1-14

Image Classification for Some Training Data Sets of Image from Overhead by Using Convolutional Neural Network

Shu SUMIMOTO Yuichi MIYATA 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 investigate the prediction of the depth of some objects, such a human and chairs, in overhead images with Fully Convolutional Residual Networks (FCRN). Further, we blend a RGB image and a depth prediction image by using OpenCV. Also, we process RGB images into edge extraction images.

We classify the images of a human and a car which are composed of depth prediction images and edge extraction images, blended images by using Convolutional Neural Network (CNN). We aim at differentiating human or other objects.

2. Proposed System

We propose to classify images of a human and a car by using a CNN. First, we process RGB images into depth prediction images by using FCRN and edge extraction images, images blended by RGB images and depth prediction images by using OpenCV in Fig. 1. Second, we classify images of a human and a car with the CNN which has 2 convolutional layers, 2 pooling layers and 2 fully connected layers.

We compare the learning and test accuracies of each data set. When the CNN learns, images are compressed. Therefore, training and test images are 32 $\times$ 32 pixels. The learning rate of this the CNN is 0.00009.

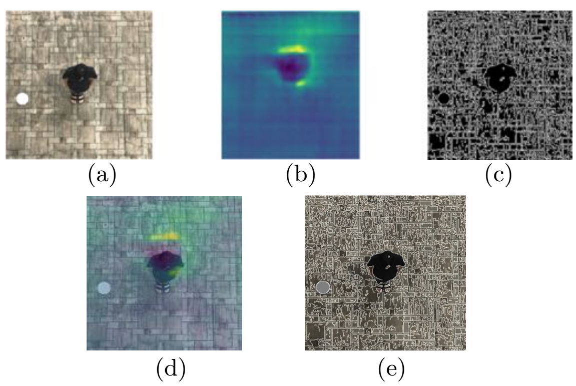

Figure 1: Images of a human from overhead.
(a) A RGB image. (b) A depth prediction image.
(c) An edge extraction image.
(d) A image blended by a RGB image and a depth prediction image.
(e) A image blended by a RGB image and an edge extraction image.

3. Simulation Results

When the CNN learns 100 RGB images, the test accuracy is 0.5 in Table 1. This value is very low, in other words the CNN can not classify a human and a car from overhead view. From this result, we prepare other three data sets, and compare thier test accuracies. First data set has 100 RGB images and 100 depth prediction images. Second data set has 100 RGB images and 100 edge extraction images. Third data set has 100 imgages blended by a RGB image and a depth prediction image, and 100 images blended by a RGB image and an edge extraction image.

The test accuracies in Table 2 and 3 are higher value than in Table 1, so we consider that it is effective for the CNN to classify images form overhead. Also, when the CNN learns blended images, the test accuracy in Table 4 is higher than in Table 1. From these results, the combination with RGB information is important.

Table 1: Average of learning and test accuracies when the CNN learn 100 RGB images.

Learning accuracy	1.00
Test accuracy	0.50

Table 2: Average of learning and test accuracies when the CNN learn 100 RGB images and 100 depth prediction images.

Learning accuracy	1.00
Test accuracy	0.80

Table 3: Average of learning and test accuracies when the CNN learn 100 RGB images and 100 edge extraction images.

Learning accuracy	1.00
Test accuracy	0.67

Table 4: Average of learning and test accuracies when the CNN learn 100 imgages blended by a RGB image and a depth prediction image, and 100 images blended by a RGB image and an edge extraction image.

Learning accuracy	1.00
Test accuracy	0.73

4. Conclusion

From these simulation results, we consider it is effective for classifying images taken by a camera which is far from an object to predict the depth and extract edge of images.

However, the test accuracies are still low, so we will make other data sets based on combination with RGB information.