Vol.33 No.11 Nov. 2021

基于 VGG-CapsNet 的零件手绘草图识别模型

杨钟亮1), 黄瑞红1), 陈育苗2), 张凇3), 毛新华4)

- 1) (东华大学机械工程学院 上海 201620)
- 2) (华东理工大学艺术设计与传媒学院 上海 200237)
- 3) (Department of Materials, The University of Manchester Manchester M13 9PL)
- 4)(北京中丽制机工程技术有限公司 北京 101111)

(yzl@dhu.edu.cn)

摘 要:针对概念设计阶段现有 CAD 系统难以通过手绘草图准确匹配对应的零件的问题,提出预训练网络(VGG)与胶囊网络(CapsNet)相结合的零件手绘草图识别模型(VGG-CapsNet). 招募了 5 名设计师绘制零件草图,构建包含标准件和非标件在内的 23 类零件手绘草图数据集;设计了组间实验与组内实验,分别构建基于 VGG-CapsNet 的零件手绘草图识别模型,并与 rVGG-13 模型、rCNN-13 模型的识别结果进行比较.实验结果表明, VGG-CapsNet 模型在组间和组内实验中的平均准确率均高于其他 2 种模型,为零件设计知识的检索与重用提供技术支持.

关键词: 手绘草图; 零件识别; 胶囊网络; 深度学习

中图法分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2021.18774

Freehand-Sketched Part Recognition Using VGG-CapsNet

Yang Zhongliang¹⁾, Huang Ruihong¹⁾, Chen Yumiao²⁾, Zhang Song³⁾, and Mao Xinhua⁴⁾

- 1) (College of Mechanical Engineering, Donghua University, Shanghai 201620)
- ²⁾ (School of Art, Design and Media, East China University of Science and Technology, Shanghai 200237)
- ³⁾ (Department of Materials, The University of Manchester, Manchester M13 9PL)

Abstract: To solve the problem that the existing CAD system is difficult to match the corresponding parts accurately through the freehand sketch in the conceptual design, a recognition model (VGG-CapsNet) for freehand sketch of part is proposed, which combining the pre-trained network (VGG) and capsule network (CapsNet). Five designers are recruited to sketch parts, and build 23 kinds of freehand sketch of parts including standard parts and non-standard parts. The between-group experiment and within-group experiment are designed, and then the recognition models of VGG-CapsNet are constructed respectively. The recognition results of the VGG-CapsNet models are compared with the rVGG-13 models and the rCNN-13 models. The experimental results show that the mean accuracy of VGG-CapsNet model is higher than the other two models, which provides technical support for the retrieval and reuse of part design knowledge.

Key words: freehand sketch; part recognition; capsule network; deep learning

收稿日期: 2020-11-11; 修回日期: 2021-04-14. 基金项目: 国家自然科学基金(51905175); 上海市浦江人才计划(2019PJC021); 浙江省健康智慧厨房系统集成重点实验室开放基金(2014E10014); 中央高校基本科研业务费专项资金(2232018D3-27). 杨钟亮(1982—), 男,博士,副教授,硕士生导师,CCF会员,主要研究方向为人工智能设计;黄瑞红(1996—),女,硕士研究生,主要研究方向为人工智能设计;陈育苗(1988—),女,博士,讲师,硕士生导师,主要研究方向为人机交互;张凇(1996—),男,博士研究生,主要研究方向为人工智能设计;毛新华(1975—),男,博士,高级工程师,主要研究方向为智能制造.

⁴⁾ (Beijing Chonglee Machinery Engineering Co., Ltd., Beijing 101111)

草图被用来记录和表达人们的想法及概念. 从原始洞穴绘画时代开始,人类就使用简单的图 表现真实世界的物体和概念[1]. 对于工业设计, 手 绘草图更是设计初期阶段设计师们快速记录灵感、 表达和交流设计思想的必不可少的手段. 然而, 在 实际设计过程中往往会遇到这样的问题: 在仅知 道产品外观或功能的情况下,设计师想要检索相 关零件的信息及其上下文构件的装配关系,受限 于自身的知识,通常难以迅速、精准地获取某些零 件的名称, 需要耗费大量时间和精力去搜寻自己 预期的零件. 现有的 CAD 系统在实体建模、仿真 和优化等设计阶段发挥了极大的作用, 但在设计 构思、概念设计等早期阶段的设计知识检索与匹配 等方面, 仍难以为设计师提供有效支持[2]. 在新产 品设计过程中, 仅有 20%的产品来自真正的创新, 40%可从现有设计中获取, 其余 40%可以通过修改 现有的设计获得[3]. 如果设计师能直接通过手绘草 图检索到预期零件作为设计参考并重用[2],将极大 地提升产品概念设计与详细设计的转换效率.

零件的手绘草图识别属于图像识别的一种特 殊形式. 近年来, 尽管卷积神经网络(convolutional neural networks, CNN)在图像识别中取得了巨大成 功,但 CNN 也存在无法检测特征之间相对位置关 系的缺陷. 对此, Sabour 等[4]提出了名为胶囊网络 (capsule network, CapsNet)的网络架构, 在未进行 数据增强的情况下,它在识别高度重叠的数字方 面表现出极好的性能. 可见, 采用 CapsNet 对于提 升零件手绘草图识别准确率将是一种新的尝试. 然而, 与手写数字相比, 本文研究的零件手绘草图 具有更复杂的特征. 即使同一个零件, 受个人绘图 习惯、水平和绘画风格的影响, 所绘草图在角度、 线型和精细程度等方面均存在较大的组内差异. 另外, 与基于自然图像的物体识别相比, 手绘零件 草图识别缺乏现成的用于模型训练的大样本数据 集,标准零件与非标零件数量本身也存在类别不 平衡的问题.

在此背景下,本文首先构建了包含标准件和非标件在内的 23 类手绘草图数据集,然后将预训练VGG和 CapsNet 相结合,提出基于 VGG-CapsNet 的零件手绘草图识别模型,并与 rVGG-13, rCNN-13模型进行比较,以获得更加精确的零件手绘草图识别效果.本文模型有助于设计师在概念设计阶段更有效地检索预期零件,实现对已有设计知识的检索与重用,对提高产品设计效率具有重要意义,也为开发面向概念设计阶段的基于知识的CAD系统提供技术支持.

1 相关工作

1.1 草图识别

随着电子设备和压力传感设备的发展,人们 开始逐渐改变传统的纸笔手绘方式,越来越多的 人选择使用触摸屏设备(iPad、手绘板等)手绘草图. 草图是产品设计中最重要的环节之一,被广泛应用 于设计认知及设计思维研究,对于设计概念定位、 发展、表达、推演和形成都有着不可替代的作用^[5].

对草图领域的研究主要包括草图识别^[6-11]、基于草图的图像检索^[12]以及基于草图的三维模型检索^[13-15]等. 早期的草图识别技术的研究对象主要是手写符号或字母^[6], 其识别常规的草图具有局限性, 且缺乏定义良好的数据集. 对此, Eitz 等^[7]构建了一个包含 250 类对象共 20000 幅草图的大型手绘草图数据集——TU-Berlin, 通过提取特征袋(bag of feature, BoF)草图表示, 并使用支持向量机(support vector machine, SVM)对草图进行分类. Schneider 等^[8]提取费舍尔向量表示草图, 并使用SVM 进行草图分类.

随着深度学习的发展, CNN 成功地应用于图 像识别中. 但最初 CNN 主要用于有色彩和纹理的 自然图像识别, 直接将其用在草图识别上表现不 佳^[9]. 直到 2015 年, Yu 等^[9]提出了一种多尺度、多 通道的深度神经网络 Sketch-a-Net, 用于手绘草图 识别, 并且其在 TU-Berlin 数据集上的识别准确率 首次超过人类. 之后, Yu 等[12]使用 Sketch-a-Net 作 为一个三元排序网络分支的基本体系结构, 并使 用边缘映射和特定的草图数据增强策略, 实现在 特定类别上的细粒度检索. 赵鹏等[10]利用手绘草 图的笔画顺序信息,提出基于深度 CNN 与递归神 经网络的手绘草图识别模型 deep-CRNN-sketch, 并使用 SVM 进行训练和分类. 柯小甜等[11]将机构 看做由图元构成的图形,提出一种面向概念设计 的机构草图识别及分析方法, 但其研究对象仅限 于机械运动机构.

虽然深度 CNN 在图像识别上取得成功,但其缺点和局限性也随之暴露.深度神经网络需要大量带标签的训练数据,以避免出现过拟合.自然图像识别所用的大量数据主要来源于网络公开可用的图片,但缺乏可以直接使用的手绘零件草图的大样本数据库.此外, CNN 的池化操作使对象特征之间重要的空间信息丢失,而特征之间的位置关系是确定其类别的关键^[16].因此, CapsNet^[4]既保留特征之间重要的空间信息,又具有旋转不变性,

可以识别不同方向的物体,而且与深度神经网络相比,所需要的训练数据少.

在数据集方面,草图识别主要以TU-Berlin^[7]、改进的TU-Berlin^[8]和QuickDraw^[17]等为基准展开.TU-Berlin 数据集涵盖了日常生活中常见的250类对象,如汽车、椅子和水壶等,每类80幅图片.改进的TU-Berlin 数据集在原有数据集的基础上删除人类识别错误的草图,并保留删除后每类至少有56张草图的类别,共有160类.QuickDraw是一个大型矢量数据集,通过在线游戏平台收集矢量草图,用户需要在20s内根据类别名称画出相应的草图,该数据集涵盖的类别也是日常生活中常见的345类对象.通过文献检索发现,当前仍缺乏专门针对零件手绘草图识别的草图数据库,但以上这些数据集的收集方法为本文构建零件草图数据集提供了一个很好的参照.

1.2 三维模型检索

随着 CAD 及三维建模技术的发展,产生了海 量的三维模型. 对三维模型的有效检索和复用需 求也急剧增加[14]. 早期的三维模型检索主要是基 于文本的模型检索, 但由于文本难以对模型精准 定义, 并且受标注者主观影响较大, 需要耗费大量 人力物力,检索效果并不理想.因此,基于内容的 三维模型检索快速发展. 基于内容检索的研究主 要分为3类:基于示例模型、二维投影视图和手绘 图的三维模型检索方法[18]. 近年来, 草图因其简 单、方便的特点被广泛用于三维模型检索中. Wang 等[13]使用 2 个孪生 CNN 分别学习草图和视图的特 征,实现了基于草图的三维形状检索. 白静等[14] 提出一种端到端三维模型草图检索框架,建立跨 域数据的联合特征分布, 提高了检索准确率. 张开 兴等[15]提出了一种基于草图的产品三维 CAD 模型 概念设计推送方法, 能从三视图检索到对应的三 维 CAD 模型,为设计人员在概念设计阶段提供了 设计参考. 可见, 大量的研究主要是以手绘草图检 索多边形模型, 或是以标准投影视图检索对应的 三维零件, 针对概念设计阶段直接以手绘草图检 索对应三维零件的研究并不多见.

1.3 零件识别

零件是机械制造中最小的单元,是产品的基本构成要素.现有的机械零件识别研究主要包括基于视图轮廓、形状特征和识别算法等.许鑫杰等[19]使用 Canny 边缘检测算法提取出零件的边缘,并进一步提取出零件轮廓的相关特征与待检测的零件轮廓特征相比较,实现对零件的识别检测.郑健红

等^[20]提出了基于 Mask R-CNN 和多特征融合的 SVM 识别算法,通过对识别结果的二次分类,实现对零件抓取区域的精确识别. Zeng 等^[21]提出了基于视图相关的拓扑图(view dependent graphs, VD图),并设计了从零件的平均形状构造 VD图、构造素描图和分层搜索相似参数零件的自动算法,实现对基于草图的三维参数化零件检索.

产品的外在造型与内部结构均在一定程度上取决于构成它们的零件.在实际设计过程中,尤其是新手工业设计师往往缺乏零件设计的相关经验,难以有效地将外观造型与内部结构设计快速衔接与转换.在零件识别中,现有的研究主要围绕机械装配过程中大量使用的螺栓、螺母和垫片等标准件展开^[19-20].现有的主流 CAD 软件,如 SolidWorks,Creo 和 Siemens NX 等,均配备了参数化的标准零件库,但这些软件尚不支持在三维建模前的概念设计阶段通过手绘草图识别出相应的零件模型.可见,研发高效、准确的零件手绘草图识别算法是亟待解决的关键技术之一.

1.4 识别算法

早期的草图识别算法主要是通过提取方向梯度直方图^[22]、尺度不变特征变换^[23]、费舍尔向量^[8] 等图像特征进行分类的. 前期研究中本课题组先后尝试将 SVM、多层感知机、自组织特征映射神经网络、基因表达式编程、融合长短期记忆网络(long short-term memory, LSTM)与卡尔曼滤波器(Kalman filter, KF)的 KF-LSTM 等算法用于基于表面肌电信号(surface electromyography, sEMG)的单笔手绘草图识别与重建中^[24].

目前草图识别中大部分工作是基于 CNN 的深度学习方法^[9-10]. 对于缺乏大量标注数据集的各类识别任务,使用在大规模数据集上预训练的 CNN 作为特征提取器,将学习到的通用图像表示用于新的识别任务,可以有效地改善各类识别任务性能^[25]. Simonyan 等^[26]提出的 VGG16 在 ImageNet 挑战赛^[27]中取得优异成绩,并且可以作为特征提取器用于其他小型数据集的识别. Devis 等^[28]使用预训练的 VGG19 网络进行迁移学习,提取草图特征,实现对输入草图图像的检索. Mouffok 等^[29]使用预训练的 ResNet-50 网络分别对二维草图和三维形状(非装配体)进行分类,并计算查询(二维草图)与目标(三维形状)的 Softmax 向量之间的成对余弦相似度,从二维草图检索出对应的三维形状.

2017年, Sabour 等^[4]颠覆性地以 CapsNet 替代 CNN, 不仅在网络设计上减少了参数, 而且有效

地避免了过拟合问题. 随后, CapsNet 在识别交通标志^[30]、甲骨文构件^[31]和医学影像检测^[32]等方面表现出良好的性能. Kumar^[30]提出了一种基于深度学习结构的 CapsNet 交通标志检测方法, 在德国交通标志数据集上的准确率达到了 97.6%. 鲁绪正等^[31]提出了一种基于 CapsNet 和迁移学习的模型 OracleNet, 可以自动识别并标记甲骨文字形中包含的构件. Jiménez-Sánchez 等^[32]比较了 CapsNet 和CNN 在医学图像分析的典型数据集约束下(即少量标注数据和类不平衡)的表现, 结果表明, CapsNet 用更少的数据量来训练, 仍可以获得相同或更好的性能, 并且对不平衡的类别分布更加稳定.

一方面, 在缺乏如 TU-Berlin 等大量标注数据

集的情况下,使用预训练的 VGG 模型作为特征提取器,可以有效地改善草图识别任务性能;另一方面,CapsNet 在处理小样本及样本不平衡的识别任务中具有良好的表现,尤其在识别高度重叠的手写数字方面表现出良好的性能,非常适合作为本文零件手绘草图的分类器.通过检索,尚未发现将此方法用于零件草图识别的文献报道.

2 本文模型

2.1 本文模型框架

本文提出面向零件手绘草图识别的 VGG-CapsNet 模型, 其概念框架如图 1 所示.

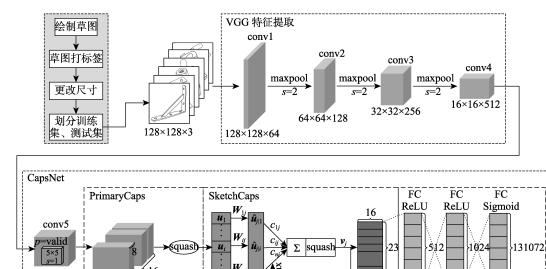


图 1 本文模型框架

 $b_{ij}=b_{ij}+\hat{\boldsymbol{u}}_{i|i}\cdot\boldsymbol{v}_{i}$

使用在大规模数据集上预训练的神经网络作为特征提取器,可以有效地提升识别任务性能,而且, CapsNet 具有空间感知性及旋转不变性的优点,本文将两者相结合实现对零件手绘草图的识别,主要模块如下.

8×8×8×16

12×12×256

- (1) 数据收集与预处理. 首先确定需要识别的标准件与非标件的类别, 收集清晰的零件图片, 接着绘制这些零件的草图, 并对手绘草图预处理.
- (2) 卷积特征提取. 在卷积特征提取中,靠近底部的层提取的是如视觉边缘、色彩、纹理等更加通用的特征图,而本文的草图数据集与 ImageNet数据集有较大的差异. 在前期实验中,分别选取在 ImageNet 上预训练的 VGG16 模型的前 7 层、前 10 层及前 13 层进行特征提取的对比实验,结果表明,前 10 层进行特征提取有更好的效果. 因此,针对

小型的手绘草图数据集,采用预先在 ImageNet 上 训练好的 VGG16 模型的前 10 层提取零件手绘草图的特征.

(3) CapsNet 训练. 将 VGG 提取的草图特征输入 CapsNet 进行训练,输出特定部分存在的概率及位置信息、变形信息等属性,并将结果封装在一个向量中^[4]. 经过路由迭代,得到所检测零件手绘草图的概率.

2.2 算法流程

Step1. 将预处理后的手绘草图输入预训练的 VGG 模型中进行特征提取,输出层为 conv4(VGG16 模型的第 10 层),图像输入尺寸为 128×128×3,经特征提取后输出图像特征尺寸为 16×16×512. 通过特征提取得到零件草图的底层通用特征.

Step2. 将 conv4 的图像特征输入 CapsNet 的 conv5

层,对输入的图像特征进行卷积运算.

Step3. 将 conv5 的卷积运算结果输入胶囊层 (PrimaryCaps). 该层本质与卷积层相似,接收来自上一层的输入并进行卷积运算,不同的是将结果重塑为 16 个 8D 胶囊,得到胶囊的输出向量图像 u_i .

Step4. 采用 squash 压缩函数将输出向量 u_i 的长度压缩到 0~1, 此时 u_i 的长度代表胶囊所检测到的螺纹、通孔和圆弧等特征的概率,并且方向不发生改变. 为了避免训练时出现梯度消失的问题,在压缩函数中加入较小的 ε 值,将压缩函数的输出向量 u_i 输入 SketchCaps

层,其计算公式为
$$u_i = \frac{\|u_i\|^2}{1 + \|u_i\|^2} \frac{u_i}{\|u_i\|}$$

Step5. 将 SketchCaps 层与 PrimaryCaps 层完全相连,在 SketchCaps 层开始运行动态路由算法. Sketch-Caps 层的高层胶囊 j 接收来自 PrimaryCaps 层的低层胶囊 i 的输出向量 u_i ,并将其与相应的权重矩阵 W_{ij} 相乘,得到预测向量

$$\hat{\boldsymbol{u}}_{j|i} = \boldsymbol{W}_{ij} \boldsymbol{u}_i \tag{1}$$

用于预测高层特征应该存在的位置. 其中, W_{ij} 在网络反向传播期间进行更新, 它编码了低层特征(螺纹、通孔和圆弧)等与高层特征之间重要的空间关系及其他关系. 这也是传统 CNN 存在的缺点, 最大池化层丢失了有价值的信息, 并没有编码特征之间的重要的空间关系, 而空间关系对于检测一个对象中各个特征分布是否正确具有重要意义.

Step6. 计算低层胶囊 i 和高层胶囊 j 的耦合系数

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_{k} \exp(b_{ik})}$$
 (2)

其代表低层胶囊输出向量 \hat{u}_{ji} 能传递到高层胶囊 j 的权重,其数量等于高层胶囊 j 的数量. 第 1 次迭代后,所有 c_{ij} 的值都相等,之后 c_{ij} 通过动态路由过程更新. b_{ij} 初始值为 0

Step7. 对预测向量 $\hat{\pmb{u}}_{ji}$ 进行加权并求和, 得到总的输入向量

$$\mathbf{s}_{j} = \sum_{i} c_{ij} \hat{\mathbf{u}}_{j|i} \tag{3}$$

Step8. 对总的输入向量 s_j 使用压缩函数,确保输出向量 v_j 长度为 0~1,且不改变输出向量的方向.输出向量

$$\mathbf{v}_{j} = \frac{\|\mathbf{s}_{j}\|^{2}}{1 + \|\mathbf{s}_{j}\|^{2}} \frac{\mathbf{s}_{j}}{\|\mathbf{s}_{j}\|} \tag{4}$$

Step9. 计算胶囊 j 的输出向量 v_j 与预测向量 \hat{u}_{ji} 的点积, 再加上旧权重, 得到更新的权重

$$b_{ij} = b_{ij} + \mathbf{v}_j \cdot \hat{\mathbf{u}}_{j|i} \tag{5}$$

若点积值为正,说明胶囊i预测的输出与胶囊j实际的

输出具有一致性,因此权重系数变大,反之则变小.如上所述,式(1)~式(5)构成一次完整的动态路由过程,通过3次迭代^[4]更新权重.

Step10. 经过 3 次路由迭代,得到最终的输出向量 v_j ,计算向量 v_j 的长度,即为所检测零件手绘草图存在的概率.

2.3 损失函数

若顶层胶囊 j 检测到零件草图,那么它的向量长度接近1,否则接近0.对每个胶囊,其边缘损失计算公式为

 $L_k = T_k \max \left(0, m^+ - \|\mathbf{v}_k\|\right)^2 + \lambda (1 - T_k) \max \left(0, \|\mathbf{v}_k\| - m^-\right)^2.$ 其中, k 为图像类别,如果检测到类别 k ,则 $T_k = 1$; $m^+ = 0.9$; $m^- = 0.1$; $\lambda = 0.5$.

2.4 评价指标

本文利用测试图像的测试数据集对模型进行了评价. 以 5 次测试数据的平均准确率 A 及方差作为评价指标. 此外, 还计算了每次实验的精确度、召回率、 F_1 值以及混淆矩阵, 作为辅助评价指标,并使用配对样本 Wilcoxon 符号秩检验方法对准确率进行统计分析. 平均准确率的计算公式为

$$A = \frac{1}{5} \sum_{i=1}^{5} \frac{n}{N} \times 100\% .$$

其中, n 为识别正确的草图数量; N 为识别草图的总数.

3 实验

3.1 数据集收集

本文用到的零件对象主要来源于化纤卷绕机 所用到的部分零件.根据机械领域划分标准,将零件分为标准件和非标件2大类.其中,标准件包括 螺栓、螺母、垫片、齿轮和销这5类,非标件共18 类.这些零件中每类零件又包含1~6个子类,详细 种类信息如表1所示.

本文实验共招募到5名参与者, 所采用的草图绘制方法借鉴了 Eitz 等^[7]方法. 在正式绘制草图之前, 首先向参与者讲解绘制草图的注意事项: (1)绘制的草图可识别, 即看到这张草图就知道它是什么零件, 不需要再去猜测; (2)绘制草图的过程中不要有反复涂抹的痕迹; (3)只需要画出所给零件的线稿图(不要有色调). 在正式实验前, 让参与者根据以上注意事项尝试3次绘图, 并对参与者绘制完成的草图进行检查. 在确定参与者完全了解绘图要求之后, 才可以正式开始绘制草图.

表 1 零件手绘草图名称及数量

	7- (11)-41-11/2/2				
序号	名称	数量	子类数量		
1	螺栓	83	6		
2	螺母	54	5		
3	齿轮	63	6		
4	垫片	59	6		
5	销	57	5		
6	齿条	36	2		
7	三角支架	37	3		
8	角码	40	5		
9	铝柱	42	3		
10	导丝头	32	1		
11	丝头座	33	1		
12	陶瓷刀	30	1		
13	定制角码	81	6		
14	护手板	36	3		
15	挡板	37	3		
16	轴承座	34	2		
17	定制支架	33	2		
18	气缸固定件	37	2		
19	光轴固定环	40	3		
20	牛眼	53	6		
21	顶丝	39	4		
22	导轨限位块	35	2		
23	合页	87	6		

参与者使用 WACOM CTL-672 手绘板和配套的手绘笔在绘图软件 Photoshop 中画草图. 草图大小统一设置为 480×480. 参与者需要认真观察每张零件图片,然后在一张新建的空白画布上画出零件的一个草图,对画草图的时间没有限制. 每个零件要画出可识别的包含透视图与三视图在内的尽可能多的草图. 最终,实验共收集到来自 5 名参与者的 23 类草图. 删除不合格的草图后,筛选出标准件草图 5 类共 312 幅,非标件草图 18 类共 762 幅.由于每类零件所包含的子类不同,而零件草图又是由 5 名参与者绘制,因此最终草图数据集存在样本不均衡的情况. 但本文收集草图是按照实际项目进行的,所以并没有刻意增删每类草图的数量以使每类样本均衡. 图 2 所示为手绘零件草图类别示例.



气缸固定件 光轴固定环 牛眼 顶丝 导轨限位块 合页图 2 零件手绘草图示例

3.2 图像预处理

在正式训练之前, 先对草图进行预处理: (1) 将每类草图随机打乱顺序并按照类别打标签, 为后续数据集划分做准备; (2) 为了保留足够的草图信息, 且考虑计算机资源的限制, 采用区域插值的方法将草图尺寸缩小为 128×128×3; (3) 将草图按训练集 80%, 验证集 20%的比例随机划分, 为之后的分层 K 折交叉验证做准备, 训练集、验证集和测试集具体草图数量如表 2 所示.

表 2 不同数据集草图数量

数据集		草图数量/个	
双加木	训练集	验证集	测试集
标准件	201	50	65
非标件	481	121	160
全部零件	682	171	225

3.3 实验设计

本文设计了2组实验方案,包括组间识别和组 内识别.

组间识别是指将标准件草图和非标件草图分开识别,建立2个针对标准件和非标件的训练和测试模型,得到对应的标准件草图和非标件草图的识别结果.该实验的特点是可以避免因标准件和非标件样本不均衡对分类结果带来的影响.

组内识别是指将标准件草图和非标件草图放在一起进行识别,构建一个训练和测试模型,模型可以同时识别标准件草图和非标件草图.该实验相对来说具有更强的普适性和泛化性.

3.4 模型构建

为了证明 VGG-CapsNet 的有效性,本文对基于卷积特征提取的深度卷积神经网络 rVGG-13 和传统的卷积神经网络 rCNN-13 进行对比. 算法比较的目的如下: (1) 确定采用预训练网络进行手绘草图特征提取是否有助于提升模型性能; (2) 确定使用 CapsNet 代替 Softmax 分类器是否有助于提升模型性能.

本文使用已经在 ImageNet 上预训练的 VGG16 模型作为特征提取器,将提取的初始特征图作为 VGG-CapsNet 模型和 rVGG-13 模型的输入,初始 特征图尺寸为 16×16×512. 而 rCNN-13 模型输入为 经过归一化处理的草图数据集,输入尺寸为 128×128×3.

VGG-CapsNet 模型由编码器和解码器 2 部分构成. 编码器部分包括一个卷积层, 一个 Primary-Caps 层和 SketchCaps 层. conv5 层设定 256 个步长

为 1 的 5×5 卷积核,使用 ReLU激活函数,填充为 valid. PrimaryCaps 层为胶囊层设定 256 个步长为 2 的 5×5 卷积核,进行卷积运算后,将输出重塑为 16 个 8D 胶囊. SketchCaps 层为 16D 胶囊,胶囊的个数取决于识别的类别数量.解码器由 2 个全连接层组成,其神经元个数分别为 512 和 1024,用于重建输入,同时解码器起到正则化作用,避免过拟合. VGG-CapsNet 网络结构参数如表 3 所示.

表 3 VGG-CapsNet 网络结构参数

卷积层	卷积核尺寸	卷积核数量	步长	输出
input				128×128×3
conv1-1	3×3	64	1	128×128×64
conv1-2	3×3	64	1	128×128×64
maxpool	3×3		2	64×64×64
conv2-1	3×3	128	1	64×64×128
conv2-2	3×3	128	1	64×64×128
maxpool	3×3		2	32×32×128
conv3-1	3×3	256	1	32×32×256
conv3-2	3×3	256	1	32×32×256
conv3-3	3×3	256	1	32×32×256
maxpool	3×3		2	16×16×256
conv4-1	3×3	512	1	16×16×512
conv4-2	3×3	512	1	16×16×512
conv4-3	3×3	512	1	16×16×512
conv5	5×5	256	2	12×12×256
PrimaryCaps	5×5	8×16	2	256×8D
SketchCaps				23×16D
FC1	1×1	512	1	1×1
FC2	1×1	1024	1	1×1

rVGG-13模型有2个完全一样的卷积层,每层有256个步长为1的5×5卷积核,使用ReLU激活函数,填充为valid.之后是2个全连接层,神经元个数分别为512和1024,每个全连接层后都使用0.5的dropout防止过拟合.rVGG-13模型网络结构设置与VGG-CapsNet完全相同,不同之处在于VGG-CapsNet中使用胶囊代替了传统的神经元,以此来验证CapsNet的作用.

rCNN-13 模型作为本文研究的基准网络,其网络结构参数与 rVGG-13 模型相同,不同之处在于rVGG-13 的卷积层保留了在 ImageNet 上预训练的 VGG16 网络的权重参数,而 rCNN-13 是常规的CNN,需要从头开始训练,以此来验证使用 VGG预训练网络特征提取的作用.rVGG-13/rCNN-13 网络结构参数如表 4 所示.

这些网络模型均在 Windows10 系统上进行训

练,使用GPU加速.系统硬件为3.4 GHz 英特尔酷 睿双核处理器和8 GB RAM, NVIDIA GeForce GTX1060 6 GB. 编程语言为 Python3.6,使用基于Tensorflow 后端的 Keras 来实现.

表 4 rVGG-13/rCNN-13 网络结构参数

世界					
conv1-1 3×3 64 1 128×128×64 conv1-2 3×3 64 1 128×128×64 maxpool 3×3 2 64×64×64 conv2-1 3×3 128 1 64×64×128 conv2-2 3×3 128 1 64×64×128 maxpool 3×3 256 1 32×32×256 conv3-1 3×3 256 1 32×32×256 conv3-2 3×3 256 1 32×32×256 conv3-3 3×3 256 1 32×32×256 maxpool 3×3 256 1 32×32×256 conv4-1 3×3 512 1 16×16×512 conv4-2 3×3 512 1 16×16×512 conv4-3 3×3 512 1 16×16×512 conv5 5×5 256 2 12×12×256 conv6 5×5 256 2 8×8×256 FC1 1×1 512 1	卷积层	卷积核尺寸	卷积核数量	步长	输出
conv1-2 3×3 64 1 128×128×64 maxpool 3×3 2 64×64×64 conv2-1 3×3 128 1 64×64×128 conv2-2 3×3 128 1 64×64×128 maxpool 3×3 2 32×32×128 conv3-1 3×3 256 1 32×32×256 conv3-2 3×3 256 1 32×32×256 conv3-3 3×3 256 1 32×32×256 maxpool 3×3 2 16×16×256 conv4-1 3×3 512 1 16×16×512 conv4-2 3×3 512 1 16×16×512 conv4-3 3×3 512 1 16×16×512 conv5 5×5 256 2 12×12×256 conv6 5×5 256 2 8×8×256 FC1 1×1 512 1 1×1	input				128×128×3
maxpool 3×3 2 64×64×64 conv2-1 3×3 128 1 64×64×128 conv2-2 3×3 128 1 64×64×128 maxpool 3×3 2 32×32×128 conv3-1 3×3 256 1 32×32×256 conv3-2 3×3 256 1 32×32×256 conv3-3 3×3 256 1 32×32×256 maxpool 3×3 2 16×16×256 conv4-1 3×3 512 1 16×16×512 conv4-2 3×3 512 1 16×16×512 conv4-3 3×3 512 1 16×16×512 conv5 5×5 256 2 12×12×256 conv6 5×5 256 2 8×8×256 FC1 1×1 512 1 1×1	conv1-1	3×3	64	1	128×128×64
conv2-1 3×3 128 1 64×64×128 conv2-2 3×3 128 1 64×64×128 maxpool 3×3 2 32×32×128 conv3-1 3×3 256 1 32×32×256 conv3-2 3×3 256 1 32×32×256 conv3-3 3×3 256 1 32×32×256 maxpool 3×3 2 16×16×256 conv4-1 3×3 512 1 16×16×512 conv4-2 3×3 512 1 16×16×512 conv4-3 3×3 512 1 16×16×512 conv5 5×5 256 2 12×12×256 conv6 5×5 256 2 8×8×256 FC1 1×1 512 1 1×1	conv1-2	3×3	64	1	128×128×64
conv2-2 3×3 128 1 64×64×128 maxpool 3×3 2 32×32×128 conv3-1 3×3 256 1 32×32×256 conv3-2 3×3 256 1 32×32×256 conv3-3 3×3 256 1 32×32×256 maxpool 3×3 2 16×16×256 conv4-1 3×3 512 1 16×16×512 conv4-2 3×3 512 1 16×16×512 conv4-3 3×3 512 1 16×16×512 conv5 5×5 256 2 12×12×256 conv6 5×5 256 2 8×8×256 FC1 1×1 512 1 1×1	maxpool	3×3		2	64×64×64
maxpool 3×3 2 32×32×128 conv3-1 3×3 256 1 32×32×256 conv3-2 3×3 256 1 32×32×256 conv3-3 3×3 256 1 32×32×256 maxpool 3×3 2 16×16×256 conv4-1 3×3 512 1 16×16×512 conv4-2 3×3 512 1 16×16×512 conv4-3 3×3 512 1 16×16×512 conv5 5×5 256 2 12×12×256 conv6 5×5 256 2 8×8×256 FC1 1×1 512 1 1×1	conv2-1	3×3	128	1	64×64×128
conv3-1 3×3 256 1 32×32×256 conv3-2 3×3 256 1 32×32×256 conv3-3 3×3 256 1 32×32×256 maxpool 3×3 2 16×16×256 conv4-1 3×3 512 1 16×16×512 conv4-2 3×3 512 1 16×16×512 conv4-3 3×3 512 1 16×16×512 conv5 5×5 256 2 12×12×256 conv6 5×5 256 2 8×8×256 FC1 1×1 512 1 1×1	conv2-2	3×3	128	1	64×64×128
conv3-2 3×3 256 1 32×32×256 conv3-3 3×3 256 1 32×32×256 maxpool 3×3 2 16×16×256 conv4-1 3×3 512 1 16×16×512 conv4-2 3×3 512 1 16×16×512 conv4-3 3×3 512 1 16×16×512 conv5 5×5 256 2 12×12×256 conv6 5×5 256 2 8×8×256 FC1 1×1 512 1 1×1	maxpool	3×3		2	32×32×128
conv3-3 3×3 256 1 32×32×256 maxpool 3×3 2 16×16×256 conv4-1 3×3 512 1 16×16×512 conv4-2 3×3 512 1 16×16×512 conv4-3 3×3 512 1 16×16×512 conv5 5×5 256 2 12×12×256 conv6 5×5 256 2 8×8×256 FC1 1×1 512 1 1×1	conv3-1	3×3	256	1	32×32×256
maxpool 3×3 2 16×16×256 conv4-1 3×3 512 1 16×16×512 conv4-2 3×3 512 1 16×16×512 conv4-3 3×3 512 1 16×16×512 conv5 5×5 256 2 12×12×256 conv6 5×5 256 2 8×8×256 FC1 1×1 512 1 1×1	conv3-2	3×3	256	1	32×32×256
conv4-1 3×3 512 1 16×16×512 conv4-2 3×3 512 1 16×16×512 conv4-3 3×3 512 1 16×16×512 conv5 5×5 256 2 12×12×256 conv6 5×5 256 2 8×8×256 FC1 1×1 512 1 1×1	conv3-3	3×3	256	1	32×32×256
conv4-2 3×3 512 1 16×16×512 conv4-3 3×3 512 1 16×16×512 conv5 5×5 256 2 12×12×256 conv6 5×5 256 2 8×8×256 FC1 1×1 512 1 1×1	maxpool	3×3		2	16×16×256
conv4-3 3×3 512 1 16×16×512 conv5 5×5 256 2 12×12×256 conv6 5×5 256 2 8×8×256 FC1 1×1 512 1 1×1	conv4-1	3×3	512	1	16×16×512
conv5 5×5 256 2 12×12×256 conv6 5×5 256 2 8×8×256 FC1 1×1 512 1 1×1	conv4-2	3×3	512	1	16×16×512
conv6 5×5 256 2 8×8×256 FC1 1×1 512 1 1×1	conv4-3	3×3	512	1	16×16×512
FC1 1×1 512 1 1×1	conv5	5×5	256	2	12×12×256
	conv6	5×5	256	2	8×8×256
FC2 1×1 1024 1 1×1	FC1	1×1	512	1	1×1
	FC2	1×1	1 024	1	1×1

采用 StratifiedKFold 方法进行 5 折交叉验证,通过分层采样,可以保证训练集、验证集中各类别样本的比例与原始数据集相同,同时避免训练数据过拟合并产生稳定的结果. 将训练数据集划分为 5 组,选取其中 4 组作为训练集,剩余一组为验证集,分别进行 5 次实验,计算平均准确率及方差.

所有模型都使用 Adam 优化算法以端到端的方式训练.由于计算资源限制,batch-size 设置为 16.对于 VGG-CapsNet 模型和 rVGG-13 模型,在经过测试 0.001, 0.0005 和 0.0001 的学习率后,选取最优值 0.000 5, VGG-CapsNet 模型的路由迭代次数为 3次,模型迭代次数为 50次,输出 K 折后的平均准确率.而 rCNN-13模型学习率最优值为 0.000 1,并且所有模型学习率都增加 0.9 的权重衰减,模型训练次数为 100次,输出 K 折后的平均准确率.

4 结果与讨论

4.1 VGG-CapsNet 识别结果

VGG-CapsNet 模型对组间识别及组内识别的结果如表 5 所示.

表 5 VGG-CapsNet 模型的识别结果及方差 %

数据集	准确率	精确度	召回率	F_1 值
标准件	96.92±0.02	97.00±0.01	97.00±0.03	97.00±0.03
非标件	93.88±0.01	94.40 ± 0.01	93.80±0.01	94.00 ± 0.01
全部零件	90.31±0.00	91.80 ± 0.01	90.40 ± 0.00	90.40 ± 0.00

(1) 组间实验结果

VGG-CapsNet 模型对标准件的识别准确率最高,平均准确率达到96.92%. 选取5折交叉验证中准确率居中的模型,其混淆矩阵如图3所示. 其中,3类标准件草图螺母、齿轮和垫片的识别准确率为100.00%,可以准确区分;而螺栓和销这2类标准件草图可能由于相似的螺纹特征,导致识别错误.

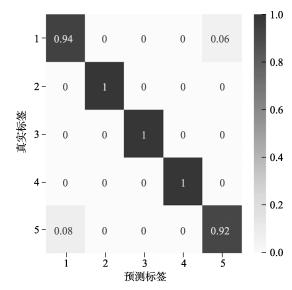


图 3 VGG-CapsNet 的组间识别标准件的混淆矩阵

VGG-CapsNet 模型对非标件的识别准确率次之,平均准确率为93.88%. 选取5次识别结果中准确率居中的模型,其混淆矩阵如图 4 所示. 其中,有 12 类非标零件草图的识别准确率为 100.00%,可以准确区分;导丝头、定制角码、挡板、牛眼和合页这 5 类零件草图识别准确率均超过 85.00%;而角码识别准确率为 62.00%,可能是由于角码与定制角码、定制支架及合页具有相似的通孔特征,容易识别错误.

(2) 组内实验结果

VGG-CapsNet 模型对全部零件的识别准确率为 90.31%. 选取 5 次识别结果中准确率居中的模型, 其混淆矩阵如图 5 所示. 其中, 11 类零件草图识别准确率为 100.00%, 可以准确区分; 螺栓、垫片、销、三角支架、铝柱、导丝头、定制角码、牛眼和导轨限位块这 9 类零件草图识别准确率均超

过 85.00%; 角码、顶丝和合页这 3 类零件草图识别准确率最低, 角码(识别准确率为 62.00%)易被误识别为定制角码、气缸固定件和合页, 可能是由于这几类零件具有相似的折弯特征; 顶丝(识别准确率为 62.00%)最易被误识别为螺栓, 可能是由于这 2 类零件具有相似的螺纹特征; 合页(识别准确率为 72.00%)最易被误识别为轴承座, 可能是由于这 2 类零件具有相似的对称的孔特征. 总体来看,组间识别效果比组内识别效果要好.

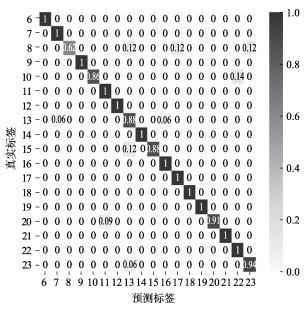


图 4 VGG-CapsNet 的组间识别非标件的混淆矩阵

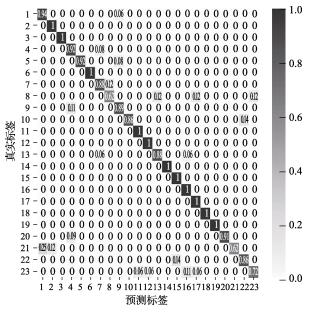


图 5 VGG-CapsNet 的组内识别的混淆矩阵

4.2 rVGG-13 识别结果

rVGG-13 模型对组间识别及组内识别的结果

如表 6 所示. 在组间实验中,模型对标准件的平均识别准确率为 95.38%; 对非标件的平均识别准确率为 86.00%. 在组内实验中,模型对全部零件的平均识别准确率为 81.24%.

表 6 rVGG-13 模型的识别结果及方差 %

数据集	准确率	精确度	召回率	F ₁ 值
 标准件	95.38±0.10	95.80±0.06	95.00±0.13	95.20±0.11
非标件	86.00±0.14	87.80 ± 0.15	86.20±0.16	86.00±0.16
全部零件	81.24±0.01	83.00±0.05	81.00±0.02	80.60 ± 0.02

4.3 rCNN-13 识别结果

rCNN-13 模型对组间识别及组内识别的结果如表 7 所示. 在组间实验中,模型对标准件的平均识别准确率为 85.23%;对非标件的平均识别准确率为 62.63%. 在组内实验中,模型对全部零件的平均识别准确率为 63.82%.

表 7 rCNN-13 模型的识别结果及方差 %

数据集	准确率	精确度	召回率	F_1 值
标准件	85.23±0.17	85.60±0.16	84.40±0.16	84.60±0.18
非标件	62.63±0.05	64.80 ± 0.02	61.40±0.07	61.00±0.10
全部零件	63.82±0.04	65.80±0.09	61.60±0.05	61.20±0.03

4.4 实验结果讨论

由表 5~表 7 可以看出, 3 种模型对标准件的识别准确率都较高,均超过 85.00%, VGG-CapsNet 比rVGG-13 高 1.54%, rVGG-13 比 rCNN-13 高 10.15%; 对于非标件, VGG-CapsNet 模型的识别准确率比 rVGG-13 高 7.88%, rVGG-13 比 rCNN-13高 23.37%; 对于全部零件, VGG-CapsNet 模型的识别准确率比 rVGG-13高 9.07%, rVGG-13比 rCNN-13高 17.42%.

采用配对样本 Wilcoxon 符号秩检验,设显著性水平为0.05,对3种模型的组间和组内识别准确率进行比较,统计结果如图6所示.

对于图 6中的 3个数据集,使用预训练的 VGG模型作为特征提取器的 VGG-CapsNet 和 rVGG-13模型的平均识别准确率均显著高于 rCNN-13模型,验证了使用预训练的 VGG16模型作为特征提取器能更好地提取零件草图的通用特征,有利于之后的图像识别分类.

在图 6b 和图 6c 中, VGG-CapsNet 模型在组间 非标件实验及组内全部零件实验中的准确率均显 著高于 rVGG-13 模型, 而两者唯一的不同之处就 是使用胶囊代替了传统的神经元, 证明了 CapsNet 在对小样本及类不均衡的草图分类中依然可以学

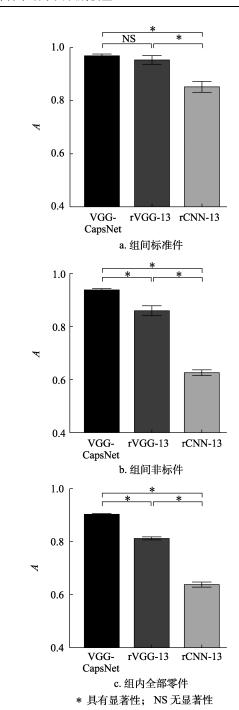


图 6 配对样本 Wilcoxon 符号秩准确率直方图

习到较多的样本特征, 具有更好的稳定性.

然而,本文模型也存在一定的局限性.一方面,实验只收集了来自 5 位参与者的零件手绘草图,由于数据有限,标准件与非标件存在样本不均衡的问题,模型可能难以学习到完整的样本特征,具有相似螺纹、通孔、折弯结构的零件易混淆,并且参与者绘画风格的差异对模型的效果可能也存在一定的影响.后续研究将进一步扩大实验样本,提高模型识别准确率.另一方面,本文采用 VGG与 CapsNet 相结合取得较好的效果,但还有很多预

训练模型,如 Inception, ResNet, Xception, Dense-Net 等,均可用于图像特征提取.后续将尝试更多的算法,进一步提高识别准确率.

5 结 语

本文提出面向零件手绘草图的识别模型 VGG-CapsNet,并开展实验与比较分析.结果表明,将预训练的 VGG16 模型作为特征提取器并与 CapsNet 相结合,在处理有限的数据量和类不平衡 时具有更强的泛化能力和稳定性.后续将进一步 扩大训练样本,优化模型,提高零件手绘草图的识别准确率,应用于 CAD 系统的零件设计知识检索 与重用.

参考文献(References):

- [1] Lu W, Tran E. Free-hand sketch recognition classification[OL].
 [2020-11-11]. http://cs231n.stanford.edu/reports/2017/pdfs/420.
 pdf
- [2] Goel A K, Vattam S, Wiltgen B, et al. Cognitive, collaborative, conceptual and creative—four characteristics of the next generation of knowledge-based CAD systems: a study in biologically inspired design[J]. Computer-Aided Design, 2012, 44 (10): 879-900
- [3] Gunn T G. The mechanization of design and manufacturing[J]. Scientific American, 1982, 247(3): 114-131
- [4] Sabour S, Frosst N, Hinton G E. Dynamic routing between capsules[C] //Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM Press, 2017: 3859-3869
- [5] Sun Lingyun, Wang Changlu, Chai Chunlei, *et al.* A computer-aided design technology based on creative point[J]. Scientia Sinica: Informationis, 2013, 43(8): 996-1011(in Chinese)
 (孙凌云, 王长路, 柴春雷, 等. 基于创意拐点的计算机辅助草图设计技术[J]. 中国科学: 信息科学, 2013, 43(8): 996-1011)
- [6] Hse H, Newton A R. Sketched symbol recognition using zernike moments[C] //Proceedings of the 17th International Conference on Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2004: 367-370
- [7] Eitz M, Hays J, Alexa M. How do humans sketch objects?[J]. ACM Transactions on Graphics, 2012, 31(4): Article No.44
- [8] Schneider R G, Tuytelaars T. Sketch classification and classification-driven analysis using fisher vectors[J]. ACM Transactions on Graphics, 2014, 33(6): Article No.174
- [9] Yu Q, Yang Y, Song Y Z, et al. Sketch-a-Net that beats humans[OL]. [2020-11-11]. https://arxiv.org/abs/1501.07873
- [10] Zhao Peng, Liu Yang, Liu Huiting, et al. A sketch recognition method based on deep convolutional-recurrent neural network[J]. Journal of Computer-Aided Design & Computer

- Graphics, 2018, 30(2): 217-224(in Chinese) (赵鹏, 刘杨, 刘慧婷, 等. 基于深度卷积-递归神经网络的手绘草图识别方法[J]. 计算机辅助设计与图形学学报, 2018, 30(2): 217-224)
- [11] Ke Xiaotian, Tian Huaiwen, Yang Tao, *et al.* Sketch for structure recognition and analysis[J]. Application Research of Computers, 2017, 34(5): 1568-1571(in Chinese) (柯小甜, 田怀文, 杨陶, 等. 基于草绘的机构识别及机构分析方法研究[J]. 计算机应用研究, 2017, 34(5): 1568-1571)
- [12] Yu Q, Liu F, Song Y Z, et al. Sketch me that shoe[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2016: 799-807
- [13] Wang F, Kang L, Li Y. Sketch-based 3D shape retrieval using convolutional neural networks[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2015: 1875-1883
- [14] Bai Jing, Kong Dexin, Zhou Wenhui, *et al.* Joint feature mapping for end-to-end sketch-based 3D model retrieval[J]. Journal of Computer-Aided Design & Computer Graphics, 2019, 31(12): 2056-2065(in Chinese) (白静, 孔德馨, 周文惠, 等. 基于联合特征映射的端到端三维模型草图检索[J]. 计算机辅助设计与图形学学报, 2019, 31(12): 2056-2065)
- [15] Zhang Kaixing, Liu Honghao, Li Jinfeng, et al. Conceptual design push method of 3D CAD models based on sketches[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(2): 365-372(in Chinese) (张开兴, 刘洪豪, 李金凤, 等. 基于草图的产品三维 CAD模型概念设计推送方法[J]. 农业机械学报, 2019, 50(2): 365-372)
- [16] Hinton G E, Krizhevsky A, Wang S D. Transforming auto-encoders[C] //Proceedings of International Conference on Artificial Neural Networks. Heidelberg: Springer, 2011: 44-51
- [17] Ha D, Eck D. A neural representation of sketch drawings[OL]. [2020-11-11]. https://arxiv.org/abs/1704.03477
- [18] Liu Zhi, Yin Shichao, Pan Xiang, *et al.* 3D model retrieval method based on feature lines[J]. Journal of Computer-Aided Design & Computer Graphics, 2016, 28(9): 1512-1520(in Chinese)
 (刘志, 尹世超, 潘翔, 等. 基于特征线条的三维模型检索方法[J]. 计算机辅助设计与图形学学报, 2016, 28(9): 1512-1520)
- [19] Xu Xinjie, Wang Xiufeng, Lu Wenqi, *et al.* Development of part contour recognition system based on edge detection[J]. Journal of Mechanical & Electrical Engineering, 2019, 36(2): 201-205(in Chinese) (许鑫杰,王秀锋,鲁文其,等.基于边缘检测的零件轮廓识别系统开发[J]. 机电工程, 2019, 36(2): 201-205)
- [20] Zheng Jianhong, Bao Guanjun, Zhang Libin, et al. Metal part recognition based on deep learning and support vector machine[J]. Journal of Image and Graphics, 2019, 24(12): 2233-2242(in Chinese)
 (郑健红,鲍官军,张立彬,等.结合深度学习与支持向量机的金属零件识别[J].中国图象图形学报,2019,24(12): 2233-2242)
- [21] Zeng L, Dong Z K, Yu J Y, et al. Sketch-based retrieval and in-

- stantiation of parametric parts[J]. Computer-Aided Design, 2019, 113: 82-95
- [22] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C] //Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2005: 886-893
- [23] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60: 91-110
- [24] Yang Zhongliang, Wen Yangliang, Chen Yumiao. Hybrid KF-LSTM model for sEMG-based handwriting numeral traces reconstruction[J]. Journal of Computer-Aided Design & Computer Graphics, 2019, 31(7): 1247-1257(in Chinese) (杨钟亮,文杨靓,陈育苗. 基于 KF-LSTM 模型的手写数字 轨迹的 sEMG 重建算法[J]. 计算机辅助设计与图形学学报, 2019, 31(7): 1247-1257)
- [25] Razavian A S, Azizpour H, Sullivan J, et al. CNN features off-the-shelf: an astounding baseline for recognition[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Los Alamitos: IEEE Computer Society Press, 2014: 806-813
- [26] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[OL]. [2020-11-11]. https:// arxiv.org/abs/1409.1556

- [27] Deng J, Dong W, Socher R, et al. ImageNet: a large-scale hierarchical image database[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2009: 248-255
- [28] Devis N, Pattara N J, Shoni S, et al. Sketch based image retrieval using transfer learning[C] //Proceedings of the 3rd International Conference on Electronics, Communication and Aerospace Technology. Los Alamitos: IEEE Computer Society Press, 2019: 642-646
- [29] Mouffok M Z, Tabia H, Elhara O A. Dual independent classification for sketch-based 3D shape retrieval[C] //Proceedings of the IEEE International Conference on Image Processing. Los Alamitos: IEEE Computer Society Press, 2020: 2676-2680
- [30] Kumar A D. Novel deep learning model for traffic sign detection using capsule networks[OL]. [2020-11-11]. https://arxiv.org/abs/1805.04424
- [31] Lu Xuzheng, Cai Hengjin, Lin Li. Recognition of oracle radical based on the capsule network[J]. CAAI Transactions on Intelligent Systems, 2020, 15(2): 243-254(in Chinese) (鲁绪正,蔡恒进,林莉. 基于 Capsule 网络的甲骨文构件识别方法[J]. 智能系统学报, 2020, 15(2): 243-254)
- [32] Jiménez-Sánchez A, Albarqouni S, Mateus D. Capsule networks against medical imaging data challenges[M]. Heidelberg: Springer, 2018: 150-160