Point cloud classification with PointNet

Abstract:

Point cloud is an important type of geometric data structure. Due to its irregular format, most researchers transform such data to regular 3D voxel grids or collections of images. This, however, renders data unnecessarily voluminous and causes issues. In this project, we design a novel type of neural network that directly consumes point clouds and well respects the permutation invariance of points in the input. Our network, named PointNet, provides a unified architecture for applications ranging from object classification, part segmentation, to scene semantic parsing. Though simple, PointNet is highly efficient and effective. Empirically, it shows strong performance on par or even better than state of the art. Theoretically, we provide analysis towards understanding of what the network has learnt and why the network is robust with respect to input perturbation and corruption.

Existing System:

During the past few years, 3D imaging technology experienced a major progress with the production of inexpensive depth sensors (e.g., Kinect). This caused a leap in the development of many successful semantic segmentation methods that use both RGB and depth. However, the 3D sensing field has recently undergone a follow up shift with the availability of mature technology for scanning large-scale spaces, e.g., an entire building. Such systems can reliably form the 3D point cloud of thousands of square meters with the number of points often exceeding hundreds of millions. This demands semantic parsing methods capable of coping with this scale, and ideally, exploiting the unique characteristics of such data. Thus, this system is incapable of classifying the point cloud.

Proposed System:

In this paper we explore deep learning architectures capable of reasoning about 3D geometric data such as point clouds or meshes. Typical convolutional architectures require highly regular input data formats, like those of image grids or 3D voxels, in order to perform weight sharing and other kernel optimizations. Since point clouds or meshes are not in a regular format, most researchers typically transform such data to regular 3D voxel grids or collections of images (e.g., views) before feeding them to a deep net architecture. This data representation transformation, however, renders the resulting data unnecessarily voluminous — while also introducing quantization artifacts that can obscure natural invariances of the data. For this reason we focus on a different input representation for 3D geometry using simply point clouds – and name our resulting deep nets PointNet.

Software Tools:

1. TensorFlow
2. Keras
3. Python3
4. TRIMESH
5. Matplotlib
6. Jupyter Notebook
7. VS Code

Hardware Tools:

1. Laptop
2. Operating System: Windows 11
3. RAM: 16GB RAM
4. GPU

Applications:

1. Classification of 3D Objects in indoor and outdoor navigation spaces.
2. Retail, Commercial Outlets where object identification is the key asset.