

# 基于分层基因优选多特征融合的图像材质属性标注

张红斌<sup>1</sup> 邱蝶蝶<sup>1</sup> 邬任重<sup>1</sup> 蒋子良<sup>1</sup> 武晋鹏<sup>1</sup> 姬东鸿<sup>2</sup>

**摘要** 图像材质属性标注在电商平台、机器人视觉、工业检测等领域都具有广阔的应用前景. 准确利用特征间的互补性及分类模型的决策能力是提升标注性能的关键. 提出分层基因优选多特征融合 (Stratified gene selection multi-feature fusion, SGSMMFF) 算法: 提取图像传统及深度学习特征; 采用分类模型计算特征预估概率; 改进有效区域基因优选 (Effective range based gene selection, ERGS) 算法, 并在其中融入分层先验信息 (Stratified priori information, SPI), 逐层、动态地为预估概率计算 ERGS 权重; 池化预估概率并做 ERGS 加权, 实现多特征融合. 在 MattrSet 和 Fabric 两个数据集上完成实验, 结果表明: SGSMMFF 算法中可加入任意分类模型, 并实现多特征融合; 平均值池化方法、分层先验信息所提供的难分样本信息、“S + G + L”及“S + V”特征组合等均有助于改善材质属性标注性能. 在上述两个数据集上, SGSMMFF 算法的精准度较最强基线分别提升 18.70%、15.60%.

**关键词** 材质属性标注, 分层基因优选, 多特征融合, 预估概率, 分层先验信息, 难分样本信息

**引用格式** 张红斌, 邱蝶蝶, 邬任重, 蒋子良, 武晋鹏, 姬东鸿. 基于分层基因优选多特征融合的图像材质属性标注. 自动化学报, 2020, 46(10): 2191–2213

**DOI** 10.16383/j.aas.c190012

## Stratified Gene Selection Multi-Feature Fusion for Image Material Attribute Annotation

ZHANG Hong-Bin<sup>1</sup> QIU Die-Die<sup>1</sup> WU Ren-Zhong<sup>1</sup> JIANG Zi-Liang<sup>1</sup> WU Jin-Peng<sup>1</sup> JI Dong-Hong<sup>2</sup>

**Abstract** Material attribute annotation can be broadly applied in many different scenarios in large-scale product image retrieval, robotics and industrial inspection. Accurately utilizing the complementarity between different image features and the decision abilities of classification models is the key factor to improve the final annotation performance. To address the problem, a novel algorithm called stratified gene selection multi-feature fusion (SGSMFF) for material attribute annotation is proposed. Both the traditional and deep learning image features are extracted firstly. Then any classification model is utilized to compute the estimated probability of each image feature. The traditional effective range based gene selection (ERGS) algorithm is modified in turn and the stratified priori information (SPI) obtained from two perspectives is integrated into the modified ERGS algorithm to dynamically compute the ERGS weight of each estimated probability. Two pooling strategies i.e. Maximum and Average are proposed to complete the final multi-feature fusion procedure. The proposed SGSMFF algorithm is validated on two different datasets: MattrSet and Fabric. Experimental results demonstrate that any classification model can be integrated into the innovative SGSMFF algorithm. Several fundamental factors such as the proposed Average pooling strategy, the hard negative information provided by the stratified priori information, and the feature combinations including “S + G + L” and “S + V” all help improve the final annotation performance. Our approach significantly outperforms state-of-the-art baseline about 18.70% and 15.60% on the above datasets respectively.

**Key words** Material attribute annotation, stratified gene selection, multi-feature fusion, estimated probability, stratified priori information (SPI), hard negative information

**Citation** Zhang Hong-Bin, Qiu Die-Die, Wu Ren-Zhong, Jiang Zi-Liang, Wu Jin-Peng, Ji Dong-Hong. Stratified gene selection multi-feature fusion for image material attribute annotation. *Acta Automatica Sinica*, 2020, 46(10): 2191–2213

收稿日期 2019-01-04 录用日期 2019-04-30  
Manuscript received January 4, 2019; accepted April 30, 2019  
国家自然科学基金 (61762038, 61861016), 教育部人文社会科学研究一般项目 (20YJAZH1142), 江西省自然科学基金 (20202BABL202044), 江西省科技厅重点研发计划 (20171BBG70093, 20192BBE50071), 江西省教育厅科学技术研究项目 (GJJ190323) 资助

Supported by National Natural Science Foundation of China (61762038, 61861016), Humanity and Social Science Foundation of the Ministry of Education (20YJAZH1142), Natural Science Foundation of Jiangxi Province (20202BABL202044), Key Re-

search and Development Plan of Jiangxi Provincial Science and Technology Department (20171BBG70093, 20192BBE50071), and Science and Technology Projects of Jiangxi Provincial Department of Education (GJJ190323)

本文责任编辑 黎铭

Recommended by Associate Editor LI Ming

1. 华东交通大学软件学院 南昌 330013 2. 武汉大学国家网络安全学院 武汉 430072

1. Software School, East China Jiaotong University, Nanchang 330013 2. School of Cyber Science and Engineering, Wuhan University, Wuhan 430072

基于视觉特性识别图像材质属性是机器视觉领域的热点研究问题, 因为物体表面的材质类型可以为人们提供大量有价值的语义信息(如柔软性、防水性、透气性与耐磨性等), 这对于准确认知不同类型的材质具有非常重要的意义. 故图像材质属性标注的应用价值较大, 主要应用包括: 在淘宝、京东等主流电子商务平台上, 用户检索指定材质类别的商品; 机器人先识别物体的材质属性, 然后根据材质特点确定对物体的夹取力度、触碰强度及可否拉拽; 在工业领域中, 机器自动完成纺织原材料分类, 以减少人工识别误差并降低相应成本.

本文提出全新的分层基因优选多特征融合算法, 以准确、高效、便捷地完成图像材质属性标注, 并挖掘材质属性所蕴含的深层语义信息. 最后, 构建一个类似视觉基因计划<sup>[1]</sup>的材质属性表示(认知)体系, 为上述电子商务、机器人视觉、工业检测等领域的材质属性识别应用奠定基础. 本文的主要贡献: 提出基于分层基因优选多特征融合(Stratified gene selection multi-feature fusion, SGSMFF)的图像材质属性标注模型. 根据数据集中样本特性, 提取图像的传统特征(LBP<sup>[2]</sup>、Gist<sup>[3]</sup>、SIFT<sup>[4]</sup>)及深度学习特征(VGG16<sup>[5]</sup>); 采用任意分类模型计算特征预估概率; 改进 ERGS (Effective range based gene selection, ERGS) 算法<sup>[6]</sup>并在其中融入分层先验信息(Stratified priori information, SPI), 设计全新的分层基因优选(Stratified gene selection, SGS)算法, 基于 SGS 逐层、动态地为预估概率计算 ERGS 权重; 最后, 池化预估概率并做 ERGS 加权, 实现分层基因优选多特征融合. SGSMFF 的主要特点: 1) 层次性. 它融入分层先验信息, 以动态计算 ERGS 权重, 进而改善标注精准度; 2) 统一性. 它可采用任意分类模型计算特征预估概率; 3) 包容性. 它能实现多特征融合, 包括传统特征和深度学习特征; 4) 高效性. 实验数据表明: 它的标注精准度大幅优于若干有竞争力的基线.

因此, 本文工作既有一定实用价值, 又是对机器学习模型的创新.

## 1 相关工作

本文工作主要涉及三个方面: 图像属性预测、材质属性标注及特征融合. 因此, 分别对这三方面的工作进行综述, 具体如下:

### 1.1 图像属性预测

图像属性位于中间语义层, 它可以给人们提供丰富的语义信息. 近年, 图像属性预测吸引了机器视觉领域研究人员的重点关注. Farhadi 等<sup>[7]</sup>抽取图像特征, 训练二元 SVM 模型, 实现属性预测;

Kumar 等<sup>[8-9]</sup>构建鼻、眼、口等属性分类器, 以实现人脸识别及搜索; Jayaraman 等<sup>[10]</sup>基于属性预测完成零样本学习(Zero-shot learning); Berg 等<sup>[11]</sup>在噪声文本中挖掘刻画商品图像关键特性的属性; Gan 等<sup>[12]</sup>创建类别无关特征, 以更准确地完成属性预测; Liu 等<sup>[13]</sup>提出一种无监督的词典学习方法, 分析属性之间的相关性, 进而改善图像分类性能; Tang 等<sup>[14]</sup>提取共享于不同类别图像间的中间层特征, 以完成场景图像属性预测; 近年来, 多任务深度学习框架开始被用于图像属性预测<sup>[15-17]</sup>, 如人脸识别<sup>[15]</sup>、服饰识别<sup>[16]</sup>、场景理解<sup>[17]</sup>等. 上述工作<sup>[7-17]</sup>主要预测图像属性存在与否(非“0”即“1”, 这类属性也称二元属性). 然而, 有时人们更关注属性的程度大小, 这就需要采用相对属性模型<sup>[18]</sup>定量评估属性, 以更准确地认知属性. 相对属性由 Parikh 等<sup>[18]</sup>提出, 并首先应用于 WhittleSearch 系统<sup>[19-20]</sup>, 以完成高质量的人机检索交互. 相对属性预测有助于改善图像分类性能<sup>[19, 21]</sup>, 甚至完成零样本学习<sup>[22]</sup>. 多任务深度学习框架也可用于相对属性预测<sup>[23-24]</sup>, 如综合局部上下文和全局样式信息的相对属性预测<sup>[23]</sup>, 基于注意力机制与排序函数的端到端相对属性预测<sup>[24]</sup>等, 由于充分利用属性间的深层共享特征, 它们<sup>[23-24]</sup>均取得了较传统二元属性模型更优的预测性能. 最近, GAN 模型<sup>[25]</sup>被用来完成图像增强, 丰富训练样本, 并提升相对属性预测精度<sup>[26]</sup>.

### 1.2 材质属性标注

由于在电商平台、机器人视觉、工业检测等领域都具有非常广阔的应用前景, 材质属性标注已成为图像属性标注中的研究热点. 材质属性标注可被视为纹理分类问题<sup>[27-39]</sup>. 第一个来自真实世界的相关数据集是 CURET<sup>[27]</sup>. Leung 等<sup>[28]</sup>围绕 CURET 采用 48 组滤波器完成材质分类, 同时引入 3D Textons, 更好地刻画材质纹理特性; Varma 等<sup>[29]</sup>采用 38 组滤波器(也称 MR8)完成材质分类, 并聚焦于解决旋转不变性问题; Heera 等<sup>[30]</sup>运用  $3 \times 3$  的图像块替代滤波器组, 获得了更好的材质识别性能. 近年来, Sharan 等<sup>[31]</sup>创建了一个来自 Flickr 的材质属性数据集 FMD (包含 10 种材质, 每种材质 100 张图像). Liu 等<sup>[32]</sup>围绕 FMD, 运用贝叶斯网络获取了 45% 的分类精准度. Hu 等<sup>[33]</sup>采用核描述子模型将 FMD 数据集上的材质分类精准度提升到 54%. Sharan 等<sup>[34]</sup>基于感知特征将 FMD 上的识别精准度进一步提升到 57%. 不同于上述粗粒度数据集<sup>[27, 30-31]</sup>, Kampouris 等<sup>[35]</sup>首次提出细粒度材质数据集 Fabric, 他综合反射特性和微几何特性完成材质分类, 其识别精度达 79.6% (细粒度数据集中噪声较小).

上述实验<sup>[27-35]</sup>均为二维表面的纹理信息分析,与之相比,三维表面的纹理分析能更好地获取相关信息,因为三维外观伴随场景反射的照明和视角而变化. Dong 等<sup>[36]</sup>使用新的照明条件和视角来重新捕获纹理. Jian 等<sup>[37]</sup>提出快速自适应(Fast self-adaptive method, FSAM)和手动校正等方法,以校正非均匀照明并消除3D表面重建失真等问题. Jian 等<sup>[38]</sup>提出一种基于小波变换的方法来捕捉和融合粗糙表面纹理,该方法不仅可以捕获和融合任意光照方向下的三维表面纹理,还能保留表面几何特性. Jian 等<sup>[39]</sup>提出一种基于光照补偿和增强的照明不敏感纹理识别方法. 综合上述工作<sup>[27-39]</sup>可知:图像材质属性标注具有很大挑战,因为物体表面的视觉特性受形状、反射特性、光照、视角等因素影响.

本文在单类别图像特征提取基础上,首先运用 t-SNE 模型<sup>[40]</sup>对图像材质数据做可视化展示,以更好地认知材质属性标注问题. 实验采用的数据集是笔者构造的 MattrSet 及通用数据集 Fabric<sup>[35]</sup>, MattrSet 是跨越异类商品的较大数据集(详见第

3.1.1 节),它来自 Web 环境,存在噪声,是一种较粗粒度的数据集. Fabric 是对衣物材质直接采样获得的数据集(详见第 3.1.2 节),它来自服装商店,材质更精细,是一种较细粒度的数据集. 因此,本文选取它们来验证 SGSMMFF 算法的鲁棒性. 图像特征选取 SIFT<sup>[2]</sup>、Gist<sup>[3]</sup>、LBP<sup>[4]</sup>及 VGG16<sup>[5]</sup>,故基于 t-SNE 的数据可视化结果分别如图 1 和图 2 所示.

如图 1 所示,在 MattrSet 数据集中,特征不同,各材质属性的分布有很大差异. 总体而言, Pu 材质在 t-SNE 特征空间中具有更清晰的分布,而 Polyester 和 Nylon 材质的样本却存在大量重叠,即出现较严重的分类混淆. 如图 2 所示,在 Fabric 数据集中,特征不同,各材质属性的分布也存在很大差异. 但相对 MattrSet 数据集,其样本的聚合效果更好. 如在 SIFT 特征空间中, Wool、Denim、Cotton 等材质属性分布较好,这是缘于 Fabric 数据集从细粒度的角度刻画材质属性,而 MattrSet 数据集中的噪声信息更多.

综上,可以发现:单类别特征<sup>[2-5]</sup>识别材质属

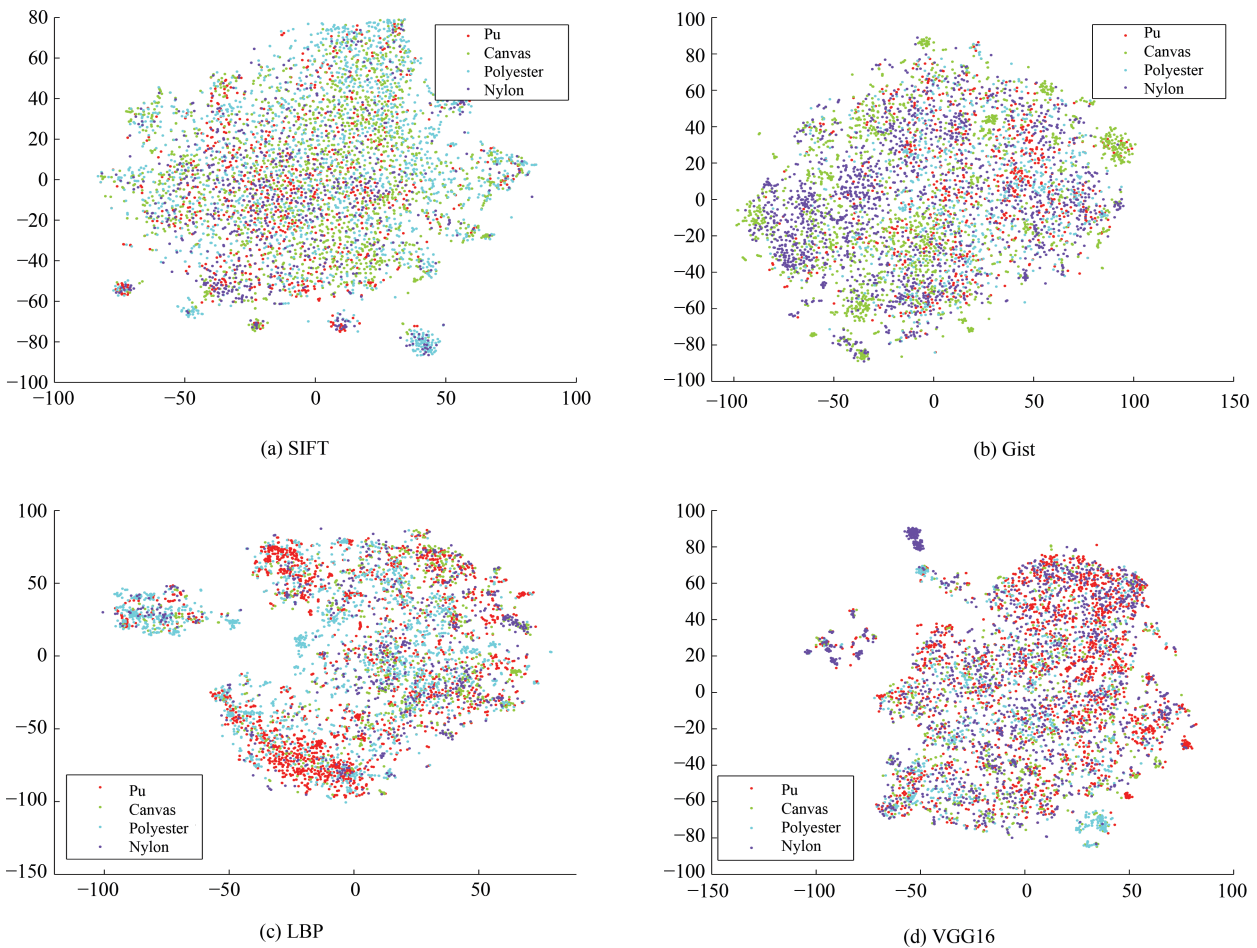


图 1 基于 t-SNE 的 MattrSet 数据集样本可视化结果(均为调参后的最佳效果)

Fig. 1 Visualization results of MattrSet dataset based on t-SNE (the best results are obtained after parameters tuning)

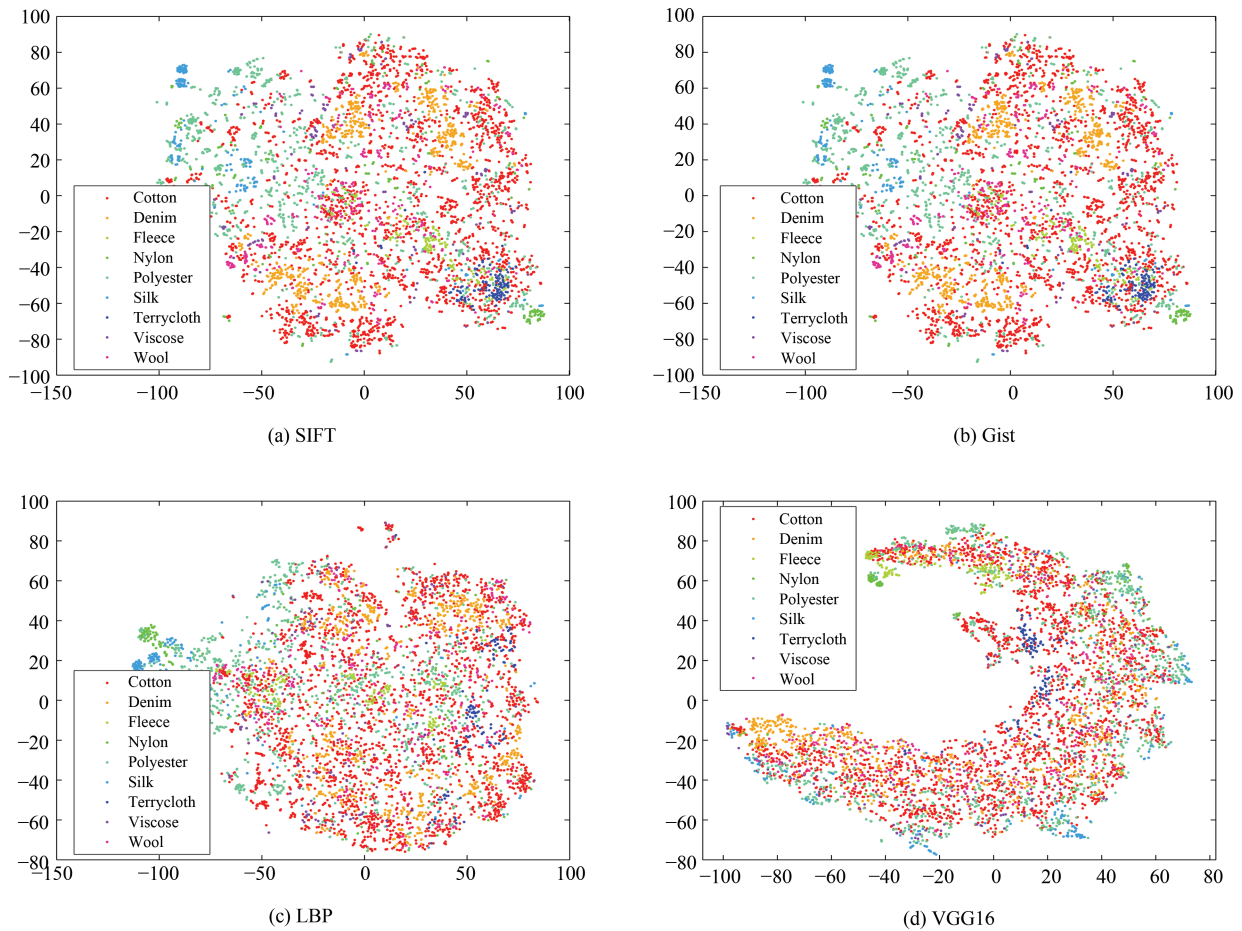


图2 基于 t-SNE 的 Fabric 数据集样本可视化结果 (均为调参后的最佳效果)

Fig. 2 Visualization results of Fabric dataset based on t-SNE (the best results are obtained after parameters tuning)

性的效果均不佳, 其中细粒度材质相对容易判别; 材质属性识别精度普遍偏低, 该研究任务具有一定挑战. 因此, 要完成高质量图像材质属性标注, 既要设计性能优越的分类模型, 又应充分利用不同图像特征之间的互补性.

### 1.3 特征融合

由于考虑不同特征对分类的影响, 特征融合一直以来都是机器学习领域的研究热点. 特征融合方法包括: 1) 提升策略. 它指将一组弱分类器集成为强分类器, 以获取更好的分类性能. 主流方法: 梯度提升决策树 (Gradient boosting decision tree, GBDT)<sup>[41]</sup>、Adaboost<sup>[42]</sup>、RankBoost<sup>[43]</sup>、极端梯度提升树 (eXtreme gradient boosting, XGBoost)<sup>[44]</sup>、分层梯度提升决策树 (Multi-layered gradient boosting decision trees, ML-GBDT)<sup>[45]</sup> 等; 2) 多特征融合. 它对特征加权以融合生成“新特征”, 基于“新特征”完成分类任务. 主流方法: 多核学习 (Multiple kernel learning, MKL)<sup>[46]</sup>、多核提升 (Multiple kernel boost, MKBoost)<sup>[47]</sup>、基于

多层感知机 (Multiple layer perception, MLP)<sup>[48]</sup> 的融合等; 3) 图模型. 它先构造多个异质子图以评估图像之间的内容相关性, 把这些子图融合为一个大图, 完成基于排序的大规模图像检索. 主流方法: 多重关联图 (Multiple affinity graphs, MAG)<sup>[49]</sup>、变分相关向量机 (Variational relevance vector machine, VRVM)<sup>[50]</sup>、图像图 (ImageGraph)<sup>[51]</sup> 等; 4) 搜索方法. 它指采用智能优化模型搜索出一组最优的特征组合. 主流方法: 有效区域基因优选 (Effective range based gene selection, ERGS)<sup>[6]</sup>、蝗虫优化算法 (Grasshopper optimization approaches, GOA)<sup>[52]</sup>、二元 Salp 群优化算法 (Binary salp swarm algorithm, BSSA)<sup>[53]</sup>、灰狼优化算法 (Gray-wolf optimization, GWO)<sup>[54]</sup> 等. 第 (1) 类方法<sup>[41-45]</sup> 属于决策层融合, 它更多依赖分类模型, 未考虑特征间的互补性; 第 (2) 类方法<sup>[46-48]</sup> 属于特征层融合, 它依赖完整的核矩阵运算, 且“新特征”的判别性有限; 第 (3) 类方法<sup>[49-51]</sup> 主要用于图像检索而非分类, 但图模型的生成效率偏低; 第 (4) 类方法<sup>[6, 52-54]</sup> 也称特征选择 (Feature selec-

tion), 它专注于优选判别性最强的特征, 而未考虑特征之间的互补性, 且多在封闭数据集上开展实验.

综合上述三个方面的相关研究, 可以发现: 图像属性预测是机器视觉领域的热点问题, 而材质属性标注隶属该研究范畴, 它在工业、商业、机器人等领域都具有重要的应用价值, 但其研究存在挑战. 单一特征及传统特征融合方法均不能较好地完成材质属性标注, 有必要设计一种新算法, 该算法无需构造“新特征”, 却能充分利用不同特征之间的互补性, 以及分类模型的决策能力. 即综合第 (1)、(2) 类特征融合算法 (提升策略与多特征融合) 的优点, 完成高质量的材质属性标注, 这正是设计分层基因优选多特征融合 (SGSMFF) 算法的原始动机.

### 2 基于分层基因优选多特征融合算法的图像材质属性标注模型

#### 2.1 模型框架

基于分层基因优选多特征融合 (SGSMFF) 算法的图像材质属性标注模型如图 3 所示 (图 3 采用

MattrSet 数据集进行说明, Fabric 数据集与之类似).

如图 3 所示: SGSMFF 算法先采用基本分类模型计算图像特征的预估概率 (由于采用 7 类基本分类模型, 且 MattrSet 数据集包含 4 种材质, 故预估概率是  $7 \times 4$  大小的矩阵. 若采用 Fabric 数据集, 则预估概率是  $7 \times 9$  大小. 颜色越深, 概率值越大); 改进 ERGS 算法并融入分层先验信息 SPI (各层中“1”表示待标注材质属性的正例,“0”表示对应负例. 例如, 层 1 中包含 4 种材质, 其中三种材质被作为负例; 层 3 中包含两种材质, 仅有一种材质被作为负例. SPI 的描述详见第 2.2.1 节), 逐层且动态地为预估概率计算 ERGS 权重 (该权重是  $3 \times 7 \times 4$  大小的张量. 若采用 Fabric 数据集, 则预估概率是  $8 \times 7 \times 4$  大小. 颜色越深, 权重越大); 池化 (平均值和最大值两种) 预估概率, 并对池化结果执行 ERGS 加权, 即基于 SUM 函数实现融合; 最后, 基于 MAX 决策函数计算融合之后样本的最大隶属度, 完成材质属性标注.

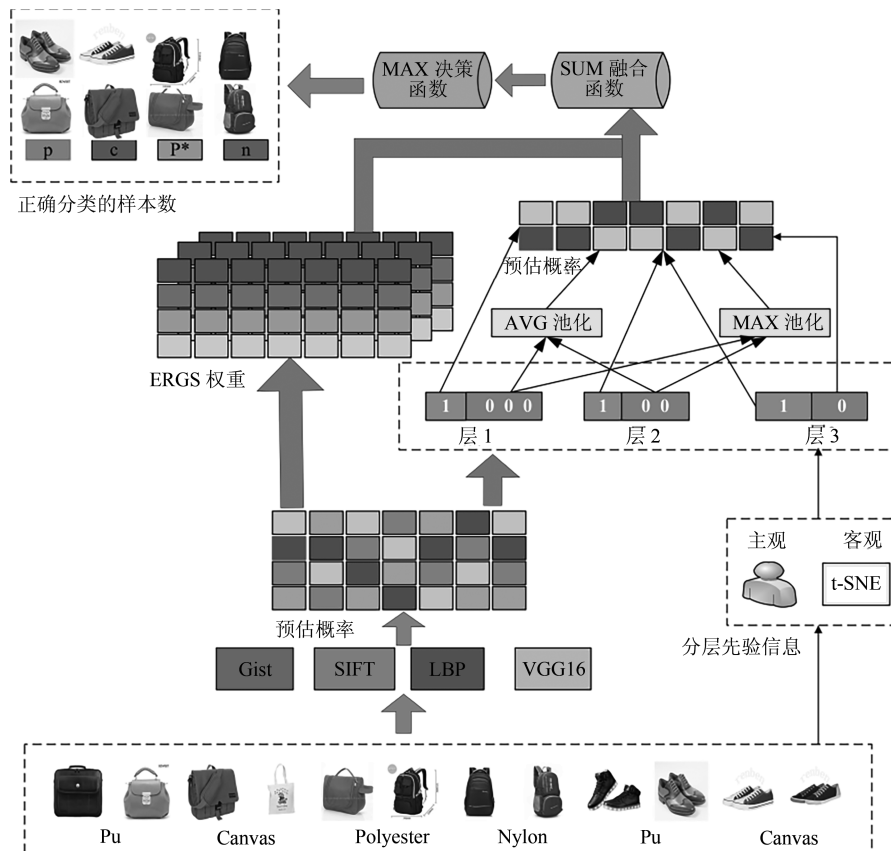


图 3 基于 SGSMFF 算法的材质属性标注模型 (以 MattrSet 数据集为例, p 表示 Pu 材质、p\* 表示 Polyester 材质、c 表示 Canvas 材质、n 表示 Nylon 材质. “1” 表示正例, “0” 表示负例)

Fig. 3 The proposed material attribute annotation model based on the SGSMFF algorithm (MattrSet is used as example,

p, p\*, c, n, “1”, and “0” represent Pu, Polyester, Canvas, Nylon, positive, and negative examples, respectively)

在本框架中, 将提取 LBP<sup>[2]</sup>、Gist<sup>[3]</sup>、SIFT<sup>[4]</sup> 及深度学习特征 (VGG16<sup>[5]</sup>), 这与研究内容有密切关系, 即本文聚焦图像材质属性识别问题, 物体表面的材质属性与图像的纹理、形态、视角等有关, 故有针对性地选取相关特征: LBP 即局部二进制模型, 它是一种描述局部纹理特征的算子, 而材质属性多表现为局部图像区域内的规律性纹理变化, 故 LBP 特征能较好地捕获这些变化; Gist 是一种基于多尺度与多方向 Gabor 滤波的特征, 它着力刻画全局纹理, 即图像中的简单上下文信息, 故 Gist 也能较好地刻画材质属性, 但相对 LBP, 它易受背景噪声干扰; SIFT 即尺度不变特征变换, 它主要描述图像中的局部形态特性, 且可以减少因光照、视角、形态等变化带来的识别误差, 上述变化通常被认为是干扰材质属性识别的关键性因素, 故 SIFT 能胜任材质属性标注; VGG16 是基于深度学习模型的非线性特征, 它经过多次卷积、池化而来, 能较好地刻画图像中的深层语义, 且它的生成方式不同于传统特征, 是对多特征融合的有利补充. 当然, 本文主要关注 SGSMMFF 算法的设计, 未来拟引入其他特征做融合, 以进一步改善材质属性标注性能. 本文第 2.2.1 节、第 2.2.2 节、第 2.2.3 节分别介绍“获取分层先验信息 SPI”、“改进的 ERGS 算法”及“SGS 算法”.

## 2.2 SGSMMFF 算法

### 2.2.1 获取分层先验信息 SPI

本文讨论图像材质属性标注问题, 如图 3 所示, 可先获取分层先验信息 SPI, 以指导各层 ERGS 权重的动态更新, 从而完成高质量的多特征融合, 最终提升材质属性标注性能. 用  $SPI_M$  表示 MattrSet 数据集的分层先验信息, 用  $SPI_F$  表示 Fabric 数据集的分层先验信息. 综合“主观 + 客观”的方法来获取材质属性标注的分层先验信息. 主观法包括: 1) 专家知识; 2) 历史数据; 3) 基于 t-SNE<sup>[33]</sup> 散点图的观察法. 客观法: 基于 t-SNE<sup>[33]</sup> 代价值法.

针对 MattrSet 数据集: 首先, 根据材料学专家知识获取  $SPI_M$ . 由材料学领域知识可知: 各类材质表面的纹理、光滑度、透明度、反射率等都有显著差异, 故它们的标注难易度不同. Pu 材质柔软性较好, 透气性较高, 类内纹理变化非常稳定, 较容易识别; Polyester 材质弹性最好, 强度和耐磨性较好, 类内纹理变化很大, 它包含粗颗粒和细颗粒两种, 相应的反射率、折射率存在差异, 较难识别; Canvas 材质是一种较粗厚的棉织物或麻织物, 它结实耐用、防水性好, 且多采用平纹组织 (少量为斜纹组织), 类内纹理变化较稳定, 较易识别. Nylon 材质强度高、耐磨性好, 类内纹理变化较稳定, 识别难度适中. 总体而言, Pu 最易识别, Polyester 最难识别, 其他两

类材质介于它们之间. 因此, 根据材料学专家知识的分层先验信息  $SPI_{M-1}$ : “Pu → Canvas → Nylon → Polyester” 及  $SPI_{M-2}$ : “Pu → Nylon → Canvas → Polyester” ( $SPI_{M-2}$  表示: 先标注 Pu, 然后标注 Nylon, 最后标注 Canvas 和 Polyester, 即微调 Canvas 和 Nylon 的顺序. 后续描述与之类似, 不再赘述).

其次, 基于历史数据获取  $SPI_{M-3}$ . 在执行 SGSMMFF 算法前, 采用了 7 类基本模型、4 种基本特征完成图像材质属性标注, 以形成历史数据 (参见第 3.2.1 节). 历史数据显示: 标注 Pu、Canvas、Nylon、Polyester 的均值平均准确率 (Mean average precision, MAP) 分别是 63.75%、58.57%、39.71%、37.45%. 该数据表明: Pu 材质最易识别, 而 Polyester 材质最难识别. 因此, 基于历史数据的分层先验信息  $SPI_{M-3}$ : “Pu → Canvas → Nylon → Polyester”, 它与  $SPI_{M-1}$  完全一致.

再次, 基于 t-SNE 散点图做样本可视化, 获取  $SPI_{M-4}$ . 4 种图像特征的散点图如图 1 所示. 在图 1 中, Pu 材质能较好地聚合在一起, Canvas 材质次之, 而 Nylon 和 Polyester 两类材质分布在广阔的特征空间内, 聚合程度最差. 显然, 仅采用单类别图像特征, 材质属性标注的效果较差 (参考上面的均值平均准确率), 需要对不同特征进行融合. 参考图 1, 基于 t-SNE 散点图的分层先验信息  $SPI_{M-4}$ : “Pu → Canvas → Nylon → Polyester”, 它与  $SPI_{M-1}$  完全一致.

最后, 基于客观法获取分层先验信息  $SPI_{M-5}$ . 采用 4 种基本图像特征<sup>[2-5]</sup> 分别完成基于 t-SNE 的样本可视化工作, 分解 t-SNE 模型输出的总代价值中针对不同材质属性的代价值分量, 计算出如表 1 所示各类材质的代价均值, 以确定分层先验信息  $SPI_{M-5}$ . 因此, 基于 t-SNE 代价均值 (越小越好) 得到的分层先验信息  $SPI_{M-5}$ : “Pu → Nylon → Polyester → Canvas”, 由于最后一层仅有两类材质, 故  $SPI_{M-2}$  与  $SPI_{M-5}$  完全一致.

综合  $SPI_{M-1}$ 、 $SPI_{M-2}$ 、 $SPI_{M-3}$ 、 $SPI_{M-4}$  及  $SPI_{M-5}$ , 可归纳出 MattrSet 数据集中图像材质属性标注的分层先验信息: “Pu → Canvas → Nylon → Polyester”、“Pu → Nylon → Canvas → Polyester”.  $SPI_M$  聚焦于先划分 Pu、Canvas 等易分材质, 然后再划分 Nylon、Polyester 等难分材质, 即由易到难完成材质属性标注. 此外, 设计其他几种不同的分层先验信息:  $SPI_{M-6}$ 、 $SPI_{M-7}$ 、 $SPI_{M-8}$ 、 $SPI_{M-9}$ , 与上述  $SPI_M$  做比较, 它们的具体描述如第 3.2.3 节表 6 所示.

针对 Fabric 数据集, 从两个角度获取分层先

表 1 各类材质的 t-SNE 代价值 (针对不同数据集, 每列最小值如 **1.3079** 等所示)Table 1 The t-SNE cost value of different material (for different dataset, the minimum value of each column is shown as **1.3079** etc.)

数据集	材质	t-SNE 代价值				
		Gist	SIFT	LBP	VGG	代价均值
MattrSet	Pu	<b>1.3079</b>	<b>1.4615</b>	<b>0.7735</b>	<b>0.9142</b>	<b>1.1143</b>
	Canvas	1.4517	1.7962	0.8653	0.9660	1.2698
	Nylon	1.4077	1.7227	0.8333	0.9360	1.2249
	Polyester	1.3948	1.7285	0.8318	0.9982	1.2383
	Cotton	1.0282	1.2109	1.2102	0.8974	1.0867
	Denim	0.4405	0.9569	0.5581	0.4354	0.5977
	Fleece	0.2267	0.5844	<b>0.1583</b>	0.1219	0.2728
Fabrc	Nylon	0.2219	<b>0.1730</b>	0.2105	0.1480	<b>0.1884</b>
	Polyester	0.7151	0.9591	0.7243	0.5471	0.7364
	Silk	<b>0.1852</b>	0.3642	0.1944	0.2078	0.2379
	Terrycloth	0.2441	0.4616	0.3116	0.1457	0.2907
	Viscose	0.2319	0.5017	0.2818	<b>0.1035</b>	0.2797
	Wool	0.4072	0.6868	0.4417	0.2565	0.4480

验信息, 即基于历史数据 MAP 和表 1 中不同材质的代价均值, 最终分别得到 Fabric 数据集的分层先验信息  $SPI_F-1$  与  $SPI_F-2$ ,  $SPI_F-1$ : “Silk  $\rightarrow$  Denim  $\rightarrow$  Terrycloth  $\rightarrow$  Cotton  $\rightarrow$  Viscose  $\rightarrow$  Polyester  $\rightarrow$  Fleece  $\rightarrow$  Nylon  $\rightarrow$  Wool” (Cotton、Denim、Fleece、Nylon、Polyester、Silk、Terrycloth、Viscose、Wool 材质的 MAP 值分别为 67.44%、74.20%、63.43%、62.23%、63.54%、76.79%、69.35%、66.07%、58.50%, 基于 MAP 做降序排列),  $SPI_F-2$ : “Nylon  $\rightarrow$  Silk  $\rightarrow$  Fleece  $\rightarrow$  Viscose  $\rightarrow$  Terrycloth  $\rightarrow$  Wool  $\rightarrow$  Denim  $\rightarrow$  Polyester  $\rightarrow$  Cotton”。

综上, 采用主观、客观两种方法可以得到材质属性标注的分层先验信息 SPI, 第 2.2.3 节的 SGS 算法将融入分层先验信息 SPI, 完成高质量的材质属性标注。

## 2.2.2 改进的 ERGS 算法

原 ERGS 算法<sup>[6]</sup> 仅完成特征选择, 即选取 ERGS 权值最大的特征, 而放弃其他特征, 这必然会丢失一些重要的判别信息, 进而影响材质属性标注性能. 本文改进原 ERGS<sup>[6]</sup> 为特征融合算法, 设计新的决策函数: 根据 ERGS 评估结果为所有特征都分配权值, 以加权分类模型输出的预估概率, 完成多特征融合. 因此, 材质属性标注由全部特征共同决定, 这符合人类认知客观世界的一般规律: 综合纹理、形态等多类视觉感知来识别外部世界. 特征  $feat_z \in F$ ,  $z \in \{1, \dots, n\}$ ,  $n$  是特征总数. 特征  $feat_z$  在  $c_q$  类样本上的有效区域  $ER_{zq}$  被定义为

式 (1), 对于 MattrSet 数据集  $q \in \{1, \dots, 4\}$ , 对于 Fabric 数据集  $q \in \{1, \dots, 9\}$ :

$$ER_{zq} = [r_{zq}^-, r_{zq}^+] = [\mu_{zq} - (1 - p_q) \gamma \sigma_{zq}, u_{zq} + (1 - p_q) \gamma \sigma_{zq}] \quad (1)$$

$r_{zq}^-$  和  $r_{zq}^+$  分别表示特征  $feat_z$  在  $c_q$  类样本上有效区域的上边界和下边界;  $u_{zq}$  和  $\sigma_{zq}$  分别表示特征  $feat_z$  在  $c_q$  类样本上的平均值和标准偏差;  $p_q$  为  $c_q$  类样本的先验概率 ( $1 > p_q > 0$ ), 因子  $1 - p_q$  可降低标准偏差  $\sigma_{zq}$  对特征有效区域上下边界影响;  $\gamma$  是通过切比雪夫不等式推导出的一个常量, 即 1.732.

$$P(|F - \mu_{ik}| \geq \gamma \sigma_{ik}) \leq \frac{1}{\gamma^2} \quad (2)$$

然后, 计算特征  $feat_z$  的重叠区域  $OA_z$ :

$$OA_z = \sum_{j=1}^{l-1} \sum_{k=j+1}^l \varphi_z(j, k) \quad (3)$$

$$\varphi_z(j, q) = \begin{cases} r_{zq}^+ - r_{zq}^-, & r_{zq}^+ > r_{zq}^- \\ 0, & \text{其他} \end{cases} \quad (4)$$

基于重叠区域  $OA_z$  计算重叠区域系数  $AC_z$ :

$$AC_z = \frac{OA_z}{\max_q(r_{zq}^+) - \min_q(r_{zq}^-)} \quad (5)$$

改进 ERGS 算法, 即根据重叠区域系数  $AC_z$  计

算特征  $feat_z$  的 ERGS 权重  $w_z$ , 如式 (6) 所示:

$$w_z = \frac{AC_z}{\text{sum}\{AC_j : 1 \leq j \leq n\}} \quad (6)$$

设计新的决策函数, 即采用 ERGS 权重加权分类模型输出的预估概率, 取其最大值生成特征融合结果, 决策函数如式 (7) 所示. 式 (7) 仅实现最基本的特征融合, 在第 2.2.3 节的 SGS 算法中将对它做进一步优化, 详见第 2.2.3 节内容.

$$\max(p_z) = \max\{\text{sum}(p_{zj} \times w_z : 1 \leq j \leq n)\} \quad (7)$$

综上, 改进后的有效区域基因优选算法如下:

**算法 1. 改进后的有效区域基因优选算法**

**输入.** 图像特征:  $F = \{feat_1, feat_2, \dots, feat_n\}$ ,  $feat_z \subseteq \mathbf{R}^k$ ,  $z \in \{1, \dots, n\}$ .

**输出.** 图像特征的 ERGS 权重矩阵  $W$ .

- 1) 初始化  $W = 0$ ;
- 2) **repeat**;
- 3) 如式 (1) 所示, 计算  $feat_z$  在  $c_q$  类样本上的有效区域  $ER_{zq}$ ;
- 4) 如式 (3) 所示, 计算  $feat_z$  重叠区域  $OA_z$ ;
- 5) 如式 (5) 所示, 计算  $feat_z$  重叠区域系数  $AC_z$ ;
- 6) 如式 (6) 所示, 计算  $feat_z$  的 ERGS 权重  $w_z$ ;
- 7) 把 ERGS 权重  $w_z$  保存到  $W$  中第  $z$  列;
- 8) **until** 处理完全部特征;
- 9) 输出图像特征的 ERGS 权重矩阵  $W$ .

### 2.2.3 SGS 算法

首先, 给出 SGS 算法中的主要参数描述, 具体如表 2 所示.

如图 3 所示, SGS 算法先抽取图像特征, 包括 SIFT、LBP、Gist 等传统特征及 VGG16 深度学习特征; 采用基本分类模型计算图像特征的预估概率; 然后, 引入分层先验信息确定材质标注顺序, 即将待标注材质的样本作为正例, 其他样本作为负例, 执行多轮二元分类. 在分类过程中, 改进传统的 ERGS 算法<sup>[6]</sup>, 逐层并动态地为预估概率计算 ERGS 权重 (参见第 2.2.2 节); 最后, 结合分层特性进一步优化式 (7): 对预估概率做平均值池化或最大值池化, 同时执行 ERGS 加权, 完成基于 SGSMFF 算法的图像材质属性标注. 对于 MattrSet 数据集, 在第 1、2、3 层分类时, 基于平均值池化策略的 ERGS 决策函数如式 (8)~(10) 所示. 此外, 对于 MattrSet 数据集, 基于最大值池化策略的 ERGS 决策函数如式 (11)~(13) 所示. Fabric 数据集与之相似, 不再赘述.

$$\max(p_{i1}) = \max\left\{\text{sum}\left(\text{avg}(p_{ij}, p_{ij+1}, p_{ij+2}) \times w_{1j} : 1 \leq j \leq cc\right)\right\} \quad (8)$$

$$\max(p_{i2}) = \max\left\{\text{sum}\left(\text{avg}(p_{ij}, p_{ij+1}) \times w_{2j} : 1 \leq j \leq cc\right)\right\} \quad (9)$$

$$\max(p_{i3}) = \max\left\{\text{sum}\left(\text{avg}(p_{ij} \times w_{3j}) : 1 \leq j \leq cc\right)\right\} \quad (10)$$

表 2 SGS 算法中的参数设置

Table 2 parameter settings of the proposed SGS algorithm

参数	$T$	$CM$	$y_i$	$F$	$C$
意义	图像数据集	分类模型集合	图像样本标签	图像特征集合	材质属性标签集合
参数值	$\left\{ \begin{matrix} (x_1, y_1), \dots, \\ (x_N, y_N) \end{matrix} \right\}$	$\left\{ \begin{matrix} Classifier_1, \dots, \\ Classifier_{cm} \end{matrix} \right\}$	$y_i \subseteq C$	$\{feat_1, feat_2, \dots, feat_n\}$ $feat_z \subseteq \mathbf{R}^k, z \in \{1, \dots, n\}$	$C = \{c_1, c_2, \dots, c_l\}$
参数	$D$	$x_i$	$l$	$N$	$n$
意义	特征组合集合	图像样本	材质属性标签数	样本总数	特征总数
参数值	$\{feat\_comb_1, \dots, feat\_comb_d\}$ $feat\_comb_i = \{feat_1, \dots, feat_{cc}\}$	\	MattrSet: 4 Fabric: 9	MattrSet: 11021 Fabric: 5064	4
参数	$k$	$d$	$cc$	$cm$	\
意义	特征 $feat_z$ 的维度	特征组合总数	组合中的特征数	分类模型数量	\
参数值	Gist: 512 LBP: 1180 SIFT: 800 VGG16: 1000	11	单类别的特征: 1 两种特征融合: 2 三种特征融合: 3 4 种特征融合: 4	7	\



$$\max(p_{i1}) = \max \left\{ \sum \left( \max(p_{ij}, p_{ij+1}, p_{ij+2}) \times w_{1j} : 1 \leq j \leq cc \right) \right\} \quad (11)$$

$$\max(p_{i2}) = \max \left\{ \sum \left( \max(p_{ij}, p_{ij+1}) \times w_{2j} : 1 \leq j \leq cc \right) \right\} \quad (12)$$

$$\max(p_{i3}) = \max \left\{ \sum \left( \max(p_{ij} \times w_{3j}) : 1 \leq j \leq cc \right) \right\} \quad (13)$$

其中,  $p_{ij}, p_{ij+1}, p_{ij+2}$  均为基本分类模型输出的预估概率,  $w_{1j}, w_{2j}, w_{3j}$  分别为第 1、2、3 层的 ERGS 权重. SGS 算法如算法 2 所示 (以 MattrSet 数据集为例, Fabric 数据集与之类似, 仅分层数不一样):

#### 算法 2. 分层基因优选多特征融合算法

**输入.**  $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$ , 即图像数据.

**输出.** 图像的材质属性.

1) 抽取图像特征: SIFT、LBP、Gist 及 VGG16, 得到  $F = \{feat_1, feat_2, \dots, feat_n\}$ ;

2) 采用基本分类模型计算图像特征的预估概率  $p_{ij}$ ;

3) 对  $F$  中的特征进行组合, 得到  $D = \{feat\_comb_1, feat\_comb_2, \dots, feat\_comb_d\}$ , 其中  $feat\_comb_i = \{feat_1, feat_2, \dots, feat_{cc}\}$ ;

4) 采用第 2.2.1 节方法获取分层先验信息 SPI;

5) **repeat**;

6) 执行改进的 ERGS 算法 (算法 1) 计算每类特征的第 1 层 ERGS 权重  $w_{1j}$ ;

7) 基于式 (8) 或式 (11), 用 ERGS 权重对预估概率  $p_{ij}$  加权, 完成第 1 层分类;

8) 执行改进的 ERGS 算法 (算法 1) 计算每类特征的第 2 层 ERGS 权重  $w_{2j}$ ;

9) 基于式 (9) 或式 (12), 用 ERGS 权重对特征预估概率  $p_{ij}$  加权, 完成第 2 层分类;

10) 执行改进的 ERGS 算法 (算法 1) 计算每类特征的第 3 层 ERGS 权重  $w_{3j}$ ;

11) 基于式 (10) 或式 (13), 用 ERGS 权重对特征预估概率  $p_{ij}$  加权, 完成第 3 层分类;

12) **until** 处理完  $D$  中全部特征组合;

13) 基于多特征融合结果, 输出图像的材质属性.

## 3 实验结果及分析

### 3.1 数据集、基线及度量标准

#### 3.1.1 MattrSet 数据集

现有材质数据集的数据量偏少, 不利于标

注模型更全面地学习鲁棒的视觉特性, 且它们也未跨越异类商品图像, 因此, 笔者从网站 <http://www.made-in-china.com/> 抓取真实商品图像, 在材料学专家指导下, 花费 2 人月对数据去噪、筛选、分类, 构建全新的材质属性数据集 MattrSet, 它有 11 021 张图像 (共 1.33 GB 数据, 数据集已公开, 可从网址 [https://drive.google.com/open?id=12xXX\\_MuwII8hghwXFLtT3sneEzgA4-](https://drive.google.com/open?id=12xXX_MuwII8hghwXFLtT3sneEzgA4-) SN 上获取), MattrSet 包含 Bags 和 Shoes 两类商品及 Pu (皮革)、Canvas (帆布)、Nylon (尼龙)、Polyester (涤纶) 4 种材质. Bags 类 7 409 张图像 (Pu 属性 1 982 张、Canvas 属性 1 948 张、Nylon 属性 1 764 张、Polyester 属性 1 715 张); Shoes 类 3 612 张图像, Shoes 类中 Nylon、Polyester 这两种材质非常稀缺, 它只包含 Pu 和 Canvas 材质, 数量分别为 1 757 张和 1 855 张. MattrSet 完全来自 Web, 噪声较多, 故它是一个粗粒度数据集. 实验随机选取 50% 的数据作为训练集, 剩余 50% 的数据作为测试集.

#### 3.1.2 Fabric 数据集

Fabric<sup>[35]</sup> 数据集为衣服材质属性数据集. Kampouris 等<sup>[35]</sup> 使用光度立体传感器 (分辨率  $640 \times 480$ ) 和笔记本电脑, 访问实体商店收集了 2 000 多件衣服和织物表面的样本, 并现场确定标签上的服装成分, 构建出 Fabric, 它是一个细粒度数据集. Fabric 数据集有 1 266 个样本, 根据不同光照、视角等变化, 每个样本又包含 4 张图像, 因此, Fabric 共有 5 064 张图像, 分辨率大小为  $400 \times 400$ , 包含 9 种材质属性, 分别是 Cotton 属性 2 352 张、Wool 属性 360 张、Terrycloth 属性 120 张、Fleece 属性 132 张、Nylon 属性 200 张、Silk 属性 200 张、Denim 属性 648 张、Viscose 属性 148 张和 Polyester 属性 904 张 (共 1.7 GB 数据, 该数据集也已公开, 下载地址: <http://ibug.doc.ic.ac.uk/resources/fabrics>). Fabric 是来自现实世界中的织物分布, 样本分布不平衡, 且 Cotton 材质和 Polyester 材质居多. 实验随机选取 50% 的数据作为训练集, 剩余 50% 的数据作为测试集.

#### 3.1.3 基线模型

本文设计的分层基因优选算法有: 分层基因优选逻辑回归 (Stratified gene selection logistic regression, SGS-LR)、分层基因优选随机森林 (Stratified gene selection random forest, SGS-RF)、分层基因优选 K 近邻 (Stratified gene selection K-nearest neighbor, SGS-KNN)、分层基因优选决策树 (Stratified gene selection decision tree, SGS-DT)、分层基因优选朴素贝叶斯 (Stratified gene

selection Naive Bayes, SGS-NB)、分层基因优选梯度提升决策树 (Stratified gene selection gradient boosting decision tree, SGS-GBDT) 与分层基因优选极端梯度提升树 (Stratified gene selection eXtreme gradient boosting, SGS-XGBoost). 对比基线有 5 大类, 其中前两类基线合称基本模型. 基线的具体描述如下:

1) 简单模型: 逻辑回归<sup>[55]</sup> (Logistic regression, LR)、随机森林<sup>[56]</sup> (Random forest, RF)、K 近邻<sup>[57]</sup> (K-nearest neighbor, KNN)、决策树<sup>[58]</sup> (Decision tree, DT)、朴素贝叶斯<sup>[59]</sup> (Naive Bayes, NB), 它们均在 scikit-learn<sup>[60]</sup> 中实现;

2) 融合类模型: 传统 ERGS<sup>[6]</sup>、GBDT<sup>[41]</sup> (用 200 个弱分类器)、XGBoost<sup>[44]</sup> (用 100 个弱分类器)<sup>[2]</sup>;

3) 深度学习模型: VGG<sup>[5]</sup>、InceptionResNetV2<sup>[61]</sup>、Densenet169<sup>[62]</sup> 及 MobileNets<sup>[63]</sup>, 它们是目前主流的深度学习模型, 本文迁移 ImageNet 实现预训练, 然后微调参数, 完成材质属性分类;

4) GS 类模型: 基于算法 1 与基本分类模型而来: 基因优选逻辑回归 (Gene selection logistic regression, GS-LR)、基因优选随机森林 (Gene selection random forest, GS-RF)、基因优选 K 近邻 (Gene selection K-nearest neighbor, GS-KNN)、基因优选决策树 (Gene selection decision tree, GS-DT)、基因优选朴素贝叶斯 (Gene selection Naive Bayes, GS-NB)、基因优选梯度提升决策树 (Gene selection gradient boosting decision tree, GS-GBDT) 及基因优选极端梯度提升树<sup>[64]</sup> (Gene selection eXtreme gradient boosting, GS-XGBoost);

5) 其他模型: Farhadi<sup>[7]</sup> 的 2 元 SVM 模型、Adaboost<sup>[42]</sup>、Kampouris<sup>[35]</sup> 的工作.

### 3.1.4 度量标准

采用精准度刻画模型的综合标注性能, 以 MatrrSet 为例, 其计算公式如式 (14) 所示:

$$\text{Accuracy} = \frac{TP_{\text{all}} + TN_{\text{all}}}{TP_{\text{all}} + FP_{\text{all}} + TN_{\text{all}} + FN_{\text{all}}} \quad (14)$$

“ $TP_{\text{all}} + TN_{\text{all}}$ ”是正确标注的图像总数,  $TP_{\text{all}}$  是正确标注的正例, 即属于材质 material 的图像数, 而  $TN_{\text{all}}$  是正确标注的负例, 即不属于材质 material 的图像数, “ $TP_{\text{all}} + FP_{\text{all}} + TN_{\text{all}} + FN_{\text{all}}$ ”是全部图像总数.  $TP_{\text{all}}$ 、 $FP_{\text{all}}$ 、 $TN_{\text{all}}$ 、 $FN_{\text{all}}$  的计算公式

如下所示:

$$\begin{aligned} TP_{\text{all}} &= TP_{\text{Pu}} + TP_{\text{Canvas}} + \\ &\quad TP_{\text{Polyester}} + TP_{\text{Nylon}} \\ TN_{\text{all}} &= TN_{\text{Pu}} + TN_{\text{Canvas}} + \\ &\quad TN_{\text{Polyester}} + TN_{\text{Nylon}} \\ FP_{\text{all}} &= FP_{\text{Pu}} + FP_{\text{Canvas}} + \\ &\quad FP_{\text{Polyester}} + FP_{\text{Nylon}} \\ FN_{\text{all}} &= FN_{\text{Pu}} + FN_{\text{Canvas}} + \\ &\quad FN_{\text{Polyester}} + FN_{\text{Nylon}} \end{aligned} \quad (15)$$

设计评估不同分层先验方案的指标: 均值平均精度 (Mean average accuracy, MAA), 如式 (16) 所示.

$\text{feat} \in \{\text{LBP}, \text{Gist}, \text{SIFT}, \text{VGG16}, \text{feat\_comb}\}$ , 各类特征组合 feat\_comb 详见表 4 (共计 11 种, 即  $N_{\text{feat}} = d = 11$ ), model 指所有 GS 类模型 (参考表 4),  $N_{\text{model}} = 7$ .  $\text{Accuracy}_{\text{SGS\_AVG}}$  表示基于 SGS 算法 (平均值池化) 的精准度,  $\text{Accuracy}_{\text{SGS\_MAX}}$  表示基于 SGS 算法 (最大值池化) 的精准度. MAA 指标评判针对全部模型及全部特征, 平均值池化方法相对最大值池化方法的均值提升幅度, 它是一个综合性指标, 可以客观、全面地评估模型标注性能.

## 3.2 实验结果及分析

第 3.2.1 节叙述基本分类模型的 Accuracy; 第 3.2.2 节叙述 GS 类模型 (第 4 类基线) 相对基本分类模型的 Accuracy 提升幅度; 第 3.2.3 节叙述 SGS 类模型 (最大值池化) 相对 GS 类模型的 Accuracy 提升幅度; 第 3.2.4 节叙述 SGS 类模型 (平均值池化) 相对 SGS 类模型 (最大值池化) 的 Accuracy 提升幅度; 第 3.2.5 对比全部模型 Accuracy 并做模型简化测试.

### 3.2.1 基本分类模型 Accuracy

表 3 是采用 7 种基本分类模型获取的材质属性标注精准度, 由式 (14) 计算. 其中, “L”、“G”、“S”、“V” 分别表示 LBP、Gist、SIFT、VGG16 这 4 种特征 (下同), “S + G” 表示 SIFT 特征与 Gist 特征做融合, “S + G + L” 表示 SIFT、Gist、LBP 这 3 种特征做融合, 其他特征组合命名以此类推.

如表 3 所示: 在 MatrrSet 数据集中, 1) “L” 特征的整体标注性能最佳 (在 7 个模型中取得 5 个最优), “G” 特征次之 (在 7 个模型中取得一个最

$$\text{MAA} = \frac{\sum_{n=1}^{N_{\text{model}}} \sum_{i=1}^{N_{\text{feat}}} (\text{Accuracy}_{\text{SGS\_AVG}}(i)(n) - \text{Accuracy}_{\text{SGS\_MAX}}(i)(n))}{N_{\text{model}} \times N_{\text{feat}}} \quad (16)$$

表 3 基本分类模型的标注精准度 (各数据集每列最优值如 **45.24** 等表示, 单位: %)Table 3 The accuracy of basic model (for each dataset, the optimal value of each column is expressed as **45.24**, etc, unit: %)

数据集	特征	基本分类模型的 Accuracy						
		DT	GBDT	KNN	LR	NB	RF	XGBoost
MatrSet	L	43.67	<b>61.28</b>	<b>56.33</b>	<b>55.84</b>	40.86	<b>59.76</b>	<b>62.93</b>
	S	34.40	52.41	43.73	50.48	<b>48.36</b>	47.96	52.90
	G	<b>45.24</b>	60.34	56.05	49.19	43.27	59.32	61.99
	V	42.11	52.17	45.45	35.52	34.51	53.55	54.64
Fabric	L	47.31	70.02	68.17	69.15	27.17	62.40	70.38
	S	<b>67.10</b>	<b>79.66</b>	37.84	<b>80.85</b>	<b>56.60</b>	<b>75.79</b>	<b>82.03</b>
	G	51.90	70.70	<b>71.37</b>	55.63	51.80	68.33	73.42
	V	49.45	65.01	57.46	58.10	46.88	64.34	66.59

优、5 个次优), 这缘于“L”和“G”特征着力刻画图像局部形态及全局纹理变化, 局部形态及全局纹理是区分不同材质属性的关键视觉特性. 而“V”特征表现不佳, 这缘于“V”特征是在 ImageNet 数据集上预训练而来, 通过分析 ImageNet, 笔者发现: 它并未对材质属性分类, 这便导致 VGG 模型在预训练后存在偏置 (Bias), 即忽略材质属性识别; 2) XGBoost 模型的整体标注性能最佳, 因为 XGBoost 模型集成若干弱分类器为一个强分类器, 该分类器判别能力较强, 能有效地识别不同材质属性. 当选“L”特征时, XGBoost 模型标注精准度最优, 达到 62.93%, 它较次优值 (GBDT 模型选取“L”特征时) 提升  $62.93\% - 61.28\% = 1.65\%$ . 此外, 传统 ERGS 算法的性能为 54.94% (基于 SIFT 特征, 选取 RBF 核). 在 Fabric 数据集中, 1) “S”特征的整体标注性能最佳 (7 个模型中取得 6 个最优), “G”特征次之 (7 个模型中取得一个最优、5 个次优), 这缘于“S”和“G”特征对图像旋转、缩放、仿射和尺度空间等具有一定不变性. 而“V”特征同样表现不佳, 原因同上; 2) XGBoost 模型的整体标注性能最佳, 原因同上. 当选取“S”特征时, XGBoost 模型的标注性能最优, 到达 82.03%, 它较次优值 (LR 模型选取“S”特征时) 提升  $82.03\% - 80.85\% = 1.18\%$ . 此外, 传统 ERGS 算法的性能为 81.87% (基于 SIFT 特征且选取 Linear 核).

综上, 表 3 仅是单类别特征的材质属性标注结果, 未考虑特征之间的关系, 故精准度高低多依赖于分类模型, 材质属性标注的 Accuracy 还有较大提升空间.

### 3.2.2 GS 类模型相对基本分类模型的 Accuracy 变化

表 4 刻画在 MatrSet 数据集上, GS 类模型精准度相对表 3 结果的变化值  $\Delta Accuracy_{M1}$  与  $\Delta Accuracy_{M2}$ ,  $\Delta Accuracy_{M1}$  或  $\Delta Accuracy_{M2}$

为正值表示 GS 类模型优于基本分类模型.  $\Delta Accuracy_{M1} = \text{avg}(\text{Accuracy}_{GS}) - \text{avg}(\text{Accuracy}_{Basic})$ , 它表示模型平均精准度差值. 表 3 中, 在 MatrSet 数据集上, DT 模型平均精准度是 41.35%, 表 4 中, GS-DT 模型平均精准度 42.80%, 故相应  $\Delta Accuracy_{M1} = 1.45\%$ .  $\Delta Accuracy_{M2} = \text{max}(\text{Accuracy}_{GS}) - \text{max}(\text{Accuracy}_{Basic})$ , 它表示模型最大精准度差值. 表 3 中, 在 MatrSet 数据集上, DT 模型最大精准度 45.24%, 表 4 中, GS-DT 模型最大精准度 46.20%, 故相应  $\Delta Accuracy_{M2} = 0.96\%$ .  $\Delta Accuracy_{M1}$ 、 $\Delta Accuracy_{M2}$  为正表示 Accuracy 有提升. 同理可得到表 5. 它刻画在 Fabric 数据集上, GS 类模型精准度相对表 3 结果的变化值  $\Delta Accuracy_{F1}$  与  $\Delta Accuracy_{F2}$  (注: 本节不采用 SPI).

如表 4 所示: 1) “S + G + L”特征组合的材质属性标注性能最佳, 这说明三类基本特征之间具有较强互补性, 即“G”特征着力刻画纹理视觉特性, “L”、“S”特征着力减少形态、光照、旋转等变化对材质属性标注的干扰. 这一互补性有助于改善标注精准度; 其他组合如“S + G”、“L + G”也有不错性能; 2) “V”特征开始扮演重要角色. 选取合适的基本模型, “S + V”特征组合的标注性能 (选 GS-XGBoost 算法, 为 56.54%) 优于特征“S” (表 3 中最优值 52.90%) 或“V” (表 3 中最优值 54.64%). 这说明: “S + V”组合不但可以准确识别图像纹理, 且对形状变化有较强“鲁棒性”; 3) GS-XGBoost 模型整体标注性能最佳, 当选“S + G + L”特征组合时, 其标注精准度达 67.67%, 相比表 3 最优值提升  $67.67\% - 62.93\% = 4.74\%$ ; 4) 无论是  $\Delta Accuracy_{M1}$  还是  $\Delta Accuracy_{M2}$ , 它们均大于 0, 这说明: 改进的 ERGS 算法 (算法 1) 的有效性, 即它确实能融合多个互补的图像特征, 以改善标注

表 4 MattrSet 数据集上, GS 类模型精准度及相对基本模型的 Accuracy 变化 (每列最优值如 **46.20** 等表示, 单位: %)   
 Table 4 The accuracies of GS models and the corresponding accuracy variations compared to the basic models in the MattrSet dataset (The optimal value of each column is expressed as **46.20**, etc., unit: %)

特征	GS 类模型的 Accuracy						
	GS-DT	GS-GBDT	GS-KNN	GS-LR	GS-NB	GS-RF	GS-XGBoost
S + G + L + V	42.33	59.83	54.89	54.57	49.12	59.14	61.23
S + G + L	<b>46.20</b>	<b>65.13</b>	<b>62.10</b>	58.41	44.87	61.70	<b>67.67</b>
S + G + V	42.40	57.67	51.52	50.72	49.25	57.47	59.27
S + L + V	42.35	57.45	51.33	53.04	39.23	57.29	58.94
L + G + V	42.42	59.09	54.37	52.59	41.66	58.98	59.87
S + G	45.31	62.48	58.32	53.22	<b>49.61</b>	60.14	64.39
S + V	42.49	53.98	41.76	46.36	47.85	55.13	56.54
S + L	37.12	62.95	57.56	<b>59.45</b>	46.05	58.92	63.75
L + G	45.65	63.79	60.90	57.34	40.88	<b>61.75</b>	65.08
L + V	42.26	56.20	50.19	50.43	40.37	56.40	57.52
G + V	42.28	56.40	50.10	47.00	43.49	56.69	57.63
$\Delta\text{Accuracy}_{M1}$	1.45	2.99	3.52	5.25	3.01	3.36	2.97
$\Delta\text{Accuracy}_{M2}$	0.96	3.85	5.77	3.61	1.25	1.99	4.74

表 5 Fabric 数据集上, GS 类模型精准度及其相对基本模型的 Accuracy 变化 (每列最优值如 **79.98** 等表示, 单位: %)   
 Table 5 The accuracies of GS models and the corresponding accuracy variations compared to the basic models in the Fabric dataset (The optimal value of each column is expressed as **79.98**, etc., unit: %)

特征	GS 类模型的 Accuracy						
	GS-DT	GS-GBDT	GS-KNN	GS-LR	GS-NB	GS-RF	GS-XGBoost
S + G + L + V	58.93	73.46	68.84	68.29	45.06	69.19	78.75
S + G + L	65.64	75.25	62.56	74.12	39.69	69.67	80.57
S + G + V	49.45	71.96	66.90	66.00	48.82	70.85	72.71
S + L + V	47.95	72.12	67.58	69.08	43.84	68.48	72.71
L + G + V	47.43	72.43	71.09	65.17	42.54	68.17	72.24
S + G	60.35	<b>79.98</b>	57.66	77.76	<b>57.42</b>	<b>76.58</b>	<b>81.95</b>
S + V	49.45	69.91	59.28	66.94	48.54	70.02	70.62
S + L	<b>64.69</b>	76.26	41.00	<b>78.28</b>	31.87	71.17	81.71
L + G	48.74	71.64	<b>73.74</b>	66.71	27.29	66.00	73.54
L + V	47.47	69.31	66.63	65.44	42.54	64.34	70.38
G + V	49.45	67.73	63.78	59.64	47.47	67.58	69.04
$\Delta\text{Accuracy}_{F1}$	-0.34	1.38	4.84	2.92	-2.42	1.56	1.82
$\Delta\text{Accuracy}_{F2}$	-1.46	0.32	2.37	-2.57	0.82	0.79	-0.08

精准度, 这为设计 SGS 算法 (算法 2) 奠定了重要基础. 此外, GS-LR、GS-KNN 等模型性能提升显著, 即引入多特征融合思想之后, 原线性类模型也能获得不错的分类结果.

如表 5 所示: 1) “S + G” 特征组合的材质属性标注性能最佳 (7 个模型中取得 4 个最优), 它表明: 这两类特征之间具有较强互补性, 即 “G” 特征着力刻画视觉纹理特性, “S” 特征着力减少形态、光照等因素变化对材质属性标注的干扰, 这一互补性有助于改善标注精准度. 该结果不同于表 4, 即不同数据集应选用不同的特征组合; 2) “V” 特征与传统特征间的互补性较弱; 3) GS-

XGBoost 模型的整体标注性能最佳, 选 “S + G” 特征组合时, 其标注精准度达 81.95%, 相比表 3 最优值下降  $82.03\% - 81.95\% = 0.08\%$ ; 4) 无论是  $\Delta\text{Accuracy}_{F1}$  还是  $\Delta\text{Accuracy}_{F2}$ , 表 5 中多数结果均大于 0, 这说明: 改进的 ERGS 算法对细粒度数据集 Fabric 同样有效, 但提升效果略低于 MattrSet 数据集.

### 3.2.3 SGS 类模型 (采用最大值池化) 相对 GS 类模型的 Accuracy 变化

表 6 表示: 在 MattrSet 数据集上, 采用 SGS 类模型 (最大值池化) 后, 标注精准度相对表 4 的变化值  $\Delta\text{Accuracy}_{M3}$ .  $\Delta\text{Accuracy}_{M3} =$

表 6 MattrSet 数据集中 SGS\_MAX 类模型相对 GS 类模型的 Accuracy 变化值  $\Delta\text{Accuracy}_{M3}$   
(性能衰减如  $-0.20$  所示, 单位: %)

Table 6 The accuracy variations of the SGS\_MAX model compared to the GS model in the MattrSet dataset:  
 $\Delta\text{Accuracy}_{M3}$  (The performance degradation indicators are marked in  $-0.20$ , unit: %)

SPI <sub>M</sub>	SGS 模型	$\Delta\text{Accuracy}_{M3} = \text{Accuracy}_{\text{SGS\_MAX}} - \text{Accuracy}_{\text{GS}}$											Avg <sub>model</sub>
		all	S + G	L S + G + V	S + L + V	L + G + V	S + G S + V	S + L L + G	L + V G + V				
pcp*n SPI <sub>M-1</sub>	DT	14.27	10.02	14.11	14.19	14.21	12.98	14.00	14.56	13.58	14.85	14.68	13.77
	GBDT	12.81	12.51	13.29	13.71	12.30	13.23	14.69	13.92	11.29	13.27	13.29	13.12
	KNN	12.45	8.93	13.01	13.16	12.34	10.60	20.50	10.46	9.84	13.83	14.23	12.67
	LR	22.96	17.03	24.25	24.29	23.83	20.06	27.81	16.04	14.89	25.06	26.72	22.09
	NB	16.21	16.61	20.44	22.00	23.92	11.49	21.65	12.18	16.17	16.83	20.53	18.00
	RF	14.81	16.84	14.94	15.17	13.83	16.77	15.08	19.21	14.89	14.16	14.34	15.46
pnp*c SPI <sub>M-2</sub>	XGBoost	12.00	10.15	12.47	12.76	12.34	11.34	12.95	13.83	10.87	12.89	12.89	12.23
	DT	14.89	10.05	14.03	14.28	15.63	11.32	12.09	15.30	13.98	15.16	14.57	13.75
	GBDT	9.71	7.75	10.76	11.00	10.00	9.20	12.40	8.22	9.18	11.45	11.53	10.11
	KNN	9.62	6.15	11.34	11.35	10.40	8.41	18.47	8.13	9.19	12.02	12.54	10.69
	LR	5.75	7.54	8.29	7.03	7.06	8.86	11.11	9.33	9.73	7.95	10.89	8.50
	NB	6.15	15.10	3.65	15.43	13.87	7.41	3.85	12.98	13.43	14.09	4.47	10.04
p*ncp SPI <sub>M-6</sub>	RF	8.22	8.04	9.33	9.35	9.02	8.09	10.87	7.01	9.28	10.76	10.44	9.13
	XGBoost	9.00	6.48	10.31	10.33	10.54	8.64	10.79	8.46	9.62	11.28	11.40	9.71
	DT	10.98	7.08	10.71	10.58	10.62	8.95	8.64	11.06	9.99	10.20	10.05	9.90
	GBDT	3.97	3.03	4.08	4.41	4.64	3.18	4.78	2.38	5.10	5.52	5.30	4.22
	KNN	3.31	2.83	4.48	5.18	4.13	5.27	10.92	5.46	5.10	6.23	5.54	5.31
	LR	<b>-0.20</b>	0.89	<b>0.00</b>	0.33	<b>-0.40</b>	0.71	<b>0.00</b>	1.83	2.43	<b>-0.20</b>	0.01	0.49
p*pcn SPI <sub>M-7</sub>	NB	3.59	13.89	2.14	8.17	2.43	10.76	4.12	13.74	10.62	2.65	<b>0.00</b>	6.56
	RF	2.27	1.51	2.81	2.78	3.34	1.43	3.57	0.85	3.27	4.32	4.21	2.76
	XGBoost	4.66	2.69	4.81	5.45	5.77	3.63	6.08	3.12	5.24	6.58	6.39	4.95
	DT	15.50	9.71	15.31	15.65	15.05	13.76	14.73	15.88	14.85	15.90	15.92	14.75
	GBDT	8.79	7.08	9.22	9.99	9.76	7.75	16.25	8.19	8.82	11.20	10.60	9.79
	KNN	8.68	5.50	10.18	10.22	9.73	8.28	17.25	6.83	8.11	12.05	12.09	9.90
cpp*n SPI <sub>M-8</sub>	LR	6.42	7.79	8.47	7.35	7.77	9.48	11.15	8.15	9.79	8.67	11.10	8.74
	NB	10.76	17.81	6.39	17.91	13.36	11.22	8.39	14.47	17.15	15.70	4.50	12.51
	RF	7.79	6.30	8.49	9.09	8.62	6.61	9.78	6.57	7.57	10.24	9.66	8.25
	XGBoost	9.02	6.55	9.60	10.22	9.92	7.49	10.57	7.99	8.25	11.19	10.91	9.25
	DT	14.32	11.16	14.25	14.52	14.29	13.34	14.27	15.09	13.11	14.23	15.10	13.97
	GBDT	13.89	12.09	15.30	15.07	13.72	13.61	16.07	13.82	11.00	14.38	15.05	14.00
npcp* SPI <sub>M-9</sub>	KNN	13.31	9.96	14.19	14.69	13.48	10.42	19.92	11.69	10.35	14.40	15.70	13.46
	LR	22.94	17.77	23.78	23.36	22.81	21.22	26.76	16.41	15.60	23.61	24.94	21.75
	NB	18.71	19.44	22.45	24.18	22.96	15.25	23.32	15.78	16.99	24.22	35.93	21.75
	RF	16.08	16.44	16.45	16.43	15.21	16.33	16.72	19.10	15.00	15.61	15.86	16.29
	XGBoost	12.89	9.60	13.77	13.72	13.03	11.76	13.89	13.81	10.87	13.55	13.89	12.80
	DT	15.30	12.20	14.67	14.70	14.90	13.25	13.94	19.79	14.12	14.86	14.99	14.79
Avg <sub>Feat</sub>	GBDT	9.33	6.75	10.47	10.89	9.80	8.60	12.33	7.24	8.82	11.02	11.38	9.69
	KNN	9.71	6.13	11.40	11.33	10.25	8.17	18.31	8.84	8.80	11.71	12.82	10.68
	LR	6.06	6.57	8.31	7.32	4.08	8.61	11.15	8.04	8.75	8.40	10.98	8.02
	NB	6.59	14.25	4.19	17.84	15.01	7.81	4.86	10.22	14.05	14.10	4.58	10.32
	RF	7.90	6.64	9.31	9.27	8.64	7.59	10.60	6.85	9.03	10.22	10.18	8.75
	XGBoost	8.71	5.83	10.31	10.40	10.22	8.09	11.00	7.66	9.16	11.04	11.15	9.42
	Avg <sub>Feat</sub>	10.48	9.54	11.09	12.26	11.49	10.02	12.99	10.73	10.66	12.27	12.18	/

$\text{Accuracy}_{\text{SGS\_MAX}} - \text{Accuracy}_{\text{GS}}$ ,  $\Delta\text{Accuracy}_{\text{M3}}$  为正值表示 SGS 类模型优于 GS 类模型. 分层先验信息  $\text{SPI}_{\text{M-1}}$ 、 $\text{SPI}_{\text{M-3}}$  和  $\text{SPI}_{\text{M-4}}$  相同: “Pu  $\rightarrow$  Canvas  $\rightarrow$  Nylon  $\rightarrow$  Polyester”, 它被缩写为  $\text{pcp}^*\text{n}$  ( $\text{p}^*$  即 Polyester,  $\text{n}$  即 Nylon,  $\text{c}$  即 Canvas,  $\text{p}$  即 Pu). 其他 5 种 SPI 方案表示为:  $\text{p}^*\text{ncp}$  (由最难到最易标注材质, 即  $\text{SPI}_{\text{M-6}}$ )、 $\text{cpp}^*\text{n}$  (微调  $\text{p}$  位置, 即  $\text{SPI}_{\text{M-8}}$ )、 $\text{npcp}^*$  ( $\text{n}$  前移, 即  $\text{SPI}_{\text{M-9}}$ )、 $\text{p}^*\text{pcn}$  ( $\text{p}^*$  前移, 即  $\text{SPI}_{\text{M-7}}$ ) 和  $\text{pnp}^*\text{c}$  (即  $\text{SPI}_{\text{M-2}}$ 、 $\text{SPI}_{\text{M-5}}$ ), 字符含义同  $\text{pcp}^*\text{n}$ . 表 6 最后一列 ( $\text{Avg}_{\text{model}}$ ) 是不同模型提升幅度均值, 表 6 最后一行 ( $\text{Avg}_{\text{Feat}}$ ) 是不同特征提升幅度均值. 同理, 可得 Fabric 数据集的实验结果, 如表 7 所示. 表 7 将计算  $\Delta\text{Accuracy}_{\text{F3}}$ ,  $\Delta\text{Accuracy}_{\text{F3}}$  为正值表示 SGS 类模型优于 GS 类模型. 对于 Fabric 数据集, 根据历史数据和 t-SNE 代价均值分别获取分层先验信息  $\text{SPI}_{\text{F-1}}$  和  $\text{SPI}_{\text{F-2}}$ ,  $\text{SPI}_{\text{F-1}}$  为  $\text{sdctcvpfnw}$ , 分层顺序: “Silk $\rightarrow$ Denim $\rightarrow$ Terrycloth $\rightarrow$ Cotton $\rightarrow$ Viscose $\rightarrow$ Polyester $\rightarrow$ Fleece $\rightarrow$ Nylon $\rightarrow$ Wool”. 同理, 可得  $\text{SPI}_{\text{F-2}}$  为  $\text{nsfvtwdpc}$ .

如表 6 所示: 1) 除极少数指标 (选取  $\text{p}^*\text{ncp}$  方案时得到的负数结果) 出现衰减之外, 绝大多数  $\Delta\text{Accuracy}_{\text{M3}}$  指标为正值, 这充分说明 SGS 算法的有效性, 即分层先验信息 SPI 在图像材质属性标注中发挥了重要作用. 相反, 先难后易 ( $\text{p}^*\text{ncp}$  方案) 的识别顺序不利于大幅提升标注精准度, 尤其是选

取 LR 模型时, 部分性能出现衰减. 因此, 可得出结论: 选择  $\text{p}^*\text{ncp}$  方案时,  $\text{SPI}_{\text{M}}$  对线性类模型帮助不大.

2) 基于 MAA 指标可得到:  $\text{cpp}^*\text{n}$  ( $\text{SPI}_{\text{M-8}}$ ) 方案平均提升 16.30%, 提升幅度最大;  $\text{pcp}^*\text{n}$  ( $\text{SPI}_{\text{M-1}}$ 、 $\text{SPI}_{\text{M-3}}$ 、 $\text{SPI}_{\text{M-4}}$ )、 $\text{p}^*\text{pcn}$  ( $\text{SPI}_{\text{M-7}}$ )、 $\text{pnp}^*\text{c}$  ( $\text{SPI}_{\text{M-2}}$ 、 $\text{SPI}_{\text{M-5}}$ )、 $\text{npcp}^*$  ( $\text{SPI}_{\text{M-9}}$ ) 等方案分别平均提升 15.35%、10.47%、10.29%、10.25%, 而  $\text{p}^*\text{ncp}$  ( $\text{SPI}_{\text{M-6}}$ ) 提升幅度最小. 这说明: 材质属性分类顺序会影响标注性能, 先识别 Pu 或 Canvas (微调  $\text{p}$  或  $\text{c}$ ), 标注精准度大幅提升, 即这两种材质相对 Polyester 和 Nylon 具有更鲜明的纹理特性, SPI 为分类模型提供了非常有效的难分样本信息<sup>[65]</sup>. 因此, SGS 算法充分利用不同特征间的互补性 (GS 思想) 及 SPI 所提供的难分样本信息, 大幅改善材质属性标注性能.  $\text{SPI}_{\text{M-1}}$  (或  $\text{SPI}_{\text{M-3}}$  或  $\text{SPI}_{\text{M-4}}$ ) 方案次优, 即材料学专家的知识也发挥重要作用, 这符合人类的客观认知.  $\text{SPI}_{\text{M-2}}$  (或  $\text{SPI}_{\text{M-5}}$ ) 方案的提升幅度有限, 即前移  $\text{n}$  后性能提升不显著.  $\text{p}^*$  前移后,  $\text{p}^*\text{ncp}$  性能变化也不显著,  $\text{SPI}_{\text{M}}$  没有给分类模型提供有效的难分样本信息 ( $\text{n}$ 、 $\text{c}$ 、 $\text{p}$  混合时, 视觉特征混淆度较大, 参见图 1).

3) 计算各分类模型的平均提升幅度, SGS-DT、SGS-GBDT、SGS-KNN、SGS-LR、SGS-NB、SGS-RF、SGS-XGBoost 等模型分别平均提升 13.49%、10.16%、10.45%、11.60%、13.20%、

表 7 Fabric 数据集中 SGS\_MAX 类模型相对 GS 类模型的 Accuracy 变化值  $\Delta\text{Accuracy}_{\text{F3}}$  (性能衰减如 **-3.47** 所示, 单位: %)

Table 7 The accuracy variations of the SGS\_MAX model compared to the GS model in the Fabric dataset:  $\Delta\text{Accuracy}_{\text{F3}}$  (The performance degradation indicators are marked in **-3.47**, unit: %)

$\text{SPI}_{\text{F}}$	SGS 模型	$\Delta\text{Accuracy}_{\text{F3}} = \text{Accuracy}_{\text{SGS\_MAX}} - \text{Accuracy}_{\text{GS}}$											
		all	S + G	S + L	S + G + L	S + G + V	S + L + V	L + G + V	S + G	S + V	S + L	L + G	L + V
	DT	8.01	6.32	16.62	18.40	13.00	5.80	36.41	9.24	39.14	11.22	10.86	15.91
	GBDT	15.36	15.31	16.35	15.99	13.98	11.25	16.78	14.02	15.80	15.92	17.06	15.26
	KNN	6.79	0.59	8.34	0.63	13.82	1.98	6.30	3.19	12.08	12.36	17.97	7.64
$\text{SPI}_{\text{F-1}}$	LR	18.87	15.97	20.18	19.07	18.36	11.73	20.54	12.79	15.83	18.37	20.85	17.51
	NB	33.47	24.49	29.70	30.65	27.92	15.41	28.71	35.55	23.46	27.68	27.57	27.69
	RF	18.33	18.72	16.55	17.74	16.86	12.99	15.45	16.55	19.19	18.60	16.82	17.07
	XGBoost	11.26	11.25	16.86	15.88	15.48	10.27	15.24	10.00	15.64	15.32	17.18	14.03
$\text{SPI}_{\text{F-2}}$	DT	7.74	5.57	15.56	18.24	15.88	8.05	17.18	9.36	12.16	14.10	12.24	12.37
	GBDT	0.87	3.23	1.70	1.73	<b>-1.18</b>	1.10	4.46	4.86	1.31	0.91	2.77	1.98
	KNN	6.24	19.00	5.61	3.63	2.33	22.55	9.91	39.45	5.64	2.37	7.19	11.27
	LR	0.94	4.12	0.82	2.05	<b>-0.52</b>	2.02	4.78	4.42	1.73	1.28	1.77	2.13
	NB	6.32	29.15	1.02	6.75	6.43	19.99	3.79	45.58	32.86	6.59	1.27	14.52
	RF	<b>-17.81</b>	5.09	0.79	2.22	0.43	1.38	2.61	4.15	1.65	2.96	1.38	0.44
	XGBoost	<b>-3.47</b>	<b>-0.16</b>	2.09	1.30	0.75	1.42	4.22	0.83	1.66	1.22	3.31	1.20
$\text{Avg}_{\text{Feat}}$	8.07	11.33	10.87	11.02	10.25	9.00	13.31	15.00	14.15	10.64	11.30	/	

10.11%、9.73%。SGS-DT 模型平均提升幅度最大, SGS-NB 模型次之, SGS-XGBoost 模型最差。显然,  $SPI_M$  对于线性类模型帮助更大。可大胆推测:  $SPI_M$  是刻画样本分布的非线性信息(难分样本), 它对线性模型是一种极好的补充。

4) 根据  $Avg_{Feat}$  值可知: “S + V” 特征组合的平均提升幅度最显著(12.99%), “G + V” 特征组合次之, “S + G + L” 组合平均提升幅度最小(9.54%)。这说明: 在 SGS 类模型中, 深度学习特征“V”较传统特征“S”、“G”、“L”等能更好地拟合  $SPI_M$ , 从而改善材质属性标注精准度。由于“V”是一种非线性深层特征, 故这从另一个侧面验证了  $SPI_M$  是一种有效的非线性判别信息。

5) 评估各方案中的最优值。在  $p^*ncp$  方案中, 选取“S + G + L”特征组合及 SGS-XGBoost 模型时标注性能最优, 达 70.36%; 在  $cpp^*n$  方案中, 选取“S + G + L”特征组合及 SGS-RF 模型时标注性能最优, 达 78.14%; 在  $npcp^*$  方案中, 选取“L + G”特征组合及 SGS-XGBoost 模型时标注性能最优, 达 74.24%; 在  $pnp^*c$  方案中, 选取“L + G”特征组合及 SGS-XGBoost 模型时标注性能最优, 达 74.70%; 在  $pcp^*n$  方案中, 选取“S + G + L”特征组合及 SGS-RF 模型时, 标注性能最优, 达到 78.54%, 它是表 6 中的最优值, 相比表 4 最优值提升  $78.54\% - 67.67\% = 10.87\%$ , 性能提升非常显著。

如表 7 所示: 1) 除极少数指标出现衰减之外, 绝大多数  $\Delta Accuracy_{F3}$  为正值, 这进一步说明 SGS 算法的鲁棒性, 即分层先验信息  $SPI$  在细粒度数据集的标注中也发挥了重要作用; 2) 基于 MAA 指标可以得到,  $SPI_{F-1}$  方案优于  $SPI_{F-2}$  方案,  $SPI_{F-1}$  的平均提升幅度最大, 达到 16.44%, 这说明基于历史数据的  $SPI$  对材质属性标注的影响更大, 且标注性能与识别顺序也息息相关; 3) 计算各分类模型的平均提升幅度, SGS-DT、SGS-GBDT、SGS-KNN、SGS-LR、SGS-NB、SGS-RF、SGS-XGBoost 等模型分别平均提升 14.14%、8.62%、9.46%、9.82%、21.11%、8.76%、7.62%。SGS-NB 模型平均提升幅度最大、SGS-DT 模型次之, SGS-XGBoost 模型最差。显然, 无论是细粒度数据集 Fabric, 还是粗粒度数据集 MattrSet,  $SPI$  对于线性类模型的帮助更大。4) 由  $Avg_{Feat}$  值可知: 特征组合“S + L”的平均提升幅度最显著(15.00%), “L + G”次之, “S + G + L + V”最小(8.07%), 这说明: 在 SGS 类模型中, 细粒度数据集的深度学习特征拟合  $SPI_F$  的效果一般; 5) 评估各方案最优值, 在  $SPI_{F-1}$  和  $SPI_{F-2}$  方案中, 选取“S + G”特征组合及 SGS-XGBoost 模型时标注性能均

为最优, 分别为 92.22% 和 83.37%,  $SPI_{F-1}$  方案的最优值相比表 5 中最优值提升  $92.22\% - 81.95\% = 10.27\%$ , 性能提升也非常显著。

### 3.2.4 SGS 类模型(采用平均值池化)相对最大值池化的 Accuracy 变化

表 8 表示: 在 MattrSet 数据集中, 采用 SGS 类模型(平均值池化)后, 材质属性标注精准度相对表 6 的变化值  $\Delta Accuracy_{M4}$ , 令  $\Delta Accuracy_{M4} = Accuracy_{SGS\_AVG} - Accuracy_{SGS\_MAX}$ , 若  $\Delta Accuracy_{M4}$  为正值则表明平均值池化优于最大值池化。由于实验结果非常多, 仅取表 6 中最佳的两种  $SPI_M$  方案  $pcp^*n$  ( $SPI_M-1$ ) 和  $cpp^*n$  ( $SPI_M-8$ ) 来展示实验结果。同理可得表 9, 它表示: 在 Fabric 数据集中, 采用 SGS 类模型(平均值池化)后, 材质属性标注精准度相对表 7 的变化值  $\Delta Accuracy_{F4}$ , 计算方法同  $\Delta Accuracy_{M4}$ 。

如表 8 所示: 1) 除 1 个  $\Delta Accuracy_{M4}$  指标衰减外, 其他均为正值。这说明在 SGS 类模型中, 平均值池化优于最大值池化(参见式(8)~(10)), 这与深度学习模型中池化的使用吻合; 2) 根据  $Avg_{model}$  值可知: SGS-NB、SGS-KNN 两类模型的平均提升幅度最显著。相反, SGS-XGBoost 模型的平均提升幅度最小。这进一步说明引入  $SPI_M$  后, 模型之间的标注精准度差异在缩小, 即使是线性类模型(KNN 或 NB)也能获取较满意的标注性能; 3) 根据  $Avg_{Feat}$  值可知: 特征组合“L + V”的平均提升幅度最显著(10.01%), “All”次之(9.57%), “L + G”最小(5.94%)。显然, 引入  $SPI_M$  之后, “V”特征的作用越来越大; 4) 基于式(16)计算 MAA 指标,  $pcp^*n$  (8.61%) 略优于  $cpp^*n$  (8.07%), 即选取平均值池化, 且遵从由简入难识别顺序的  $SPI_M-1$  方案较优。5) 评估各方案中最优值。在  $pcp^*n$  方案 ( $SPI_M-1$ ) 中, 选取“S + V”组合及 GS-NB 模型时标注性能最优, 达 86.37%; 在  $cpp^*n$  方案中, 选取“S + V”特征组合及 SGS-NB 模型时标注性能最优, 达 86.37%。该最优值相比表 6 最优值提升  $86.37\% - 78.54\% = 7.83\%$ 。此外, “V”特征的作用愈加显著, 而线性类模型 SGS-LR、SGS-NB 在材质属性分类中发挥了关键作用。

如表 9 所示: 1) 除 2 个  $\Delta Accuracy_{F4}$  指标衰减外, 其他均为正值。这说明在 SGS 类模型中, 平均值池化优于最大值池化; 2) 根据  $Avg_{model}$  值可知: SGS-KNN 提升幅度最为显著, 为 14.99%, SGS-GBDT 提升幅度最小, 为 6.63%, 这进一步说明引入  $SPI_F$  后, 即使是线性类模型(KNN)也能获取较满意的标注性能; 3) 根据  $Avg_{Feat}$  值可知: “All”特征组合的平均提升幅度最显著(13.01%), “S + L +

表 8 MattrSet 数据集中, SGS\_AVG 类模型相对 SGS\_MAX 类模型 Accuracy 变化值  $\Delta\text{Accuracy}_{M4}$  (性能衰减用  $-3.82$  表示, 单位: %)

Table 8 The accuracy variations of the SGS\_AVG model compared to the SGS\_MAX model in the MattrSet dataset:  $\Delta\text{Accuracy}_{M4}$  (The performance degradation indicators are marked in  $-3.82$ , unit: %)

SPI <sub>M</sub>	SGS 模型	$\Delta\text{Accuracy}_{M4} = \text{Accuracy}_{\text{SGS\_AVG}} - \text{Accuracy}_{\text{SGS\_MAX}}$											
		all	S + G + L	S + G + V	S + L + V	L + G + V	S + G	S + V	S + L	L + G	L + V	G + V	Avg <sub>model</sub>
SPI <sub>M-1</sub>	DT	7.75	14.72	5.32	3.56	6.88	9.54	0.18	15.28	10.11	0.14	0.13	6.69
	GBDT	8.39	5.62	9.13	8.95	8.28	6.26	8.80	6.58	5.19	8.77	8.60	7.69
	KNN	10.95	9.17	11.13	11.59	10.76	8.28	9.11	9.62	7.66	10.40	10.11	9.89
	LR	7.91	6.95	8.91	8.17	8.33	7.30	10.65	6.54	7.08	9.73	9.53	8.28
	NB	18.30	6.72	16.61	15.26	18.94	7.28	16.87	7.48	0.37	27.23	20.48	14.14
	RF	8.50	5.78	8.60	8.95	7.93	6.28	9.22	6.31	6.02	8.73	8.31	7.69
	XGBoost	7.04	4.75	6.80	6.74	6.35	5.56	6.11	5.81	4.16	5.77	5.94	5.91
SPI <sub>M-8</sub>	DT	8.92	11.26	4.00	3.49	5.06	9.57	0.20	14.83	10.44	7.12	0.22	6.83
	GBDT	7.37	7.26	7.14	7.78	7.15	6.65	7.73	6.97	6.15	9.51	7.17	7.35
	KNN	11.32	9.77	11.13	10.53	10.26	9.51	10.67	9.69	7.98	11.78	9.04	10.15
	LR	8.30	6.68	9.78	9.42	9.72	6.65	12.01	6.70	6.69	5.59	12.16	8.52
	NB	15.11	6.36	14.60	13.03	19.90	7.84	15.20	<b>-3.82</b>	0.14	19.25	5.08	10.24
	RF	7.55	6.78	7.35	7.71	7.13	7.21	7.77	6.80	6.50	8.29	7.30	7.31
	XGBoost	6.62	5.76	6.07	6.54	6.37	5.30	5.65	6.16	4.68	7.87	5.61	6.06
Avg <sub>Feat</sub>	9.57	7.68	9.04	8.69	9.50	7.37	8.58	7.50	5.94	10.01	7.83	/	

表 9 Fabric 数据集中, SGS\_AVG 类模型相对 SGS\_MAX 类模型的 Accuracy 变化值  $\Delta\text{Accuracy}_{F4}$  (性能衰减用  $-16.36$  表示, 单位: %)

Table 9 The accuracy variations of the SGS\_AVG model compared to the SGS\_MAX model in the Fabric dataset:  $\Delta\text{Accuracy}_{F4}$  (The performance degradation indicators are marked in  $-16.36$ , unit: %)

SPI <sub>F</sub>	SGS 模型	$\Delta\text{Accuracy}_{F4} = \text{Accuracy}_{\text{SGS\_AVG}} - \text{Accuracy}_{\text{SGS\_MAX}}$											
		all	S + G + L	S + G + V	S + L + V	L + G + V	S + G	S + V	S + L	L + G	L + V	G + V	Avg <sub>model</sub>
SPI <sub>F-1</sub>	DT	20.94	14.41	17.74	19.12	19.90	15.88	<b>-4.58</b>	6.56	<b>-16.36</b>	10.86	12.75	10.66
	GBDT	7.19	5.14	7.42	7.39	8.42	4.27	8.29	4.63	6.08	8.45	8.73	6.91
	KNN	19.31	29.86	18.99	24.29	11.06	30.25	24.03	40.25	9.28	14.89	13.06	21.39
	LR	10.31	7.46	11.06	9.32	10.98	6.93	9.36	6.17	11.93	10.54	12.32	9.67
	NB	15.94	25.04	16.39	18.64	13.11	17.85	16.31	6.55	24.37	10.31	6.00	15.50
	RF	9.48	8.41	9.64	10.58	11.65	7.39	11.22	8.49	10.66	13.03	11.89	10.22
	XGBoost	7.34	5.10	7.71	8.49	8.49	4.42	10.63	4.66	6.08	9.24	8.53	7.34
SPI <sub>F-2</sub>	DT	8.88	8.21	5.61	4.98	4.03	4.19	2.25	4.07	4.82	2.17	2.84	4.73
	GBDT	7.98	8.01	7.62	7.90	5.61	5.69	5.25	5.97	5.76	4.74	5.33	6.35
	KNN	10.23	7.46	10.47	11.02	9.08	7.51	6.21	6.24	7.35	9.48	9.32	8.58
	LR	12.56	11.73	12.01	12.84	8.02	7.98	8.61	9.01	9.29	7.29	7.03	9.67
	NB	13.15	10.27	12.80	13.80	8.65	0.91	10.51	3.47	0.24	10.23	3.12	7.92
	RF	29.74	8.38	9.72	10.64	8.93	6.60	7.31	8.25	8.85	8.81	8.57	10.53
	XGBoost	9.08	8.93	8.61	8.85	6.59	6.44	6.12	6.99	6.67	5.41	5.89	7.23
Avg <sub>Feat</sub>	13.01	11.32	11.13	11.99	9.61	9.02	8.68	8.67	6.79	8.96	8.24	/	

V” 次之 (11.99%), “L + G” 最小 (6.79%)。显然, 在平均值池化中, 引入 SPI<sub>F</sub> 之后, “V” 特征的作用越来越大; 4) 基于式 (16) 计算 MAA 指标, SPI<sub>F-1</sub>(11.67%) 优于 SPI<sub>F-2</sub>(7.86%), 即选取平均值池化, 且遵从由简入难识别顺序的 SPI<sub>F-1</sub> 方案最优; 5) 评估各方案中最优值。在 SPI<sub>F-1</sub> 方案中,

选取 “S + G + L” 特征组合及 SGS-LR 算法时标注性能最优, 为 97.55%; 在 SPI<sub>F-2</sub> 方案中, 选取 “S + L” 特征组合及 SGS-LR 算法时标注性能最优, 为 91.71%。SGS\_AVG 类模型最优值相比 SGS\_MAX 类模型最优值提升  $97.55\% - 92.22\% = 5.33\%$ 。“V” 特征的作用愈加显著。



图 4 表示: 在 MattrSet 数据集中, SGS 类模型选取两种池化方法后所得到的精准度均值折线, 它从两个角度对比不同的池化策略: 图 4(a) 和图 4(b) 刻画 SGS 类模型的精准度均值; 图 4(c) 和图 4(d) 刻画特征组合的精准度均值. Fabric 数据集与之类似, 不再赘述. 如图 4 所示: 1) 无论采用何种分类模型, 或何种特征组合, 平均值池化的折线均在最大值池化折线之上, 即平均值池化方法更优, 这与表 8 结论吻合. 其次, pcp\*n、cpp\*n 优于其他方案, p\*npc 方案最差, 即有针对性地先标注易识别材质 (Pu 或 Canvas), 可以得到难分样本信息, 促进标注精准度大幅提升, 这与表 6 结论一致. 2) 由图 4(a) 和图 4(b) 可知: SGS-LR 模型整体标注性能最优, SGS-RF 模型次之, 即加入分层先验信息后, 线性分

类模型能获取更优性能; 基于图 4(c) 可知: “S + L” 特征组合的标注性能最优 (pcp\*n 方案), “S + G + L + V” 特征组合次之 (cpp\*n 方案), 它优于 “S + G + L” 特征组合, 这表明 “V” 特征的加入有助于改善材质属性的标注性能. 基于图 4(d) 可知: “S + G + L” 特征组合的标注性能最优 (cpp\*n 方案), “S + G” 特征组合次之 (cpp\*n 方案).

### 3.2.5 全部模型的 Accuracy 比较

在 MattrSet 数据集中, SGS-MFF 算法的最优精准度为 86.37% (选取 pcp\*n 方案及 “S + V” 特征组合, 且执行平均值池化的 SGS-NB 模型), 将它与各基线最优值比较, 结果如表 10 所示. 表 4 中, GS-XGBoost 模型在选取 “S + G + L” 特征组合时标注性能最优, 故表 10 中它被表示为

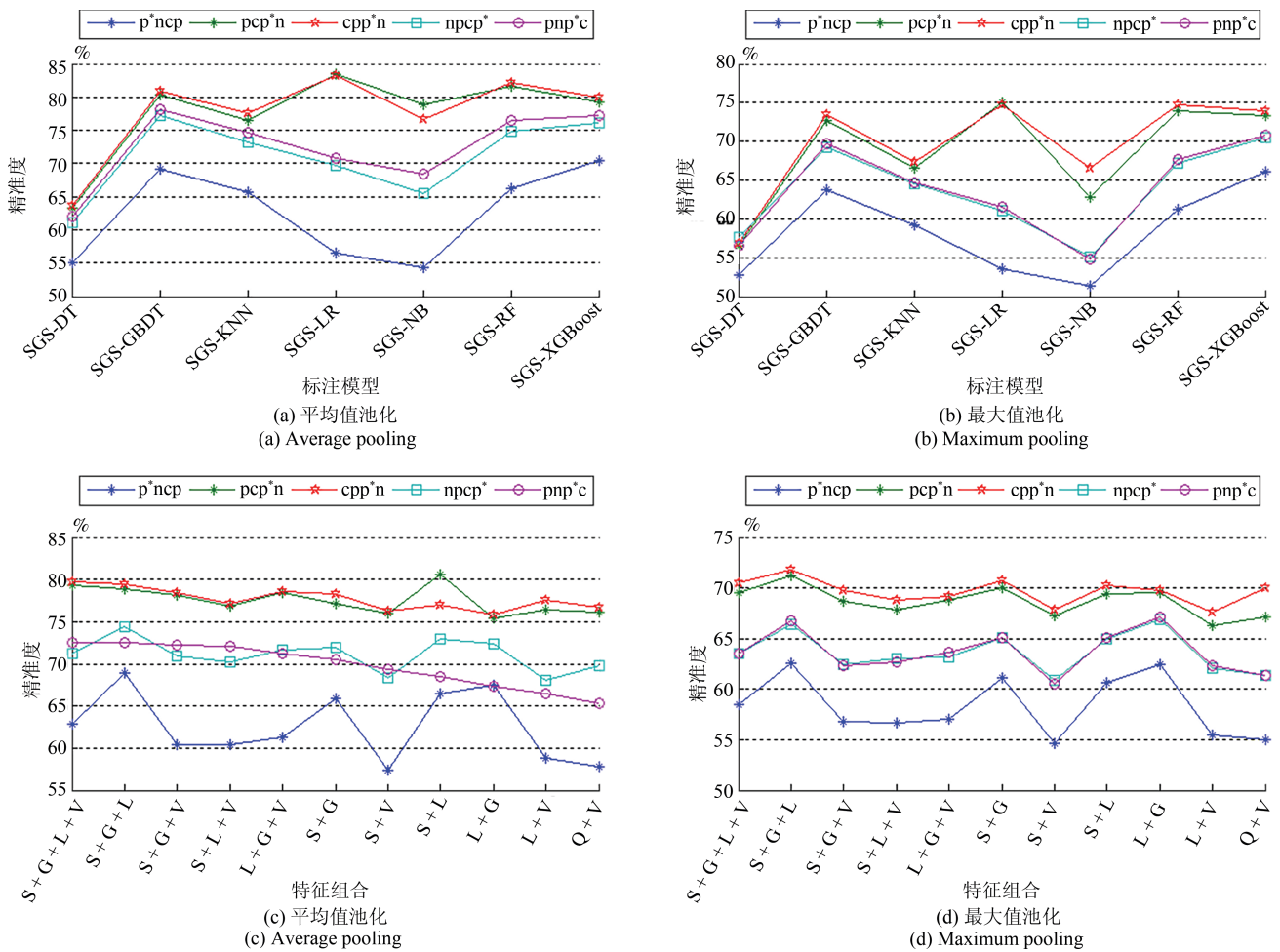


图 4 面向不同的 SGS 模型或不同特征组合的精准度均值比较

(a) 和 (b) 是不同模型, (c) 和 (d) 是不同特征组合

Fig. 4 The mean accuracy comparisons of different SGS models or feature combinations

(a) and (b) are different models, (c) and (d) are different feature combinations

表 10 MattrSet 数据集中, 各基线最优值与本文模型的 Accuracy 比较 (最优值如 **86.37** 等表示, 单位: %)Table 10 The best accuracy of each baseline in the MattrSet dataset is compared with the proposed model (The best value is marked as **86.37**, etc., unit: %)

Model	Accuracy	Model	Accuracy
1) SVM-S	50.83	2) GS-DT-SGL	46.20
3) GBDT-L	61.28	4) GS-RF-LG	61.75
5) Adaboost-L	61.54	6) GS-KNN-SGL	62.10
7) XGBoost-L	62.93	8) GS-LR-SL	59.45
9) VGG16	33.98	10) GS-NB-SG	49.61
11) InceptionResNetV2	52.09	12) GS-GBDT-SGL	65.13
13) Densenet169	59.77	14) GS-Adaboost-SGL	66.11
15) MobileNets	33.98	16) GS-XGBoost-SGL <sup>[64]</sup>	67.67
17) p*ncp-a-SGS-XGBoost-SGL	75.71	18) p*ncp-m-SGS-XGBoost-SGL	70.36
19) cpp*n-a-SGS-NB-SV	<b>86.37</b>	20) cpp*n-m-SGS-RF-SGL	78.14
21) p*pcn-a-SGS-XGBoost-SGL	80.00	22) p*pcn-m-SGS-XGBoost-SGL	74.22
23) pcp*n-a-SGS-NB-SV	<b>86.37</b>	24) pcp*n-m-SGS-RF-SGL	78.54
25) pnp*c-a-SGS-GBDT-SGL	80.05	26) pnp*c-m-SGS-XGBoost-LG	74.70
27) npcpc*-a-SGS-GBDT-SGL	79.71	28) npcpc*-m-SGS-XGBoost-LG	74.24

“GS-XGBoost-SGL”. “m”表示最大值池化, “a”表示平均值池化. 例如, 选取 cpp\*n 方案及 “S + V” 特征组合, 执行平均值池化的 SGS-NB 模型被表示为 “cpp\*n-a-SGS-NB-SV”, 其他模型命名以此类推. 同理, 在 Fabric 数据集中, SGSMFF 算法的最优精准度为 97.55% (选取 SPI<sub>F</sub>-1 方案及 “S + G + L” 特征组合, 且执行平均值池化的 SGS-LR 模型), 将它与各基线最优值比较, 结果如表 11 所示.

由表 10 可以发现: 1) 在 SGS 类模型中, “m”和 “a” 都各有 6 组最优, 但对比精准度绝对值, 平均值池化方式更优, 这与图 4、表 8 结论吻合; 2) 在 SGS 类模型中, XGBoost 模型在 6 组 (6/12 = 50%) 实验中最优, 它与 SGSMFF 框架能更好结合. 当然, 本文所提 “pcp\*n-a-SGS-NB-SV” 和 “cpp\*n-a-SGS-NB-SV” 模型获取了最佳标注性能, 精准度达 86.37%, 它较最强基线提升 (10.89% + 7.81%) = 18.70%. “pnp\*c-a-SGS-GBDT-SGL” 模型次优, 即引入最佳的分层先验信息后, 简单模型也能获得令人满意的结果, 经过推测: SPI 为分类提供非线性的难分样本信息, 这些信息极大地改善了标注性能; 3) “S + G + L”、“S + V” 等特征组合的标注性能最优. “V” 特征与 “S” 特征有较强互补性, 它们不但可以准确识别图像纹理, 且对形状变化也有较强 “鲁棒性”; 4) 深度学习模型的效果不好 (50% vs 50% 的划分效果更差, 故采用 70% 的训练样本, 30% 的测试样本. 迁移学习后加入两层全连接层进行参数微调, 完成材质分类), 这是因为它们多为数据驱动型模型, 对于中小规模数据集的效果较差. 未来拟通过 ACGAN 模型<sup>[66]</sup> 增强样本, 再运用深度

学习模型完成材质属性标注; 5) cpp\*n 和 pcp\*n 这两类分层先验信息有助于更好地完成图像属性标注. 综上, 合理地利用材质属性分类顺序, 能极大地改善粗粒度数据集的标注精准度.

由表 11 可知: 1) 在 SGS 类模型中, 平均值池化优于最大值池化, 这与表 9 结论一致; 2) “sdctcvpfnw-a-SGS-LR-SGL” 模型的标注性能最优, 即引入最佳的先验分层信息后, 简单模型也能获得令人满意的结果, 经过推测: SPI<sub>F</sub> 为分类提供了非线性的难分样本信息, 这些信息极大地改善了标注性能; 3) “S + G + L”、“S + G” 和 “S + L” 等特征组合的标注性能最优, 这说明: 对于细粒度数据集, 传统特征之间具有更强互补性; 4) 深度学习模型的效果不好 (50% vs 50% 的划分效果更差, 故采用 70% 的训练样本, 30% 的测试样本, 迁移学习后加入两层全连接层进行参数微调, 完成材质分类), 原因同上. 综上, 合理地利用材质属性分类顺序能极大地改善细粒度数据集的标注性能.

### 3.2.6 模型简化测试

在 MattrSet 数据集中, 对 SGSMFF 算法做模型简化测试, 以确定 SGSMFF 算法中各结构 (GS 算法、SGS 算法、AVG 池化、SPI) 的相对重要性. 由表 10 可发现: 基本分类模型变为 GS 类模型后, 标注性能提升 4.74% ((16) - (7)); GS 类模型变为 SGS\_MAX 类模型后, 标注性能提升 10.87% ((24) - (16)); SGS\_MAX 类模型变为 SGS\_AVG 类模型后, 标注性能提升 7.83% ((23) - (24)); 当分层先验方案由 p\*ncp (最差) 变为 pcp\*n (最优) 后, 标

表 11 Fabric 数据集中, 各基线最优值与本文模型的 Accuracy 比较 (最优值如 **97.55** 等表示, 单位: %)Table 11 The best accuracy of each baseline in the Fabric dataset is compared with the proposed model (The best value is marked as **97.55**, etc., unit: %)

Model	Accuracy	Model	Accuracy
1) SVM-S	77.92	2) GS-DT-SGL	65.64
3) GBDT-S	79.66	4) GS-RF-SG	76.58
5) AdaBoost-S	76.86	6) GS-KNN-LG	73.74
7) XGBoost-S	82.03	8) GS-LR-SL	78.28
9) VGG16	46.22	10) GS-NB-SG	57.42
11) VGG-M <sup>[35]</sup>	79.60	12) GS-GBDT-SG	79.98
13) Densenet169	46.22	14) GS-AdaBoost-SL	78.16
15) MobileNet	46.22	16) GS-XGBoost-SG <sup>[64]</sup>	81.95
17) sdtcvpfnw-m-SGS-XGBoost-SG	92.22	18) sdtcvpfnw-a-SGS-LR-SGL	<b>97.55</b>
19) nsfvtwdpc-m-SGS-XGBoost-SG	83.37	20) nsfvtwdpc-a-SGS-LR-SL	91.71

注性能提升了 10.66% ((23) - (17)). 综上, 在 MattrSet 数据集上, SGS<sub>MFF</sub> 算法中各结构的重要性降序排列: SGS > SPI > AVG > GS.

在 Fabric 数据集中, 对 SGS<sub>MFF</sub> 算法做模型简化测试. 由表 11 可以发现: 基本模型变为 GS 类模型后, 标注性能提升 -0.08% ((16) - (17)); GS 类模型变为 SGS<sub>MAX</sub> 类模型后, 性能提升 10.27% ((17) - (16)); SGS<sub>MAX</sub> 类模型变为 SGS<sub>AVG</sub> 类模型后, 标注性能提升 5.33% ((18) - (17)); 分层先验方案由 nsfvtwdpc (最差) 变为 sdtcvpfnw (最优) 后, 标注性能提升 5.84% ((18) - (20)). 综上, 在 Fabric 数据集上, SGS<sub>MFF</sub> 算法中各结构的重要性降序排列: SGS > SPI > AVG > GS.

因此, 无论是细粒度数据集, 还是粗粒度数据集, 分层思想 SGS 最重要, 其次是先验信息 SPI, 然后是平均值池化 AVG, 最后是多特征融合思想 GS. 这进一步验证了 SGS<sub>MFF</sub> 算法的鲁棒性.

### 3.3 拓展材质属性所蕴含的实用属性

由材质基本特性可知: 每一种材质属性都蕴含丰富的深层语义, 包括防水性、透气性、柔软性、可水洗性及耐磨性. 这些深层语义也称实用属性, 相比图像材质属性, 它们更贴近人类客观认知, 实用价值也更大, 故受用户的关注程度也较高. 在材料学专家

建议下并借鉴相对属性<sup>[12-14]</sup>特点, 构建如表 12 和表 13 所示的材质属性与实用属性之间的映射关系:

如表 12 和表 13 所示: 在二元映射关系中, 属性描述为“0”或“1”. 例如, 在 MattrSet 数据集中, Pu 材质的防水性标记为“1”表示 Pu 材质具有防水性, 二元映射关系也称二元属性. 在相对映射关系中, 针对某材质的实用属性, 标记数值越大, 则该材质的对应实用属性程度就越高. 由于材质数量不同, 在 MattrSet 数据集中, 最高程度值为“4”, 在 Fabric 数据集中最高程度值为“9”. 例如, 在 MattrSet 数据集中, Canvas 材质的透气性标记为“4”表示: Canvas 材质的透气性在表 12 的 4 种材质中最好, 而 Polyester 材质的透气性最差, 仅为“1”, “1 < 4”, 这就是著名的相对属性<sup>[12-14]</sup>标注, 它着力刻画属性的程度大小, 相比二元属性它蕴含更客观、真实的属性描述, 也更接近人类认知, 可以为人们提供更高质量的交互体验<sup>[10-20]</sup>.

综上, 无论是二元映射关系还是相对映射关系, 它们均刻画材质属性与实用属性间的语义关联, 这符合视觉基因计划<sup>[1]</sup>的重要思想, 有助于人们更全面地认知材质属性. 因此, 正确标注图像材质属性是挖掘其深层语义 (实用属性) 的前提. 伴随实用属性的挖掘, 人们便可架构一个全新的、层次化的材质属

表 12 MattrSet 数据集中材质属性与实用属性之间的映射关系

Table 12 The relationship between the material attributes and their utility attributes in the MattrSet dataset

材质属性	二元映射关系					相对映射关系				
	防水性	透气性	柔软性	水洗性	耐磨性	防水性	透气性	柔软性	水洗性	耐磨性
Pu (皮革)	1	0	0	0	0	4	2	1	1	1
Canvas (帆布)	0	1	0	1	1	1	4	2	4	4
Polyester (涤纶)	1	0	1	1	0	3	1	4	2	3
Nylon (尼龙)	0	0	1	1	0	2	3	3	3	2

表 13 Fabric 数据集中材质属性与实用属性之间的映射关系

Table 13 The relationship between the material attributes and their utility attributes in the Fabric dataset

材质属性	二元映射关系				相对映射关系			
	防水性	透气性	柔软性	耐磨性	防水性	透气性	柔软性	耐磨性
Wool (羊毛)	0	1	1	0	1	8	8	1
Denim (牛仔布)	0	0	0	1	8	1	1	8
Viscose (粘胶纤维)	0	1	1	1	7	7	6	9
Cotton (棉花)	0	1	1	0	4	5	5	3
Silk (丝绸)	0	1	1	0	5	9	9	5
Polyester (涤纶)	1	0	1	0	9	2	3	7
Nylon (尼龙)	0	0	1	0	6	3	2	6
Terrycloth (毛巾布)	0	1	1	1	3	4	4	4
Fleece (摇粒绒)	0	1	1	0	2	6	7	2

性认知体系, 它是对 ImageNet 的有益补充, 且可以指导相关专业的工程技术人员, 在大规模图像检索、机器人视觉、工业检测等领域中发挥重要作用。

#### 4 结语及展望

图像材质属性标注在电子商务、机器人视觉、工业检测等领域都具有非常重要的现实意义. 提出基于分层基因优选多特征融合 (SGSMFF) 的图像材质属性标注机制: 采用主、客观相结合的方法捕获分层先验信息 SPI, 改进传统 ERGS 算法, 设计分层基因优选算法, 实现多特征融合. 实验表明: SGSMFF 算法先利用 SPI 获取非线性难分样本信息, 即使线性类模型也能获取非常不错的分类性能, 其中, 在 MattrSet 数据集上, pcp\*n、cpp\*n 这两类 SPI 最优; 在 Fabric 数据集上, sdtevpfnw 方案最优. 未来, 可采用折半查找及相对比较法将新材质加入到 SPI 中; 其次, SGSMFF 考虑对图像特征进行融合, 在 MattrSet 数据集中, “S + G + L”、“S + V” 等特征组合在分类中扮演关键角色; 在 Fabric 数据集上, “S + G + L”、“S + G” 等特征组合更重要, 不同的数据集应选取对应特征组合; 经过模型简化测试可知: 在 SGSMFF 算法中, 分层思想最重要, 其次是先验信息, 然后是平均值池化, 多特征融合思想最次要. 对于定量评价, 在 MattrSet 数据集上, SGSMFF 算法较最强基线精准度提高 18.70%; 在 Fabric 数据集上, SGSMFF 算法较最强基线的精准度提高 15.60%. 最后, 基于材质属性可进一步挖掘其蕴涵的深层语义 (实用属性), 以构建全新的、层次化的材质属性认知体系, 它符合视觉基因计划重要思想, 是对 ImageNet 的有益补充.

综上, SGSMFF 算法具有层次性、统一性、包容性及高效性等 4 大优秀特性. 未来研究方向: 1) 在 SGSMFF 算法中继续引入注意力机制, 以聚焦图像中关键视觉内容, 降低来自图像背景的噪声干扰; 2)

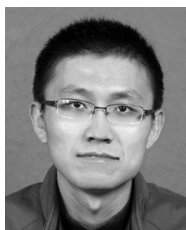
考虑到 “V” 特征的作用 (如 pcp\*n-m-SGS-NB-SV 模型), 尝试采用 ResNet、DenseNet 等更优秀的深度学习特征, 更准确地刻画图像视觉内容, 并与传统特征做深度融合, 以提升材质属性标注性能; 3) 将 SGSMFF 算法应用到自然语言处理、视频分析、肿瘤图像识别等领域, 进一步扩展其应用范围.

#### References

- 1 Krishna R, Zhu Y K, Groth O, Johnson J, Hata K, Kravitz J, et al. Visual genome: Connecting language and vision using crowdsourced dense image annotations. *International Journal of Computer Vision*, 2017, **123**(1): 32–73
- 2 Pietikäinen M, Hadid A, Zhao G Y, Ahonen T. *Computer Vision Using Local Binary Patterns*. London: Springer Berlin, 2011.
- 3 Oliva A, Torralba A. Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, 2006, **155**: 23–36
- 4 Lowe D G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004, **60**(2): 91–110
- 5 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: *Proceedings of the 3rd International Conference on Learning Representations*. San Diego, CA, USA, 2015. 1–14
- 6 Chandra B, Gupta M. An efficient statistical feature selection approach for classification of gene expression data. *Journal of Biomedical Informatics*, 2011, **44**(4): 529–535
- 7 Farhadi A, Endres I, Hoiem D, David F. Describing objects by their attributes. In: *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009: 1778–1785
- 8 Kumar N, Belhumeur P, Nayar S. FaceTracer: A search engine for large collections of images with faces. In: *Proceedings of the 10th European Conference on Computer Vision*. Marseille, France: Springer, 2008. 340–353
- 9 Kumar N, Berg A C, Belhumeur P N, Nayar S K. Attribute and simile classifiers for face verification. In: *Proceedings of the 12th International Conference on Computer Vision*. Kyoto, Japan: IEEE, 2009. 365–372

- 10 Jayaraman D, Grauman K. Zero-shot recognition with unreliable attributes. In: Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Quebec, Canada: MIT Press, 2014. 3464–3472
- 11 Berg T L, Berg A C, Shih J. Automatic attribute discovery and characterization from noisy web data. In: Proceedings of the 11th European Conference on Computer Vision. Heraklion, Greece: Springer, 2010. 663–676
- 12 Gan C, Yang T B, Gong B Q. Learning attributes equals multi-source domain generalization. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 87–97
- 13 Liu Z W, Luo P, Wang X G, Tang X O. Deep learning face attributes in the wild. In: Proceedings of the 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015. 3730–3738
- 14 Tang P, Zhang J, Wang X G, Feng B, Roli F B, Liu W Y. Learning extremely shared middle-level image representation for scene classification. *Knowledge and Information Systems*, 2017, **52**(2): 509–530
- 15 Bradley C, Boulton T E, Ventura J. Cross-modal facial attribute recognition with geometric features. In: Proceedings of the 12th IEEE International Conference on Automatic Face & Gesture Recognition. Washington, USA: IEEE, 2017. 891–896
- 16 Liu Z W, Luo P, Qiu S, Wang X G, Tang X O. DeepFashion: Powering robust clothes recognition and retrieval with rich annotations. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 1096–1104
- 17 Qi G J, Hua X S, Rui Y, Tang J H, Mei T, Zhang H J. Correlative multi-label video annotation. In: Proceedings of the 15th ACM International Conference on Multimedia. Augsburg, Germany: ACM, 2007. 17–26
- 18 Parikh D, Grauman K. Relative attributes. In: Proceedings of the 2011 International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011. 503–510
- 19 Kovashka A, Grauman K. Attribute adaptation for personalized image search. In: Proceedings of the 2013 IEEE International Conference on Computer Vision. Sydney, Australia: IEEE, 2013. 3432–3439
- 20 Kovashka A, Parikh D, Grauman K. WhittleSearch: Interactive image search with relative attribute feedback. *International Journal of Computer Vision*, 2015, **115**(2): 185–210
- 21 Yu A, Grauman K. Just noticeable differences in visual attributes. In: Proceedings of the 2015 International Conference on Computer Vision. Santiago, Chile: IEEE, 2015. 2416–2424
- 22 Cheng Y H, Qiao X, Wang X S, Yu Q. Random forest classifier for zero-shot learning based on relative attribute. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, **29**(5): 1662–1674
- 23 Yuan B D, Tu J, Zhao R W, Zheng Y B, Jiang Y G. Learning part-based mid-level representation for visual recognition. *Neurocomputing*, 2018, **275**: 2126–2136
- 24 Liu X, Wang J, Wen S L, Ding E R, Lin Y Q. Localizing by describing: Attribute-guided attention localization for fine-grained recognition. In: Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco, California, USA: AAAI, 2017. 4190–4196
- 25 Ledig C, Theis L, Huszar F, Caballero J, Cunningham A, Acosta A, et al. Photo-realistic single image super-resolution using a generative adversarial network [Online], available: <https://arxiv.org/abs/1609.04802>, September 15, 2016.
- 26 Singh K K, Lee Y J. End-to-end localization and ranking for relative attributes. In: Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016. 753–769
- 27 Liu X G, Yu Y Z, Shum H Y. Synthesizing bidirectional texture functions for real-world surfaces. In: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques. New York, USA: ACM, 2001. 97–106
- 28 Leung T, Malik J. Representing and recognizing the visual appearance of materials using three-dimensional textons. *International Journal of Computer Vision*, 2001, **43**(1): 29–44
- 29 Varma M, Zisserman A. A statistical approach to texture classification from single images. *International Journal of Computer Vision*, 2005, **62**(1–2): 61–81
- 30 Heera M M, Divya J K, Varma M S, Divya R A, Agrawal D V K. Minimum variance optimal filter design for a 3x3 MEMS gyroscope cluster configuration. *IFAC-Papersonline*, 2016, **49**(1): 639–645
- 31 Sharan L, Rosenholtz R, Adelson E. Material perception: What can you see in a brief glance? *Journal of Vision*, 2009, **9**(8): 784
- 32 Liu C, Sharan L, Adelson E H, Rosenholtz R. Exploring features in a Bayesian framework for material recognition. In: Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE, 2010. 239–246
- 33 Hu D E, Bo L F, Ren X F. Toward robust material recognition for everyday objects. In: Proceedings of the 2011 British Machine Vision Conference. Dundee, UK: BMVA Press, 2011. 1–11
- 34 Sharan L, Liu C, Rosenholtz R, Adelson E H. Recognizing materials using perceptually inspired features. *International Journal of Computer Vision*, 2013, **103**(3): 348–371
- 35 Kampouris C, Zafeiriou S, Ghosh A, Malassiotis S. Fine-grained material classification using micro-geometry and reflectance. In: Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016. 778–792
- 36 Dong J Y, Chantler M. Capture and synthesis of 3D surface texture. *International Journal of Computer Vision*, 2005, **62**(1–2): 177–194
- 37 Jian M W, Yin Y L, Dong J Y, Zhang W Y. Comprehensive assessment of non-uniform illumination for 3D heightmap reconstruction in outdoor environments. *Computers in Industry*, 2018, **99**: 110–118
- 38 Jian M W, Dong J Y. Capture and fusion of 3d surface texture. *Multimedia Tools and Applications*, 2011, **53**(1): 237–251

- 39 Jian M W, Lam K M, Dong J Y. Illumination-insensitive texture discrimination based on illumination compensation and enhancement. *Information Sciences*, 2014, **269**: 60–72
- 40 Van Der Maaten L, Hinton G. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 2008, **9**(11): 2579–2605
- 41 Friedman J H. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 2001, **29**(5): 1189–1232
- 42 Hou Jie, Mao Yao-Bin, Sun Jin-Sheng. Online boosting algorithms based on exponential and 0-1 loss. *Acta Automatica Sinica*, 2014, **40**(4): 635–642  
(侯杰, 茅耀斌, 孙金生. 基于指数损失和 0-1 损失的在线 Boosting 算法. 自动化学报, 2014, **40**(4): 635–642)
- 43 Vu H T, Gallinari P. Using RankBoost to compare retrieval systems. In: Proceedings of the 14th ACM International Conference on Information and Knowledge Management. Bremen, Germany: ACM, 2005. 309–310
- 44 Chen T Q, Guestrin C. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, USA: ACM, 2016. 785–794
- 45 Feng J, Yu Y, Zhou Z H. Multi-layered gradient boosting decision trees. In: Proceedings of the 32nd Conference on Neural Information Processing Systems. Montréal, Canada: 2018.
- 46 Vedaldi A, Gulshan V, Varma M, Zisserman A. Multiple kernels for object detection. In: Proceedings of the 2009 IEEE 12th International Conference on Computer Vision. Kyoto, Japan: IEEE, 2009. 606–613
- 47 Xia H, Hoi S C H. MKBoost: A framework of multiple kernel boosting. *IEEE Transactions on Knowledge and Data Engineer*, 2013, **25**(7): 1574–1586
- 48 Zhang Z Y, Lyons M, Schuster M, Akamatsu S. Comparison between geometry-based and gabor-wavelets-based facial expression recognition using multi-layer perceptron. In: Proceedings of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition. Nara, Japan: IEEE, 1998. 454–459
- 49 Bai S, Sun S Y, Bai X, Zhang Z X, Tian Q. Smooth neighborhood structure mining on multiple affinity graphs with applications to context-sensitive similarity. In: Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016. 592–608
- 50 Xu Dan-Lei, Du Lan, Liu Hong-Wei, Hong Ling, Li Yan-Bing. Joint feature selection and classification design based on variational relevance vector machine. *Acta Automatica Sinica*, 2011, **37**(8): 932–943  
(徐丹蕾, 杜兰, 刘宏伟, 洪灵, 李彦兵. 一种基于变分相关向量机的特征选择和分类结合方法. 自动化学报, 2011, **37**(8): 932–943)
- 51 Liu Z Q, Wang S J, Zheng L, Tian Q. Robust ImageGraph: Rank-level feature fusion for image search. *IEEE Transactions on Image Processing*, 2017, **26**(7): 3128–3141
- 52 Mafarja M, Aljarah I, Heidari A A, Hammouri A I, Faris H, Al-Zoubi A M, et al. Evolutionary population dynamics and grasshopper optimization approaches for feature selection problems. *Knowledge-Based Systems*, 2018, **145**: 25–45
- 53 Faris H, Mafarja M M, Heidari A A, Aljarah I, Al-Zoubi M, Mirjalili S, et al. An efficient binary salp swarm algorithm with crossover scheme for feature selection problems. *Knowledge-Based System*, 2018, **154**: 43–67
- 54 Emary E, Zawbaa H M, Grosan C, Hassenian A E. Feature subset selection approach by gray-wolf optimization. In: Proceedings of the 1st International Afro-European Conference for Industrial Advancement. Cham, Germany: Springer, 2014. 1–13
- 55 Cox D R. The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1958, **20**(2): 215–232
- 56 Ho T K. Random decision forests. In: Proceedings of the 3rd International Conference on Document Analysis and Recognition. Montreal, Canada: IEEE, 1995. 278–282
- 57 Altman N S. An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 1992, **46**(3): 175–185
- 58 Quilan J R. Decision trees and multi-valued attributes. *Machine Intelligence 11*. New York, USA: Oxford University Press, 1988. 305–318
- 59 Kononenko I. ID3, sequential Bayes, naive Bayes and Bayesian neural networks. In: Proceedings of European Working Session on Learning. 1989. 91–98
- 60 Garreta R, Moncecchi G. *Learning Scikit-Learn: Machine Learning in Python*. Birmingham, England: Packt Publishing, 2013.
- 61 Szegedy C, Ioffe S, Vanhoucke V, Alemi A A. Inception-v4, inception-ResNet and the impact of residual connections on learning. In: Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco, California, USA: AAAI, 2017. 4278–4284
- 62 Huang G, Liu Z, Van Der Maaten L, Weinberger K Q. Densely connected convolutional networks. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 2261–2269
- 63 Howard A G, Zhu M L, Chen B, Kalenichenko D, Wang W J, Weyand T, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications [Online], available: <https://arxiv.org/abs/1704.04861>, April 17, 2017.
- 64 Zhang H B, Qiu D D, Wu Rv Z, Deng Y X, Ji D H, Li T. Novel framework for image attribute annotation with gene selection XGBoost algorithm and relative attribute model. *Applied Soft Computing*, 2019, **80**: 57–79
- 65 Chen W H, Chen X T, Zhang J G, Huang K Q. Beyond triplet loss: A deep quadruplet network for person re-identification. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 1320–1329
- 66 Odena A, Olah C, Shlens J. Conditional image synthesis with auxiliary classifier GANs. In: Proceedings of the 34th International Conference on Machine Learning. Sydney, Australia, 2017.



**张红斌** 华东交通大学软件学院副教授, 主要研究方向为计算机视觉, 自然语言处理, 推荐系统. 本文通信作者.

E-mail: zhanghongbin@whu.edu.cn

**(ZHANG Hong-Bin** Associate professor at the Software School, East China Jiaotong University. His research

interest covers computer vision, natural language processing, and recommendation system. Corresponding author of this paper.)



**邱蝶蝶** 华东交通大学软件学院硕士研究生. 主要研究方向为图像理解, 计算机视觉, 机器学习.

E-mail: diedie88888@163.com

**(QIU Die-Die** Master student at the Software School, East China Jiaotong University. Her research interest

covers image understanding, computer vision, and machine learning.)

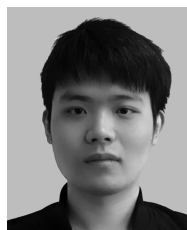


**邬任重** 华东交通大学软件学院硕士研究生. 主要研究方向为肿瘤图像识别, 计算机视觉, 机器学习.

E-mail: renchw5@163.com

**(WU Ren-Zhong** Master student at the Software School, East China Jiaotong University. Her research interest

covers tumor image recognition, computer vision, and machine learning.)

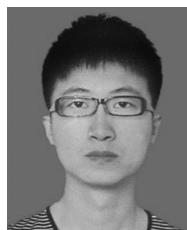


**蒋子良** 华东交通大学软件学院硕士研究生. 主要研究方向为肿瘤图像识别, 图像内容生成, 机器学习.

E-mail: a320jzlnf430a@163.com

**(JIANG Zi-Liang** Master student at the Software School, East China Jiaotong University. His research interest

covers tumor image recognition, image captioning, and machine learning.)



**武晋鹏** 华东交通大学软件学院硕士研究生. 主要研究方向为问答系统, 自然语言处理, 机器学习.

E-mail: wujinpeng920421@163.com

**(WU Jin-Peng** Master student at the Software School, East China Jiaotong University. His research interest

covers question answering, natural language processing, and machine learning.)



**姬东鸿** 武汉大学国家网络安全学院教授. 主要研究方向为舆情分析, 自然语言处理, 机器学习.

E-mail: dhji@whu.edu.cn

**(JI Dong-Hong** Professor at the School of Cyber Science and Engineering, Wuhan University. His research interest

covers public opinion analysis, natural language processing, and machine learning.)