1-7

Classification of 3D Models Made from Images from Overhead by Using Spherical Convolutional Neural Networks

Shu SUMIMOTO　Yoko UWATE　Yoshifumi NISHIO
(Tokushima University)

1. Introduction

Drones are recently infiltrating various fields, for example delivery, rescue, guard and so on. It is necessary that drones fly safely. Then image recognition by deep learning is becoming important for that drones fly safely. However, standing people in overhead images from the view of drones are not able to be recognized by using YOLOv3 which is an object detection algorithm.

In this study, we make 3D models, such a human and cars, in overhead images with Pixel2Mesh[1]. It is the neural network algorithm which can make 3D models from 2D images to extract features of object apexes and reshape a elliptical sphere to a 3D model.

We classify the 3D models of a human and a car by using Spherical Convolutional Neural Networks on Unstructured Grids (UGSCNN)[2]. It has spherical convolutional layers which are icosahedral spherical meshes. It can classify 3D models to project 3D plots to a sphere and extract 3D features.

We aim at differentiating human or other objects to make 3D models of them from 2D images.

2. Proposed system

We propose to classify objects of a human and a car taken from overhead.

First, we process RGB images from overhead into 3D models by using Pixel2Mesh in Fig. 1. It learned ShapeNet data set for 3D models and our data sets are made by transfer learning. One of our data sets have each 100 3D models of a human and a car were made from images with scenery. Another data set of 3D models made from images without scenery.

Second, we classify 3D models of a human and a car with the UGSCNN. It has 7 spherical convolutional layers. They are made with 20 meshes, and 3D plots are projected to them. Convolutional processes work on sphere and features of 3D models are extracted.

(a)　(b)　(c)　(d)

Figure 1: RGB images and 3D models.
(a) A RGB image with scenery.
(b) A 3D model with scenery.
(c) A RGB image without scenery.
(d) A 3D model without scenery.

3. Results

We compare the learning and test accuracies of 3 data sets of images or 3D models of a human and a car. The first data set has only RGB images which are 100 human images and 100 car images taken from overhead. The second data set has 3D models, which are 100 human models and 100 car models with scenery made from the first data set by using Pixel2Mesh. The third data set has 3D models as well as second data set, but they are made from first data set images without scenery.

The first data set is classified into a human or a car by using Convolutional Neural Networks which has 2 convolutional layers. The other data sets are classified by using UGSCNN.

The test accuracy of the classification for the first data set is 0.50. This low score is contrasted by higher clasification accuracy values for the second and third dataset, which respectively contain 3D models data with and without scenery. It is considered that 2D images have only RGB information of a head or a shoulder. On the other hands, 3D models have more detailed information than 2D images.

Table 1: Test accuracies.

data set	Test accuracy
RGB imaegs	0.50
3D models with scenery	0.60
3D models without scenery	0.70

4. Conclusion

In this study, we have investigated comparison of the classification into a human or a car taken from overhead with the 3 data sets. The first data set has only 2D images taken from overhead. The other data sets has 3D models made from the first data set.

From these simulation results, we consider it is effective for classifying objects taken from overhead to make 3D models from 2D images.

However, the test accuracies are still low, so we will let UGSCNN learn another 3D data made by the sensors.

References

[1] Nanyang Wang, Yinda Zhang, Zhuwen Li, Yanwei Fu, Wei Liu, and Yu-Gang Jiang, “Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images”, ECCV 2018.

[2] Chiyu “ Max ” Jiang, Jingwei Huang, Karthik Kashinath, Prabhat, Philip Marcus, Matthias NieBner, “Spherical CNNs on Unstructured Grids”, ICLR 2019.