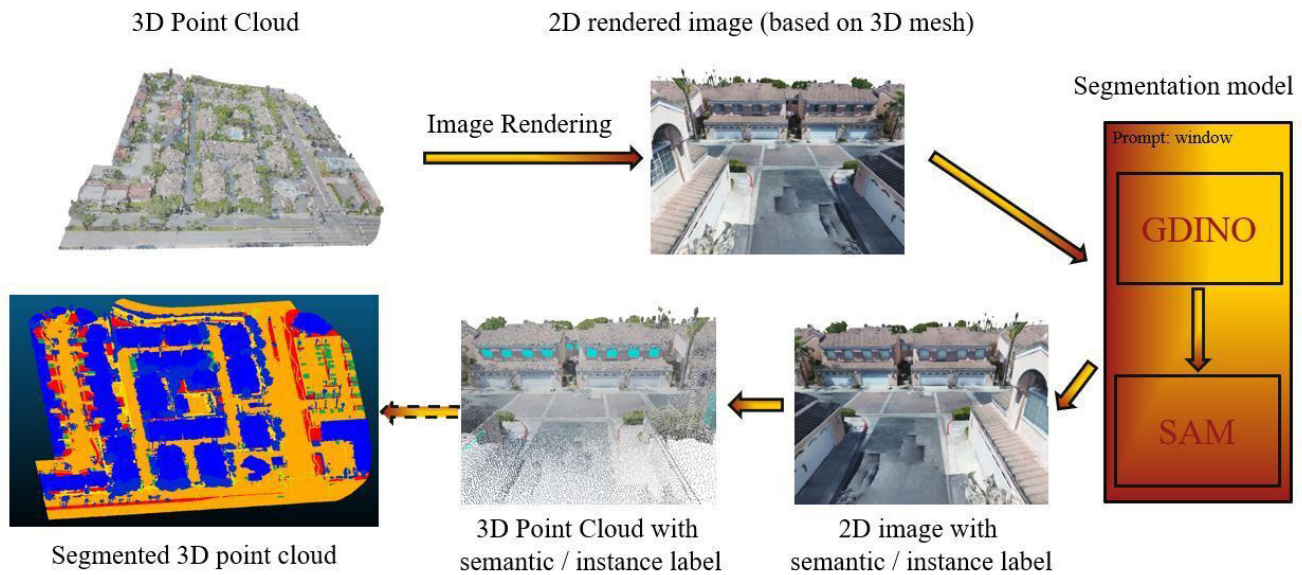# Open-Vocabulary High-Resolution 3D (OVHR3D) Data Segmentation and Annotation Framework

2023 - Current
Project Leader: Andrew Feng, Meida Chen

OVHR3D – INTERACTIVE DATA ANNOTATION PIPELINE

Background

The U.S Army and Navy have exhibited a strong interest in the field of rapid 3D reconstruction of battlefields, alongside virtual training and simulation tasks. 3D data, particularly when annotated, is essential for these applications. The use of unmanned aerial vehicles (UAVs) for the collection of aerial images and the application of photogrammetry techniques enables the swift collection and reconstruction of high-fidelity and geo-specific 3D terrain data. However, traditional manual annotation methods (KpConv, RandLA, Mask3D) are considered both costly and time-intensive and the annotation of 3D data consistently presents significant challenges, given the inherent scarcity of manually annotated 3D datasets, particularly for the military use cases.

Objectives

Recognizing this gap, our previous research leverages the data repository's manually annotated databases, as showcased at I/ITSEC 2019 and 2021, to enrich the training dataset for deep learning (DL) models. However, collecting and annotating large scale 3D data for specific tasks remains costly and inefficient. To this end, the objective of this research is to design and develop a comprehensive and efficient framework for 3D segmentation tasks to assist in 3D data annotation. This framework integrates Grounding DINO (GDINO) and Segment-anything Model (SAM), augmented by an enhancement in 2D image rendering via 3D mesh. Furthermore, the authors have also developed a user-friendly interface (UI) that facilitates the 3D annotation process, offering intuitive visualization of rendered images and the 3D point cloud.

Results

To evaluate the proposed annotation framework, outdoor scenes from collected by using unmanned aerial vehicles (UAVs) and indoor scenes collected by using NavVis VLX and RGB-D camera in USC-ICT office building were used to conduct comparative experiments between manual methods and the proposed framework, focusing on 3D segmentation efficiency and accuracy. The results demonstrate that our proposed framework surpasses manual methods in efficiency, enabling faster 3D annotation without compromising on accuracy. This indicates the potential of the framework to streamline the annotation process, thereby facilitating the training of more advanced models capable of understanding complex 3D environments with satisfactory precision.

Our results demonstrate the framework's efficiency and potential in creating high-quality 3D annotated datasets across various settings. By optimizing this process, the framework minimizes reliance on manual labor, thereby boosting overall efficiency and productivity. In practical terms, the framework's ability to annotate 3D data with an open vocabulary not only enhances accuracy but also expands the scope of applications in military simulations.

Next Steps

The capability of the proposed framework to annotate objects accurately in military-related scenes suggests its potential in automating tasks in the U.S. Army M&S. By reducing manual effort and enhancing annotation efficiency, the framework can significantly contribute to improving human operation and decision-making within military modeling and simulation applications. By automating and refining the annotation process, the framework empowers military personnel to focus on higher-level tasks, ultimately enhancing decision-making capabilities and operational readiness.

Published academic research papers are available from https://ict.usc.edu/research/publications
(Search engine keyword: USC ICT Publications)

Project Leader: Andrew Feng, Meida Chen