

A Deep Neural Network For Learning Hierarchical Multi-scale Point-Cluster Features Towards Large Scale Point Cloud Classification

Abstract—室外大场景激光雷达点云的自动精确分类是遥感和计算机视觉等领域的难题。为了利用深度学习强大的特征表达能力提高分类精度，同时避免点云栅格化带来精度损失，本文提出了一种直接用于原始激光雷达点云分类的深度学习方法，该方法构建了一种具有特征结构性和空间结构性的多层次多尺度特征的深度神经网络模型。该深度神经网络模型先分别处理单点的不同特征再进行特征融合，有效挖掘单点特征的结构性，构建显著性点集特征，再充分利用点集中点的空间相邻关系，使得点集特征对噪声不敏感，通过学习多层次多尺度点集特征，分别对不同层次点集分类，有效提高了点云分类精度。该方法在处理复杂室外场景具有很好的性能，无需对点云进行栅格化，无需保证点云划分的点集是否大小相同，并且不受点集中点排列顺序的影响。

Index Terms—Point cloud classification, deep neural network, point-cluster features, sparse representation.

I. INTRODUCTION

室外大场景激光雷达点云的自动精确分类是遥感和计算机视觉等领域的难题，而特征提取的优劣决定了分类结果好坏。当前已经有大量特征算子被用于点云分类并取得良好的分类效果，比如 spin 图[1]，曲率，法向量等。而以这些特征算子为基础，进行进一步特征提取，能获取更有效的特征，进一步提高分类精度。当前特征提取方法有 bag of words (BoW)，稀疏性约束等方法，这些特征提取方法能被看做一种单层神经网络，而这种单层结构限制了这些方法获得的特征的表达能力。当前深度学习在特征提取领域得到广泛的应用，因为深度学习能够建立多层神经网络结构，能够学习到相对单层结构更好的特征表达形式。

深度学习用于点云分类的研究还较少，主要原因是点云数据和图像数据存在较大差别。点云数据缺乏组织性，无法像图像一样通过窗口移动在图像中获取大小相同且具有明确相邻关系的图像块进行训练。正因为此，当前采用深度学习进行点云分类或三维目标提取的方法中，首先要栅格化点云或三维目标，再利用深度学习进行下一步处理。但栅格化会带来三维信息丢失，并且栅格大小也会影响最后分类精度。

针对上述问题，我们提出了一种具有特征结构性和空间结构性的多层次多尺度特征的深度学习网络。从原始无序的点云多层次分割形成尺寸不同的点集，利

用点集中所有单点特征，学习点集特征并进行分类，避免栅格化带来精度损失的同时，又能利用深度学习强大的特征表达能力提高分类精度。虽然单点特征算子能直接利用深度学习进行特征提取，但是单点特征算子维数较少时，深度学习无法获得很好的效果，且在大量点云中进行逐点分类效率很低。采用本文的深度神经网络也能够基于单点特征进行分类，只需要把点集看做单点周围一定邻域的点即可。本文的主要贡献是：

i) 提出一种直接用于原始激光雷达点云分类的深度学习方法。该方法在处理复杂室外场景具有很好的性能，无需对点云进行栅格化，无需保证点云划分的点集是否大小相同，并且不受点集中点排列顺序的影响。深度神经网络通过学习多层次多尺度点集特征，分别对不同层次点集分类，能有效提高点云分类精度。

ii) 有效挖掘单点特征的结构性，构建显著性点集特征。点云中单点特征由多个特征算子构成，各算子提取特征的方式和目的不相同，深度神经网络先单独处理各特征后再进行特征融合，精确综合了特征内及各特征间的关联关系。

iii) 充分利用点集内点的空间结构性，使得点集特征对噪声不敏感。利用 DMst [2] 赋予每个点空间相邻关系，并在 pooling 层中对 DMst 进行合并以及剪枝操作，有效利用点集中点的空间相邻关系，降低边缘不良单点特征对点集特征的影响，同时降低了每一层点的数量。

II. RELATED WORK

A. 特征表达

在直接使用特征算子或语义关系的基础上，分类过程中如果通过特征表达对已有特征进行重新组织和提取，能获得更优秀的特征，以此分类能够获得更好的分类精度。

字典学习在特征表达中获得了广泛应用。BoW[12]通过 kmeans 的聚类中心建立单词，所有单词的集合形成字典，每一个特征通过字典中的一个单词来表示，得到了原始特征的重新表达。在 BoW 的基础上，Plsa[13]和 Lda[14]把单词组成语义，特征通过语义来表达。稀疏编码中每一个特征不再只由单一单词来表示，而是由少数几个单词联合表示，通过稀疏约束进行字典学习。Yang et al.[15]提出了一种 sparse coding spatial pyramid matching (ScSPM)的方法，结果显示采用稀疏编码以及 max pooling 要比矢量量化和 averaging pooling 识别效果更好。Yang et al.[16]提出了一种结构性的有监督的字典学习方法 Fisher discrimination dictionary learning (FDDL) model, 这种方法获得的字典具有类内差异小类间差异大的特点。在点云分类中，字典学习方法也得到了应用，Wang et al.[17]和 Zhang et al.[18]分别采用了 LDA 和 latent Dirichlet allocation and sparse coding(SCLDA)的方法进行了字典学习并进行分类。

流形学习方法在特征表达中也得到了广泛的应用,流形学习方法假设高维空间中存在低维流形,而通过寻找相应的嵌入映射完成特征提取。比如 ISOMAP [19], KPCA[20], LE [21]等。Weston et al.[22]采用了深度学习的框架进行了半监督的内嵌。

Bengio et al.[23]在综述中回顾了深度学习和稀疏编码,流形学习等表达方式之间的联系,并用深度学习的角度对这些表达方式进行了解释。Krizhevsky et al.[24]提出了一种 7 个特征提取层的卷积神经网络能够有效提取高层的视觉信息。Girshick et al.[25]通过 high-capacity 的卷积神经网络进行由下而上的特征提取并采用 domain-specific fine-tuning, 相对于 VOC 2012 提高了 30%的平均精度。这些神经网络以及训练方法的提出,显示了强大的特征表达能力。

B. 基于的深度学习的 3D 物体识别

近年来深度学习在 3D 物体识别和检索中也得到了应用。Wu et al. [26]通过设置模型内部栅格值为 0, 外部值为 1 的方式栅格化 3D 模型,并设计了 Convolutional Deep Belief Network (CDBN)分类栅格化后的 3D 模型。Zhu et al.[27]利用 3d 模型不同角度的深度图像,采用 DBN 预训练,自动编码器最后微调得到 3d 模型特征。Xie et al. [28]提取 3D 物体多尺度的形状表达,提出了一种用 Fisher discrimination criterion 约束隐层的自动编码器,提取了高层次的形状特征。Socher et al. [29]采用 convolutional and recursive neural networks 对 RGB-D 图像进行目标识别。点云分类中,也有少量利用深度学习分类的研究。Guan et al. [30]采用 DBM 对分割出的每一棵树点云的垂直廓线进行分类,分类了 10 个树种。Maturana et al. [31]采用体密度图的方式对机载点云进行栅格化,设计了 3 维 CNN 学习分类机载点云。

C. 点云分类

在过去数年中点云分类得到广泛研究,大量特征提取以及分类方法被提出,大量方法采用特征算子或语义关系结合分类器进行分类。首先大量基于点的方法被提出,这些方法采用大量特征算子的组合进行分类。Chehata et al. [3]用随机森林对 5 类总计 21 种机载激光雷达点云特征进行分类,并通过迭代的特征选择获得 6 种最好的特征。Guo et al. [4]采用了 26 种特征以及 JointBoost 在复杂地面对建筑,植被,地面,电线以及架线塔五种地物进行分类。Kragh et al. [5]采用 13 种特征以及 svm 对车载点云进行分类,并针对不同车载点云中点密度变化采用变邻域半径的方法提取特征。Brodu et al. [6]提出了一种多尺度特征提取方法描述地物特征,并对两个区域内植被,岩石,水和地面进行分类并取得很好的精度。其次一些基于点集的方法获得了研究,这些方法利用了点集特征以及点集之间的关系进行分类。Zhang et al. [7]对机载点云采用区域生长进行分割,然后提

取分割后点集的几何，回波，辐射度以及拓扑特征，并用 svm 进行分类。Pu et al. [8]对车载激光雷达点云过分割后，通过位置，形状，方向等因素设计特征识别平面，柱状结构。Yang et al. [9]把车载激光雷达点云过分割成语义的超像元，然后建立了一系列规则对这些超像元进行合并，并对合并后物体设计规则进行识别。Fukano et al. [10]通过拉普拉斯收缩突出线状和面状特征，并根据线状和面状特征组合方式提取杆状地物。还有一些方法通过图像赋予点云一定空间语义关系进行分类。Zhang et al. [11]通过道路场景中地物在图像和点云中的一系列弱判别条件，然后通过一种过滤算法以及条件随机场的语义分析方法完成点云分类。

综上，深度学习进行 3D 物体识别和点云分类时都需要把 3D 物体进行栅格化,通过栅格化给予 3D 物体和点云空间结构性,获得规则的输入以及相邻关系,而我们的深度神经网络结构通过所有单点特征的权值共享,层次的结构设计,以及 DMst 建立点集内部结构,赋予了原始点云规则性以及相邻关系。

III. 深度神经网络结构

这一章将详细描述我们提出的网络结构，整个网络总体结构如图 1 所示。网络包含一个输入层，连接两种特别设计的特征提取网络 Net1 和 Net2，它们各自包含若干个特征提取层以及后续的 pooling 层，Net3 是一个独立的 pooling 层，Net4 是多层次联合分类层，所有层都是由自动编码器形成。Net1 和 Net2 中每一个蓝色方框表示一个点集，在神经网络中是一个 map，包含多个点特征，每一列包含一个特征提取层以及一个 pooling 层。经 Net3 pooling 后，所有点特征平均起来形成一个点集特征，所以，Net3 和 Net4 中的绿色方框表示点集的特征，棕色框是输出层赋予每一个点集类别。下面将论述 Net1、Net2、Net3 和 Net4 的具体结构。

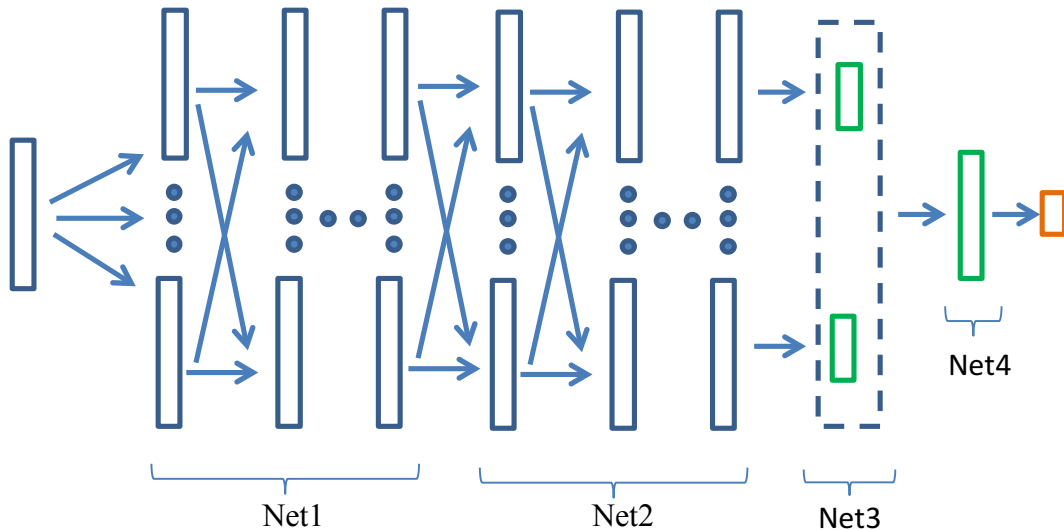


图 1. 神经网络总体结构图

A. 网络的输入及 Net3 的结构

网络输入是点集的点特征，经 Net3 pooling 后得到点集特征，并且正是 Net3 的设计使得神经网络能够处理大小不一的点集，并且对点集中点顺序不敏感。

首先和文献[17]一样去除地面点，进行多层次多尺度分割得到多层次多尺度的点集。本文采用和文献[18]相同的单点特征提取方法，获得三个不同邻域范围内的单点特征，这样每一个点在一个邻域内有 19 维特征向量，三个不同的范围，每一个点形成 57 维的特征。每个点集中所有单点特征合并起来形成 $57 \times m$ 维的特征矩阵， m 是每个点集中点数量。归一化所有点集特征矩阵的每一维后，这些点集的特征矩阵就是神经网络的输入。

因为输入点集大小不一，无法像图像一样以一个固定长度的向量表示这个输入，并且点集中点排列顺序是随机的，虽然点排列顺序变化不会改变点集中点在空间中分布情况，但会改变输入的特征矩阵，若直接用叠堆自动编码器对这些特征矩阵进行训练，得到的神经网络无法泛化使用。

本文采用所有单点权值共享策略，每一个点特征乘以相同权值后，点的顺序不会影响点特征和权值相乘的结果。由于权值只针对单点特征，所以与点集中点的数量无关。通过权值共享，点集的特征矩阵能够输入到神经网络中，所以后续描述的 Net1 和 Net2 中权值是针对单点的。因为最终需要合并这些单点特征，并对点集进行分类，所以，在 Net3 中把点集中所有单点输出进行平均 pooling，这样每个点集只输出一个特征向量。平均 pooling 过程保证了整个深度网络描述的特征不会受点顺序的影响。同时，Net2 可能输出多个 map，将每个 map 经 pooling 得到的特征向量合并成一个向量，这个向量就是 Net3 的输出。

B. Net1 和 Net2 的特征提取层

单点特征由基于特征值的特征和 spin 图特征两部分构成，这两种特征从不同角度对点云进行描述，特征内部的关联性比不同特征间的关联性要密切。为了深层次挖掘这两种特征进行，分别对其进行训练。图 2 显示了 2 个 map 的情况下，Net1 中特征提取层结构中的单个点特征与权值间关系。

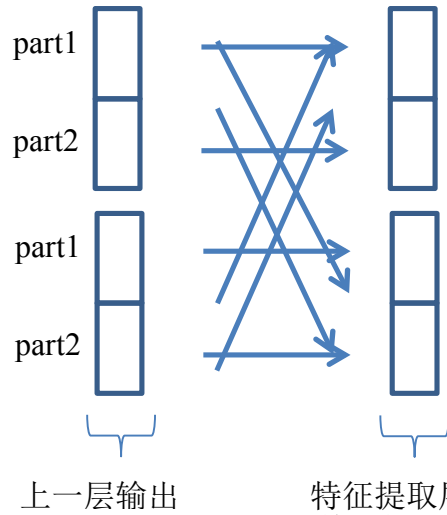


图 2 Net1 特征提取层的基本结构

经过 Net1 学习以后，分别获得了基于特征值的特征和 spin 图特征的表达。为了能够学习到这两种特征之间关系，在 Net2 中，合并这两种特征进行学习。Net2 的结构类似于 Net1 的结构，只是特征不再分成两块。通过 Net1 和 Net2，完成对不同特征内部以及特征之间的表达。当单点特征由多个特征部分构成时，那么 Net1 就由多个部分构成。

C. Net1 和 Net2 的 pooling 层

Net1 和 Net2 的特征提取层后连接 pooling 层，Pooling 层中采用均值 pooling，把每一个点邻近几个点上一层的输出进行平均得到 pooling 层的值，通过相邻点 pooling 减弱噪声影响。每个点的邻近点通过文献[2]DMst 对点集内所有点进行组织得到，DMst [2]能有效组织点云，并获得局部结构同时兼顾物体空间延展能力。使用 DMst 时，根节点是点集中最接近中心的点。

激光雷达点云在地物边界会出现扰动现象，地物边界点分布散乱，造成这些点的单点特征不稳健，需要减少这些点对整个点集特征的影响。因为 DMst 结合了 Dijkstra 算法，考虑了物体延展性，大部分边界点位于 DMst 的叶子节点，可以通过剪枝操作进行去除。

Pooling 层中只对非叶子节点进行均值 pooling，叶子节点被删除，是剪枝操作，多个 pooling 层形成一个逐层剪枝过程。经过一次 pooling 后，由于删除了原始叶子节点，位于原始叶子节点的上层节点变成了新的叶子节点，所以，在后面的 pooling 层中，这些叶子节点会被继续删除。值得注意的是，根节点只存在

于 DMst 中, 在剪枝过程中, 不再存在根节点。如果根节点只有一个点和它相连, 那么这个点是叶子节点会被删除。如果点集中只剩下一个点时, 这个点不被删除, 确保点集中至少存在一个点。图 3 用一个 2 维点集显示了两层 pooling 的剪枝过程。图中所有点都是点集中的点, 红色点是构建 DMst 时的根节点。经过两层 pooling 图 (a) 变成了图 (c)。经 pooling 后, 层中点数量下降, 有效提高计算效率。

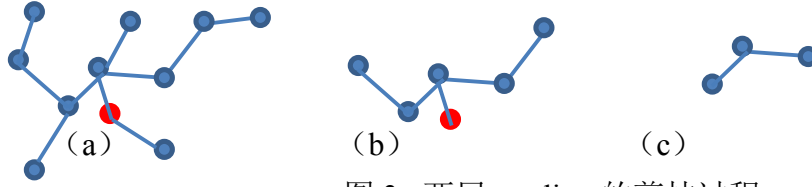


图 3. 两层 pooling 的剪枝过程.

当输入点集中点数量较少时, 只在最初的两个特征提取层进行 pooling, Net1 和 Net2 中后续的特征提取层不再连接 pooling 层。Net1 和 Net2 中的特征提取以及 pooling 层是能够进行叠堆的。

D. Sparse representation

因为点集中相邻点的特征是相似的, 且同一类点特征也是相似的, 所以单点特征具有一定稀疏性。对神经网络进行稀疏性约束, 稀疏性约束方式和自动编码器进行稀疏约束的方式相似。但是稀疏性约束是对分别对每一个 map 进行的, map 中含有多个点, 所以, 平均活跃度是当前层所有参与训练的样本中所有点的平均。

IV. POINT CLOUD CLASSIFICATION

因为采用单一层次的点集难以获得良好的分类效果, 进行多层次多尺度的分类能有效提高分类精度及鲁棒性[17], 所以整个深度网络采用能处理多层次多尺度点集的结构, 这个过程通过 Net4 实现。训练时, 把多层次多尺度点集作为输入, 同时标记每一个点集属于哪个层次, 所有点集都通过 Net1, Net2 和 Net3, 但是, 在 Net4 中不同层次的点集通过不同的 map 进行分类。当存在 2 个层次时, Net4 的结构如图 4 所示。这样, 能够同时训练对不同层次的点集, 并统一的进行多层次多尺度点集特征的提取, 避免每个层次的点集分别进行特征提取, 增加

提取的特征的鲁棒性。在训练样本不足时，不同层次的点集同时进行特征提取增加了样本数量。

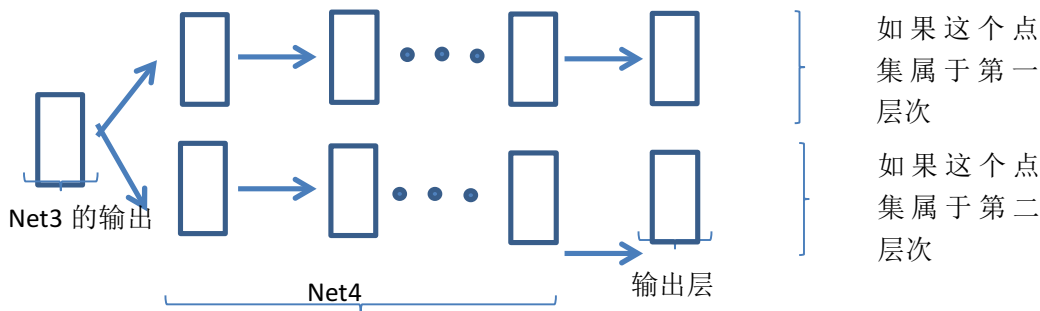


图 4 Net4 的结构

在泛化阶段，每一个未知的多层次多尺度点集均能通过深度神经网络获得该点集对应类别的概率。因为较粗层次的点集会包含其他类别的点，所以只对最高层次的点集进行分类，而最高层次点集对应每一类别的概率是通过该点集以及其他层次包含这个点集的点集概率连乘获得，最后这个点集的类别为概率最大的那个类别。

III. EXPERIMENTAL RESULTS

为了验证本文方法在室外场景点云分类的性能，我们用机载激光雷达点云对本文提出的深度神经网络进行了测试。因为深度学习需要大量样本，点云中除了植被、建筑外，其他类别的样本数量很少，因此，我将点云分成植被和建筑两类，在测试时所有不是这两类的地物已经被剔除了。

机载点云

我们采用 Leica ALS50 system 在 2010 年对中国天津扫描得到的一部分数据进行测试。扫描时的平均飞行高度是 500m，视场角是 45 度，区域内平均点密度是 20–30 points/m²。表 1 显示了训练样本以及泛化样本各类别的点云数量以及分类的精度。图 5 显示了分类的结果。

表 1 样本内点的数量以及点集数量

	Building	Tree
Training data	64952	39743
Test data	156186	73207

表 2 分类精度

	Building(precision/recall)	Tree(precision/recall)	Accuracy
Test data	100/99.8	99.6/1	99.9%

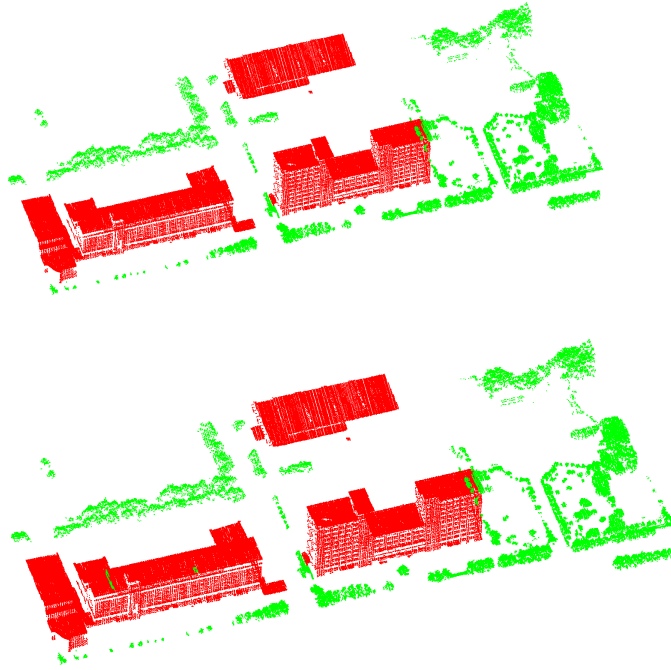


图 5 分类结果图

整个训练样本由多层次分割后的点集组成,经过分割后训练点云形成总共形成 2598 个点集,其中建筑的点集有 1580 个,树的点集有 1018 个。

分类过程中,Net1 和 Net2 各采用了两个特征提取层,Net1 中,两层都采用了 10 个 map,每个 map 包含 10 个神经元,Net2 中,两层都采用了 5 个 map,每个 map 包含 5 个神经元。因为点集中点数量较少只在第一个特征提取层后连接了 pooling 层。Net4 中只包含了 1 个特征提取层,包含 5 个神经元。因为机载点云中点云分布相对比较均匀,且发现点云噪声较小,所以平均活跃度设成 0.01,稀疏约束设成 0.2。因为训练样本较少,所以训练时多次迭代后最后训练精度已经是 100%,深度网络存在一定过拟合现象。但是因为泛化的数据和训练的数据结构较为相似,泛化精度较高。

我们测试了 Net1 和 Net2 不同的 map 数量对分类精度的影响,结果发现,如果都是 1,那么总体分类精度为 97.6%,随着 map 数量的增加而变大,但是增加到 10 时,继续增加精度不再提高反而会有所下降。这也许是因为我们是对每个 map 做稀疏约束的,所以 map 的改变会影响最后的结果。

同时我们测试了稀疏的作用,结果显示采用稀疏约束以后,分类精度会提高 1-2%,显示了点集中点特征具有一定的稀疏性。

IV. CONCLUSION

本文提出了一种直接从原始激光雷达点云中进行多层次多尺度点集特征提取的深度神经网络结构。该结构能将大小不一的点集中的单点特征构建成点

集特征，并能提取多层次点集公共特征，同时进行多层次点集的分类，增加分类精度。该结构具有特征结构性，先分别处理点云特征中不同部分，然后，再综合处理各部分输出，能够充分挖掘各特征内部及特征间的关联关系，使提取的点集特征更显著，同时该结构具有空间结构性，通过 DMst 构建点集中点与点之间联系，并通过剪枝的 pooling 层去除 DMst 中叶子节点，保证特征不受不良单点特征影响，提高分类鲁棒性。

在神经网络中，采用了自动编码器作为基本结构，当前还有其他比如 DBN，RNN 等深度网络结构，在后续的工作中我们将尝试其他的深度网络结构。随着后续样本的增加及计算效率的改进，我们将进行多类别的分类以及采用更多的层和 map 提高点云分类精度。通过提出的能够处理大小不一的点集的深度学习分类模型，我们将利用该模型对点云进行目标识别和检索工作。

REFERENCES

- [1] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 21, pp. 433-449, 1999.
- [2] Z. Wang, L. Zhang, T. Fang, P. T. Mathiopoulos, H. Qu, D. Chen, et al., "A structure-aware global optimization method for reconstructing 3-D tree models from terrestrial laser scanning data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 52, pp. 5653-5669, 2014.
- [3] N. Chehata, L. Guo, and C. Mallet, "Airborne lidar feature selection for urban classification using random forests," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 38, p. W8, 2009.
- [4] B. Guo, X. Huang, F. Zhang, and G. Sohn, "Classification of airborne laser scanning data using JointBoost," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 100, pp. 71-83, 2015.
- [5] M. Kragh, R. N. Jørgensen, and H. Pedersen, "Object Detection and Terrain Classification in Agricultural Fields Using 3D Lidar Data," in *Computer Vision Systems*, ed: Springer, 2015, pp. 188-197.
- [6] N. Brodu and D. Lague, "3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 68, pp. 121-134, 2012.
- [7] J. Zhang, X. Lin, and X. Ning, "SVM-based classification of segmented airborne LiDAR point clouds in urban areas," *Remote Sensing*, vol. 5, pp. 3749-3775, 2013.

- [8] S. Pu, M. Rutzinger, G. Vosselman, and S. O. Elberink, "Recognizing basic structures from mobile laser scanning data for road inventory studies," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, pp. S28-S39, 2011.
- [9] B. Yang, Z. Dong, G. Zhao, and W. Dai, "Hierarchical extraction of urban objects from mobile laser scanning data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 99, pp. 45-57, 2015.
- [10] K. Fukano and H. Masuda, "Detection and Classification of Pole-Like Objects from Mobile Mapping Data," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 1, pp. 57-64, 2015.
- [11] H. Zhang, J. Wang, T. Fang, and L. Quan, "Joint Segmentation of Images and Scanned Point Cloud in Large-Scale Street Scenes With Low-Annotation Cost," *Image Processing, IEEE Transactions on*, vol. 23, pp. 4763-4772, 2014.
- [12] E. Nowak, F. Jurie, and B. Triggs, "Sampling strategies for bag-of-features image classification," in *Computer Vision—ECCV 2006*, ed: Springer, 2006, pp. 490-503.
- [13] T. Hofmann, "Unsupervised learning by probabilistic latent semantic analysis," *Machine learning*, vol. 42, pp. 177-196, 2001.
- [14] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *the Journal of machine Learning research*, vol. 3, pp. 993-1022, 2003.
- [15] J. Yang, K. Yu, Y. Gong, and T. Huang, "Linear spatial pyramid matching using sparse coding for image classification," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, 2009, pp. 1794-1801.
- [16] M. Yang, L. Zhang, X. Feng, and D. Zhang, "Sparse representation based fisher discrimination dictionary learning for image classification," *International Journal of Computer Vision*, vol. 109, pp. 209-232, 2014.
- [17] Z. Wang, L. Zhang, T. Fang, P. T. Mathiopoulos, X. Tong, H. Qu, et al., "A multiscale and hierarchical feature extraction method for terrestrial laser scanning point cloud classification," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 53, pp. 2409-2425, 2015.
- [18] Z. Zhang, L. Zhang, X. Tong, Z. Wang, B. Guo, X. Huang, et al., "A Multilevel Point-Cluster-Based Discriminative Feature for ALS Point Cloud Classification."
- [19] J. B. Tenenbaum, V. De Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *science*, vol. 290, pp. 2319-2323, 2000.
- [20] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural computation*, vol. 10, pp. 1299-1319, 1998.

- [21] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural computation*, vol. 15, pp. 1373-1396, 2003.
- [22] J. Weston, F. Ratle, H. Mobahi, and R. Collobert, "Deep learning via semi-supervised embedding," in *Neural Networks: Tricks of the Trade*, ed: Springer, 2012, pp. 639-655.
- [23] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, pp. 1798-1828, 2013.
- [24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.
- [25] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580-587.
- [26] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, et al., "3d shapenets: A deep representation for volumetric shapes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1912-1920.
- [27] Z. Zhu, X. Wang, S. Bai, C. Yao, and X. Bai, "Deep learning representation using autoencoder for 3d shape retrieval," in *Security, Pattern Analysis, and Cybernetics (SPAC), 2014 International Conference on*, 2014, pp. 279-284.
- [28] J. Xie, Y. Fang, F. Zhu, and E. Wong, "Deepshape: Deep learned shape descriptor for 3d shape matching and retrieval," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1275-1283.
- [29] R. Socher, B. Huval, B. Bath, C. D. Manning, and A. Y. Ng, "Convolutional-recursive deep learning for 3d object classification," in *Advances in Neural Information Processing Systems*, 2012, pp. 665-673.
- [30] H. Guan, Y. Yu, Z. Ji, J. Li, and Q. Zhang, "Deep learning-based tree classification using mobile LiDAR data," *Remote Sensing Letters*, vol. 6, pp. 864-873, 2015.
- [31] D. Maturana and S. Scherer, "3D convolutional neural networks for landing zone detection from lidar," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, 2015, pp. 3471-3478.