1,因此这两者的值仅为一个变量;纵轴为推荐算法的性能,左纵轴为标签任务的损失函数值,右纵轴为分组任务的损失函数值。实验结果表明,当  $\beta < 0.1$  时,随着均衡化参数  $\beta$  的增加,分组推荐的性能迅速提高; $\beta$  值在 0.1 到 0.4 之间时,分组任务的性能相对稳定;然而,当  $\beta$  值超过 0.4 时,分组任务的性能随着  $\beta$  值的进一步增加而下降。综合考虑标签推荐和分组推荐的性能, $\alpha$  和  $\beta$  分别取 0.7 和 0.3 时,推荐系统的整体性能达到了理想状态。从另外一个角度考虑,如果使用较小的  $\beta$  值,在进行推荐时,多任务推荐算法的性能主要取决于于图像的标签信息;反之,如果使用较大的  $\beta$  值,推荐则以分组信息为主。但是,由于分组信息相比标签信息更加稀疏且粒度更粗,采用较小的  $\beta$  值能够更有效地补充分组信息中的不足。

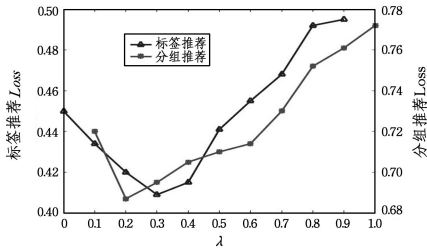


图7 不同权衡参数下的标签和分组推荐性能的结果

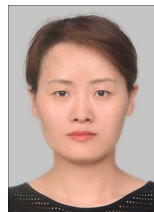
Fig. 7 Tag and group recommendation performance results for different parameters

上述实验验证了本文方法的有效性,即利用同一图像在社交网络中的不同推荐任务之间的相互关联,并融合不同任务的信息,能够使单一推荐任务的性能有所提高。

**结束语** 社交网络中的社交实体推荐是推荐领域的热点问题。本文提出的基于深度多任务学习的社交图像标签和分组联合推荐模型,从一个新视角对这些问题进行了研究。本文基于偏序学习的深度网络求解单任务推荐,有效缓解了单任务中的数据稀疏性问题;同时,使用深度多任务学习将标签和分组推荐任务的中间层特征进行连接和自动融合,使得适用于不同任务的图像特征得以共享。实验结果证明,相比基于交互数据的经典协同过滤方法和单任务的深度学习方法,本文的基于深度多任务学习的社交图像标签和分组联合推荐方法在推荐效果上有很大提升。在未来的研究中,将进一步考虑不同任务的数据量和学习难易程度的不同。此外,超参数可以设计成根据不同任务的完成程度而动态变化,从而加快计算,避免过拟合,以进一步提升标签和分组推荐的性能。再者,相比传统矩阵分解,基于图卷积的方法更具一般性,因此可考虑使用图卷积进行矩阵填充,并结合视觉信息解决冷启动问题,相信这样会有更好的推荐效果。

## 参考文献

- [1] GAO Y D, HOU L Y, YANG D L. Automatic Image Annotation Method using Multi-Label Learning Convolutional Neural Network [J]. Journal of Computer Applications, 2017, 37(1): 228-232.
- [2] CHEN F, HE Y, TANG L P. Emotional Recognition of Visual-perception Oriented Images and Its Application in the Recommendation System [J]. Journal of the China Society for Scientific and Technical Information, 2019, 38(4): 420-431.
- [3] QI C. An Approach of Image Tag Recommendation based on Subspace Learning Model [J]. Computer and Modernization, 2016, 3: 68-73.
- [4] ZHAO T L, LIU Z, HAN H J, et al. A Personalized Image Tag Recommendation Algorithm based on Bipartite Graph Model [J]. Journal of Nanjing University (Natural Science), 2018, 54(6): 1193-1205.
- [5] NGUYEN H T H, WISTUBA M, GRABOCKA J, et al. Personalized Deep Learning for Tag Recommendation [C]// Pacific Asia Conference on Knowledge Discovery and Data Mining. Springer, Cham, 2017: 186-197.
- [6] XU Y B, SU S M, FAN L Q. User-Specific Method for Aesthetic Images Recommendation [J]. Application Research of Computers, 2019, 36(12): 3853-3856.
- [7] LI J C, YUAN C, SONG Y. Multi-label Image Annotation Based on Convolution Neural Network [C]// The 11th Harmonious Human-Machine Environment, 2015: 56-59.
- [8] LIU L. A Sparse Image Recommendation Model using Content and User Preference Information [C]// IEEE International Conference on Web Intelligence. Piscataway, 2016: 232-239.
- [9] ZHANG J, YANG Y, TIAN Q, et al. Personalized Social Image Recommendation Method based on User Image Tag Model [J]. IEEE Transaction on Multimedia, 2017, 19(11): 2439-2449.
- [10] LONG M, CAO Z, WANG J, et al. Learning Multiple Tasks with Multilinear Relationship Networks [C]// Advances in Neural Information Processing Systems, 2017: 1595-1604.
- [11] LU Y, KUMAR A, ZHAI S, et al. Fully-Adaptive Feature Sharing in Multi-Task Networks with Applications in Person Attribute Classification [C]// IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1131-1140.
- [12] MISRA I, SHRIVASTAVA A, GUPTA A, et al. Cross-stitch Networks for Multi-Task Learning [C]// IEEE Conference on Computer Vision and Pattern Recognition, 2016: 3994-4003.
- [13] LEI C, LIU D, LI W, et al. Comparative Deep Learning of Hybrid Representations for Image Recommendations [C]// IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2545-2553.
- [14] CHEN M M, XUE K J. Personalized Tag Recommendation Algorithm based on Improved Tensor Decomposition Model [J]. Data Analysis and Knowledge Discovery, 2017, 3: 38-45.
- [15] ZHANG J, YANG Y, TIAN Q, et al. Personalized Social Image Recommendation Method based on User-Image-Tag Model [J]. IEEE Transactions on Multimedia, 2017, 19(11): 2439-2449.
- [16] ZHANG W, LIU F, JIANG L, et al. Recommendation Based on Collaborative Filtering by Convolution Deep Learning Model Based on Label Weight Nearest Neighbor [C]// The 10th International Symposium on Computational Intelligence and Design (ISCID). IEEE, 2018: 504-507.



**GENG Lei-lei**, born in 1984, Ph.D, associate professor, is a member of China Computer Federation. Her main research interests include machine learning and computer vision.



**CUI Chao-ran**, born in 1987, Ph.D, professor, is a member of China Computer Federation. His main research interests include information retrieval, computer vision and machine learning.