Starter Project Writeup

Instructions

Predicting canonical NOCS map from images similar to this paper.

Project Overview

Used the ShapenNetPlain dataset (car images only).

There are 2461 shapes in the training dataset, 350 shapes in the validation dataset and 703 shapes in the test dataset.

Trained by SegNet with 9 output channels in 50 epochs. The first three channels are the predictions of NOCS maps. The second three channels are the predictions of X-NOCS maps. The last three channels are the predictions of Depth Peeling images. I did not predict masks, since shapes in ShapenNetPlain dataset do not contain backgrounds.

The model was only trained on single views.

Model summaries

SegNet: A decoder upsamples its input by using the pool indices from it encoder.

```
Model: "seg_net"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
encoder (Encoder)            multiple                  14731584
_________________________________________________________________
decoder (Decoder)            multiple                  14732525
=================================================================
Total params: 29,464,109
Trainable params: 14,725,083
Non-trainable params: 14,739,026
_________________________________________________________________
```

To improve the training speed, I used vgg16 pre-trained weight as the weight of encoder.
Output channels: 9,
Optimizer: adam,
Learning rate: 0.0001,
Loss function: L2.

Performance

The model was tested on a test dataset involving 703 shape images. The result is evaluated by 2-way Chamfer distance multiplied by 100. Due to limited hardware and time, I only trained on car images (single view) for 50 epochs and I did not use masks.

Loss	Output	Cars (Chamfer distance)
L2	(X-)NOCS+Peel	1.7195

The following graphs show some sample results.