

Segmentasi Objek 3D dengan Edge Computing melalui iPhone = 3D Object Segmentation with Edge Computing via iPhone

Adrian Kaiser, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20529027&lokasi=lokal>

Abstrak

Segmentasi semantik adalah sebuah task pada bidang computer vision yang dewasa ini menjadi semakin penting. Segmentasi semantik sendiri dapat dipakai untuk memisahkan satu benda dengan benda yang lainnya, baik pada dua dimensi maupun tiga dimensi. Segmentasi semantik tiga dimensi umumnya mengutilisasikan sebuah point cloud yang dapat diambil menggunakan sensor Light Detection and Ranging (LIDAR). Sejak 2020, Apple menyertakan sensor LIDAR pada beberapa model iPhone. Hal tersebut memungkinkan orang awam untuk merekonstruksi berbagai objek dan keadaan di sekitarnya. Berdasarkan hal tersebut, dapat dirumuskan sebuah aplikasi yang dapat membantu penggunanya untuk melakukan scan terhadap benda rumah tangga untuk mengetahui panjang, lebar, tinggi, dan volume melalui kombinasi dari segmentasi semantik dan beberapa metode lainnya. Dibandingkan juga performa beberapa model yang menjadi kandidat integrasi dengan aplikasi tersebut, yaitu Dynamic Graph Convolutional Neural Network (DGCNN), Kernel Point Convolutional Neural Network (KPCConv), Point Transformer, dan Point Transformer dengan Contrast Boundary Learning (CBL). Hasil pengujian menunjukkan bahwa Point Transformer dengan CBL memiliki Intersection over Union yang paling baik. Didapatkan juga bahwa DGCNN adalah model yang paling baik untuk diimplementasikan sepenuhnya pada iPhone untuk edge computing.

.....Semantic segmentation is a computer vision task that has become increasingly important in recent years. Semantic segmentation can be utilized to separate one object from another in a two dimensional or three dimensional environment. Semantic segmentation normally utilizes a point cloud that can be obtained using a Light Detection and Ranging (LIDAR) sensor. As of 2020, Apple has packaged a built-in LIDAR sensor on a few iPhone models. This allows everyday users to reconstruct all sorts of objects around them. Owing to that fact, there can be formulized an application that helps its users to find the length, width, height, and volume of an object through a combination of semantic segmentation along with a few other methods. We also compared the performance of different models as candidates to be integrated into the application, which are Dynamic Graph Convolutional Neural Network (DGCNN), Kernel Point Convolutional Neural Network (KPCConv), Point Transformer, and Point Transformer with Contrast Boundary Learning (CBL). We found that Point Transformer with CBL has the best Intersection over Union result. We also found that DGCNN is the best model to be fully implemented on an iPhone for edge computing.