1 Guided Gated Convolution Neural Network for Normal Inference

1.1 Light Net Evaluation

We evaluate the GCNN architecture based light net on the test dataset. The light net parameters are further used as the initial parameter of light branch of the Trignet.

A qualitative evaluation of light net is shown in Figure 1. The average angular error are lower than 0.3 degree on all of test cases. In this case, to distinguish the output and the ground truth is already not easily only use naked eyes.
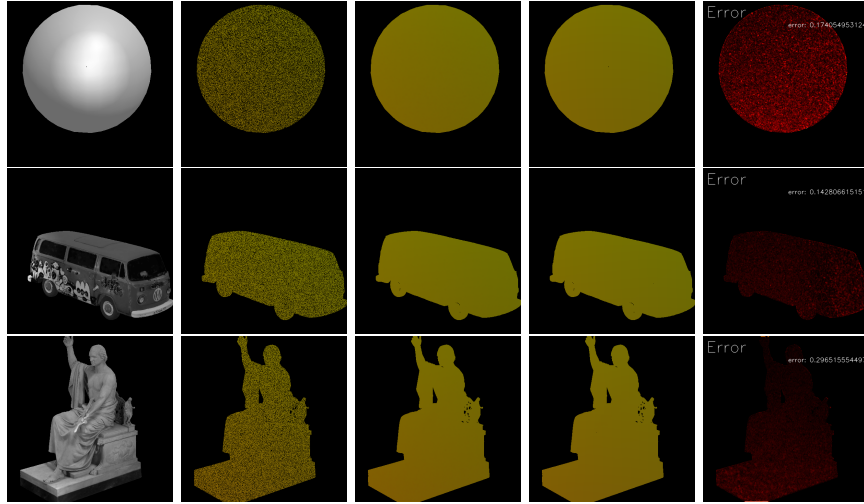

Figure 1: Qualitative evaluation of light net on Sphere, bus and Washington statue. From left to right, grayscale image, semi-dense light map(input), full-dense light map(output), full-dense light map(ground-truth), error map.

Figure 2 uses a scatter plot to show the average angular error on 100 different test cases. The average pixel-wise angular error is 0.17 degree as shown in Table 1. An regression line has been added in the plot to analysis the tendency of the errors. The angular error slightly goes up when valid point number in the scene increase. It is make sense since the valid pixels are usually connected and concentrate in a single patch, the less of the area of the patch, which corresponding the number of the valid points, the less variation of the light direction among the pixels, thus better the evaluation.
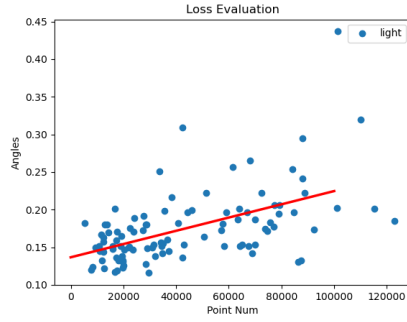
Figure 2: The evaluation on 100 test cases of Light Net.

1.2 Trignet Net Evaluation

The qualitative evaluation of Trignet uses the model trained based on masked L2 loss.
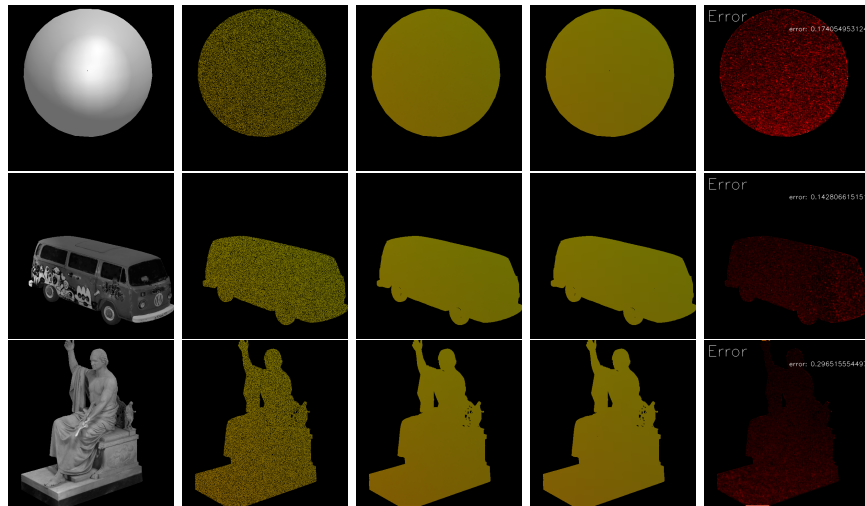
1.2.1 title

Figure 3: Qualitative evaluation of Trignet net on Sphere, bus and Washington statue. From left to right, semi 3D vertex map, full-dense normal map(output), full-dense normal map(ground-truth), error map.
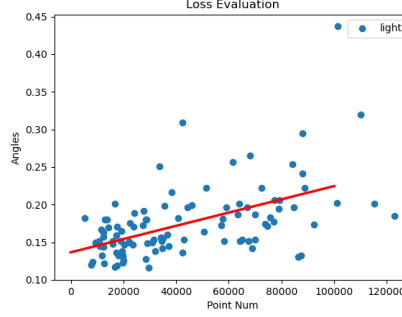
Figure 4: The evaluation on 100 test cases of Trignet.

1.3 comparison

From the Figure 5 we can observe the normal difference between ground-truth and GCNN predicted normals in another dimension. It separates the interval $[-1, 1]$, which is exactly the range of normal vector, to 256 sections. Then it counts the number of points locates in each section for 3 axes. The 3 axes are fitted quite well in most of interval but other than $[-0.25, 0.25]$ for x and y axes and interval close to -1 for z axis. Therefore a further constraint can be considered to the loss function related to the normal difference shown in this figure.

It is faulty that almost no normal has -1 z-component in GCNN predicted normal map. The reason?

Model	Angle	Time /ms	bz	lr-schedule	lr-df	l/i. Nr.
LightNet	0.17	4.72				
GCNN	10.57	5.25	8	200,1600	0.5	0
an3	9999	9999	8			1
VIL-1	9999	9999	8	10,1000	0.5	1
VIL-1	9999	9999	8	3,12,1000	0.5	1
VIL-3	10.83	53.95	8	100,1000	0.1	3
VIL-10	11.20	134.28	8	100,1000	0.1	10

Table 1: The error of models. The angle error is the average angle error of all valid pixels in the test case. The time unit is millisecond. bz is the batch size, lr-schedule is learning rate schedule. lr-df is learning rate decay factor, l/i. Nr is the number of light-image maps used for each scene
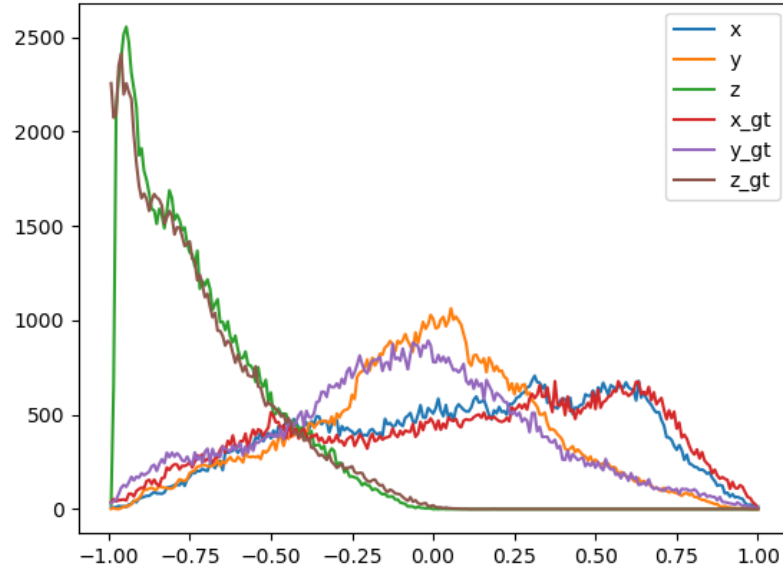
3

Figure 5: The normal difference of between GCNN and ground-truth in x, y, z-axis respectively. The y axis indicates the number of points, x axis indicates the value of normal in x/y/z axis. (The chart is based on the "dragon" scene showing above)
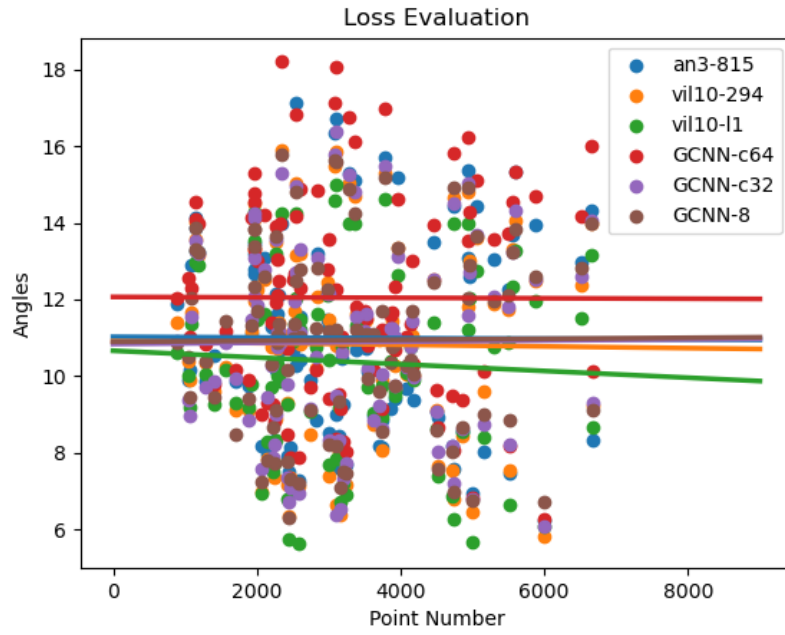
Figure 6: The normal difference of between GCNN and ground-truth in x, y, z-axis respectively. The y axis indicates the number of points, x axis indicates the value of normal in x/y/z axis. (The chart is based on the "dragon" scene showing above)