Human detection method using MobileNetV2 as an indoor safety management system

Soo Jeong Choi
Dept. of IT Engineering
Sookmyung Women's University
Seoul, South Korea
suzzang77@sookmyung.ac.kr

Yujin Lim
Dept. of IT Engineering
Sookmyung Women's University
Seoul, South Korea
yujin91@sookmyung.ac.kr

***Abstract*— With the recent development of artificial intelligence technology, artificial intelligence technology is being actively applied to various fields. Among them, in the field of safety management, it is necessary to continuously develop a safety management system through object detection. However, in the indoor safety system, the objects are densely populated and the performance of object detection rapidly decreases. To solve this problem, we analyze the performance of the existing object detection technologies in the indoor environment where people are densely populated. The performance was analyzed by applying the soft-NMS technique based on SSDLite and MobileNetV2 models.**

I. INTRODUCTION

With the recent development of artificial intelligence technology, it is being applied to various fields. In particular, it is widely applied in the field of safety management. In addition, a safety management system through automated human detection is constantly required in modern society. However, when objects, especially people, are densely populated in a limited indoor space, the overlapping of objects in the image sharply increases, which decreases the performance of object detection.

Therefore, to solve this problem, we analyze the performance of existing object detection techniques in a limited indoor environment with high density of objects as a system for indoor safety management. We assume that detection objects were limited to people. In addition, SSDLite[1] and MobileNetV2[1] with relatively low computing resources requirements were adopted. The performance of object detection was analyzed in various experimental scenarios by applying soft-NMS[2] to improve the detection rate of densely populated objects in the image.

II. PERFORMANCE ANALYSIS

In this paper, using Crowdhuman Dataset[3], a dataset with high object redundancy, we experimented by dividing it into four levels based on the degree of people's redundancy. As the level goes up, the number of people overlapped in the image increases. We analyzed the performance according to the degree of object redundancy in the image by dividing it into a soft-NMS Gaussian technique and a Linear technique.

Average Precision (AP) is one of indicators for evaluating the performance of an object detection system. The detection rate was the best in the level 1 data, when the degree of overlap was the lowest. And the AP when using the Gaussian technique was about 35.49%, and the AP when using the linear technique was about 33.97%. Overall, it was confirmed that the Gaussian technique shows higher detection performance.

III. CONCLUSION AND FUTURE RESEARCH

We analyze the performance of human detection using soft-NMS in a limited indoor space with high density of people. An experiment was conducted by applying soft-NMS, which is known to have better performance than NMS. However, it was found that the detection rate significantly decreased as the level of human density increased.

Accordingly, referring to various NMS techniques[4], we have a plan to propose an algorithm that increases human detection rate when people are concentrated in a limited space.

ACKNOWLEDGEMENT

This research was supported by stage 4 BK21 project in Sookmyung Women's Univ of the National Research Foundation of Korea Grant. This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2021R1F1A1047113)

REFERENCES

[1] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. -C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4510-4520, Salt Lake City, USA, 18-23 Jun. 2018.

[2] N. Bodla, B. Singh, R. Chellappa and L. S. Davis, "Soft-NMS—Improving Object Detection with One Line of Code," IEEE International Conference on Computer Vision (ICCV), pp. 5562-5570, Venice, Italy, 22-29 Oct. 2017.

[3] CrowdHuman - A Benchmark for Detecting Human in a Crowd, https://www.crowdhuman.org

[4] P. Zhou, C. Zhou, P. Peng, J. Du, X. Sun, X. Guo, and F. Huang, "Noh-NMS: Improving Pedestrian Detection by Nearby Objects Hallucination," ACM International Conference on Multimedia, pp. 1967–1975, Seattle, USA, 12 Oct, 2020.