Bulletin of Networking, Computing, Systems, and Software – www.bncss.org, ISSN 2186-5140
Volume 9, Number 1, pages 46–47, January 2020

Cart Path Recognition in a Golf Course Using Deep Fully Convolutional Networks

Gukjin Son, Junkang Kim, Wooyoung Jung and Youngduk Kim*
Research Division of Future Automotive Technology
DGIST
Daegu, Korea
{sudopop, kjk1208, wyjung, ydkim}@dgist.ac.kr

***Abstract*— Autonomous driving vehicles are a growing reality. Autonomous driving related industries are also growing. We believe that an unmanned golf cart is an appropriate application area for such autonomous driving. So, we have commenced the project of the golf cart the purpose of which is to provide an assistant to golf players. The golf cart includes almost all the functions the Autonomous driving should have. Image processing and recognition capability are of course one of the key functions that a golf cart needs. To develop such a kind of a golf cart would take much time. Thus, we have started the research and development of a demonstration prototype, which can perform some of the main tasks of the golf cart. In this paper, we will deal with the recognition of a cart path by themselves, trained end-to-end, pixels-to-pixels. We have collected and annotated the cart path of challenging scenes captured in a golf course. And we developed a deep learning network for a cart path. As a result, we achieve 93% accuracy and 50ms inference time.**

Keywords—fully convolutional network, golf cart, cart path, Semantic Segmentation, deep learning, autonomous driving

I. Introduction

Autonomous driving is a large system that consists of various vision system and sensors. Autonomous driving has a major impact on other industries. An unmanned golf cart (UGC) is an appropriate application area for such autonomous driving. There are several reasons for this. One is that the functions required for a UGC include almost all the functions the Autonomous driving should have. Image processing and recognition capability are of course one of the key functions that a UGC needs. Furthermore, it should be able to sense environmental conditions such as gap with the golf player and distance with the obstacle using various kinds of sensors and it should possess sophisticated communication capability with golf players. These capabilities allow UGC to carry golf clubs and follow the golf player. On the other hand, it is difficult for many golf company to hire good caddies and it would be most helpful for them if UGC that can work for golf player could be realized. Based on the above considerations we have commenced the project of the UGC the purpose of which is to provide an assistant to golf players. The UGC will have to deal with its environment (obstacles, hazardous areas) and gap from the golf player. To develop such a kind of UGC would take much time. Thus, we have started the research and development of a demonstration prototype, which can perform some of the main tasks of the UGC. In this paper, we will deal with the recognition of the cart path.

Recently, a few convolutional neural network (CNN) based approaches have been developed to tackle the problem in an end-to-end fashion including learning-based algorithms. They demonstrate good performance on benchmarks and in

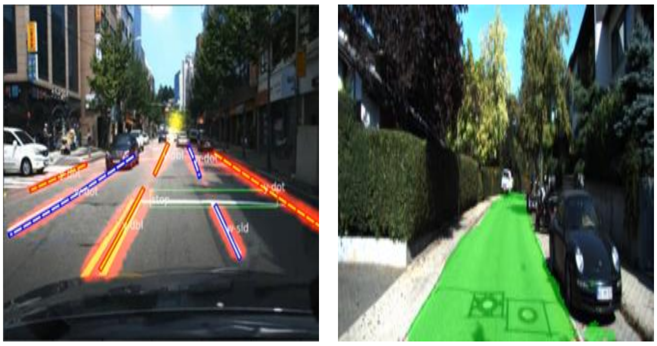

Figure 1. The left image shows the output of VPGNet [1]. The right image shows output of FCN [2]

real road scenes. It has several applications in in computer vision – road detection [1], [2] in Figure 1. CNN is very effective in semantic segmentation, that is labeling each region or pixel with a class of objects/non-objects. Semantic segmentation plays an important role in image understanding and essential for image analysis tasks. Because of these advantages, we use semantic segmentation of deep learning to recognize cart path. Our network is inspired by the work of Fully Convolutional Networks(FCN) [2]. Available datasets are often limited and insufficient for a golf course road. Since no proper datasets available for a golf course road, we have collected and annotated cart paths of challenging scenes captured in a golf course.

This paper is organized as follows. Section 2 explains our network architecture and training scheme, experimental results. Finally, Section 3 concludes our work.

II. METHOD

A. Data Collection and Annotation

We have collected the dataset in the Ciel golf course, South Korea. Since our dataset is captured under fine weather conditions, we mount a camera outside a UGC. The camera is directed to the f ront view of the UGC. Image resolution is 640×480. The number of images is shown in Table 1. We manually annotate a cart path and background in a pixel-level for each object. Each pixel contains a class label.

TABLE I. Number of frames in the dataset.

	Total frames
Training set	1,383
Validation set	345
Test set	562

TABLE II. PROPOSED NETWORK STRUCTURE.

Layer	Conv 1	Conv 2	Conv 3	Conv 4	Conv 5	Conv 6	Conv 7
Kernel size, stride, pad	3, 1, 100	3, 1, 1	3, 1, 1	3, 1, 1	3, 1, 1	1, 1, 0	1, 1, 0
Pooling size, stride	2, 2	2, 2	2, 2	2, 2	2, 2		
Addition						Dropout	Dropout

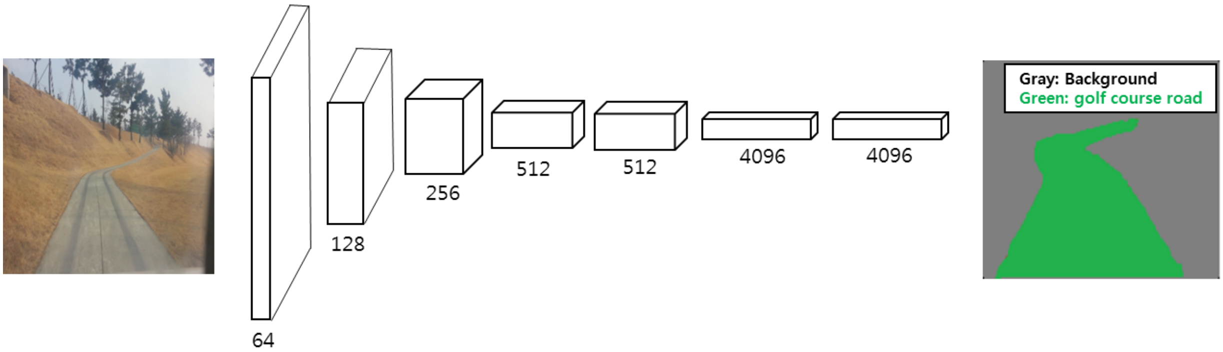

Figure 2. Our network can efficiently learn to make dense predictions for per pixel task like semantic segmentation.

B. *Architecture*

Our network is inspired by the work of [2]. Each layer of data in a convnet is a three-dimensional array of size h × w × d, where h and w are spatial dimensions, and d is the feature or channel dimension. The first layer is the image, with pixel size h × w, and d color channels. The network consists of a total of eight layers. Convolutional Neural Networks are built on basic components (convolution, max pooling and Relu activation functions). The overall architecture is described in Table 2 and Figure 2.

C. *Training*

We train by Stochastic Gradient Descent optimization with momentum. We use a minibatch size of 4 images and fixed learning rates of 10^{-5} for FCN-VGG16. We use momentum 0.99, weight decay of 5^{-4}. Dropout was included where used in the original classifier nets. It can be seen that the final loss of Iterate 6,000 converges to 0.002.

D. *Result*

We test our Network on semantic segmentation and scene parsing with 562 scene of test data. Although these tasks have distinguished between a cart path and background. We report one metric from common semantic segmentation and scene parsing evaluations that are region intersection over union(IU). Let n_{ij} be the number of pixels of class i predicted to belong to class j, where there are n_{cl} different classes, and let $t_i = \sum_j n_{ij}$ be the total number of pixels of class i. as in:

$$\textit{Mean IU: } (1/n_{cl} \sum_i n_{ii} / (t_i + \sum_j n_{ji} - n_{ii}) \qquad (1)$$

TABLE III. PERFORMANCE OF NETWORKS ON THE TEST SETS.

	accuracy	inference time
VPGNet [1]	97.3%	~ 90ms
Our	93%	~ 50ms

Figure 3. The left image shows the output of our network. The right image shows the output of the VPGNet.

If IU has a value of more than 50%, the image is classified as correct. We show the learning and inference of networks in Table 3. VPGNet outperforms our approach. But our network is faster inference time than VPGNet. Therefore, UGC is suitable for our network, which is capable of real-time processing.

III. CONCLUSION

In this work, we introduced how to recognize the cart paths in a golf course. The evaluation shows that our network is suitable for cart paths recognition. In the future study, we will study how to recognize additional objects such as obstacles, hazardous areas and seek ways to increase the accuracy.

ACKNOWLEDGMENT

This work was supported by the DGIST R&D Program of the Ministry of Science and ICT (19-FA-07)

REFERENCES

[1] S. Lee *et al.*, “VPGNet: Vanishing Point Guided Network for Lane and Road Marking Detection and Recognition.”

[2] E. Shelhamer, J. Long, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2017.