Traffic Light Recognition for Autonomous Driving Vehicle: Using Mono Camera and ITS

Mun-Kyu Lee, Jeong-Won Pyo, Sang-Hyeon Bae, Sung-Hyeon Joo, and Tae-Yong Kuc
Department of Electrical and Computer Engineering, College of Information and Communication Engineering, Sungkyunkwan University, Suwon, Korea
Email: munkyu@g.skku.edu, {jungwon900, shbae.skku, sh.joo, tykuc}@skku.edu

***Abstract*—This paper introduces a system of real-time Traffic Light (TL) recognition, which is an essential element for unmanned autonomous vehicles when driving through urban city. The system of TL recognition is an integrating system that fuses a binary processing method, a network model method, and an ITS information. By using two algorithms that processes results through a mono camera, we enhance our recognition accuracy when ITS information is not confirmed properly. We evaluated the individual and integrated TL recognition performance of the system in actual testbed to ensure that our system is satisfied the performance for our autonomous driving scenario. The result of the experiment met our autonomous driving scenario conditions.**

***Index Terms*—autonomous driving, traffic light recognition, ITS, image processing, deep learning**

I. INTRODUCTION

Unmanned autonomous driving research has developed rapidly, such as mobile robots, unmanned shuttle taxis, and cluster autonomous trucks. This research is mainly affected by the reliability of sensor information and the accuracy of algorithms that process this sensor information. For example, recognition performance can greatly depend on the sensor reception rate and recognition algorithm that we used. If one of the recognitions for autonomous driving is poorly processed, it may lead to major accidents.

In this paper, in particular, we proposed a system for performing TL recognition, which is the most important for autonomous driving of vehicles. The system is composed and tested in a real environment to ensure stable and efficient TL recognition. In this system, we perceive the status of TL through Intelligent Transport Systems (ITS), and in addition, we process a separate TL recognition using an image obtained from mono camera. By fusing this two information, the target status of the TL is received and used for autonomous driving. Here, we used Global Navigation Satellite System (GNSS) information and High-definition map (HD map) information to determine that the vehicle reaches near the TL. We verified the effectiveness and safety of this system by performing some scenarios on intersections and TL.

Manuscript received November 29, 2021; revised May 5, 2022.

II. METHOD

A. Proposed Method Overview

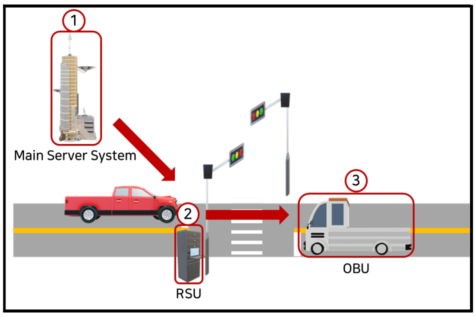

Figure 1. ITS communication method.

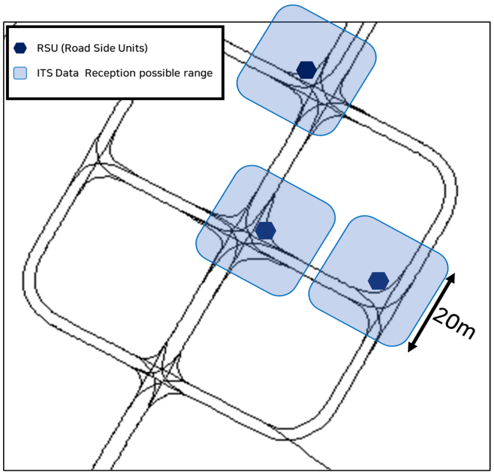

Figure 2. The reception range of ITS communication on the HD map.

Fig. 1 is a reception method of ITS information. ITS information is transmitted from the main server system to the Road Side Unit (RSU) installed at each intersection, and such information may be received from a vehicle equipped with a nearby On-board Unit (OBU). [1]-[4]

ITS communication is a form of transmitting signal system information encrypted based on the J2735 standard, by communication between OBU of the vehicle and RSU installed at the intersection area to communicate within a certain range. And Fig. 2 shows the range that data can be receive from each RSU.

The ITS method and the mono camera-based method each have pros and cons.

doi: 10.18178/joig.10.3.102-108

As shown in Fig. 1 and Fig. 2, the ITS-based TL recognition is close-range communication. This recognition method covers some cons of the mono camera as follows:

1) Differences in recognition accuracy according to differences in the amount of light over season, weather, and time.
2) Problems that cannot be recognized when the recognition range for a camera is covered by an object.
3) Unstable mono camera based-on computer vision perception itself.

However, ITS-based method still has cons. Because of the communication from the main server to the RSU and the OBU, there are relatively unstable than wired LAN communication. Besides such communication issues, it is difficult to understand which intersection data is required to drive a vehicle autonomously. Therefore, we use a location information from GNSS sensor and HD map to distinguish these intersections. This may cause a problem of correcting position of the GNSS even before a problem with the ITS itself occurs, and if proper location data is not obtained, a wrong decision may be made. In order to remedy such shortcomings of ITS, mono camera vision-based recognition methods are double-configured in this research.

B. *Device Configuration*

Before explaining the proposed method, we describe the devices we have configured for use in TL recognition.

1) *Device configuration for our system*

- GNSS and antenna.
- OBU and antenna.

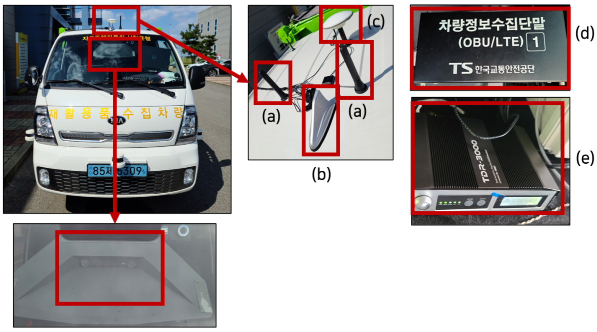

Figure 3. Device configurations on vehicle for ITS.

First, an OBU module (Fig. 3d) needs to be attached to a vehicle in case of ITS, a part of our system. Second, an LTE antenna (Fig. 3a) and a data transmission/reception antenna (Fig. 3b) need to be attached to both outside and center of the vehicle in order to use ITS efficiently. Finally, GNSS (Fig. 3e), which we use to locate the vehicle accurately, needs to have an antenna (Fig. 3c) at the same location as the above antennas to receive satellite data.

2) *Device configuration for computer vision*

As shown in Fig. 3f, the camera is configured at a height of 1.8m ahead the vehicle. The camera has an FoV as shown in Fig. 4, with a horizontal angle of 96 degrees and a vertical angle of 64 degrees. Camera hardware configuration is 2K/10hz in accordance with the environmental characteristics of our scenario that is slow speed when entering the stop line, as high definition is more advantageous than high scanning rate due to the characteristics of TL recognition.

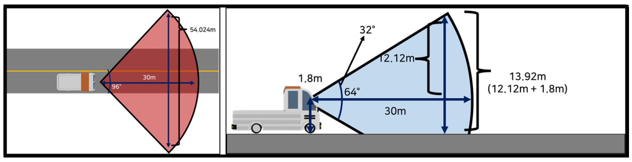

Figure 4. Camera FoV (vertical, horizontal).

C. *Proposed Method*

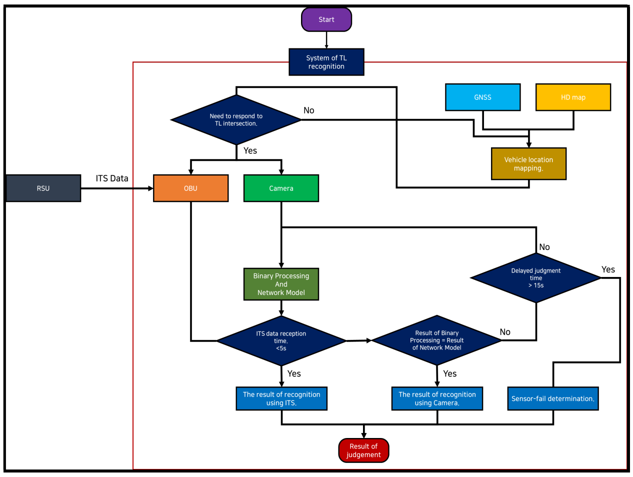

Figure 5. Flow chart of the overall system.

The overall flowchart of the integrated algorithm is as shown in Fig. 5. Our system identifies the location of the vehicle through mapping of GNSS and HD map, then starts recognizing TLs in intersections. We first set the information of ITS as the most reliable in the system. However, if the reception of ITS information is delayed for more than 5 seconds, the reliability of the result obtained through the camera sensor-based recognizer is set to the highest result and used as the final determination result. In the case of camera sensor-based recognizers, as mentioned above, the binary processing result is trusted when the network model method produces the same result. And if the overall system has a delay of more than 15 seconds, it is determined that the system is not functioning normally and halted.

1) *Computer vision-based recognition*

A series of data preprocessing processes are performed to be used efficiently for binary classifiers and object recognition. [5]-[10] The process is as follows:

a) *Color and shape-based classifier*

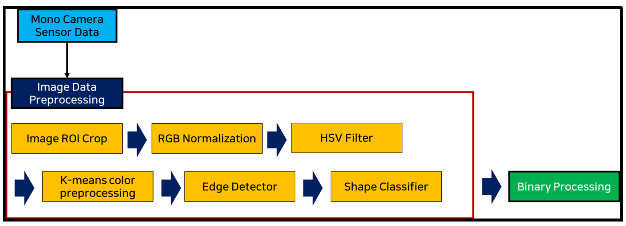

Figure 6. Color and shape-based classifier configuration chart.

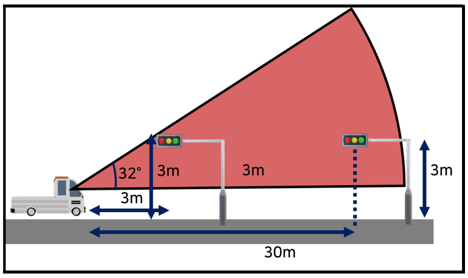

Figure 7. Compare TL's height and camera FoV.

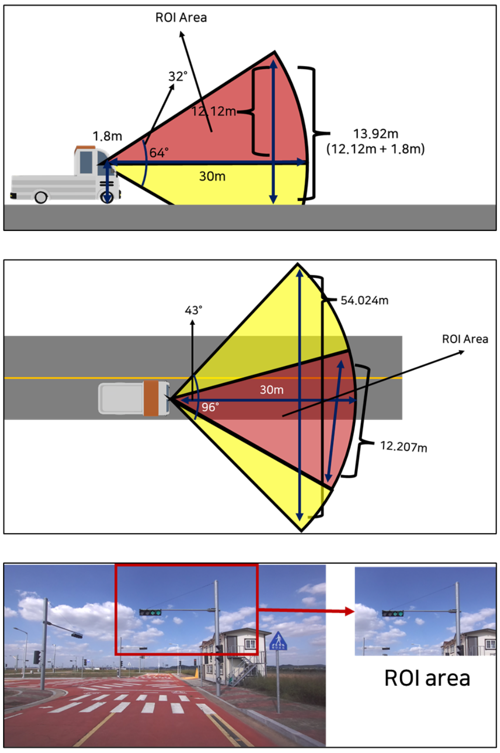

Figure 8. ROI FOV of the camera original FOV.

First, Fig. 6 is a flowchart of a color and shape-based classifier. Recognition values can be finally obtained through color and shape-based classifiers through processes such as image crop, RGB normalization, HSV filter, etc.

As shown in Fig. 7, since TL of testbed is installed at a height of 2.5 to 3m and 3 to 30m away from the stop line, the upper 32 degrees of the vertical angle of view 64 degrees were set as Region of Interest (ROI) in Fig. 8, and the lower 32 degrees were cropped. In addition, since the TL is installed on the right side of direction of traffic, 43 degrees, except for the unnecessary part of the horizontal angle of view 96 degrees, were set as ROI as shown in Fig. 8 and rest of them were cropped. Therefore, we could save the resource and reduce–interference in data other than TL.

Before the next step, color-based recognition, we applied RGB normalization to reduce color deviation in the image and Minmax Normalization by setting the color range from 50 to 200. Then, the result of comparing the differences before and after using data from bright and cloudy days are shown in Fig. 9.

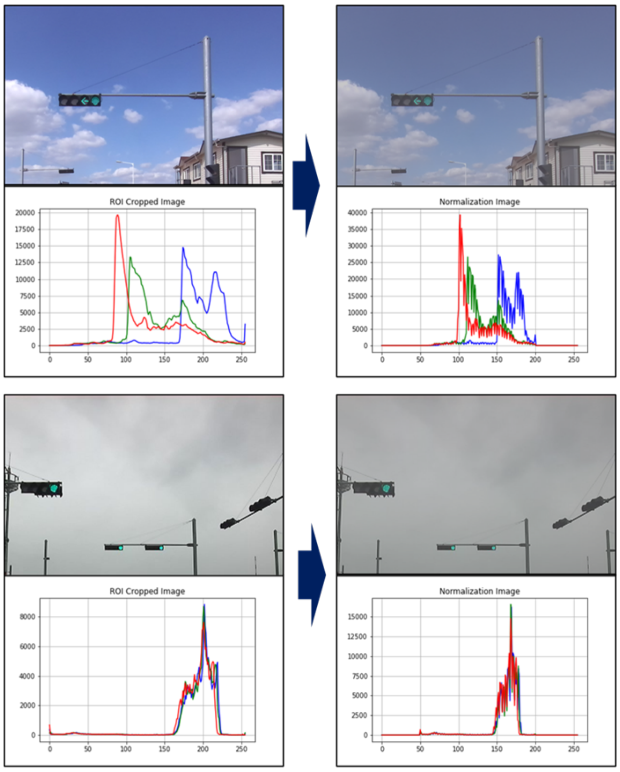

Figure 9. Comparing existing data with RGB normalization in a bright and cloudy environment.

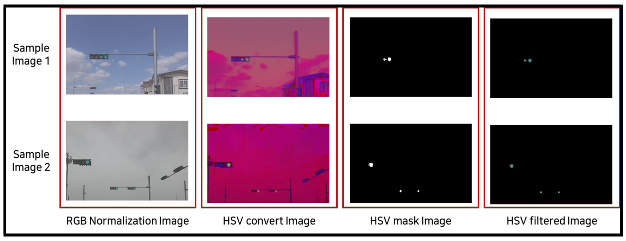

Figure 10. Compare original and HSV filtered data.

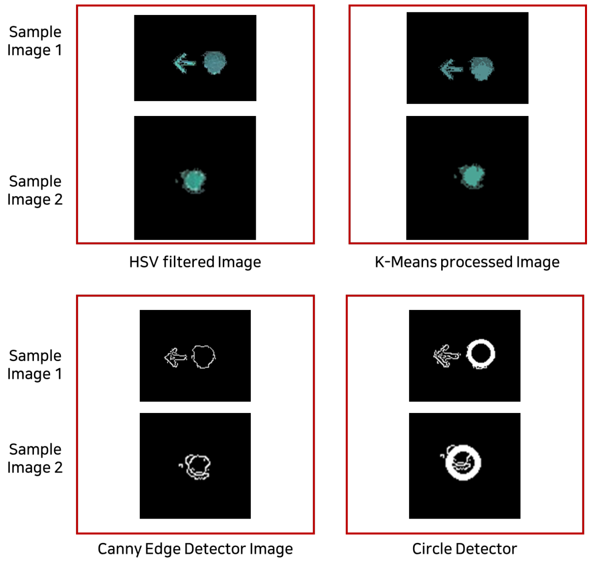

Figure 11. Difference in image preprocessing step by step.

In the next step, HSV filtering was used to use only a specific color range. Since we set green as a target color for recognition here, the range of H, S, and V values corresponding to the range of green should be used. We used colors within this range and set the rest of colors to 0. Fig. 10 is an example of the image using such an HSV filter. This figure includes the mask based on the range of HSV and the image to which the mask is applied.

After HSV filtering, color normalization using K-Means was applied for uniformity in color, which is the end of (color-related) preprocessing. Such preprocessing methods maximize the effect of HSV filtering. Then, we used Canny edge detection algorithm for Edge Detection due to shape classification.

As the last step, we need to detect circles and arrows in TL. We used Hough gradient method to detect the circular shape. (Fig. 11)

The following description of an arrow shape classifier is for left turn recognition. The very first step to detect the shape of the arrow is to detect a line, as the arrow consists of a line. We solely targeted arrows facing left to exclude arrows other than TL as much as possible. As shown in Fig. 12, we drew a straight line from the beginning to the end of the arrow, classified the direction of the arrow by calculating the angle. Then which allows to recognize whether a left turn can be made.

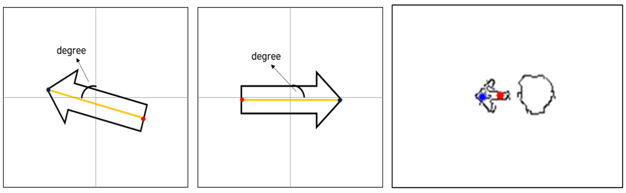

Figure 12. Data detected using a arrow shape detector.

b) *Object detector*

Our object detector (Fig. 13) worked with data applied up to RGB normalization, which is one of the preprocessing procedures we previously introduced. In this research, *Bosch Small Traffic Light Dataset* (http://k0b.de/bstld) was trained through Transfer Learning using YOLO v3[10]. Due to the difference between the form of TL and datasets in testbed, we increased cognitive performance by placing rotation at 90 degrees for use in horizontal TL. Right arrow data was used as left arrow data as well, because it is not present in the testbed environment.

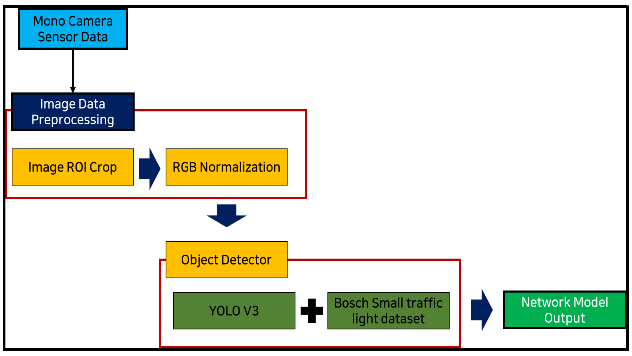

Figure 13. Object detector configuration chart.

III. TESTBED

As shown in Fig. 14 and Fig. 15, 1 The Central City Zone and 2 The Community Zone were selected as experimental areas among the various sections of testbed that represent the urban environment. In addition, as shown in Fig. 2, the shape of the road or the location of the intersection of the database HD MAP could be accurately identified.

Figure 14. The overall structure of testbed.

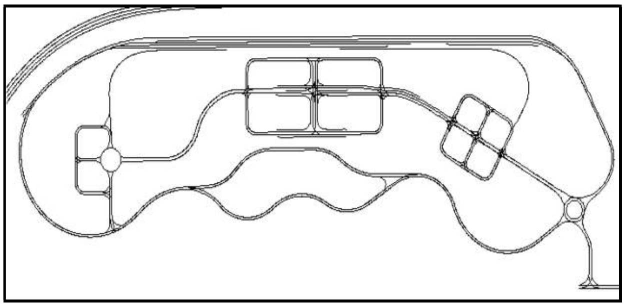

Figure 15. HD map of testbed.

One of the characteristics of this testbed is that there are four TL intersections in Zone 1 and three TL intersections in Zone 2. The distance from the stop line to TL is a minimum of 3-metre and a maximum of 30-metre and has light and shade, which allows us to conduct various cases of the TL recognition test. Most of all, we can construct and test complementary systems as this testbed is the ITS pilot area, which made it unique, by using the technology and vision-based TL recognition technology together.

IV. EXPERIMENT

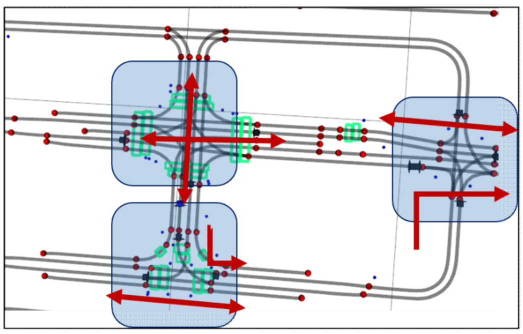

Figure 16. Example for direction in which vehicle can drive at the intersection.

TABLE I. INFORMATION ABOUT DIRECTIONS OF INTERSECTIONS

No.	Zone	Intersection	Direction (Towards North, south, east and west)	Straight TL.	left turn TL.	Distance (stop line ~ TL.)(m)
1	Community	1	W	X	O	3.5
2	Community	1	E	X	X	-
3	Community	2	N	O	O	6
4	Community	2	S	O	O	5
5	Community	3	W	O	O	5
6	City	4	W	X	O	3
7	City	5	N	O	O	30
...	...	...	...	...	...	...
42	City	7	E	X	O	3

Fig. 16 shows that there is a limit on the direction of driving at the intersection of the testbed. Table I shows the presence or absence of a straight and left turn TL for each zone, intersection, and direction, and the distance from the stop line to the TL. In this content, Intersection 5 of the central city zone has the farthest TL distance in the testbed, and in the case of Intersection 1 of the community zone, which is a three-pronged road, both straight and left turns are X, showing that it is impossible to go from east to west.

There are seven intersections with TLs on our testbed, but except on blocked areas, straight movement is possible in 23 directions and left turn is possible in 19 directions. If the direction of entering the intersection changes, various variables change, which means that our test conditions vary. We calculated the recognition rate of each individual system and an integrated system when performing multiple directions of straight and left turn recognition tests at each intersection.

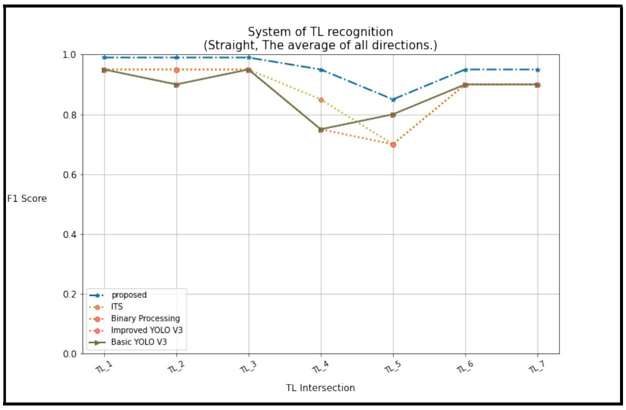

Figure 17. F1 score for each intersection in the straight scenario.

In Table I, Fig. 17, can see the score of the straight scenario. Tests were conducted in 23 straight directions for 7 intersections. It was confirmed that the accuracy of intersections 4 and 5 where RSU and TLs were relatively far from the stop line decreased a lot. However, in the case of straightforwardness, most of the F1 score was high.

In the case of left rotation, the accuracy of the camera sensor-based system decreased a lot, and the results were dependent on ITS. This showed that the reliability of the arrow shape recognition itself was low. However, it can be seen that it complements a single ITS system (Fig. 18).

Fig. 19 shows the average value of F1 score for left turn and straight scenario. It was confirmed that the system we proposed had a higher F1 score than each individual system.

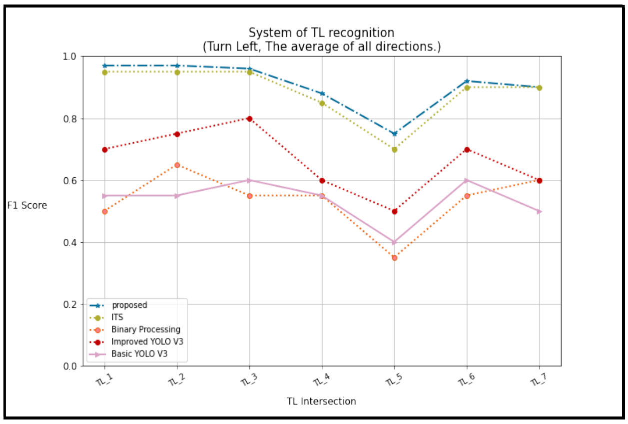

Figure 18. F1 score for each intersection in the left-turn scenario.

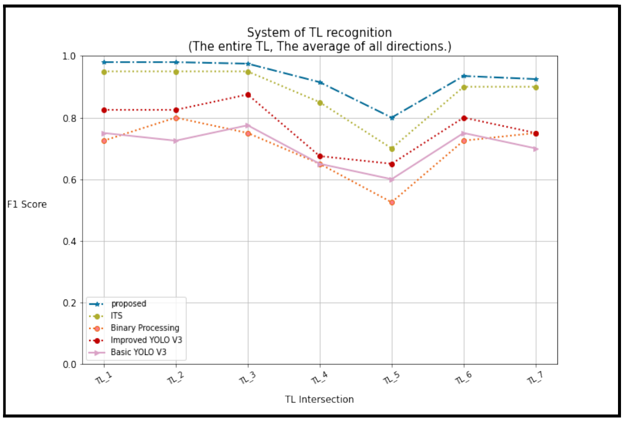

Figure 19. The average F1 score value of the straight and left turn scenarios.

V. CONCLUSION

This research proposed the TL recognition algorithm for unmanned autonomous vehicles in city environments. We integrated ITS with camera sensor data processing algorithms and used them together. The method we recognize TL in the following scenarios with less difficulty than a single method.

1) When an object (ex. vehicle) in front covers the angle of view of the camera.
2) Various environmental situations (amount of sunlight and clouds, differences in seasons, etc.)
3) A situation where it is recognized outside the RSU range.

In consequence, the testbed and scenarios we had defined showed appropriate performance for autonomous driving.

However, when the distance grew, and the sunlight got too strong or the cloud too thick, the recognition rate of the left turn sign might decrease. And it was also difficult to distinguish it from the actual TL, when there was a light with a green HSV color range other than the TL. This may be seen as a limitation of the camera sensor itself, but it is thought that it can be overcome to some extent if the overall layout of the TL is used for recognition.

Our research was conducted in a specific testbed, which shows such differences in conditions from general roads as the consistency of height of the TL and non-existence of high-rise building.

It is thought that several additional adjustments will be needed to be used on general roads. For example, it is considered necessary to commercialize ITS on the general roads, adjust parameters such as camera angle of view and ROI, and expand datasets.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization: Mun-Kyu Lee, Jeong-Won Pyo, Sang-Hyeon Bae, and Tae-Yong Kuc; validation: Mun-Kyu Lee, Sang-Hyeon Bae, Sung-Hyeon Joo, S.M., and Tae-Yong Kuc; investigation: Mun-Kyu Lee, Jeong-Won Pyo, Sang-Hyeon Bae, and Sung-Hyeon Joo; resources: Mun-Kyu Lee, Jeong-Won Pyo, Sang-Hyeon Bae, and Sung-Hyeon Joo; data curation: Jeong-Won Pyo, Sang-Hyeon Bae, and Mun-Kyu Lee; writing—original draft preparation, Mun-Kyu Lee; writing—review and editing: Mun-Kyu Lee; visualization: Mun-Kyu Lee; supervision: Tae-Yong Kuc; project administration: Tae-Yong Kuc; funding acquisition: Tae-Yong Kuc.

All authors have read and agreed to the published version of the manuscript.

Acknowledgment

This research was supported by the SungKyunKwan University and the BK21 FOUR (Graduate School Innovation) funded by the Ministry of Education (MOE, Korea) and National Research Foundation of Korea (NRF).

The research was conducted through the K-City test bed and infrastructure support of the Korea Transportation Safety Authority's Automobile Safety Research Institute.

References

[1] B. Williams, *Intelligent Transport Systems Standards*, Artech House, 2008.
[2] J. Kim, H. Cho, M. Hwangbo, and J. Choi, "Deep traffic light detection for self-driving cars from a large-scale dataset," in *Proc. 21st International on Intelligent Transportation*, 2018.
[3] H. Nguyen, "A deep learning framework for small traffic light recognition in traffic scene image," *Journal of Theoretical and Applied Information Technology*, 2020.
[4] Z. Ennahhal and I. Berrada, "Real time traffic light detection and classification using deep learning," in *Proc. International Conference on Wireless Networks and Mobile Communications*, 2019.
[5] U. Franke, D. Pfeiffer, C. Rabe, C. Knoeppel, M. Enzweiler, F. Stein, and R. G. Herrtwich, "Making bertha see," in *Proc. the Workshop on Computer Vision for Autonomous Driving*, Sydney, Australia, 8-12 December 2013.
[6] A. Broggi, P. Cerri, S. Debattisti, M. C. Laghi, P. P. Porta, M. Panciroli, and A. Prioletti, "PROUD-public road urban driverless test: Architecture and results," in *Proc. the IEEE Intelligent Vehicles Symposium*, Dearborn, MI, USA, 8-11 June 2014, pp. 648-654.
[7] S. Kato, E. Takeuchi, Y. Ishiguro, Y. Ninomiya, K. Takeda, and T. Hamada, "An open approach to autonomous vehicles," *IEEE Micro*, vol. 35, no. 5, pp. 60-69, 2015.
[8] J. P. Powell and R. Palacin, "Passenger stability within moving railway vehicles: Limits on maximum longitudinal acceleration," *Urban Rail Transit*, vol. 1, pp. 95-103, 2015.
[9] Y. Keisuke, K. Akisuke, S. Naoki, A. Toru, M. Aldibaja, and R. Yanase, "Robust traffic light and arrow detection using digital map with spatial prior information for automated driving," *Sensors*, vol. 20, no. 4, 2020.
[10] J. R. A. Farhadi, "YOLOv3: An incremental improvement," *Tech Report on Computer Vision and Pattern Recognition (cs.CV)*, 2018.

Mun-Kyu Lee received the B.S. degree from the school of Department of Economics, Chosun univ, Gwang-ju, South Korea, in 2020. He is currently pursuing a Master's degree with the School of Electrical and Electronics Engineering, SungKyunKwan University, Suwon, South Ko- rea. His research interests include artificial intelligence, deep learning, and data analysis.

Jeong-Won Pyo received his B.S. degree from the School of Department of Mechatronics Engineering, Korea Polytechnic University, Siheung, South Korea, in 2014. He is currently pursuing a Ph.D. degree with the School of Electrical and Electronics Engineering, SungKyunKwan University, Suwon, South Korea. His research interests include autonomous driving, artificial intelligence, and image processing.

Sang-Hyeon Bae received his B.S. degree from the School of Electrical and Electronics Engineering, Sungkyunkwan University, Suwon, South Korea, in 2017, where he is currently pursuing his Ph.D. degree. He has been with the School of Electrical and Electronics Engineering, Sungkyunkwan University. His research interests include control and planning.

Sung-Hyeon Joo received his B.S. degree from the School of Electrical and Electronics Engineering at Sungkyunkwan University in 2017. He has been with the School of Electrical and Electronics Engineering at Sungkyunkwan University, Suwon, Korea, where he is currently a Ph.D. candidate. His research interests include mobile robot navigation and semantic slams.

Tae-Yong Kuc received a B.S. degree in control and instrumentation engineering from Seoul National University, South Korea, in 1988, and M.S. and Ph.D. degrees from the Pohang University of Science and Technology, South Korea, in 1990 and 1993, respectively. From April to Au- gust 1993, he worked as a chief research engineer at the Precision Machinery Institute of Samsung Aerospace Company and from September 1993 to

February 1995 as a senior lecturer with the Department of Electrical Engineering, Mokpo National University, South Korea. Since March 1995, he has been with the School of Electrical and Electronics Engineering, Sungkyunkwan University, Suwon, South Korea, where he is currently a professor. His research interests include intelligent robotics, adaptive and learning control, and visual sensor processing for computer-aided control systems.