基于优化多视角图像采集的点云分类

何瑞函[1,2],蔡 勇[1,2],张建生[1,2]
(1.西南科技大学 制造科学与工程学院,四川 绵阳 621010;
2.制造过程测试技术省部共建教育部重点实验室,四川 绵阳 621010)

摘 要: 基于二维多视角3D识别方法中,可使用多个2D投影图表示三维模型特征信息,但不同视角投影图像的特征不同,神经网络对其学习效率也有所差异。卷积神经网络能够映射图像的特征,可用此方法分析这个问题。混合视角数据集分析不同视角投影特征在卷积神经网络中的重要性,根据重要性的不同优化混合视角数据集的采集密度。最终实验结果表明,不同视角产生的二维图像分类准确率不一样,其中俯视角度投影的分类准确率最差,优化后的数据集在相同神经网络模型下达到了最优分类准确率。

关键词: 三维点云;多视角图像;卷积神经网络;图像分类
中图分类号: TP391.7 **文献标识码:** A DOI:10.16157/j.issn.0258-7998.200892

中文引用格式: 何瑞函,蔡勇,张建生. 基于优化多视角图像采集的点云分类[J].电子技术应用,2021,47(10):82-85.
英文引用格式: He Ruihan,Cai Yong,Zhang Jiansheng. Point cloud classification based on optimized multi-view image acquisition[J]. Application of Electronic Technique,2021,47(10):82-85.

Point cloud classification based on optimized multi-view image acquisition

He Ruihan[1,2],Cai Yong[1,2],Zhang Jiansheng[1,2]
(1.School of Manufacturing Science and Engineering,Southwest University of Science and Technology,Mianyang 621010,China;
2.Key Laboratory of Testing Technology for Manufacturing Process,Mianyang 621010,China)

Abstract: In the 3D recognition method based on 2D multi-perspective, multiple 2D projected images can be used to represent the feature information of 3D model. However, the features of projected images from different perspectives are different, and the learning efficiency of the neural network is also different. The convolutional neural network can map the features of images, and this method can be used to analyze this problem. The importance of projection features of different perspectives in the convolutional neural network was analyzed in the mix-view data set, and the collection density of the mix-view data set was optimized according to the different importance. The final experimental results show that the classification accuracy of 2D images generated from different perspectives is different, among which the classification accuracy of overhead projection is the worst, and the optimized data set achieves the optimal classification accuracy in the same neural networks model.

Key words: 3D point cloud;multi-view image;convolution neural network;image classification

0 引言

随着激光扫描技术的发展,点云作为最能表现物体三维特征的数据深受研究者热爱。深度学习、卷积神经网络(Convolutional Neural Network,CNN)[1-2]近几年引领计算机视觉领域的研究趋势,并且CNN网络在二维图像分类与识别上显得高效。点云在空间中的无序性、旋转不变性、非结构化数据,导致其不能直接作为CNN网络的输入[3]。因此,使用深度学习对点云进行研究的方法有3种:多视图[4]、体素法[5]、直接对点云[6-7]。

基于二维多视角3D识别的方法,本文通过优化采集数据集的方式提升CNN神经网络[8]的分类效果。本文对点云模型进行不同视角投影,得到多组2D图像数据集。首先用多个VGG16卷积模型[9-10]提取单独视角数据集,得到多个映射了图像特征的卷积神经网络模型;然后将多个包含特征的VGG16模块加上自定义层后"并联"连接分类层作为分析网络模型,混合视角图像数据集作为网络输入;最后通过分析多个特征提取模块的权重,优化多视角图像的采集密度,提升二维多视角3D识别效率,即分类效果。

1 相关工作

文献[4]提出的多视角卷积神经网络(Multi-View CNN,MVCNN)将物体的三维数据通过多视角的"观察"得到二维图像,再将二维图像作为输入数据带入传统的二维图像卷积网络中进行训练。相比于一些直接用三维数据训练,所用的二维图像卷积网络非常成熟,其训练效果还能超过大量三维数据训练的结果。文献[5]的体素法的

核心是基于三维数据的“立体栅格化”，将点云数据表示为 30×30×30 大小的“立体栅格”，通过将空间中的点云划分为体素，用体素来表示三维物体，然后再对空间中的体素进行三维卷积。其卷积类似二维图像深度学习的卷积神经网络，所以计算在空间和时间上都很复杂。文献[6]直接将三维点云放入网络进行训练，基本思想是对输入点云每个点的空间编码进行学习。

多视图方法最大化利用了 CNN 的优点，文献[4]中提出 MVCNN 模型最有代表性。作者指出使用二维图像表示三维数据会丢失三维特征，为了减少特征的丢失，用多张不同角度的二维图像来代替三维模型。尽管将三维模型转为二维图像的这种方式很巧妙，但也存在未知因素，例如：是否某种视角的图像最重要，或者是否有某种图像包含了主要的特征信息。在文献[4]中，作者尚未阐述这些问题，因此本文将针对此问题进行分析。

2 数据集

观察三维物体有几种典型的视角，即从正上方俯视、斜上方俯视和侧面侧视，本文通过使用 Blender 软件在空间中设置虚拟相机对 ModelNet10 数据集的点云模型(.off)依次进行“拍照”提取 2D 图像，相机工作时指向 3D 网格数据的中心。

ModelNet10 为标记朝向的 ModelNet[11]数据集子集，共 10 类，训练集共有 3 991 个点云文件，测试集 908 个点云文件，本文数据集对 ModelNet10 所有点云文件进行采集。

如图 1 所示，每个虚拟相机分别以 $\alpha=0°,45°,90°$ 不同角度绕 Z 轴每隔 36°生成一个 2D 视角渲染图，每一个视角三维模型获得 10 个图像，一共得到 3 个子数据集 Data_0、Data_45、Data_90。子数据集用单独视角的 10 张二维图像表示一个点云。3 个子数据集微调 3 个 VGG16 卷积神经网络，分别储存不同 α 角度投影图像的特征信息。将 3 个子数据集打包为混合数据集(Data_mix)，作为调试模型的数据集，用共计 30(即 3×10)张二维图像表示一个点云模型。

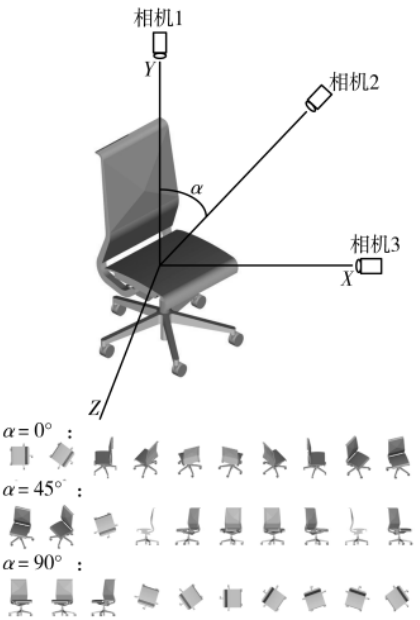

图 1 数据集采集

通过对 Blender 软件生成 2D 图像多次对比，最终在软件渲染中选择快照材质光照，颜色为纹理，像素为 300×300。用“resize_image_with_crop_or_pad”对数据集预处理，将图片裁剪为 224×224 的图像。

本文研究目的为，通过改变 30 张二维图像在 3 个不同视角中的采集密度，提高神经网络对点云的分类准确率。

3 卷积神经网络模型

以文献[9]中 VGG16 模型作为基础网络模型，包含 13 个卷积层、3 个全连接层，一共 16(即 13+3)层。在本文中使用到的包含 13 个卷积层、5 个池化层和 ReLU 激活函数。

通过卷积层、池化层和激活函数等操作，将原始数据集的特征信息映射到隐含层，再通过全连接层将隐含层的特征信息映射到样本标记空间进行分类[12]。本文用 Imagenet 数据集权重作为 VGG16 的预训练参数，再使用采集的数据集微调，最后用 Sotfmax 回归算法与交叉熵代价函数配合调节网络中的权重等参数[13]。

3.1 VGG16 卷积

以 VGG16 特征提取网络为基础，提取图像特征信息，将特征信息保存在 VGG16 的卷积层中。模型结构如图 2(a)所示，采集的 3D 模型子数据集输入到隐含层 VGG16(除全连接层)，图中的全连接层为分类空间，最后两层施加权重正则化防止网络过拟合。使用单视角投影数据集对 VGG16_α 模型进行微调，将单组视角 2D 图像的特征映射到 VGG16_α 的卷积层空间，3D 模型子数据集图像的特征信息储存在该卷积层空间，本文称之为特征匹配模块。

3.2 用于分析的网络模型

图 2(b)中用混合角度数据集(Data_mix)作为特征分析模型(VGG16_dbg)的输入；隐含层为特征匹配与调试空间，由 3 个 VGG16 卷积空间加上自定义层通过并联的方式组成。3 个特征匹配模块(VGG16_α)的权重等参数，通过迁移学习[14]的方式载入特征分析模型(VGG16_dbg)中。

自定义层(如图 2(c)所示)对每个特征匹配模块的输出上施以权重，在微调中评判每个特征匹配模块的权重大小，神经元表达式形如下：

$$(w_0x_0, w_0x_1, \cdots, w_0x_9) \tag{1}$$

concatenate 层(如图 2(d)所示)将 3 个 1×10 的神经元连接为一个(1×30)的网络层，如式(2)所示：

$$(w_0x_0, \cdots, w_1x_{12}, \cdots, w_2x_{29}) \tag{2}$$

最后与分类层相连。

为了使得自定义层中权重，与 3 个特征匹配模块在特征分析模型(VGG16_dbg)中的权重等价。将自定义层中权重初始值设置为相等(0 或 1)，3 个特征匹配模块(VGG16_0、VGG16_45、VGG16_90)在初始工作时具有一

《电子技术应用》http://www.chinaaet.com

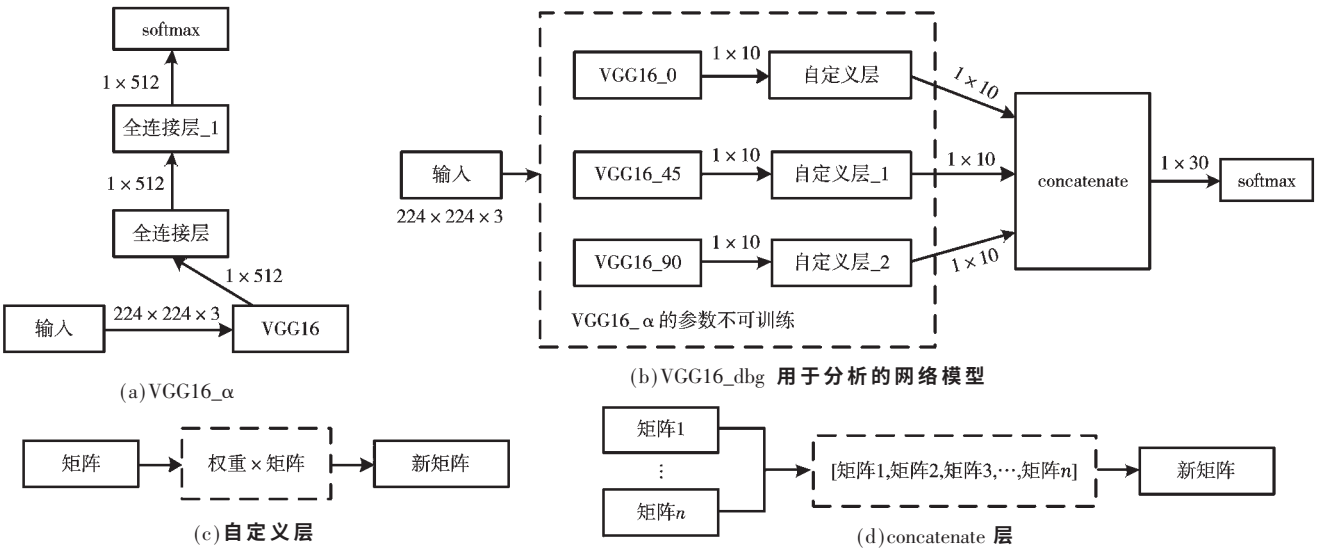

图 2 卷积网络模型

样条件，通过用“concatenate”连接层对自定义层的输出进行拼接时，自定义层中的权重不会被其他层的激活函数、损失函数影响[15]，自定义层中权重初始值设置为相等不会造成梯度消失、梯度爆炸或神经元输出始终不变等问题[16]。

4 实验

该实验计算机环境为：Intel－酷睿 i7－9750H 的 CPU，16 GB DDR4 内存，在 Windows 10 64 位操作系统下使用 GTX2060 6 GB 的 GPU，使用 TensorFlow 框架实现。

4.1 评价方式

文中采用两种评价标准，第一种为单张图像分类准确度，第二种为点云投影集合分类准确度。其中，第二种方式在点云识别中更具有说服性，将投影集合中预测的最大类(label)作为预测结果。本文以第二种评价方式作为最主要的评价标准。

4.2 微调结果与分析

从表 1 中可见，当相机与竖直方向夹角为 45°时，二维投影识别效果最好，在图 1 中 α=45°的图像也具有明显特征；相反 α=0°视角投影图像中包含较少特征，分类预测结果也是最差。

调试 VGG16_dbg 模型共有 313 个可训练参数，包括 3 个自定义层权重和分类层 310 个参数。多次实验发现，调试模型在第 5 个 epoch 时准确率达到峰值 66%，避免引起过拟合，特征分析模型 epoch=5。

从表 2 中的平均权重发现 w_0 的平均值最小，w_{45} 最大。α=45°视角的模型特征较为重要，按照权重的比例，对不同角度图像的数量进行增减，通过计算得到每一个视角的投影数 I，公式如下：

$$I=I_{\text{all}}\frac{\Delta\overline{w}_{\alpha}}{\sum\Delta\overline{w}_{i}} \tag{3}$$

式中，I_{all} 为总投影数，$\Delta\overline{w}_{\alpha}$ 为 α 视角特征匹配模块的权重绝对值均值，$\sum\Delta\overline{w}_{i}$ 为权重绝对值均值总和。

由式(3)得到视角 0°、45°、90°的视图分别为 7、12、11 张，共 30 张二维图像。按照此数量优化本文的训练集，最终 ModelNet10 数据集在 VGG16 网络模型上识别准确率准约为 93.6%，对比未优化的数据集约提高 1.8%。表 3 中，本实验方法与其他网络相比也达到了良好的分类准确率。

表 1 不同视角数据集准确率比较

网络模型	训练数据集			测试图像数据集视角/(°)	分类准确率/%	
	预训练	微调	视角 α/(°)(投影数)		单张图像	投影集合
VGG16	–	ModelNet10	45 (10)	45(10)	77.8	78.6
VGG16	ImageNet	ModelNet10	0 (10)	0 (10)	75.6	77.4
VGG16	ImageNet	ModelNet10	45(10)	45(10)	89.8	91.3
VGG16	ImageNet	ModelNet10	90(10)	90(10)	82.7	84.8
VGG16	ImageNet	ModelNet10	mix(30)	mix(30)	83.2	91.8
VGG16_dbg	ImageNet	ModelNet10	mix(30)	mix(30)	64.4	–
VGG16	ImageNet	ModelNet10	opt(30)	mix(30)	82.6	92.2
VGG16	ImageNet	ModelNet10	opt(30)	opt(30)	85.6	93.6

注：(1)VGG16_dbg 为文中分析权重模型，(2)mix 为 3 个视角的投影各 1 张，(3)opt 为优化后的混合数据集 30 张。

表 2 分析模型中各视角卷积网络的权重变化

视角	权重初始值为 0 的 3 次实验			权重初始值为 1 的 3 次实验			权重绝对值均值 $\Delta\overline{w}_\alpha$	权重占比 $\frac{\Delta\overline{w}_\alpha}{\sum\Delta\overline{w}_i}$/%	优化后投影数量/张
0°(w_0)	−0.37	0.49	0.37	0.38	0.49	0.31	0.53	24.1	7
45°(w_{45})	0.64	0.69	0.86	0.67	0.57	0.73	0.89	40.7	12
90°(w_{90})	0.54	0.75	0.74	0.64	0.64	−0.39	0.77	35.2	11

表 3 本文方法与其他算法的分类准确率比较

网络	数据集	分类准确率/%
Ma[17]	ModelNet10	95.29
VoxNet[5]	ModelNet10	92
3DShapeNets[18]	ModelNet10	83.5
PointNet[19]	ModelNet10	77.6
VGG16	ModelNet10(mix)	91.8
VGG16	ModelNet10(opt)	93.6

5 结论

本文优化 3D 模型投影图的采取方式，使 VGG16 网络模型对指定点云测试集的识别效果得到提升。构建特征网络权重，分析不同视角 3D 投影图的特征在同一个 CNN 网络下的差异性，并根据权重比例优化数据集的采集密度，使得 VGG16 网络学习到 3D 模型投影图的更多特征。实验结果证明，通过优化输入数据采集可以使网络模型学习到更多特征，达到更好的分类效果。

参考文献

[1] 周飞燕，金林鹏，董军.卷积神经网络研究综述[J].计算机学报，2017，40(6)：1229-1251.

[2] 孙志军，薛磊，许阳明，等.深度学习研究综述[J].计算机应用研究，2012，29(8)：2806-2810.

[3] 李海生，武玉娟，郑艳萍，等.基于深度学习的三维数据分析理解方法研究综述[J].计算机学报，2020，43(1)：41-63.

[4] SU H，MAJI S，KALOGERAKIS E，et al.Multi-view convolutional neural networks for 3D shape recognition[C]//2015 IEEE International Conference on Computer Vision，2015：2380-7504.

[5] MATURANA D，SCHERER S.VoxNet：a 3D convolutional neural network for real-time object recognition[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS)，2015：922-928.

[6] CHARLES R Q，SU H，KAICHUN M，et al.PointNet：deep learning on point sets for 3D classification and segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR)，2017：77-85.

[7] QI C R，YI L，SU H，et al.PointNet++：deep hierarchical feature learning on point sets in a metric space[C]//Proceedings of the IEEE International Conference on Intelligent Robots and Systems，2017：922-928.

[8] KRIZHEVSKY A，SUTSKEVER I，HINTON G.ImageNet classification with deep convolutional neural networks[C]//Proceedings of the 25th Conference on NIPS，2012，1：1097-1105.

[9] SIMONYAN K，ZISSERMAN A.Very deep convolutional networks for large-scale image recognition[J]. Computer Vision and Pattern Recognition，arXiv：1409.1556，2015.

[10] 字云飞，李业丽，孙华艳.基于深度神经网络的个性化推荐系统研究[J].电子技术应用，2019，45(1)：14-18，22.

[11] Wu Zhirong，Song Shuran，KHOSLA A，et al.3D ShapeNets：a deep representation for volumetric shapes[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR)，2015：1912-1920.

[12] 许少尉，陈思宇.基于深度学习的图像分类方法[J].电子技术应用，2018，44(6)：116-119.

[13] 黄睿，陆许明，邬依林.基于 TensorFlow 深度学习手写体数字识别及应用[J].电子技术应用，2018，44(10)：6-10.

[14] YOSINSKI J，CLUNE J，BENGIO Y，et al.How transferable are features in deep neural networks?[C]//Advances in Neural Information Processing Systems，2014：3320-3328.

[15] 仝卫国，李敏霞，张一可.深度学习优化算法研究[J].计算机科学，2018，45(S2)：155-159.

[16] 费建超，芮挺，周遊，等.基于梯度的多输入卷积神经网络[J].光电工程，2015，42(3)：33-38.

[17] MA C，GUO Y，YANG J，et al.Learning multi-view representation with LSTM for 3-D shape recognition and retrieval[J].IEEE Transactions on Multimedia，2018，21(5)：1169-1182.

[18] WU Z，SONG S，KHOSLA A，et al.3D ShapeNets：a deep representation for volumetric shapes[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR)，2015：1912-1920.

[19] GARCIA-GARCIA A，GOMEZ-DONOSO F，GARCIA-RODRIGUEZ J，et al.PointNet：a 3D convolutional neural network for real-time object class recognition[C]//International Joint Conference on Neural Networks(IJCNN)，2016：1578-1584.

(收稿日期：2020-09-09)

作者简介：

何瑞函(1995-)，通信作者，男，硕士研究生，主要研究方向：反求工程与快速制造、图像处理，E-mail：574549063@qq.com。

蔡勇(1962-)，男，博士，教授，主要研究方向：反求工程、计算机视觉、智能制造等。

张建生(1980-)，男，博士研究生，讲师，主要研究方向：反求工程、计算机图形学等。

扫码下载电子文档

《电子技术应用》http://www.chinaaet.com

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所