NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

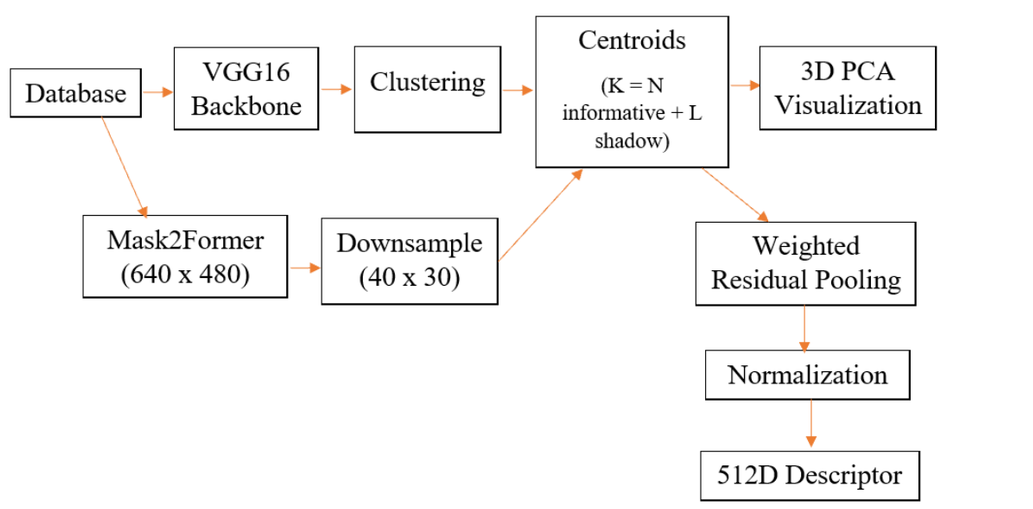
School of Computer Science and Engineering
College of Engineering

Visual Localization at NTU Campus

SCSE22-0277

Student: Kesarimangalam Srinivasan Abhinaya Supervisor: Prof. Lin Weisi

Comparison of different illuminations



Implemented APPSVR Architecture

Sample Image of SCSE Carpark

Project Objectives:

The aim of the project was to study the effect of semantic segmentation in visual localization; NetVLAD and APPVSR as potential solutions for visual localization in an indoor location like the Nanyang Technological University (NTU) Campus. NetVLAD was appended as a pooling layer on top of the VGG-16 Architecture. Soft-assignment to different clusters created a trainable layer for the creation of VLAD descriptors. Semantic information, generated through state-of-the-art semantic segmentation model, Mask2Former was integrated into the NetVLAD pipeline to identify informative and shadow areas. Semantic-based attention was used to tackle dynamic foregrounds and changing weather conditions. Utilizing semantic information to generate attention has shown to be helpful with an increase in Recall@1 rates from 0.8381 to 0.8563.

Potential applications include visual localization based indoor navigation systems for GPS inaccessible locations such as NTU, and a game-based trail to motivate users to explore unknown locations in the metaverse. A proof of concept was developed for the same.

Recall Rates on Pittsburgh30k

Recall @ N	NetVLAD	Proposed Model
N = 1	0.8381	0.8563
N = 5	0.9468	0.941
N = 10	0.9666	0.9607
N = 20	0.9813	0.9792