JOURNAL OF IMAGE AND GRAPHICS

主办：中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2022
04
VOL.27

ISSN1006-8961
CN11-3758/TB

口罩人脸姿态分类 P1125

中国图象图形学报

刊名题字：宋健 | 月刊（1996年创刊）

第27卷第4期（总第312期）
2022年4月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

主管单位 中国科学院
主办单位 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

主　　编 吴一戎
编辑出版 《中国图象图形学报》编辑出版委员会
通信地址 北京市海淀区北四环西路19号
邮　　编 100190
电子信箱 jig@aircas.ac.cn
电　　话 010-58887035
网　　址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号
总 发 行 北京报刊发行局
订　　购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
（邮政信箱：北京399信箱　邮编：100048）
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian | Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China
Zip code 100190
E-mail jig@aircas.ac.cn
Telephone 010-58887035
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048,P.R.China))
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

2022年4月 第27卷 第4期（总第312期）
中国图象图形学报
Zhongguo Tuxiang Tuxing Xuebao

目次

残差密集结构的东巴画渐进式重建(第1084页)

结合双模板融合与孪生网络的鲁棒视觉目标跟踪(第1191页)

注意力机制的曲面沉浸式投影系统补偿(第1238页)

Volume 27, Number 4
Published April 16, 2022

CONTENTS
JOURNAL OF IMAGE AND GRAPHICS

Gradual model reconstruction of Dongba painting based on residual dense structure(P1084)

Double template fusion based siamese network for robust visual object tracking(P1191)

An attention mechanism based inter-reflection compensation network for immersive projection system(P1238)

E-mail: jig@aircas.ac.cn
Website: www.cjig.cn
Tel: 010-58887035

中国图象图形学报
JOURNAL OF IMAGE AND GRAPHICS
©中国图象图形学报版权所有

中图法分类号:TP391.4　　**文献标识码**:A　　**文章编号**:1006-8961(2022)04-1226-12

论文引用格式:Yang B, Xiang X Q, Kong W Z, Shi Y and Yao J L. 2022. Automatic art analysis based on adaptive multi-task learning. Journal of Image and Graphics,27(04):1226-1237(杨冰,向学勤,孔万增,施妍,姚金良.2022.自适应多任务学习的自动艺术分析.中国图象图形学报,27(04):1226-1237)[DOI:10.11834/jig.200648]

自适应多任务学习的自动艺术分析

杨冰[1,3],向学勤[2*],孔万增[1,3],施妍[4],姚金良[1,3]

1. 杭州电子科技大学计算机学院,杭州　310018;2. 浙江省脑机协同智能重点实验室,杭州　310018;
3. 杭州凌感科技有限公司,杭州　310051;4. 杭州电子科技大学人文艺术与数字媒体学院,杭州　310018

摘　要:**目的**　艺术品数字化为从计算机视觉角度对艺术品研究提供了巨大机会。为更好地为数字艺术品博物馆提供艺术作品分类和艺术检索功能,使人们深入理解艺术品内涵,弘扬传统文化,促进文化遗产保护,本文将多任务学习引入自动艺术分析任务,基于贝叶斯理论提出一种原创性的自适应多任务学习方法。**方法**　基于层次贝叶斯理论利用各任务之间的相关性引入任务簇约束损失函数模型。依据贝叶斯建模方法,通过最大化不确定性的高斯似然构造多任务损失函数,最终构建了一种自适应多任务学习模型。这种自适应多任务学习模型能够很便利地扩展至任意同类学习任务,相比其他最新模型能够更好地提升学习的性能,取得更佳的分析效果。**结果**　本文方法解决了多任务学习中每个任务损失之间相对权重难以决策这一难题,能够自动决策损失函数的权重。为了评估本文方法的性能,在多模态艺术语义理解 SemArt 数据库上进行艺术作品分类以及跨模态艺术检索实验。艺术作品分类实验结果表明,本文方法相比于固定权重的多任务学习方法,在"时间范围"属性上提升了 4.43%,同时本文方法的效果也优于自动确定损失权重的现有方法。跨模态艺术检索实验结果也表明,与使用"作者"属性的最新的基于知识图谱模型相比较,本文方法的改进幅度为 9.91%,性能与分类的结果一致。**结论**　本文方法可以在多任务学习框架内自适应地学习每个任务的权重,与目前流行的方法相比能显著提高自动艺术分析任务的性能。

关键词:自动艺术分析;自适应多任务学习;贝叶斯理论;艺术分类;跨模态艺术检索

Automatic art analysis based on adaptive multi-task learning

Yang Bing[1,3], Xiang Xueqin[2*], Kong Wanzeng[1,3], Shi Yan[4], Yao Jinliang[1,3]

1. College of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310018, China;
2. Key Laboratory of Brain Machine Collaborative Intelligence of Zhejiang Province, Hangzhou 310018, China;
3. uSens Incorporated Company, Hangzhou 310051, China;
4. School of Media and Design, Hangzhou Dianzi University, Hangzhou 310018, China

Abstract: **Objective** To improve learning efficiency and prediction accuracy, multi-task learning aims to tackle multiple tasks based on the generic features assumption those are prior to task-related features. Multi-task learning technique has been applied in a variety of computer vision applications on the aspects of object detection and tracking, object recognition,

收稿日期:2020-11-17;**修回日期**:2021-02-18;**预印本日期**:2021-02-25

*通信作者:向学勤　xxueq@aliyun.com

基金项目:国家自然科学基金项目(61633010, U1909202);浙江省基础公益研究计划(LGG22F020027);浙江省重点研发计划(2020C04009);浙江省脑机协同智能重点实验室项目(2020E10010)

Supported by: National Natural Science Foundation of China (61633010, U1909202); Zhejiang Basic Public Welfare Research Program (LGG22F020027); Key Research and Development Program of Zhejiang Province (2020C04009); Key Laboratory of Brain Machine Collaborative Intelligence of Zhejiang Province (2020E10010)

human-based identification and human facial attribute classification. The worldwide digitization of artwork has called to art research from the aspect of computer vision and further facilitated cultural heritage preservation. Automatic artwork analysis has been developing the art style, the content of the painting, or the oriented attributes analysis for art research. Our multi-task learning for automatic art analysis application is based on the historical, social and artistic information. The existing multi-task joint learning methods learn multiple tasks based on a labor cost and time consuming weighted sum of losses. Our method illustrates art classification and art retrieval tools for the application of Digital Art Museum, which is convenient for researchers to deeply understand the connotation of art and further harness traditional cultural heritage research. **Method** A multiple objectives learning method is based on Bayesian theory. In terms of Bayesian analyzed results, we use the correlation between each task and introduce task cluster (clustering) to constrain the model. Then, we formulate a multi-task loss function via maximizing the Gaussian possibility derived of homoscedastic uncertainty via task-dependent uncertainty in Bayesian modeling. **Result** In order to slice into art classification and art retrieval missions, we identify the SemArt dataset, a recent multi-modal benchmark for understanding the semantic essence of the art, which is designed to retrieve the art paginating cross different modal, and could be readily modified for the classification of art paginating. This dataset contains 21 384 art painting images, which is randomly split into training, validation and test sets based on 19 244, 1 069 and 1 069 samples, respectively. First, we conduct art classification experiments on the SemArt dataset, and then evaluate the performance through classification accuracy, i. e. , the proportion of properly predicted paintings to the total amount of paintings in test procedure. The art classification results demonstrate that our model is qualified based on proposed adaptive multi-task learning technique while in the previous multi-task learning model, the weight of each task in fixed. For example, in "Timeframe" classification task, the improvement is about 4. 43% with respect to the previous model. In order to calculate the task-specific weighting, the previous model barriers are limited to twice back forward tracing. The art classification results also validate the importance of introducing weighting constraints in our model. Next, we also evaluate our model on cross-modal art retrieval tasks. Experiments are conducted through Text2Art Challenge Evaluation where painting samples are sorted out based on their similarity to an oriented text, and vice versa. The calculated ranking results are evaluated by median rank and recall rate at K, with K being 1, 5 and 10 on the test dataset and performances. Median rank denotes the value separating the higher half of the relevant ranking position amount all samples, whereas recall at rate K represents the rate of samples for which its relevant image is in the top K positions of the ranking. Compared with the most recent knowledge-graph-based model in the context of author attribute, the improvement is about 9. 91% in average which is consistent of classification results. Finally, we compare our model with manual evaluators. Following an artistic text, which contains comment, title, author, type, school and time schedule, participants are required to pick the most proper painting image out from a collection of 10 images. There are two distinct levels in this task as mentioned below: the collection of painting images are easy to random selected from the test set, and the difficulty is where the 10 collected images have the identical attribute category (i. e. , portraits, landscapes) . All participants are required to conduct the task for 100 artistic texts in each level. The performance is reported as the proportion of clear feedbacks over all responses. Our demonstrated results also illustrate that our modeling accuracy is quite closer to human evaluators. **Conclusion** We harness an adaptive multi-task learning method to weight multiple loss functions based on Bayesian theory for automatic art analysis tasks. Furthermore, we conduct several experiments on the public available art dataset. The synthesized results on this dataset include both art classification and art retrieval challenges.

Key words: automatic art analysis; adaptive multi-task learning; Bayesian theory; art classification; cross-modal art retrieval

0 引 言

多任务学习(张钰 等,2020)理论认为通用特征相对于任务相关特征具有更强的表达能力,尝试同时解决多个任务提高学习效率和预测准确性。目前,多任务学习技术已经成功应用于众多计算机视觉任务,例如目标跟踪、目标检测和识别、面部特征点检测和面部属性分类等。多任务学习方法大多使用简单的加权损失总和共同学习多个任务,各个损

失之间权重一般是统一设置或者手动调整(Garcia等,2019),然而实际应用中手工寻找最优的损失权重非常耗时。Cipolla 等人(2018)提出一种基于同类不确定性原理组合多个损失函数同时学习多个目标的方法,针对每个任务选择正确的损失权重,有效提升了多任务学习的最终性能,Chen 等人(2018)提出一种梯度归一化方法,通过动态调整梯度幅度自动平衡多任务学习网络模型的训练。

艺术品数字化促进了文化遗产的保护和传播,自动艺术分析作为最具有发展前途的方向之一得到了广泛关注。自动艺术分析旨在借助人工设计(Khan 等,2014)的描述符或者深度学习方法识别艺术绘画中的特定属性。自动艺术分析绝大部分早期工作都集中于通过人工设计的描述符提取艺术绘画最具代表性的视觉特征。Johnson 等人(2008)使用小波分解方法通过分析笔迹识别作者。Khan 等人(2014)将颜色、边缘或纹理特征组合在一起以实现作者、类型和学校分类。Carneiro 等人(2012)利用SIFT(scale-invariant feature transform)特征将绘画分组为不同的属性。随着深度学习技术的发展,卷积神经网络(convolutional neural network,CNN)已成功应用于艺术绘画的内容和风格分析,例如在艺术风格迁移领域(Sanakoyeu 等,2018)取得了令人惊讶的视觉效果。杨秀芹和张华熊(2020)为了充分提取版画、中国画、水彩画和水粉画等艺术图像的整体风格和局部细节特征,提出通过双核压缩激活模块和深度可分离卷积搭建卷积神经网络对艺术图像进行分类。一般来说,这些方法(Chu 和 Wu,2018)首先从预先训练的卷积神经网络中提取深度特征,然后使用艺术绘画图像对提取的深度特征进行微调,以期获取更好的效果。

大多数自动艺术分析方法主要是通过分析艺术绘画的风格和类型描述艺术作品的视觉本质。例如盛家川和李玉芝(2018)首先将国画风格阐述为一系列的艺术目标,如马、人物等,然后使用深度卷积神经网络描述这些艺术目标的高级语义特征,并通过支持向量机对各种艺术目标的分类结果进行融合。然而,从艺术专家的角度来看(李立红,2019),艺术研究不仅涉及艺术绘画的视觉信息,而且包括蕴含的社会、历史和艺术信息。尽管自动艺术分析研究已经取得了显著的进步,但大多数研究者都将各种艺术目标作为独立的个体进行分析,本文将多任务学习引入自动艺术分析领域,同时解决多任务学习中每个任务损失之间相对权重难以决策这一难题,在贝叶斯理论框架下,本文提出一种自适应多任务学习模型,同时利用视觉外观和视觉上下文信息,通过学习合适的损失权重提升自动艺术分析性能。在艺术图像数据库上的实验验证了本文方法的优越性。

1 自适应多任务学习方法

本文方法采用硬参数共享形式的多任务学习框架,流程图如图 1 所示。各个任务先共享一个编码器网络,再使用各自的解码器网络提取任务相关特征,随后在贝叶斯理论框架下对多任务损失进行建模并完成训练,最终实现自动艺术分析任务(艺术图像分类、跨模态艺术图像检索等)。

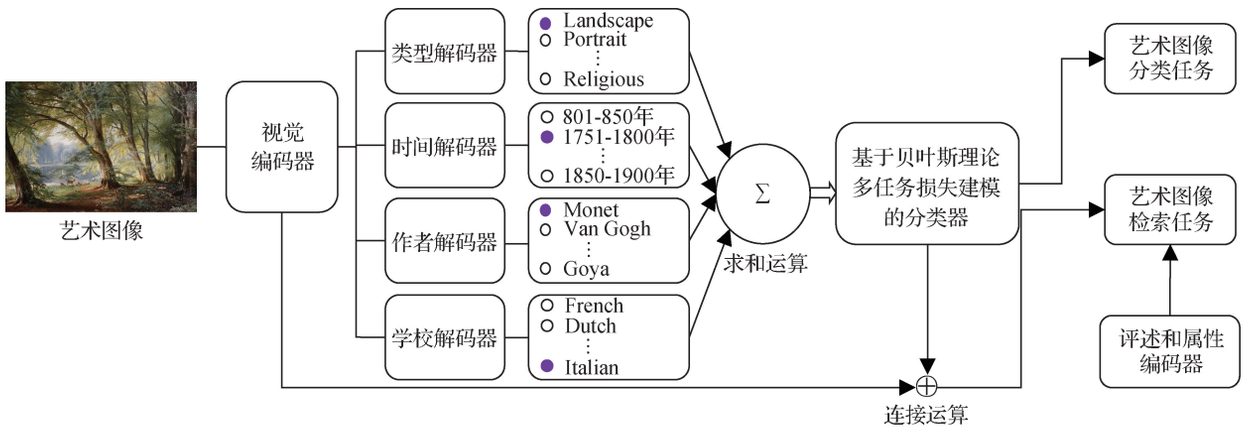


图 1 本文方法流程

Fig. 1 The framework of our proposed method

1.1　多任务学习基本模型

目前在深度学习应用中流行的多任务学习方法试图解决针对多个目标的模型优化问题。具体而言，在多任务学习中，给定 H 个学习任务以及 M 个训练样本 $\{x_i^h, y_i^h\}_{i=1}^M$，其中 $x_i^h \in \mathbf{R}^d$ 和 y_i^h 分别表示第 i 个训练样本及其在第 h 个任务中的标签，多任务学习的优化目标（张钰 等，2020）为

$$\arg\min_{\{\boldsymbol{w}^h\}_{h=1}^H} \sum_{h=1}^{H}\sum_{i=1}^{M} \psi^h L_h(g(x_i^h;\boldsymbol{w}^h), y_i^h) \tag{1}$$

式中，g 是以 $\boldsymbol{w}^h$ 为参数的映射函数，L_h 代表第 h 个任务的损失函数，ψ^h 为损失函数的权重参数，起着平衡各个任务损失的作用。

本文采用硬参数共享形式的多任务学习框架，结合图 1 的流程图和式（1），优化目标可表示为

$$\arg\min_{\boldsymbol{w}_E,\{\boldsymbol{w}_D^h\}_{h=1}^H} \sum_{h=1}^{H}\sum_{i=1}^{M} \psi^h L_h(g_D(g_E(x_i;\boldsymbol{w}_E);\boldsymbol{w}_D^h), y_i^h) \tag{2}$$

式中，g_E 表示以 $\boldsymbol{w}_E$ 为参数的视觉编码器，g_D 表示以 $\boldsymbol{w}_D^h$ 为参数的任务相关解码器，其输出为维度大小 O_L 的向量（O_L 为每个任务的类别数目）。本文采用的交叉熵损失函数（Garcia 等，2019）定义为

$$L_h(z_i^L, y_i^L) = -\log\left(\frac{\exp(z_i^L[y_i^L])}{\sum_o \exp(z_i^L[\boldsymbol{o}])}\right) \tag{3}$$

式中，$z_i^L = g_D(g_E(x_i;\boldsymbol{w}_E);\boldsymbol{w}_D^h)$，exp 为所有的标签集合。多任务学习基本模型一般采用损失函数简单求和进行训练，其中损失函数权重是固定不变的（即对所有任务 ψ^h 取值都一样）。然而，如 Cipolla 等人（2018）指出，多任务学习模型的性能与选择合适的损失函数权重密切相关。为解决上述问题，本文提出一种贝叶斯框架下多任务损失建模方法。

1.2　基于贝叶斯理论的自适应多任务学习模型

Cipolla 等人（2018）率先提出运用贝叶斯建模方法自动调节多任务学习优化目标。受此启发，本文方法对上述方法进行了扩展，统一在贝叶斯理论框架下对多任务学习损失函数进行建模。

本文方法中，g_E 代表以 $\boldsymbol{w}_E$ 为参数的视觉编码器，各个任务均共享该视觉编码器；g_D 代表以 $\boldsymbol{w}_D^h$ 为参数的任务相关解码器。多任务学习模型认为各个学习任务之间存在相互关联性，因此依据层次贝叶斯（hierarchical Bayes）理论，可以假定参数 $\boldsymbol{w}_D^h$ 来源于特定的概率分布，如高斯分布。故而对于所有的任务（Bakker 和 Heskes，2003），$\boldsymbol{w}_D^h$ 可表示为

$$\boldsymbol{w}_D^h = \boldsymbol{w}_D^0 + \boldsymbol{\psi}_h \tag{4}$$

式中，$h = 1, 2, \cdots, H$，当各个任务相似时，向量 $\boldsymbol{\psi}_h$ 表示一个小的扰动量。换言之，本文假定由于任务存在相关性，其任务相关参数 $\boldsymbol{w}_D^h$ 接近于特定的参数 $\boldsymbol{w}_D^0$（在层次贝叶斯理论中可视为高斯均值）。

因此，借助层次贝叶斯理论（Evgeniou 和 Pontil，2004），本文多任务学习优化目标可扩展为

$$\arg\min_{\boldsymbol{w}_E,\{\boldsymbol{w}_D^h\}_{h=1}^H} \Big\{ \sum_{h=1}^{H}\sum_{i=1}^{M} \psi^h L_h(g_D(g_E(x_i;\boldsymbol{w}_E);\boldsymbol{w}_D^h), y_i^h) + \sum_{h=1}^{H} \|\boldsymbol{w}_D^h\|^2 + \frac{\gamma}{H}\sum_{h=1}^{H} \|\boldsymbol{w}_D^h - \boldsymbol{w}_D^0\|^2 \Big\} \tag{5}$$

式中，γ 为超参数。相比于式（2），式（5）增加了两项正则项：任务相关解码器参数 $\boldsymbol{w}_D^h$ 向量模值以及这些参数向量和参数向量平均值之间差异的模值。式（5）实质上借助层次贝叶斯理论，引入任务簇（clustering）约束模型，通过增加不同任务向量模值以及它们之间方差这两个正则项，促使不同任务向量趋向于任务的平均向量。因此在训练过程中，不同任务向量受到任务平均向量的约束，其向量权值并不是独立产生的，这体现了多任务学习模型中各个任务之间的相关性。

再者，依据贝叶斯建模方法，多任务损失函数可以通过最大化不确定性的高斯似然进行构造（Cipolla 等，2018），这种不确定性可以看成是任务相关的不确定性。本文采用 softmax 函数进行艺术图像分类，则多任务学习模型的损失函数最小化目标 $L(\boldsymbol{w}, \sigma_1, \cdots, \sigma_h)$ 定义为

$$-\log P(z_1, \cdots, z_h = c \mid g^w(x)) = -\log[\operatorname{softmax}(z_1 = c; g^w(x), \sigma_1) \times \cdots \times \operatorname{softmax}(z_h = c; g^w(x), \sigma_h)] \tag{6}$$

式中，参数 $\boldsymbol{w} = [\boldsymbol{w}_E, \boldsymbol{w}_D]$ 代表网络整体参数。进一步将 softmax 函数展开，可以得到

$$-\log P(z_1, \cdots, z_h = c \mid g^w(x)) = \frac{1}{\sigma_1^2}L_1(\boldsymbol{w}) + \log\frac{\sum_{c'}\exp\left(\frac{1}{\sigma_1^2}g_{c'}^w(x)\right)}{\sum_{c'}\exp(g_{c'}^w(x))^{\frac{1}{\sigma_1^2}}} + \cdots + \frac{1}{\sigma_h^2}L_h(\boldsymbol{w}) + \log\frac{\sum_{c'}\exp\left(\frac{1}{\sigma_h^2}g_{c'}^w(x)\right)}{\sum_{c'}\exp(g_{c'}^w(x))^{\frac{1}{\sigma_h^2}}} \approx \frac{1}{\sigma_1^2}L_1(\boldsymbol{w}) + \cdots + \frac{1}{\sigma_h^2}L_h(\boldsymbol{w}) +$$

$$\log(\sigma_1) + \cdots + \log(\sigma_h) \tag{7}$$

根据式(7)能够学习到每种损失的相对权重。σ 取值较大时,其损失贡献较小;反之,其损失贡献较大。并且,式(7)最后的正则项能避免 σ 取值过大。最终,联合式(5)和式(7),本文多任务学习优化目标定义为

$$\arg\min_{\boldsymbol{w}_E,\{\boldsymbol{w}_D^h\}_{h=1}^H}\left\{\sum_{h=1}^{H}\left[\frac{1}{\sigma_h^2}L_h(\boldsymbol{w})+\log(\sigma_h)\right]+\sum_{h=1}^{H}\|\boldsymbol{w}_D^h\|^2+\frac{\gamma}{H}\sum_{h=1}^{H}\|\boldsymbol{w}_D^h-\boldsymbol{w}_D^0\|^2\right\} \tag{8}$$

式中,γ 是超参数。

本文在贝叶斯理论基础下,构建了一种自适应多任务学习模型,其优化目标定义如式(8)所示。这种自适应多任务学习模型能够很便利地扩展至任意同类学习任务,例如基于多任务学习的艺术分析任务,从而使本文可以有效地学习到每种损失合适的相对权重,同时正则项的引入能够提升模型的最终性能。值得注意的是,本文方法的性能和任务数量没有必然关联。在多属性商品图像分类、面部属性分类等领域,本文方法也具备较大的应用潜力。

Cipolla 等人(2018)采用类似式(7)的形式构造多任务损失函数。图 2 描述了本文方法和 Cipolla 等人(2018)方法训练过程中的训练损失变动情况。从图 2 中可以明显看出,使用 Cipolla 等人(2018)方法进行训练时,训练损失会处于波动状态且无法进一步收敛。而采用本文方法进行训练时,训练损失会快速下降到某一收敛值,这说明通过层次贝叶斯理论引入的正则项起着较好的约束作用。在艺术图像数据库上的实验验证了本文方法的有效性。

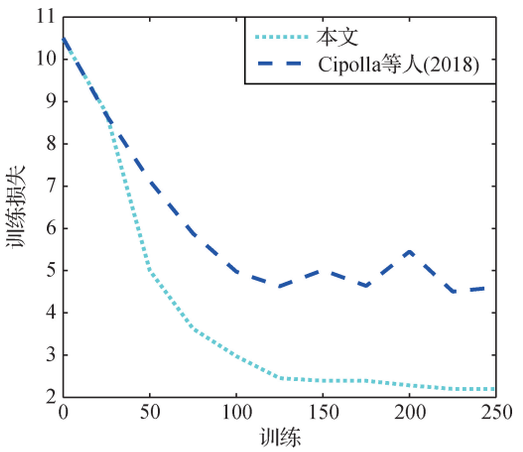


图 2 训练过程中训练损失比较

Fig. 2 Training loss comparisons during training procedure

2 自动艺术分析实验

为验证本文提出的自适应多任务学习模型的性能,在艺术图像数据库上进行自动艺术分析实验。为与 Garcia 等人(2019)的方法进行公平比较,进行类似具有代表性的艺术图像分类、跨模态艺术图像检索实验。

2.1 艺术图像数据库

现有多种艺术图像数据库,然而一些数据库存在如下缺陷:数据集样本太小因而无法训练深度学习模型(Crowley 和 Zisserman,2014)、每个样本并不具备多种属性(Karayev 等,2014)或者无法公开下载(Mao 等,2017)。本文选取 Garcia 和 Vogiatzis(2018)发布的 SemArt 艺术图像数据库作为实验对象。该数据库是一种面向艺术语义理解的多模态数据库,可以用于跨模态检索以及分类任务。SemArt 艺术图像数据库共有 21 382 幅艺术绘画图像,本文将其随机分为训练集、验证集和测试集,分别包括 19 244、1 069 和 1 069 幅图像。每幅艺术绘画图像都含有详细的艺术文本注释以及多种属性(例如作者(author)、标题(title)、日期(data)、技术(technique)、类型(type)、学校(school)和时间范围(TF)等)。SemArt 艺术图像数据库中样本如图 3 所示。

图 3 SemArt 艺术图像数据库中绘画样本

Fig. 3 Painting samples in SemArt database

考虑到 SemArt 艺术图像数据库中不是所有的绘画样本都具有“日期”(data)和“技术”(technique)属性,本文选取 5 种属性进行衡量。1)“作者”(author)属性。SemArt 艺术图像数据库包含3 281 位不同的作者,类似 Garcia 和 Vogiatzis(2018)的做法,本文只选择训练集中至少有 10 幅艺术绘画的作者,而将作品少于 10 幅的作者设置为未知画家。最后共计收集 350 位不同的画家;2)“标题”(title 属性)。数据库中共有 14 902 个不同的标题,其中 38.8% 的绘画具备

一个非唯一性的标题。在所有标题中,静物和自画像是最常见的标题;3)“类型”(type)属性。类型字段根据肖像、风景、宗教、研究、流派、静物、神话、室内、历史以及其它等 10 种不同流派对绘画进行分类;4)“学校”(school)属性。数据库中有 26 所艺术学校,根据“学校”属性,本文丢弃训练集中出现不超过 10 幅绘画的学校,并将这些学校命名为未知学校。共收集 25 所不同的艺术学校,学校所在地包括意大利、德国和法国等;5)“时间范围”(TF)属性。根据每幅画的创作时间构建时间范围属性。在公元 801—1900 年之间按 50 年一个周期均匀采样,得到 22 个不同的时间范围。类似地,本文只选择训练集中至少包含 10 幅绘画的时间范围,总计收集 18 个类别的时间范围集合,其中包括一个未知类别的时间范围。上述 5 种属性中,标题属性和类型属性描述了艺术图像的种类特征;作者属性反映了艺术图像的风格特征;学校属性表达了艺术图像的社会属性;时间范围属性阐述了艺术图像的历史属性。因此本文采用的多任务学习方法能够全面地刻画艺术品的内在属性特征。

2.2　艺术图像分类实验

2.2.1　分类实现细节及对比模型

本文采用去除最后一个全连接层的 ResNet50(residual neural network)(He 等,2016)作为图 1 所示的视觉编码器。ResNet50 使用其标准的面向自然图像分类任务的预训练权重进行初始化,其余各层的权重则随机初始化。输入的艺术绘画图像统一缩放为 256 × 256 像素,并随机裁剪成边长为 224 像素的小图。训练网络模型时,对绘画图像随机水平翻转以增加样本数量。ResNet50 输出为一个 2 048 维度的向量,随后分别送入几种不同类型的任务解码器中,该任务解码器采用一个全连接层进行构造,解码器的大小与每个任务中的类别数相对应。因此,任务解码器参数大小为一个 2 048 × (O_L) 的权重矩阵(O_L 为每个任务的类别数目),具体实现时为了方便式(8)的计算,对上述权重矩阵按列进行平均计算,最终得到一个 2 048 维度的任务相关向量代入多任务学习目标式(8)。本文选择动量为 0.9,学习率为 0.002 5 的随机梯度下降器作为优化器。训练时,批处理(mini-batch)的大小为 28,最大迭代次数为 300。

为了衡量本文方法的性能,选用如下模型比较:

1)预训练模型(pre-trained model)。选用面向自然图像分类任务的带有预训练权重的 VGG16(Visual Geometry Group network 16-layer)(Simonyan 和 Zisserman,2015)、ResNet50 以及 ResNet152 模型(He 等,2016),随后修改最后一个全连接层的输出大小,使之与每个任务类别数目相等,最后一层全连接层权重随机初始化,模型中其他层权重保持不变。

2)微调模型(fine-tuned model)。类似预训练模型,选用 VGG16、ResNet50 以及 ResNet152 模型,并且修改最后一个全连接层,模型各层中权重参数在训练过程进行微调。

3)内容感知多任务学习模型(context-aware multi-task learning model,MTL)。为了实现内容感知,Garcia 等人(2019)提出一种多任务学习模型(multi-task learning,MTL)共同学习多个艺术任务,发掘任务之间的视觉相似性。

4)内容感知知识图谱模型(context-aware knowledge graph model,KGM)。Garcia 等人(2019)采用知识图谱用于学习艺术属性之间的独特关系,首先将一组绘画与其艺术相关的属性联系起来,生成特定的艺术知识图谱,随后以图中的节点邻域和位置编码为向量表示上下文内容。

5)权重不确定性多任务学习模型(weight uncertainty multi-task learning model,WU)。通过考虑每个任务的权重不确定性,Cipolla 等人(2018)试图在每个任务的损失函数之间学习到合适的权重,从而提高模型的实际性能。

6)梯度归一化多任务学习模型(gradient normalization multi-task learning model, GradNorm)。Chen 等人(2018)提出一种梯度归一化(GradNorm)算法,通过动态调整梯度幅度自动平衡多任务学习训练。该方法可以提高准确性并减少多个任务之间的过度拟合。

上述 6 种模型中,预训练模型和微调模型为传统的深度学习网络模型;内容感知多任务学习模型和内容感知知识图谱模型为有代表性的固定权重的多任务学习模型;权重不确定性多任务学习模型和梯度归一化多任务学习模型为有代表性的可自动调节权重的多任务学习模型。与该 6 种模型进行比较,可全面衡量本文方法的性能。

2.2.2　分类结果

本文在 SemArt 数据库上进行艺术图像分类实验,通过衡量分类准确性(即正确分类的绘画占绘

画总数的比率)评估性能,结果如图 4 所示。可以看出,1)除了内容感知知识图谱模型(KGM)在“类型”(type)任务上效果最优以外,本文方法在大多数分类任务中都获得了最佳的准确性。2)预训练模型在所有模型中表现最差,这是因为它们关注于自然图像分类,因此在艺术图像分类领域没有很好的判别能力。而通过对这些模型进行微调,则可以提高预训练模型的性能。3)相比于微调模型,MTL 或者 KGM 模型通过额外获取各种艺术属性中的关系,获得了更高的准确性。

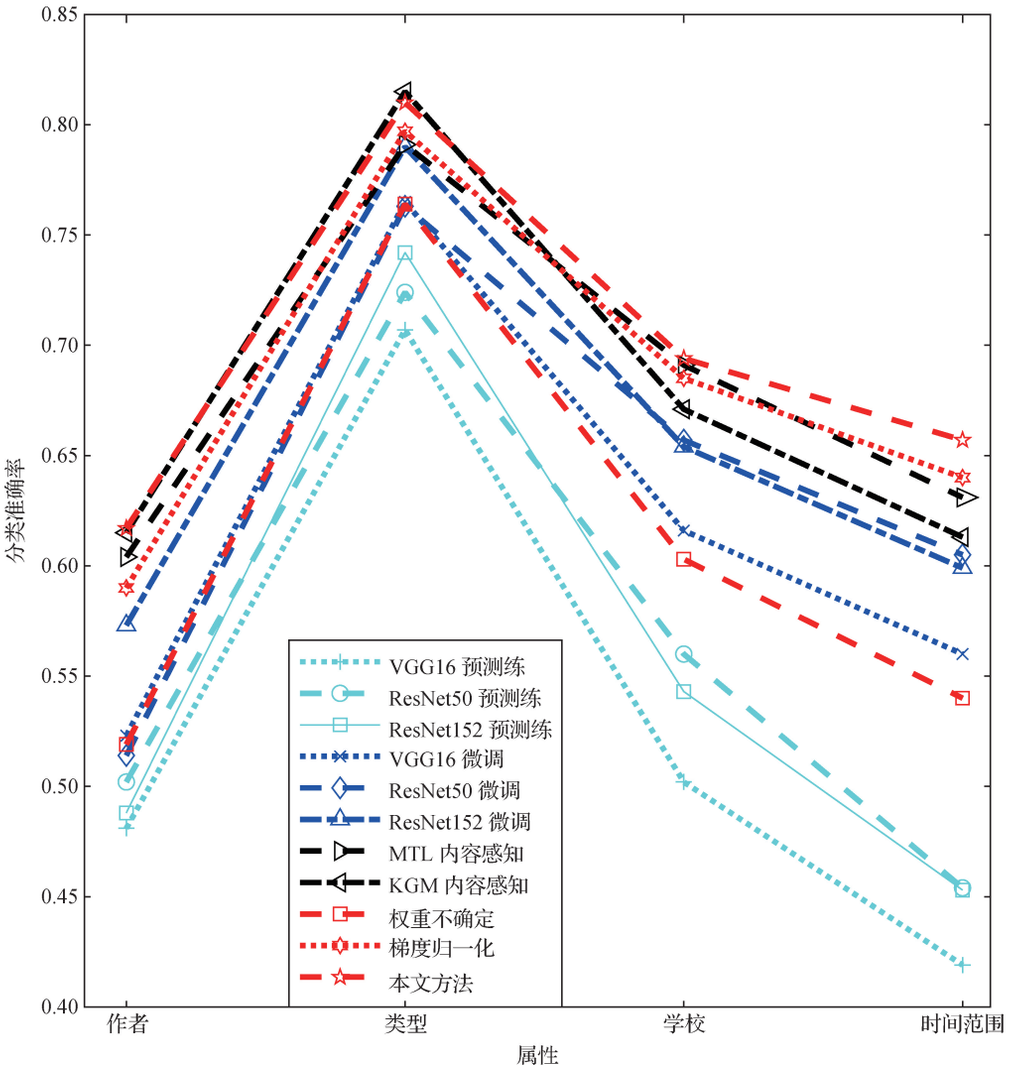


图 4　SemArt 艺术图像数据库分类实验

Fig. 4　Classification experiments of SemArt database

MTL 模型中每个任务的权重均保持固定,而本文提出的自适应多任务学习方法通过自动调节权重提高分类性能。例如,在“时间范围”分类任务中,相对于 MTL 模型,本文方法性能提升了 4.43%。在“学校”(school)和“时间范围”(TF)分类任务中,MTL 模型的性能优于 KGM 模型,而 KGM 模型在“类型”(type)和“作者”(author)分类任务中表现更佳。与之对应,本文方法的性能保持一致且更加稳定。此外,本文方法也优于目前流行的自动确定多任务损失权重方法(Cipolla 等,2018;Chen 等,2018)。值得注意的是,为了计算特定的任务权重,Chen 等人(2018)方法需要两次反向传播,既费时又不易于实现。分类实验中,本文方法性能优于 Cipolla 等人(2018)方法,验证了本文方法借助层次贝叶斯理论加入正则项约束的必要性。在分类实验中还发现,本文方法“时间范围”(TF)分类任务的权重略大于“类型”(type)分类任务的权重,说明本文方法可自适应地调节权重来关注相对较难的分类

任务。

2.3　艺术图像检索实验

2.3.1　检索实现细节

本文在 SemArt 艺术图像数据库上进行全面的实验以衡量各种模型的性能，类似 Garcia 和 Vogiatzis(2018)提出的 Text2Art 的做法，进行根据艺术文本检索相关绘画以及根据绘画检索相关艺术文本实验。为了与 Garcia 等人(2019)模型进行公平比较，在视觉编码器中添加艺术分类器，将视觉表征信息集成到跨模态的检索模型中。跨模态的检索模型详细构建过程如下：

1)视觉编码器。输入的艺术绘画图像统一缩放为 256 × 256 像素大小，并随机裁剪成边长为 224 像素的小图。随后将其作为 ResNet50 输入图像，该 ResNet50 使用标准的预训练权重进行初始化，输出大小为 1 000 维的向量 $\boldsymbol{a}_{\text{cnn}}$。另一方面，绘画图像同时输入到已训练完毕的分类器中(例如本文提出的模型、微调模型、MTL 模型或者 KGM 模型)，以期获得与属性相关的向量 $\boldsymbol{a}_{\text{att}}$，该向量含有 c 个分量，c 表示分类器的输出类别数目。视觉编码器最终通过两个向量的级联来表示，即 $\boldsymbol{a}=\boldsymbol{a}_{\text{cnn}}\oplus\boldsymbol{a}_{\text{att}}$。

2)评述和属性编码器。艺术绘画每个评述通过词频—逆文档频率(tf-idf)策略进行编码，评述字典库大小为 9 708，由训练数据集中至少重复出现 10 次的字母词构成，随后获取与评述相关的 tf-idf 向量 $\boldsymbol{b}_{\text{com}}$。类似地，通过使用大小为 9 092 的词汇表，可以将艺术绘画的标题编码为另一向量 $\boldsymbol{b}_{\text{tit}}$。进一步，采用一位有效(one-hot)向量 $\boldsymbol{b}_{\text{att}}$ 编码 author、type、school 或 TF 属性。最终将这 3 个向量级联起来得到 $\boldsymbol{b}=\boldsymbol{b}_{\text{com}}\oplus\boldsymbol{b}_{\text{tit}}\oplus\boldsymbol{b}_{\text{att}}$。

3)跨模态投影。为了衡量跨模态数据的相似性，分别使用非线性函数 D_a 和 D_b 将 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 转换到 128 维空间中。非线性函数通过一个全连接层、tanh 激活函数和 L2 归一化来构造。一旦 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 投影完毕，本文采用余弦相似度计算和匹配排序结果。

为了训练提出的检索模型，本文收集了绘画样本的正负匹配对。然后基于余弦边际损失函数，训练检索模型的权重(艺术分类器的权重保持不变)。具体为

$$\chi(\boldsymbol{a}_k,\boldsymbol{b}_i)=\begin{cases}1-Csim(D_a(\boldsymbol{a}_k),D_b(\boldsymbol{b}_i)) & k=i\\ \max(0,Csim(D_a(\boldsymbol{a}_k),D_b(\boldsymbol{b}_i))-\theta) & k\neq i\end{cases} \tag{9}$$

式中，$Csim$ 表示余弦相似度比较函数，θ 是取值为 0.1 的固定阈值。

2.3.2　检索实验结果

本文通过在 SemArt 数据库上进行 Text2Art 任务来评估各种模型的性能，其中绘图图像根据与其给定的文本相似性进行排序，反之亦然。检索结果报告为中位数排名(median rank，MR)以及 K 的召回率($R@K$)，K 取值为 1、5 和 10。MR 取值越低，K 的召回率越高，检索结果越好。为了全面衡量检索效果，与以下几种方法进行比较：1) Garcia 和 Vogiatzis(2018)提出的 CML(context-aware multi-task learning model)模型。该模型对评述和标题信息进行了编码，而没有用到属性信息；2) CML*。CML* 是 CML 模型的重实现版本，效果上稍微有所提升；3) AMD 模型。该模型在训练过程中利用属性推断视觉和文字映射，微调模型采用 ResNet152(He 等，2016)结构。4) Garcia 等人(2019)方法。该方法采用各种内容感知相关分类器，如 MTL-author、MTL-type、MTL-school、MTL-timeframe、KGM-author、KGM-type，KGM-school 和 KGM-timeframe。检索实验结果如表 1 和图 5 所示，可以看出，与其他模型相比，本文方法取得最佳性能。例如，与采用“作者”(author)属性的 KGM 模型相比，本文方法提升了 9.91%。使用“学校”(school)属性的 MTL 模型性能优于 KGM 模型，而 KGM 模型在“类型”(type)和“作者”(author)属性方面表现更佳。本文方法性能保持一致并且比较稳定，因此更适合于实际的跨模态检索应用。同时，与目前流行的自动确定多任务损失权重方法(Cipolla 等，2018；Chen 等，2018)进行了比较，本文方法在艺术检索任务中取得了更好的效果。

从表 1 和图 5 还可以看出，将输出与指定属性连接起来的模型(ResNet152、MTL、KGM 和本文方法)较 AMD 模型结果有较大改善和提升，然而不同属性之间的性能差异较大，在 ResNet152、MTL、KGM 和本文方法中，与“类型”(type)、“学校”(school)和“时间范围”(TF)属性相比，“作者”(author)属性具有最佳性能。本文推测这种现象可能起源于每个属性的类别数目有所不同。

表 1　SemArt 艺术图像数据库上 Text2Art 任务的中位数排名

Table 1　Median rank results on the Text2Art challenge of SemArt database

模型	属性	中位数排名	
		文本检索绘画	绘画检索文本
CML	未使用	14	14
CML*	未使用	10	12
AMD	author	17	16
AMD	type	17	16
AMD	school	19	16
AMD	TF	20	17
ResNet152	author	7	8
ResNet152	type	9	11
ResNet152	school	10	12
ResNet152	TF	18	16
MTL	author	7	9
MTL	type	12	12
MTL	school	8	10
MTL	TF	9	12
KGM	author	6	7
KGM	type	10	10
KGM	school	12	11
KGM	TF	10	12
WU	author	7	10
WU	type	11	11
WU	school	8	11
WU	TF	10	12
GradNorm	author	6	7
GradNorm	type	9	10
GradNorm	school	7	9
GradNorm	TF	8	9
本文	author	5	6
本文	type	8	9
本文	school	6	7
本文	TF	7	7

本文采用 Garcia 和 Vogiatzis(2018)做法,将 CCA(Garcia 和 Vogiatzis,2018)、CML(Garcia 和 Vogiatzis,2018)、KGM(author)、WU(author)、GradNorm(author)和本文方法(author)等模型与人类识别效果进行比较。对于给定的艺术描述,诸如评述(comment)、标题(title)、作者(author)、类型(type)、学校(school)和时间范围(TF),要求人类评估者从 10 幅图像中选择最合适的图像。此任务包含两种不同的人为定义的难度级别:1)容易级别,即从测试集中的所有绘画中随机选择图像;2)困难级别,即挑选的数据集中 10 幅图像含有相同的属性类型(例如肖像、风景等)。每种级别中评估人员对 100 幅艺术作品进行识别,并统计最终的准确度,结果如表 2 所示。从表中可以看出,本文方法较其他方法性能更优,识别效果接近于人工评估的结果。

3　讨　论

本文方法中唯一需要确定的参数是超参数 γ(式(8)),表 3 列举了 γ 不同取值对于艺术图像分类任务最终正确率的影响。显然,γ 在不同取值下,本文方法分类精度基本保持稳定,其中 $\gamma = 1$ 时效果略好。因此本文所有试验中,γ 参数都取固定值。

为了探索不同模型如何挖掘艺术作品信息,本文通过测量各种艺术属性的可分离性进行更深一步研究。具体而言,本文从艺术图像分类任务中收集测试数据,并使用 Davies-Bouldin(DB)指数(Garcia 等,2019)计算聚类 i 和聚类 k 之间的可分离性,具体为

$$DB_{\text{index}} = \frac{1}{N}\sum_{i=1}^{N}\left(\max_{i\neq k}\left(\frac{\varepsilon_i + \varepsilon_k}{\varepsilon_{ik}}\right)\right) \tag{10}$$

式中,N 代表聚类的数量,散度(dispersion)ε_i 和分离度(separation)ε_{ik} 分别定义为

$$\varepsilon_i = \left(\frac{1}{|T_i|}\sum_{x\in T_i}\|x - C_i\|^p\right)^{1/p},\quad \varepsilon_{ik} = \|C_i - C_k\|_p \tag{11}$$

式中,C_k 代表聚类 k 的质心,C_i 代表聚类 i 的质心。计算时从训练集中收集艺术绘画样本,同时设置 $P=2$。本文利用“作者”(author)、“类型”(type)、“学校”(school)和“时间范围”(TF)属性评估不同模型的性能。显然,DB 指数的值越小,聚类分离趋势就越好。

图 6 显示不同模型不同属性的 DB 指数结果。显而易见,ResNet152 模型结果最差,而经过微调的 ResNet152 模型性能与 MTL 或 KGM 模型的性能相

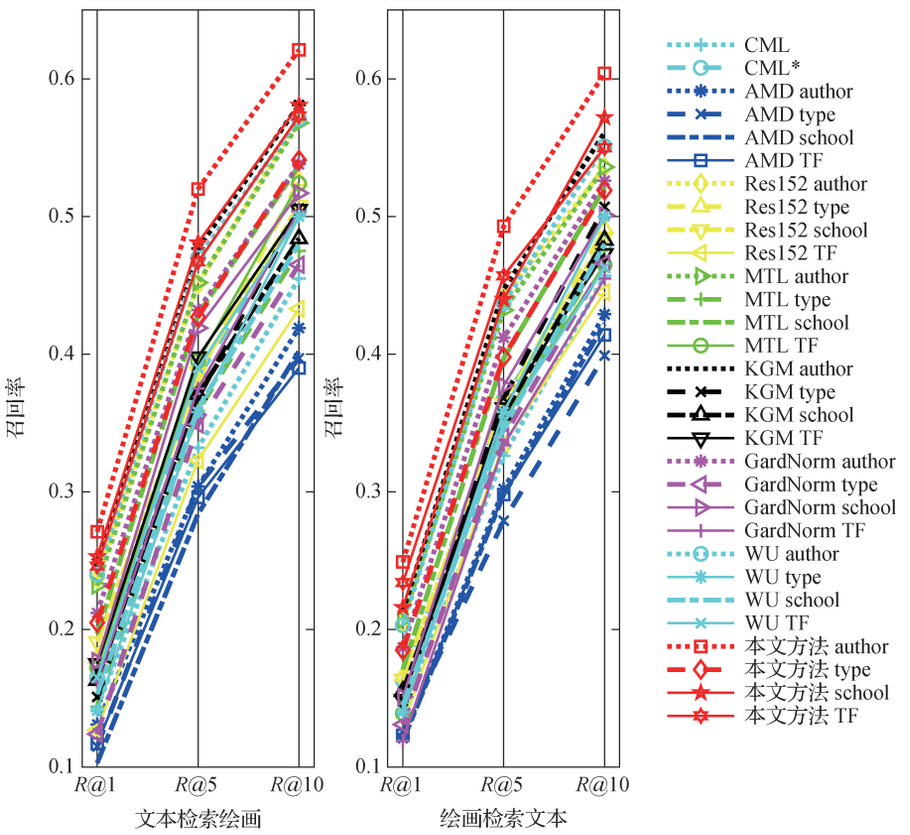


图 5　SemArt 艺术图像数据库上 Text2Art 任务召回率结果

Fig. 5　Recall results on the Text2Art challenge of SemArt database

表 2　Text2Art 任务中各种模型及人类识别性能比较

Table 2　Model performance and human evaluation on the Text2Art challenge

模型	容易数据集						困难数据集					
	风景	宗教	神话	体裁	肖像	总计	风景	宗教	神话	体裁	肖像	总计
CCA	0. 708	0. 609	0. 571	0. 714	0. 615	0. 650	0. 600	0. 525	0. 400	0. 300	0. 400	0. 470
CML	**0. 917**	0. 683	0. 714	**1**	0. 538	0. 750	0. 500	**0. 875**	0. 600	0. 200	0. 500	0. 620
KGM author	0. 875	0. 805	0. 857	0. 857	0. 846	0. 830	0. 600	0. 825	0. 700	0. 400	0. 650	0. 680
WU author	0. 854	0. 783	0. 814	0. 823	0. 819	0. 801	0. 582	0. 809	0. 667	0. 381	0. 624	0. 645
GradNorm author	0. 886	0. 809	0. 853	0. 880	0. 84	0. 842	0. 601	0. 831	0. 692	0. 404	0. 651	0. 675
本文方法 author	0. 913	**0. 817**	**0. 878**	0. 896	**0. 882**	**0. 868**	**0. 608**	0. 862	**0. 711**	**0. 421**	**0. 661**	**0. 697**
人工	0. 918	0. 795	0. 864	1	1	0. 889	0. 579	0. 744	0. 714	0. 72	0. 674	0. 714

注：加粗字体表示模型在各列（除人工行）的最优结果。

表 3　不同 γ 取值对分类准确率的影响

Table 3　Classification accuracy of different values of γ

γ	平均准确率
10	0. 661
2	0. 675
1	0. 682
0. 5	0. 672
0. 1	0. 658

当。除了“作者”（author）属性，对于大多数属性例如“类型”（type）、“学校”（school）以及“时间范围”（TF）属性，KGM 模型的性能要优于 MTL 模型。相比其他模型，本文方法由于对于艺术作品具有高度的区分性和一致的表现能力，故而在所有属性上都取得了最佳性能。值得注意的是，“作者”（author）属性由于其类别众多而具有最高的分散性，因此其对应的 *DB* 指数最低。

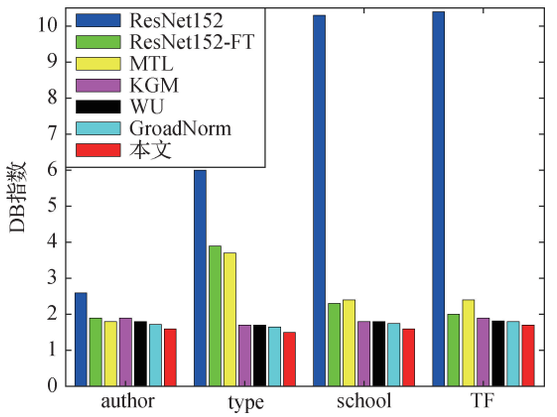


图 6　不同模型不同属性的 Davies-Bouldin 指数

Fig. 6　Davies-Bouldin index for each different attribute of different models

4　结　论

自动艺术分析作为当前文化遗产保护和传播领域研究热点之一得到了广泛关注。本文在贝叶斯理论框架下,针对同类型任务(例如分类任务)提出一种原创的自适应多任务学习方法完成自动艺术分析任务,并通过实验验证了方法的有效性和优越性,可得到以下结论:1)本文方法可以在多任务学习框架内自适应地学习每个任务的权重,从而提高自动艺术分析任务的性能。2)在公开艺术图像数据库上的艺术图像分类、跨模态艺术图像检索实验表明,由于本文方法具备较高的和稳定的艺术信息判别能力,因此较目前流行的固定权重或者可自动调节权重的多任务学习方法都取得了更好的识别效果。3)本文方法仅选用经典的深度学习模型构建多任务学习模型,下一步计划研究如何为共享参数的多任务学习选择合适的模型,如何将知识图谱模型与多任务学习模型结合进一步提高性能。

参考文献(References)

Bakker B and Heskes T. 2003. Task clustering and gating for Bayesian multitask learning. The Journal of Machine Learning Research, 4: 83-99 [DOI: 10.1162/153244304322765658]

Carneiro G, da Silva N P, Del Bue A and Paulo Costeira J. 2012. Artistic image classification: an analysis on the PRINTART database// Proceedings of the 12th European Conference on Computer Vision. Florence, Italy: Springer: 143-157 [DOI: 10.1007/978-3-642-33765-9_11]

Chen Z, Badrinarayanan V, Lee C Y and Rabinovich A. 2018. GradNorm: gradient normalization for adaptive loss balancing in deep multitask networks//Proceedings of the 35th International Conference on Machine Learning. Stockholm, Sweden: PMLR: 794-803

Chu W T and Wu Y L. 2018. Image style classification based on learnt deep correlation features. IEEE Transactions on Multimedia, 20(9): 2491-2502 [DOI: 10.1109/TMM.2018.2801718]

Cipolla R, Gal Y and Kendall A. 2018. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 7482-7491 [DOI: 10.1109/CVPR.2018.00781]

Crowley E and Zisserman A. 2014. The state of the art: object retrieval in paintings using discriminative regions//Proceedings of British Machine Vision Conference. Nottingham, UK: BMVA Press: 1-12 [DOI: 10.5244/C.28.38]

Evgeniou T and Pontil M. 2004. Regularized multi——task learning// Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Seattle, USA: ACM: 109-117 [DOI: 10.1145/1014052.1014067]

Garcia N and Vogiatzis G. 2018. How to read paintings: semantic art understanding with multi-modal retrieval//Proceedings of the European Conference on Computer Vision. Munich, Germany: Springer: 676-691 [DOI: 10.1007/978-3-030-11012-3_52]

Garcia N, Renoust B and Nakashima Y. 2019. Context-aware embeddings for automatic art analysis//Proceedings of 2019 on International Conference on Multimedia Retrieval. Ottawa, Canada: ACM: 25-33[DOI: 10.1145/3323873.3325028]

He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]

Johnson C R, Hendriks E, Berezhnoy I J, Brevdo E, Hughes S M, Daubechies I, Li J, Postma E and Wang J Z. 2008. Image processing for artist identification. IEEE Signal Processing Magazine, 25(4): 37-48 [DOI: 10.1109/MSP.2008.923513]

Karayev S. Trentacoste M, Han H, Agarwala A, Darrell T, Hertzmann A and Winnemoeller H. 2014. Recognizing image style//Proceedings of 2014 British Machine Vision Conference. Nottingham, UK: BMVA Press: 1-11 [DOI: 10.5244/C.28.122]

Khan F S, Beigpour S, Van de Weijer J and Felsberg M. 2014. Painting-91: a large scale database for computational painting categorization. Machine Vision and Applications, 25(6): 1385-1397 [DOI: 10.1007/s00138-014-0621-6]

Li L H. 2019. Artistic Conception Composition of Chinese Freehand Brushwork Oil Painting from the Perspective of Integration of Chinese and Western. Nanjing: Nanjing University of the Arts (李立红. 2019. 中西融合视域下中国写意油画的意境构成. 南京: 南京艺术学院)

Mao H, Cheung M and She J. 2017. DeepArt: learning joint representations of visual arts//Proceedings of the 25th ACM international conference on Multimedia. Mountain View, USA: ACM: 1183-1191 [DOI: 10.1145/3123266.3123405]

Sanakoyeu A, Kotovenko D, Lang S and Ommer B. 2018. A style-aware content loss for real-time HD style transfer//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 715-731 [DOI: 10.1007/978-3-030-01237-3_43]

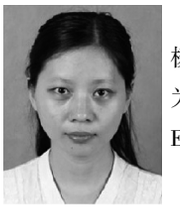
Sheng J C and Li Y Z. 2018. Learning artistic objects for improved classification of Chinese paintings. Journal of Image and Graphics, 23(8): 1193-1206 (盛家川, 李玉芝. 2018. 国画的艺术目标分割及深度学习与分类. 中国图象图形学报, 23(8): 1193-1206) [DOI: 10.11834/jig.170545]

Simonyan K and Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition//Proceedings of the 3rd International Conference on Learning Representations. San Diego, USA: Elsevier: 1-14

Yang X Q and Zhang H X. 2020. Art image classification with double kernel squeeze-and-excitation neural network. Journal of Image and Graphics, 25(5): 967-976 (杨秀芹, 张华熊. 2020. 双核压缩激活神经网络艺术图像分类. 中国图象图形学报, 25(5): 967-976) [DOI: 10.11834/jig.190245]

Zhang Y, Liu J W and Zuo X. 2020. Survey of multi-task learning. Chinese Journal of Computers, 43(7): 1340-1378 (张钰, 刘建伟, 左信. 2020. 多任务学习. 计算机学报, 43(7): 1340-1378) [DOI: 10.11897/SP.J.1016.2020.01340]

作者简介



杨冰,1985 年生,女,副教授,主要研究方向为模式识别和计算机视觉。
E-mail: yb@hdu.edu.cn

向学勤,通信作者,男,工程师,主要研究方向为人工智能。
E-mail: xxueq@aliyun.com

孔万增,男,教授,主要研究方向为脑机信号处理和机器学习。E-mail: kongwanzeng@hdu.edu.cn
施妍,女,副教授,主要研究方向为数字媒体处理。
E-mail: yanshi@hdu.edu.cn
姚金良,男,副教授,主要研究方向为模式识别和计算机视觉。yaojinl@hdu.edu.cn