remote sensing

Article

GeoSparseNet: A Multi-Source Geometry-Aware CNN for Urban Scene Analysis

Muhammad Kamran Afzal [1,2,3], Weiquan Liu [1,2,*], Yu Zang [2], Shuting Chen [4], Hafiz Muhammad Rehan Afzal [5,6], Jibril Muhammad Adam [2], Bai Yang [3], Jonathan Li [7] and Cheng Wang [2]

1 College of Computer Engineering, Jimei University, Xiamen 361021, China
2 Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Informatics, Xiamen University, Xiamen 361005, China; kamran@stu.xmu.edu.cn (M.K.A.); zangyu7@xmu.edu.cn (Y.Z.); 23020170155980@stu.xmu.edu.cn (J.M.A.); cwang@xmu.edu.cn (C.W.)
3 Center for Integrative Conservation & Yunnan Key Laboratory for Conservation of Tropical Rainforests and Asian Elephants, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, China; kamran@xtbg.ac.cn (M.K.A.); baiyang@xtbg.ac.cn (B.Y.)
4 Mathematics and Digital Science School, Chengyi College, Jimei University, Xiamen 361021, China; chenst2016@jmu.edu.cn
5 The School of Life Sciences, Northwestern Polytechnical University, Xi'an 710072, China; rehan@nwpu.edu.cn (H.M.R.A.)
6 School of Electrical Engineering and Computing, University of Newcastle, Callaghan, NSW 2308, Australia; c3249813@uon.edu.au (H.M.R.A.)
7 Departments of Geography and Environmental Management and Systems Design Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada; junli@uwaterloo.ca
* Correspondence: wqliu@jmu.edu.cn

Abstract: The convolutional neural networks (CNNs) functioning on geometric learning for the urban large-scale 3D meshes are indispensable because of their substantial, complex, and deformed shape constitutions. To address this issue, we proposed a novel Geometry-Aware Multi-Source Sparse-Attention CNN (GeoSparseNet) for the urban large-scale triangular mesh classification task. GeoSparseNet leverages the non-uniformity of 3D meshes to depict both broad flat areas and finely detailed features by adopting the multi-scale convolutional kernels. By operating on the mesh edges to prepare for subsequent convolutions, our method exploits the inherent geodesic connections by utilizing the Large Kernel Attention (LKA) based Pooling and Unpooling layers to maintain the shape topology for accurate classification predictions. Learning which edges in a mesh face to collapse, GeoSparseNet establishes a task-oriented process where the network highlights and enhances crucial features while eliminating unnecessary ones. Compared to previous methods, our innovative approach outperforms them significantly by directly processing extensive 3D mesh data, resulting in more discerning feature maps. We achieved an accuracy rate of 87.5% when testing on an urban large-scale model dataset of the Australian city of Adelaide.

Keywords: deep learning; 3D meshes; urban-scale; remote sensing; Geometry-Aware; attention

check for updates

Citation: Afzal, M.K.; Liu, W.; Zang, Y.; Chen, S.; Afzal, H.M.R.; Adam, J.M.; Yang, B.; Li, J.; Wang, C. GeoSparseNet: A Multi-Source Geometry-Aware CNN for Urban Scene Analysis. *Remote Sens.* **2024**, *16*, 1827. https://doi.org/10.3390/rs16111827

Academic Editor: Massimiliano Pepe

Received: 7 March 2024
Revised: 13 April 2024
Accepted: 22 April 2024
Published: 21 May 2024

1. Introduction

The 3D scene or shape analysis of a concoction of dense, complex, curved surfaces, or irregular geometries of meshes in large urban environments including buildings, trees, cars, and other elements has been made possible by recent developments in 3D computer vision and photogrammetry [1]. To enable the use of these meshes in a variety of applications, such as smart urban planning, navigation systems, virtual reality, radiation estimation, noise modeling, and photovoltaic perspective, it is essential to extract semantic information from the mesh models [2,3].

For semantic classification of 3D mesh data, a trivial number of machine learning-based algorithms are available; and their primary focus is to process 3D point clouds [4,5].