Viewpoint Equivariance for Multi-View 3D Object Detection

Abstract:

Multi-view 3D object detection plays a vital role in various computer vision applications, including autonomous driving, robotics, and augmented reality. Achieving accurate and robust object detection across multiple views requires addressing the challenges posed by changes in viewpoint. In this paper, we propose a novel approach to tackle the viewpoint variability problem, leveraging the concept of viewpoint equivariance.

Our method is based on a deep learning architecture that incorporates a viewpoint equivariant module. By introducing a dedicated equivariant module into the detection pipeline, we enable the network to learn viewpoint transformations that are invariant to the camera perspective. This viewpoint equivariance allows our model to generalize well across different views, enhancing detection accuracy and robustness.

To train our model, we develop a large-scale dataset consisting of multi-view 3D object annotations. We carefully design a data augmentation strategy that synthesizes diverse viewpoints while maintaining object-level consistency. By leveraging this dataset, we train our model to effectively capture viewpoint-invariant features and generalize to unseen views.

We evaluate our approach on several benchmark datasets for multi-view 3D object detection. Experimental results demonstrate that our viewpoint-equivariant model outperforms existing state-of-the-art methods, achieving significant improvements in detection accuracy across various viewpoints. Additionally, our approach demonstrates robustness to changes in camera parameters, such as focal length and sensor resolution.

Furthermore, we conduct extensive ablation studies to analyse the effectiveness of our viewpoint equivariant module and provide insights into its impact on performance. The results highlight the importance of considering viewpoint equivariance for multi-view 3D object detection tasks.

In summary, our proposed approach addresses the challenge of viewpoint variability in multi-view 3D object detection through the incorporation of a viewpoint equivariant module. The experimental results demonstrate the effectiveness and robustness of our method, providing a significant step forward in enhancing the accuracy and generalization capabilities of multi-view 3D object detection systems.

Learning Object-level Point Augmentor for Semi-supervised 3D Object Detection

Abstract:

Three-dimensional (3D) object detection plays a crucial role in various applications such as autonomous driving and robotics. However, obtaining annotated 3D point clouds for training purposes is often challenging and time-consuming. Semi-supervised learning techniques have emerged as a promising solution to address the scarcity of labeled data, leveraging both labeled and unlabeled samples. In this paper, we propose a novel approach called the Object-level Point Augmentor (OPA) for semi-supervised 3D object detection.

OPA aims to enhance the performance of 3D object detection by generating realistic synthetic data to augment the limited labeled dataset. Unlike previous methods that augment individual points, OPA operates at the object level, enabling the generation of semantically coherent objects. Our approach leverages the power of generative models, specifically a conditional generative adversarial network (cGAN), to synthesize realistic object instances. By utilizing the available labeled data, OPA effectively learns the underlying structure and characteristics of the objects, enabling the generation of high-quality augmented samples.

To ensure the generated samples are consistent with the real-world distribution, we introduce an iterative refinement mechanism that progressively improves the quality of the augmented objects. This refinement process leverages both labeled and unlabeled data, enhancing the generalization capability of the model and enabling better performance on unseen objects during testing.

We extensively evaluate the proposed OPA approach on several benchmark datasets for 3D object detection. Experimental results demonstrate that our method outperforms existing state-of-the-art techniques in terms of detection accuracy and robustness, especially in scenarios with limited labeled data. The OPA approach significantly reduces the dependence on expensive and time-consuming annotation processes, making it a valuable tool for practical applications.

Overall, our work introduces a novel solution for semi-supervised 3D object detection, leveraging object-level augmentation and iterative refinement. By enabling the generation of realistic synthetic data, OPA enhances the generalization capability of the model and improves detection performance in scenarios where labeled data is limited. Our research contributes to the advancement of 3D object detection and offers potential for broader adoption in real-world applications.