E-mail: csa@iscas.ac.cn http://www.c-s-a.org.cn Tel: +86-10-62661041

# Point-GBLS: 结合深宽度学习的三维点云分类 网络<sup>①</sup>

张国有<sup>1</sup>, 左嘉欣<sup>1</sup>, 潘理虎<sup>1</sup>, 郝志祥<sup>1</sup>, 郭 伟<sup>1</sup>, 张雪楠<sup>2</sup>

<sup>1</sup>(太原科技大学 计算机科学与技术学院,太原 030024) <sup>2</sup>(中国联合网络通信有限公司 政企客户事业群,北京 100033) 通信作者:张国有, E-mail: zhangguoyou@tyust.edu.cn

**摘 要**: 基于点云的三维物体识别和检测是计算机视觉和自主导航领域的一个重要研究课题. 如今, 深度学习算法 大大提高了三维点云分类的准确性和鲁棒性. 然而, 深度学习网络通常存在网络结构复杂、训练过程耗时等问题. 本文提出了一种三维点云分类网络 Point-GBLS, 它将深度学习和宽度学习系统结合在一起. 网络结构简单, 训练时 间短. 首先通过基于深度学习的特征提取网络提取点云特征, 然后用改进的宽度学习系统对其进行分类. ModelNet40 和 ScanObjectNN 数据集上的实验表明, Point-GBLS 识别准确率分别达到 92% 以上和 78% 以上, 训练 时间低于同类深度学习方法的 50% 以上, 优于具有相同骨干的深度学习网络.

关键词:三维模型分类;点云;深度学习;宽度学习系统

引用格式:张国有,左嘉欣,潘理虎,郝志祥,郭伟,张雪楠.Point-GBLS:结合深宽度学习的三维点云分类网络.计算机系统应用,2025,34(3):1-13. http://www.c-s-a.org.cn/1003-3254/9796.html

# Point-GBLS: 3D Point Cloud Classification Network Combined with Deep-broad Learning

ZHANG Guo-You<sup>1</sup>, ZUO Jia-Xin<sup>1</sup>, PAN Li-Hu<sup>1</sup>, HAO Zhi-Xiang<sup>1</sup>, GUO Wei<sup>1</sup>, ZHANG Xue-Nan<sup>2</sup>

<sup>1</sup>(Computer Science and Technology Academy, Taiyuan University of Science and Technology, Taiyuan 030024, China) <sup>2</sup>(Government and Enterprise Customer Business Group, China United Network Telecommunications Co. Ltd., Beijing 100033, China)

Abstract: 3D object recognition and detection based on point clouds is an important research topic in the fields of computer vision and autonomous navigation. Nowadays, deep learning algorithms have greatly improved the accuracy and robustness of 3D point cloud classification. However, deep learning networks usually have problems such as complex network structure and time-consuming training. This study proposes a three-dimensional point cloud classification network named Point-GBLS, which combines deep learning and a broad learning system. The network structure is simple and the training time is short. Firstly, point cloud features are extracted by a deep learning-based feature extraction network. Then, an improved broad learning system is used to classify them. Experiments on the ModelNet40 and ScanObjectNN dataset show that the recognition accuracy of Point-GBLS is more than 92% and 78% respectively. The training time is less than 50% of that of similar deep learning methods. It is superior to deep learning networks with the same backbone.

Key words: 3D model classification; point cloud; deep learning; broad learning system

3D 点云模型分类的研究意义首先在于提高计算 机视觉和机器人领域的自动化和智能化水平,使得机 器能够更好地理解和识别环境中的物体,从而做出更准确的决策.例如,在自动驾驶和机器人导航领域,通

基金项目:山西省自然科学基金 (202203021221145);国家自然科学基金 (62072325) 收稿时间: 2024-09-07;修改时间: 2024-09-30;采用时间: 2024-10-14; csa 在线出版时间: 2025-01-16 CNKI 网络首发时间: 2025-01-17

过三维点云分类技术,可以实时检测和识别道路上的 交通标志和障碍物,如车辆、行人、建筑物等.为环 境感知和决策提供重要的信息,从而实现安全、高效 的导航和路径规划.其次,可以推动三维模型重建技 术的发展.通过三维点云分类技术,为三维模型重建 提供了新的方法和思路.不仅优化了三维模型重建的 效率,提高了重建模型的质量,还促进了自动化重建 流程的发展.因此可以实现精确建模和重建,可用于 地图构建和重建领域.并且对增强现实 (AR) 和虚拟 现实 (VR) 具有重要意义. 通过对现实世界中的点云 模型进行分类,可以实现对虚拟对象的准确定位和识 别,为增强现实和虚拟现实应用提供更真实、逼真的 体验.此外,有助于提高场景理解和语义分割的准确 性. 通过特征提取、分类分割等步骤, 可以对点云数 据进行精细分类,实现对环境中不同物体的识别和理 解. 这不仅有助于机器更好地识别出复杂的结构细节, 提高空间布局分析的准确度,还能更好地理解整体场 景的组成、布局和动态交互. 综上所述. 3D 点云模型 分类的研究不仅对推动计算机视觉、机器人技术、 自动驾驶等领域的发展具有重要意义,还有助于提高 三维模型重建技术的效率和准确性,从而推动相关领 域的进步和创新.

传统的识别方法是通过视觉传感器 (如普通相机 和 RGB-D 相机) 获得二维图像信息, 从而进行分类检 测. 但其易受其工作条件的影响, 如光线条件、视角、 遮挡等. 而激光雷达通常具有更强的抗干扰能力, 与二 维图像相比, 通过激光雷达得到的点云包含更多物体 和环境的三维信息. 近年来, 基于深度学习的物体分类 和检测得到了前所未有的发展. 深度学习网络能更有 效地提取特征, 取得更准确的结果. 因此, 基于深度学 习的三维点云处理算法已成为计算机视觉领域的重要 研究课题, 并被广泛应用于自动驾驶和机器人领域.

基于深度学习的三维点云处理算法一般包括基于 体素的算法<sup>[1-3]</sup>、基于视图的算法<sup>[4-6]</sup>和基于点的算 法<sup>[7-11]</sup>.基于体素的算法通常将点云体素化,并使用三 维卷积神经网络 (3D-CNN) 提取三维特征.而基于视 图的算法通常将点云投影到二维图像上,并使用二维 深度学习方法进行处理.这两种算法都需要对原始点 云数据进行预处理,不可避免地会造成三维信息丢失. 然而,基于点的算法直接将点坐标作为输入,通过端到 端网络完成分类和回归任务.此外,基于深度学习的算 法大大提高了三维物体分类、检测的效率和准确性. 然而,它们仍然存在网络结构复杂、计算量大、训练 耗时长等缺点.

为避免深度学习网络的上述缺点,提出了两种改 进方法. 一种是极限学习机 (extreme learning machine, ELM), 由 Huang 等<sup>[12]</sup>提出的 ELM 是一种具有单隐层 的简单神经网络. Huang 等<sup>[13]</sup>和 Zhang 等<sup>[14]</sup>验证了 ELM 的有效性.在 ELM 中, 输入权重和隐层偏差是随机生 成的. 然后, 通过 Moore-Penrose 广义反演[15]分析计算 隐层和输出层之间的连接权重. 与传统的训练方法相 比,该方法具有学习速度快、泛化性能好等优点.但极 限学习机在解决复杂大规模高维数据问题时,因为深 度有限、结构简单(通常只包含一层隐藏层),而限制 了模型的复杂性和表达能力,往往容易出现函数表征 能力不足的情况.并且鲁棒性不足,对于输入特征的变 化敏感,如果数据分布发生变化或者噪声较大,可能导 致性能下降.同时,虽然 ELM 的权重参数可以在训练 过程中随机初始化,但这意味着它们缺乏传统的正则 化手段,容易导致过拟合.

另一种是 Chen 等<sup>[16]</sup>于 2017 年提出的宽度学习系统 (broad learning system, BLS). BLS 的网络结构相对简单, 仅由 3 层神经元组成, 其训练过程通过伪逆计算实现. BLS 具有网络结构简单、网络参数少、训练时间短等优点. 但 BLS 仍有不足之处, 其权重和偏差是随机生成, 不能保证所形成节点的质量, 这意味着 BLS可能无法学习到足够有用的原始数据表征. 因此, Liu 等<sup>[17]</sup>提出了 GBEAE-BLS, 权重是计算出来的, 在整个训练过程中无需进行微调. 投影的增强节点会考虑类内和类间的关系. 然而, 与二维图像不同, 三维点云通常是稀疏和无序的. 这些特点使得 GBEAE-BLS 无法直接处理点云数据.

为充分利用深度学习和宽度学习算法的优势,本 文提出一种新的三维点云分类网络 Point-GBLS.在 Point-GBLS 中,首先使用深度学习模型 PointNet++<sup>[8]</sup> 提取点云特征.然后,将提取的点云特征作为输入,用 GBEAE-BLS 完成分类任务.在 ModelNet40 数据集<sup>[18]</sup> 和 ScanObjectNN 数据集上的实验表明,所提出的 Point-GBLS 的最高准确率分别为 92.4% 和 78.3%,训练时间 也大幅缩短. Point-GBLS 大大加快了三维分类算法的 训练过程,同时保证了物体识别的准确性.

本文的主要贡献为:1) 首次提出了深宽度相结合

2 专论•综述 Special Issue

的 Point-GBLS 网络, 通过改进的 PointNet++网络模型 对点云特征进行聚合, 以提高网络的特征提取能力. 2) 使用 GBEAE-BLS 网络模型进行分类以缩短网络训 练时间, 保证分类准确率. 3) 本文算法在 ModelNet40 和 ScanObjectNN 分类数据集上表现优异.

本文第1节对三维点云处理方法进行相关研究, 第2节对 Point-GBLS 分类算法展开具体介绍,第3节 进行实验验证,第4节总结全文并讨论发展趋势.

## 1 相关研究

最初,通常采用传统方法从点云中提取特征,如 PFH<sup>[19]</sup>、FPFH<sup>[20]</sup>、SHOT<sup>[21]</sup>等.这些方法在点云分类 和配准方面取得了初步成效.然而,手工特征通常具有 计算复杂、准确率低和鲁棒性差等缺点.目前,随着人 工智能的发展,基于深度学习的点云分类和物体检测 方法开始出现.与传统方法相比,深度学习可以更好地 研究点云的高层次特征,从而显著提高识别准确率.基 于深度学习的三维点云处理方法基本上可分为6类: 基于体素的方法、基于视图的方法、基于点的方法、基 于卷积的方法、基于注意力的方法和基于 Transformer 的方法.

## 1.1 基于体素的方法

基于体素的方法是将三维空间平均划分为小的三 维体素网格,这类似于二维图像中的像素.通过对 3D 点云数据进行体素化处理,3D-CNN 可用于进一步 从点云中提取特征<sup>[1-3]</sup>.然而,三维卷积操作的复杂度 非常高.为了降低 3D-CNN 的计算成本,基于体素的方 法通常分辨率较低,从而限制了物体识别的准确性.

## 1.2 基于视图的方法

基于视图的方法是将点云投影到某个平面上,从 而获得多视图投影图像. 然后,可以使用二维深度学习 网络对多视角投影进行进一步处理<sup>[4-6]</sup>. 这些方法可以 通过减少数据维度来有效降低计算成本,但也不可避 免地会造成三维信息丢失. 因此,这些方法的鲁棒性通 常较低.

# 1.3 基于点的方法

上述两类方法都需要对原始点云数据进行复杂的 预处理.点云数据只有在转换为其他数据格式后,才能 由深度学习网络进行处理.在数据转换过程中,信息丢 失是不可避免的.此外,转换后的数据不能完全反映原 始点云数据的所有特征.为解决这个问题,出现了基于 点的分类方法<sup>[7,8]</sup>. PointNet++直接将点云的三维坐标 作为输入,无需任何预处理操作,从而实现点云的端到 端分类. PointNet++解决了点云的无序问题、特征提取 不充分、采样密度不均匀问题. 自 PointNet++提出以 来,许多其他三维点云分类和物体检测网络都是以 PointNet++为骨干构建的<sup>[22]</sup>.

## 1.4 基于卷积的方法

由于直接用与点相关的特征对核进行卷积会导致 丢弃形状信息和点排序的差异, Li 等<sup>[23]</sup>提出 PointCNN 来解决这个问题,这证实了发展局部结构对于点云分 类网络的重要性.由于点云的稀疏性、不规则性和无 序性, 很难直接对其进行卷积运算. Wu 等<sup>[24]</sup>提出将动 态滤波器应用于卷积运算,称为 PointConv,这种方法 既简单又能减轻计算机存储压力. 单铉洋等[25]提出了 用于三维点云分类的卷积神经网络 RFNet, 通过学习 中心点特征与近邻点特征之间的关系,为不规则的近 邻点分配不同权重,以此构建局部结构.然后,使用注 意力思想,提出加权平均池化 (weighted average pooling, WAP), 通过自注意力方式, 学习每个高维特征的注意 力分数,在应对点云无序性的同时,可以有效地聚合冗 余的高维特征.最后,利用交叉熵损失与中心损失之间 的互补关系,提出联合损失函数 (joint loss function, JL), 在增大类间距离的同时, 减小类内距离, 进一步提 高了网络的分类能力.

## 1.5 基于注意力的方法

Chen 等<sup>[26]</sup>提出的 GAPointNet 将自注意机制与图 卷积相结合,通过在堆叠的 MLP 层中嵌入图注意机制 来学习局部信息表示,并使用并行机制来聚合不同 GAPLayer 层的注意特征,其中 GAPLayer 层和注意层 可以嵌入到现有的训练模型中,从而更好地从无序点 云中提取局部上下文特征. 史豪斌等<sup>[27]</sup>提出的融合注 意力机制与神经网络的三维点云分类算法,使用 K 近 邻方法将点云划分为局部区域,对点与点之间的关系 进行建模以同时获取局部特征与全局特征. 注意力机 制为特征向量进行权重分配,使得最终提取的特征向 量更加鲁棒且更具有区分性,对有效特征增大权重,对 相对无效的特征进行抑制. 最后,通过混合池化层进行 聚合,混合池化由最大池化和平均池化按照不同的比 例融合,从而降低了特征的信息损失.

## 1.6 基于 Transformer 的方法

自 2017 年首次提出 Transformer 以来, 它在计算

机视觉领域取得了举世瞩目的成果. Guo 等<sup>[28]</sup>提出一种基于 Transformer 的点云学习框架-点云变换器 (point cloud Transformer, PCT),并提出带有隐式拉普拉斯算 子和归一化细化的偏移注意,该框架具有包覆不变性,更适合点云学习. He 等<sup>[29]</sup>引入了基于体素的集合关注 模块 (VSA),建立了体素集合转换器 (VoxSeT)架构. VoxSeT 可通过 VSA 模块管理点簇,并以线性复杂度 并行处理点簇. 这种方法结合了 Transformer 的高性能 和基于体素模型的高效率,在点云建模中具有良好的 性能.

深度学习方法提高了三维点云分类的准确性和鲁 棒性. 然而, 这些方法的训练过程耗时较长、网络结构 复杂、计算量大. 为了克服上述缺点, 提出了一些改进 方法. 一种是极限学习机 (ELM), 该算法随机产生输入 层与隐含层间的连接权值及隐含层神经元的阈值, 且 在训练过程中无需调整, 只需要设置隐含层神经元的 个数, 便可以获得唯一的最优解. 为了学习有意义的特 征, Kasun 等<sup>[30,31]</sup>引入了基于 ELM 的自动编码器 (ELM-AE) 作为表征学习技术, 并在 ELM-AE 的基础上给出 了多层 ELM (ML-ELM). 在 ML-ELM 中, 最后一个隐 藏层和输出层之间的连接权重由正则化最小二乘法确 定, 其他连接权重由 ELM-AE 初始化. 上述所有连接权 重都无需微调. 目前, ELM-AE 已成为构建具有深层结 构的 ELM 相关模型的基本方法.

另一种是以随机向量函数链接神经网络 (random vector functional-link neural network, RVFLNN)<sup>[32]</sup>为灵感, 设计了宽度学习系统 (BLS)<sup>[16]</sup>. 与深度学习网络相比, BLS 简化了训练过程, 减轻了计算负担. 但 BLS可能无法充分学习原始数据的有用表示, 而 GBEAE-BLS 是基于图的 ELM-AE (GBEAE), 然后将其应用于初始化连接权重, 用于获取映射特征. 初始化的权重是直接计算出来的, 在整个训练过程中无需对这些权重进行微调. 在获得映射特征后, GBEAE-BLS 中形成的增强节点会被投射到输出空间, 其中会考虑类内的紧凑性和类间的可分离性. 目前, BLS 网络已在多个二维图像分类任务中得到应用<sup>[33-38]</sup>, 但在三维点云分类方面却鲜有建树. BLS 在三维数据处理中仍有广阔的应用前景.

# 2 Point-GBLS 分类算法

Point-GBLS 的网络结构如图 1 所示,由图可知

4 专论•综述 Special Issue

Point-GBLS 网络由两部分组成,分别是横向的特征提 取网络和纵向的分类网络.图1中,特征提取网络采用 改进的 PointNet++的特征提取网络部分. 先使用 MLP 模块对点的特征升维,也就是图中的输入嵌入.然后通 过采样、分组和点网3个操作来提取点云的高维特征, 这3个操作统称为集合抽象层,其中,采样是通过迭代 选择距离当前已选点最远的点作为新的采样点,从而 保证了采样点的均匀分布.这有助于模型更好地学习 到点云数据的空间结构;分组是通过逐层抽样的方式, 实现了对点云数据的分层特征学习.这种方式使得模 型能够同时捕捉点云的局部和全局信息,提高了特征 表达的丰富性和准确性; 点网是通过局部特征聚合方 式使得模型能够充分利用邻域信息,提高了特征的鲁 棒性.特征提取网络具体操作将在第2.1节中具体展开 阐述. 分类网络采用改进的 GBEAE-BLS 网络, 将提取 的点云特征送入该网络,首先根据 GBEAE 算法学习 映射层权重,通过映射层形成映射特征,将映射特征送 入到增强层得到增强节点后,我们将增强节点投影到 投影增强层得到投影增强节点.最后,将映射特征和投 影增强节点共同输入到输出层(脊回归模型)得到分类 结果.其中,映射层初始化的权重是计算得出的,而不 是随机产生,更加合理.在图的右侧小正方形内,用相 同颜色标记的圆圈表示来自同一类别,用来自两个类 别的样本来展示样本的分布.在投影增强层中考虑到 之前从未考虑的类内和类间的关系,使得分类结果更 加准确. 分类网络的具体操作将在第 2.2 节中具体展开 阐述.

## 2.1 特征提取网络

特征提取网络如图 1 左边所示, 特征提取网络基 于 PointNet++骨干网构建而成<sup>[8]</sup>. 一个点云 P 是由 N 个点组成. 每个点都有其三维坐标特征 (x, y, z), 因此 在本文中 d=3, 对于分类来说点云数据的点数 N 通常为 1 024, 因为在处理大规模点云数据时, 直接使用所有 点可能会导致计算复杂度过高, 影响处理效率和实时 性. 通过选取一定数量的点 (如 1 024 个), 可以保证模 型输入的一致性和稳定性, 降低计算成本, 减少存储需 求, 简化网络结构设计和优化过程, 满足实时性要求, 关键的信息可以通过较少数量的点来捕获, 减少超参 数调整的难度, 同时使得模型在不同场景下的部署更 为方便和统一. 因此输入到网络中的一个点云数据 P 表示如式 (1):

$$P = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_N & y_N & z_N \end{bmatrix}$$
(1)

将原始点云数据P∈ℝ<sup>N×3</sup>作为输入,送入到 MLP 模块进行升维,使得原有特征升到更高维度上去,这样 做在后续处理过程中有利于减少特征在传递过程中的 损失,升维后 C=128. MLP 之后进行归一化批处理,激 活函数设置为 ReLU. 然后送入到集合抽象层进行相应 操作.集合抽象层由采样层、分组层、点网层组成.

分组层: 以采样层得到的采样点为中心, 在给定半径 R 下进行球查询 (ball query), 如图 2 所示. 球查询就 是在给定半径范围内选取 K 个点, 若给定范围内不足 K 个点, 则选择距离中心点最近的点, 缺少几个就相应 复制几次, 若给定范围内点数多余 K 个, 则选取距离中 心点最近的前 K 个. 由这些中心点和相应邻居点形成 了局部邻域. 其局部邻域保证了固定的区域尺度, 从而 使局部邻域特征在空间上更具泛化性.



## 图 1 Point-GBLS 网络结构图

考虑到点集在不同区域的密度不均匀现象,采用 多分辨率分组 (multi-resolution grouping, MRG)方法, 如图 3 所示. 它是指某个层级的区域特征是两个向量 的连接,其中一个向量 (图中左侧)是利用集合抽象层 对下层中每个子区域的特征进行汇总后得到的特征. 另一个向量 (图中右侧)是使用单个点网直接处理局部 区域内所有原始点得到的特征. 当局部区域的密度较 低时,第1个向量的可靠性会低于第2个向量,因为计 算第1个向量的子区域包含的点更少,采样不足的情 况更严重. 在这种情况下,第2个向量的权重应该更高. 另一方面,当局部区域的密度较高时,第1个向量的可 靠性会高于第2个向量,因为第1个向量能提供更详 细的信息, 它具有在较低层次递归检测更高分辨率的 能力. 在经过 MRG 之后, 提取到的特征更丰富.

采样层:对升维后的点云数据进行最远点采样 (farthest point sampling, FPS), FPS 算法是一种有效的 点云采样方法,它通过迭代选择距离当前已选点最远 的点作为新的采样点,从而保证了采样点的均匀分布. 这种采样策略有助于模型更好地学习到点云数据的空 间结构.

点网层:对分组层得到的每一个局部邻域运用 PointNet 的操作,先升维,再进行最大池化操作,从而得 到局部邻域的全局特征.

在经过第1个集合抽象层后,原始的点云数据

 $P \in \mathbb{R}^{1024\times3}$ 变为 $P_1 \in \mathbb{R}^{512\times256}$ ,点云数据的点数减少,特征维度提高.这样的集合抽象层操作可以进行多次,在经过第 2 个集合抽象层之后,输出的是高维点云特征  $P_2 \in \mathbb{R}^{256\times512}$ .最后通过点网层的操作从 $P_2$ 中提取点云特征,得到考虑局部邻域的点云特征 $F \in \mathbb{R}^{1024}$ .



## 2.2 分类网络

分类网络如图 1 右边所示,将前述的特征提取网络的输出特征作为输入,使用 GBEAE-BLS<sup>[17]</sup>网络完成分类任务.由图可知分类网络首先通过 GBEAE 算法学习权重,之后经过映射层、增强层、投影增强层,最后通过输出层输出分类结果.

原始的 BLS 网络中映射层的权重和偏差是随机 生成的,并通过稀疏自动编码器进行微调,但稀疏自动 编码器只考虑构建良好的表示来重建原始数据,而忽 略了基本事实的信息.所以在这里使用 GBEAE 算法 来初始化连接权重,这些权重用于替换随机生成的权 重和偏差.这些权重是通过算法计算得到的,之后不需 要微调.在 GBEAE 中,关注类内的紧凑性和类间的可 分离性.这意味着在所需的类内图中,来自同一类别的 所有样本都是相连的,如果两个样本来自不同类别,则 连接权重为 0.同时,在类间图中,来自不同类别的样本 都是相连的,来自同一类别的所有样本连接权重为 0. 类内图和类间图的权重定义为:  $G_{\text{intra}}^{ij} = \begin{cases} 1, & \pi(x_i) = \pi(x_j) \\ 0, & \text{otherwise} \end{cases}$ (2)

$$G_{\text{inter}}^{ij} = \begin{cases} 1, & \pi(x_i) \neq \pi(x_j) \\ 0, & \text{otherwise} \end{cases}$$
(3)

其中,  $\pi(x_i) = \pi(x_j)$ 表示  $x_i$  和  $x_j$  来自同一类别, 而 $\pi(x_i) \neq \pi(x_i)$ 表示两个样本来自不同类别.

通过式 (4) 计算类内的紧凑性:

$$\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} G_{intra}^{ij} \left\| g(x_i) - g(x_j) \right\|_F^2 = \operatorname{tr}(S^{\mathrm{T}} L_1 S) \qquad (4)$$

类间的可分离性的计算公式为式(5):

$$\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} G_{\text{inter}}^{ij} \left\| g(x_i) - g(x_j) \right\|_F^2 = \text{tr}(S^{\mathsf{T}} L_2 S)$$
(5)

其中,  $g(x_i)$  是样本  $x_i$  的预测结果,  $S=[g(x_1); \dots; g(x_N)]$ 和 tr(·) 是矩阵的迹.  $L_1=D_1-G_{intra}, L_2=D_2-G_{inter}. D_1 = diag(d_1^1, \dots, d_N^1), D_2 = diag(d_1^2, \dots, d_N^2), d_i^1 = \sum_{j=1}^N G_{intra}^{ij}, d_i^2 = \sum_{j=1}^N G_{inter}^{ij}.$ 式 (4) 和式 (5) 可以合并为:

$$\frac{\eta}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} G_{intra}^{ij} \|g(x_i) - g(x_j)\|_F^2$$
$$- \frac{1 - \eta}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} G_{inter}^{ij} \|g(x_i) - g(x_j)\|_F^2$$
$$= \operatorname{tr}(S^{\mathrm{T}}LS)$$
(6)

其中,  $L = \eta L_1 - (1 - \eta)L_2$ ,  $\eta$ 是用于平衡  $L_1$  和  $L_2$  的关系的系数. 因此, GBEAE 的损失函数表述为:

 $O_{\text{GBEAE}} = \frac{1}{2} \|HW - X\|_{F}^{2} + \frac{\lambda_{1}}{2} \|W\|_{F}^{2} + \frac{\alpha}{2} \operatorname{tr}(S^{\mathsf{T}}LS)$ (7)  $+ \sum_{F} S = HW, \lambda_{1} \ge \|W\|_{F}^{2} \text{ bre} m$ 

空间中类内的紧凑性和类间的可分离性的平衡参数.

式(7)的梯度为:

$$\nabla O_{\text{GBEAE}} = H^{\mathrm{T}} H W + \lambda_1 W - H^{\mathrm{T}} X + \alpha H^{\mathrm{T}} L H W \quad (8)$$
  
令  $\nabla O_{\text{GBEAE}} = 0.$  有:

$$W = (H^{\mathrm{T}}H + \lambda_1 I + \alpha H^{\mathrm{T}}LH)^{-1}H^{\mathrm{T}}X$$
(9)

如果隐藏节点数量大于训练样本数量,则应将式(9) 改写为:

$$W = H^{\mathrm{T}} (\lambda_1 I + H H^{\mathrm{T}} + \alpha L H H^{\mathrm{T}})^{-1} X$$
(10)

因此可以得到 GBEAE 算法, GBEAE 算法如算法 1 所示.

6 专论•综述 Special Issue

#### 算法 1. GBEAE 算法

输入: 训练集 $\{X, Y\}=\{x_{i,y_{j}}\}_{i=1}^{N}$ , 隐藏节点数 l, 超参数 $\eta, \lambda_{1}, \alpha$ . 输出: 所需的连接权重 W.

首先,使用 GBEAE 算法得到权重. 然后,在映射 层,使用得到的权重将 $F \in \mathbb{R}^{1024}$ 的点云特征映射到  $A_1 \in \mathbb{R}^{N_1 \times N_2}$ 中(其中,  $N_1$ 是每个窗口的节点数,  $N_2$ 是映 射层的窗口数),得到映射特征:

$$Z_i = \phi(XW_i), \, i = 1, \cdots, n \tag{11}$$

其中, *W<sub>i</sub>* 是由 GBEAE 算法求得的权重, *φ*(·)为非线性 函数, 所有的映射节点为 *Z*<sup>*n*</sup>=[*Z*<sub>1</sub>,···,*Z<sub>n</sub>*].

与原始的 BLS 操作类似, 在增强层中, 将 $A_1 \in \mathbb{R}^{N_1 \times N_2}$ 作为输入, 用随机生成的 $\tilde{W}_{hj}$ 得到增强节点 $A_2 \in \mathbb{R}^{N_3}$ .

第j组增强节点的公式为:

$$H_j = \xi_j(Z^n \tilde{W}_{hj}), j = 1, 2, \cdots, m$$
 (12)

其中,  $H_j \in Z^n$  的高维特征, 所有的增强节点为 $H^m = [H_1, H_2, \dots, H_m]$ . 由于 $\tilde{W}_{hj}$ 是随机生成的, 无法保证属于同一类别或不同类别的样本都能得到适当的表示. 因此, 在投影空间中, 假定同一类别的样本在几何上相互接近,

而不同类别的样本在几何上分布相互远离. 所以, 在投 影增强层, 使用  $E_j$  表示第 j 组投影增强后的特征,  $E_j = H^m \tilde{W}_j$ .  $E^m = [E_1, E_2, \dots, E_m]$ , 从而得到投影增强节 点 $A_3 \in \mathbb{R}^{N4}$ . 其中 $\tilde{W}_j$ 由式 (13) 计算得到:

$$\widetilde{W}_{j} = \underset{\widetilde{W}, (\widetilde{W})^{\mathrm{T}} \widetilde{W} = I}{\operatorname{argmin}} \left[ \frac{\eta}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} G_{\mathrm{intra}}^{ij} \|E_{i} - E_{j}\|_{F}^{2} - \frac{1 - \eta}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} G_{\mathrm{inter}}^{ij} \|E_{i} - E_{j}\|_{F}^{2} \right] \\
= \underset{\widetilde{W}, (\widetilde{W})^{\mathrm{T}} \widetilde{W} = I}{\operatorname{argmin}} \operatorname{tr}[(\widetilde{W})^{\mathrm{T}} (H^{m}) L H^{m} \widetilde{W}], \\
i = 1, 2, \cdots, m \tag{13}$$

式 (13) 可通过  $(H^m)^T L H^m$  的广义特征值分解来求解,  $\tilde{W}_i$  由 M 个除 0 以外的最小特征值对应的特征向量组成.

最后,在输出层,将A = [A<sub>1</sub>|A<sub>3</sub>]作为输入,并输出分 类分数.此外,映射层和输出层的输出没有任何激活函 数,只有增强层有激活函数.

在训练 Point-GBLS 时, 改进的 PointNet++特征提 取网络和 GBEAE-BLS 分类网络是分开训练的. 如图 4 所示, 在训练特征提取网络时, 将 GBEAE-BLS 分类网 络替换为全连接层. 层的大小设置为 (512, 256, *K*), 其 中 *K* 是类的数量. 该网络可以使用经典的深度学习训 练方法进行训练. 具体来说, 使用 Adam 作为优化方法, 交叉熵损失作为损失函数.



图 4 改进的 PointNet++分类网络结构图

至于 GBEAE-BLS 分类网络的训练, 映射层的参数由 GBEAE 算法生成, 增强层的权重和偏差随机设置为正交基, 然后在投影增强层通过计算得到的权重

进一步形成投影增强节点.由于输出层没有激活函数, 该层的参数可以通过伪逆变换确定.具体来说,在 GBEAE-BLS 的训练过程中,伪逆问题是通过脊回归算

法来解决的.

用于决策的线性模型的损失函数可写成:

$$O_L = \frac{1}{2} \|Q\beta - Y\|_F^2 + \frac{\lambda_2}{2} \|\beta\|_F^2$$
(14)

其中, $Q = (Z^n, E^m)$ ,  $\beta$ 是线性模型的权重,  $\lambda_2$ 是超参数, 是权重 $\beta$ 的约束条件,防止出现奇异性问题.

令
$$\frac{\partial O_L}{\partial \beta} = 0$$
,可以得到:  
 $\beta = (\lambda_2 I + Q^T Q)^{-1} Q^T Y$  (15)

综上所述, Point-GBLS 分类算法如算法 2 所示.

算法 2. Point-GBLS 分类算法

- 输入:训练集{X,Y},用于获取Zn的隐藏节点数1,增强节点数j=1, 2,…, m 用于获取 H<sub>i</sub>. 超参数η, λ<sub>1</sub>, λ<sub>2</sub>, α. 输出: 样本 x 的预测标签.
- 1) 将训练集送入到改进后的 PointNet++特征提取网络进行特征提 取,将点云特征提取到更高维度; 2) 根据算法 1 获得 W<sub>i</sub>, 参数为 l, η, λ<sub>1</sub>, α; 3) 根据式(11) 计算所有的映射特征 Z"; 4) 根据式 (12) 计算所有的增强节点 H<sup>m</sup>; 5) 根据式 (13) 计算权重  $\tilde{W}_{j}$ , 然后得到投影增强节点  $E^{m}$ ; 6) 将所有的特征设为 Q=(Z<sup>n</sup>, E<sup>m</sup>); 7) 根据式 (15) 计算β; 8) 给定测试样本 x, 经过特征提取网络进行特征提取, 并且得到了对 应的  $Q=(Z^n, E^m);$ 9) 计算预测结果  $f(x)=Q\beta, f(x)=[f_1(x), f_2(x), \dots, f_m(x)];$ 10) 返回 f(x) 的最大元素索引, 从而得到样本 x 的预测标签

3 实验验证

## 3.1 实验环境和相关参数

为验证所提算法的分类性能,在 Nvidia GeForce RTX 2080 Ti GPU 上使用 PyTorch 1.4.0 进行了实验. 训练 PointNet++特征提取网络时使用学习率为 0.001、 批量大小为 32 的 Adam 优化器对特征提取器进行了 150 次训练. GBEAE-BLS 则参照文献[17]的训练步骤 和训练参数,使用数据集 ModelNet40<sup>[18]</sup>,其中包括 9843 个训练样本和 2468 个测试样本, 40 个类别. 图 5 显示 了数据集中三维点云对象的示例.在对比训练时间时 使用 ScanObjectNN 数据集, 它旨在解决真实扫描场景 下的点云分类问题.提供了约15000个对象的真实样 本,覆盖15种类别,每类含2902个独特实例.数据不 仅包括全球与局部坐标、法线、颜色属性和语义标签, 还首次在真实数据上提供了部分注释.图6显示了数 据集中三维点云对象的示例. 在训练和推理过程中, 每 个点云对象都被重新采样为10000个点.



person

car



图 6 ScanObjectNN 点云样本

## 3.2 不同 GBEAE-BLS 参数的结果

为了检验 GBEAE-BLS 分类网络的最佳网络参数, 在测试数据集上用不同的 GBEAE-BLS 参数测试了 PointGBLS 的推理准确率,包括映射层的节点数(即特征映 射节点数 N1×N2, 其中, N1 是每个窗口的节点数, N2 是 映射层的窗口数)和投影增强层的节点数(即投影增强

8 专论•综述 Special Issue

节点数 N<sub>4</sub>). 结果见表 1 和表 2.

如表 1 和表 2 所示, 当  $N_1=N_2=10$  和  $N_4=1800$  时, 推理准确率最高, 达到 92.40%, 超过了纯深度学习网 络 PointNet++的 91.90%, 证明所提网络的有效性. 从 表 1 和表 2 中可以得到, 当  $N_1$  和  $N_2$  固定为 10,  $N_4$  增 加时,准确率并没有单调增加. 当  $N_4$  固定为 1800,  $N_1$  和  $N_2$  增加时,准确率也并没有单调增加,这说明并 不是节点数越多,准确率就越高. 网络中的投影增强节 点或特征映射节点过多会导致冗余和过拟合问题,因 而准确率下降. 此外,网络中节点过多会造成沉重的计 算负担,所需时间和资源变多,准确率却没有提升. 此 外,虽然  $N_1=N_2=10$  和  $N_1=40$ ,  $N_2=10$  时的推理准确率均 为 92.40%,但 GBEAE-BLS 的平均训练时间分别为 21.04 s 和 27.24 s,相比来说,同样准确率选择训练时间 短的. 因此,将  $N_1=N_2=10$ 、 $N_4=1800$  作为次优参数设置 来进行进一步实验.

#### 表1 不同投影增强节点数下的推理准确率 (当 N=N=10 时)

$(\exists N_1 - N_2 -$	10 H)
投影增强节点 (N4)	准确率 (%)
1 000	91.14
1 200	91.43
1 400	91.55
1 600	91.75
1 800	92.40
2 000	91.99
2 2 0 0	92.36
2400	91.91
2 600	92.17
2800	91.99
3 000	91.79

表 2 不同特征映射节点数下的推理准确率 (当 N4=1800 时)

特征映射节点 (N1×N2)	准确率 (%)
5×10=50	91.59
10×10=100	92.40
15×10=150	92.16
20×10=200	91.55
25×10=250	92.32
30×10=300	92.07
35×10=350	92.07
40×10=400	92.40
10×5=50	91.74
10×10=100	92.40
10×15=150	91.91
10×20=200	91.79
10×25=250	91.79
10×30=300	91.83
10×35=350	91.79
10×40=400	92.00

## 3.3 不同预训练深度学习模型的结果

在确定了 GBEAE-BLS 的网络参数后, 进一步研究 PointNet++特征提取器的性能如何影响整个 Point-GBLS 的推理准确率. 在 PointNet++特征提取器的训练 过程中, 在每个训练周期后保存模型. 然后, 用所有这 些预训练模型训练并测试了 GBEAE-BLS 分类网络, 结果如表 3 所示. 表 3 中, "Epoch"表示 PointNet++特 征提取器的训练周期, FC-layers 表示图 4 中使用全连 接层分类网络的推理准确率, GBEAE-BLS 表示图 1 中 使用 GBEAE-BLS 分类网络的推理准确率. 时间表示 训练所需时间.

表 3 全连接层与 GBEAE-BLS 在不同特征提取器训练时间 下的分类准确率比较 (当 N<sub>1</sub>=N<sub>2</sub>=10, N<sub>4</sub>=1800 时)

BLS
时间 (s)
18
21.04
42.08
105.2
189.36
441.84
1 0 3 0.96
1241.36
2167.12
3156

由表 3 前 6 行准确率对比可知,即使未对特征提取器进行充分的预训练,提出的网络也具有良好的推理性能.尤其是即使在没有任何训练操作的情况下,用随机网络参数初始化特征提取器,GBEAE-BLS分类的准确率仍能达到 86.00%.虽然特征提取器完全没有经过训练,但它可以被视为从ℝ<sup>N×3</sup>到ℝ<sup>1024</sup>的固定映射函数,解决了点云的无序问题,并使在ℝ<sup>1024</sup>域上应用 GBEAE-BLS 成为可能.没有任何训练操作的情况下,能达到 86.00%的准确率,这反映了 GBEAE-BLS 强大的学习能力,它利用了基于图的极限学习机自动编码器和脊回归训练方法的优势.

如表 3 所示, 采用 GBEAE-BLS 分类网络与采用 相同预训练 PointNet++提取器的全连接层相比, 准确 率更高, 在训练网络收敛后达到了 92.40% 的最高准确 率, 而 PointNet++的最高准确率只能达到 91.90%. 因 此, GBEAE-BLS 可以从提取的特征中进一步挖掘细 节, 提高分类性能.

表 3 中也显示了提出的网络在训练时间上的优势.

对于 GBEAE-BLS 分类网络, 当  $N_1=N_2=10$ ,  $N_4=1800$ 时, 训练过程平均耗时 21.04 s. 然而, 基于 PointNet++的深度学习网络在 GPU 上训练一个单次 Epoch 平均 耗时 44 s, 相比之下, GBEAE-BLS 分类网络训练时间 减少了 52%. 在测试数据集上, Point-GBLS 进行分类 所需时间相比于 PointNet++也有明显减少.

对比来看,为了达到 86% 的准确率,深度学习网 络在 GPU 上花费了 21 个 Epoch,训练时间为 15.4 min, 而提出的网络只需约 18 s. 纯粹的深度学习网络在训 练了 150 个 Epoch 之后,准确率还达不到 92%,但训练时间却消耗了 110 多分钟. 然而,提出的网络在训练总时间不超过 53 min 的情况下,在第 59 个周期的准确率超过了 92%. 总之,与纯粹的基于深度学习的分类网络相比,提出的方法可以大大缩短训练时间,并具有更好的推理性能.

## 3.4 不同归一化方法的结果

为了进一步研究影响 GBEAE-BLS 分类网络推理 结果的因素, 对映射层的输出向量采用了不同的数据 归一化方法. 一般来说, 从样本中提取的不同特征通常 具有不同的数值范围, 而数值范围较大的特征通常在 分类结果中占主导地位. 为了平衡每个维度上的特征 贡献, 通常会使用 Min-max 和 Z-score 归一化算法, 将 不同的特征缩放到相同的数值范围内. 表 4 列出了这 两种归一化算法的实验结果.

表 4 不同归一化方法对映射特	征的推理准确率 (	%)
-----------------	-----------	----

归一化方法	准确率
None	91.91
Min-max	92.40
Z-score	91.71

表 4 结果表明,不使用归一化方法时准确率为 91.91%, Min-max 归一化使准确率提高了 0.49%,而 Zscore 归一化导致准确率下降了 0.2%. 这是因为, Z-score 归一化算法假定非标准化数据服从高斯分布. 但是,如 果实际数据不服从高斯分布,归一化结果可能就不再 符合预期,所以准确率有所下降. 相反, Min-max 归一 化算法不对数据分布做任何假设,更适合数值范围有 限的数据. 因此,最终采用了 Min-max 归一化算法.

## 3.5 不同激活函数的结果

为了探索不同的非线性激活函数对网络的影响, 在增强层分别应用了 ReLU、tanh 和无激活函数,结果 如表 5 所示.

10 专论•综述 Special Issue

由表 5 可知,当增强层经过投影增强层直接连接 到输出层而没有任何激活函数时,推理准确率大幅降 至 53.22%,因为激活函数决定了网络对非线性函数的 逼近能力.

表 5	一個一個一個一個一個一個一個一個一個一個一個一個一個一個一個一個一個一個一個	
125	相见広区而个问版伯函数时们也注电师半(/0)	

准确率
53.22
88.79
92.40

由于 GBEAE-BLS 的网络结构, 增强层是唯一具 有非线性特性的部分. 如果去掉激活函数, GBEAE-BLS 分类网络就会变成纯线性映射函数, 会导致性能下降.

在增强层应用 ReLU 和 tanh 函数时, 网络的推理 准确率分别为 88.79% 和 92.40%, 网络准确率都有大 幅提升, 说明激活函数很重要. 与 ReLU 函数相比, tanh 函数的性能提升了 3.61%. 这是因为 tanh 函数零中心 化输出, 有助于数据的稳定性和收敛性, 可以减少学习 过程中的偏移, 非线性度较高.

#### 3.6 在不同数据集上训练时间的比较

为了验证所提方法的有效性,我们在 ModelNet40 数据集和 ScanObjectNN 数据集上对所提出的方法进行训练时间的比较,如表 6 和表 7 所示.

方法	准确率 (%)	训练时间 (min)
PointNet	89.20	82
PointNet++	91.90	110
PointCNN	92.20	130
Point Cloud Transformer	93.20	143
Ours	92.40	52.6

	表 7	在 ScanObjectNN 数据集上训练时间的比较	
--	-----	----------------------------	--

方法	准确率 (%)	训练时间 (min)
PointNet	68.20	105
PointNet++	77.90	141
PointCNN	78.50	162
Ours	78.30	87

由表 6 可知, 在 ModelNet40 数据集上, 所提方法 在准确率上优于大部分网络, 且训练时间最短, 比纯深 度学习的 PointNet++训练时间缩短了一半以上, 准确 率却提高了 0.5%. 由表 7 可知, 在 ScanObjectNN 数据 集上, 所有网络的准确率都相对较低, 训练时间也都有 加长. 这是因为 ScanObjectNN 数据集相对于 Model-Net40 数据集更加复杂, 更难训练. 所提出的方法基本 能达到最高准确率,但是所需训练时间却缩短将近一半.因此,通过两个数据集上训练时间的比较,可以得到所提出的方法在保证准确率的情况下,所需时间大幅减少.

## 3.7 与其他 3D 点云分类方法的比较

最后,将本文所提出网络的推理结果与其他具有代 表性的一些三维点云分类方法进行了比较,结果如表 8 所示.

表 8 与其他 3D 点云	分类方法的比较 (%)
方法	准确率
VB-Net	83.99
VoxNet	85.90
Subvolume	89.20
MVCNN	90.10
PointNet	89.20
Point-BLS	89.69
PointNet++	91.90
Ours	92.40

由表 8 可知,即使采用更简化的骨干网架构,提出 的方法也超过了原始 PointNet++的性能,推理准确率 高达 92.40%,这证明了 GBEAE-BLS 出色的学习能力. 此外,在 Subvolume 中,研究人员使用了基于体素的方 法来提取点云特征,其在 ModelNet40 数据集上的最高 分类准确率为 89.20%.在 MVCNN中,研究人员使用 了基于视图的方法来提取点云特征,其在 ModelNet40 数据集上的最高分类准确率为 90.10%.而本文基于 PointNet++的特征提取方法将分类准确率分别提高了 3.2% 和 2.3%,这表明与基于体素和基于视图的方法相 比,基于点的深度学习方法能更好地保留三维形状信 息,学习到更好的点云特征.

## 4 结论与展望

本文提出了一种新颖的三维点云分类网络 Point-GBLS, 它将深度学习和宽度学习系统结合在一起. 利 用基于深度学习的网络 PointNet++提取点云特征, 然 后利用宽度学习系统进行分类. 实验表明, 与 PointNet、 PointNet++相比, 深度学习网络与宽度学习系统的结合 提高了推理的准确性, 而且训练时间也大大缩短. 此外, 使用深度学习方法训练特征提取器也提高了网络的整 体准确性. 因此, 提出的方法能够实现高质量三维点云 分类. 下一步研究是如何将宽度学习系统用于更细粒 度的点云分割任务.

## 参考文献

- Maturana D, Scherer S. VoxNet: A 3D convolutional neural network for real-time object recognition. Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Hamburg: IEEE, 2015. 922–928. [doi: 10.1109/IROS.2015.7353481]
- 2 Zhou Y, Tuzel O. VoxelNet: End-to-end learning for point cloud based 3D object detection. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 4490–4499. [doi: 10.1109/CVPR.2018.00472]
- 3 Liu ZS, Song W, Tian YF, *et al.* VB-Net: Voxel-based broad learning network for 3D object classification. Applied Sciences, 2020, 10(19): 6735. [doi: 10.3390/app10196735]
- 4 Su H, Maji S, Kalogerakis E, *et al.* Multi-view convolutional neural networks for 3D shape recognition. Proceedings of the 2015 IEEE International Conference on Computer Vision. Santiago: IEEE, 2015. 945–953. [doi: 10.1109/ICCV.2015. 114]
- 5 Yang Z, Wang LW. Learning relationships for multi-view 3D object recognition. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019. 7504–7513. [doi: 10.1109/ICCV.2019.00760]
- 6 Wang WJ, Cai Y, Wang T. Multi-view dual attention network for 3D object recognition. Neural Computing and Applications, 2022, 34(4): 3201–3212. [doi: 10.1007/s00521-021-06588-1]
- 7 Qi CR, Su H, Mo K, *et al.* PointNet: Deep learning on point sets for 3D classification and segmentation. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 652–660. [doi: 10.1109/ CVPR.2017.16]
- 8 Qi CR, Yi L, Su H, *et al.* PointNet++: Deep hierarchical feature learning on point sets in a metric space. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: ACM, 2017. 5105–5114.
- 9 Xu QG, Sun XD, Wu CY, *et al.* Grid-GCN for fast and scalable point cloud learning. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 5660–5669. [doi: 10.1109/CVPR42600. 2020.00570]
- 10 Xu MT, Zhang JB, Zhou ZP, *et al.* Learning geometrydisentangled representation for complementary understanding of 3D object point cloud. arXiv:2012.10921v3, 2021.
- 11 Zhang RR, Wang LH, Wang YL, et al. Starting from nonparametric networks for 3D point cloud analysis.

Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023. 5344–5353. [doi: 10.1109/CVPR52729.2023.00517]

- 12 Huang GB, Zhu QY, Siew CK. Extreme learning machine: Theory and applications. Neurocomputing, 2006, 70(1–3): 489–501. [doi: 10.1016/j.neucom.2005.12.126]
- 13 Huang GB, Zhou HM, Ding XJ, et al. Extreme learning machine for regression and multiclass classification. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2012, 42(2): 513–529. [doi: 10.1109/TSMCB. 2011.2168604]
- 14 Zhang R, Lan Y, Huang GB, *et al.* Universal approximation of extreme learning machine with adaptive growth of hidden nodes. IEEE Transactions on Neural Networks and Learning Systems, 2012, 23(2): 365–371. [doi: 10.1109/TNNLS.2011. 2178124]
- 15 Liu Z, Jin W, Mu Y. Variances-constrained weighted extreme learning machine for imbalanced classification. Neurocomputing, 2020, 403: 45–52. [doi: 10.1016/j.neucom. 2020.04.052]
- 16 Chen CLP, Liu ZL. Broad learning system: An effective and efficient incremental learning system without the need for deep architecture. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(1): 10–24. [doi: 10.1109/ TNNLS.2017.2716952]
- 17 Liu Z, Huang SL, Jin W, *et al.* Graph-based broad learning system for classification. Neurocomputing, 2021, 463: 535–544. [doi: 10.1016/j.neucom.2021.07.065]
- 18 Wu ZR, Song SR, Khosla A, *et al.* 3D ShapeNets: A deep representation for volumetric shapes. Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 1912–1920. [doi: 10.1109/ CVPR.2015.7298801]
- 19 Rusu RB, Blodow N, Marton ZC, et al. Aligning point cloud views using persistent feature histograms. Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems. Nice: IEEE, 2008. 3384–3391. [doi: 10. 1109/IROS.2008.4650967]
- 20 Rusu RB, Blodow N, Beetz M. Fast point feature histograms (FPFH) for 3D registration. Proceedings of the 2009 IEEE International Conference on Robotics and Automation. Kobe: IEEE, 2009. 3212–3217. [doi: 10.1109/ROBOT.2009. 5152473]
- 21 Salti S, Tombari F, Di Stefano L. SHOT: Unique signatures of histograms for surface and texture description. Computer Vision and Image Understanding, 2014, 125: 251–264. [doi:

10.1016/j.cviu.2014.04.011]

- 22 Zhou H, Li QD, Xie QJ. Individual pig identification using back surface point clouds in 3D vision. Sensors, 2023, 23(11): 5156. [doi: 10.3390/s23115156]
- 23 Li YY, Bu R, Sun MC, *et al.* PointCNN: Convolution on Xtransformed points. Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal: ACM, 2018. 828–838.
- 24 Wu WX, Qi ZG, Li FX. PointConv: Deep convolutional networks on 3D point clouds. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 9613–9622. [doi: 10. 1109/CVPR.2019.00985]
- 25 单铉洋, 孙战里, 曾志刚. RFNet: 用于三维点云分类的卷积 神经网络. 自动化学报, 2023, 49(11): 2350-2359. [doi: 10. 16383/j.aas.c210532]
- 26 Chen C, Fragonara LZ, Tsourdos A. GAPointNet: Graph attention based point neural network for exploiting local feature of point cloud. Neurocomputing, 2021, 438: 122–132. [doi: 10.1016/j.neucom.2021.01.095]
- 27 史豪斌, 王少熙, 黄睿茜. 融合注意力机制与神经网络的三维点云分类算法. 微电子学与计算机, 2023, 40(5): 12-19. [doi: 10.19304/J.ISSN1000-7180.2022.0575]
- 28 Guo MH, Cai JX, Liu ZN, *et al.* PCT: Point cloud Transformer. Computational Visual Media, 2021, 7(2): 187–199. [doi: 10.1007/s41095-021-0229-5]
- 29 He CH, Li RH, Li S, *et al.* Voxel set Transformer: A set-toset approach to 3D object detection from point clouds. Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022. 8407–8417.
- 30 Kasun LLC, Zhou H, Huang GB, *et al.* Representational learning with ELMs for big data. IEEE Intelligent Systems, 2013, 28(6): 31–34.
- 31 Kasun LLC, Yang Y, Huang GB, *et al.* Dimension reduction with extreme learning machine. IEEE Transactions on Image Processing, 2016, 25(8): 3906–3918. [doi: 10.1109/TIP.2016. 2570569]
- 32 Pao YH, Park GH, Sobajic DJ. Learning and generalization characteristics of the random vector functional-link net. Neurocomputing, 1994, 6(2): 163–180. [doi: 10.1016/0925-2312(94)90053-1]
- 33 Chu YH, Lin HF, Yang L, *et al.* Hyperspectral image classification with discriminative manifold broad learning system. Neurocomputing, 2021, 442: 236–248. [doi: 10. 1016/j.neucom.2021.01.120]

<sup>12</sup> 专论•综述 Special Issue

- 34 Chen GM, Zhang D, Xian ZY, *et al.* Facial expressions classification based on broad learning network. Proceedings of the 10th International Conference on Information Systems and Computing Technology (ISCTech). Guilin: IEEE, 2022. 715–720. [doi: 10.1109/ISCTech58360.2022.00118]
- 35 Liu WT, Zuo Y, Li TS, *et al.* Application of image recognition for on-board sailor behavior based on broad learning system. Proceedings of the 4th International Conference on Data-driven Optimization of Complex Systems (DOCS). Chengdu: IEEE, 2022. 1–5. [doi: 10.1109/ DOCS55193.2022.9967770]
- 36 刘万军, 李雨萌, 曲海成. 基于频谱增强和卷积宽度学习的 音乐流派分类. 计算机系统应用, 2023, 32(10): 85–95. [doi: 10.15888/j.cnki.csa.009272]
- 37 Xue H, Wang XY, Meng Y, *et al.* Fine-grained classification of YOLOv5 remote sensing aircraft targets incorporating broad learning system. Proceedings of the 9th International Conference on Cloud Computing and Intelligent Systems (CCIS). Dali: IEEE, 2023. 414–418. [doi: 10.1109/ CCIS59572.2023.10262912]
- 38 Li HM, Ma JW, Zang SF, et al. Hyperspectral image classification based on cross-domain adaptive broad learning system. Proceedings of the 39th Youth Academic Annual Conference of Chinese Association of Automation (YAC). Dalian: IEEE, 2024. 1499–1503. [doi: 10.1109/YAC63405. 2024.10598527]

(校对责编: 王欣欣)

