International Journal of Advanced Culture Technology Vol.11 No.1 320-324 (2023)
DOI https://doi.org/10.17703/IJACT.2023.11.1.320

IJACT 23-3-37

Development of YOLOv5s and DeepSORT Mixed Neural Network to Improve Fire Detection Performance

[1]Jong-Hyun Lee, [2]Sang-Hyun Lee

[1]CEO, ESHEL Tree Co., Gwangju, Korea
[2]Associate Professor, Department of Computer Engineering, Honam University, Korea
[1]saintvk@naver.com, [2]leesang64@honam.ac.kr

Abstract

As urbanization accelerates and facilities that use energy increase, human life and property damage due to fire is increasing. Therefore, a fire monitoring system capable of quickly detecting a fire is required to reduce economic loss and human damage caused by a fire. In this study, we aim to develop an improved artificial intelligence model that can increase the accuracy of low fire alarms by mixing DeepSORT, which has strengths in object tracking, with the YOLOv5s model. In order to develop a fire detection model that is faster and more accurate than the existing artificial intelligence model, DeepSORT, a technology that complements and extends SORT as one of the most widely used frameworks for object tracking and YOLOv5s model, was selected and a mixed model was used and compared with the YOLOv5s model. As the final research result of this paper, the accuracy of YOLOv5s model was 96.3% and the number of frames per second was 30, and the YOLOv5s_DeepSORT mixed model was 0.9% higher in accuracy than YOLOv5s with an accuracy of 97.2% and number of frames per second: 30.

Keywords: *YOLOv5s, DeepSORT, CNN, Fire Detection, Fire Tracking*

1. INTRODUCTION

Currently, many environmental problems and property losses are occurring due to fires. In 2010, with the declaration of a policy called "War Against Fire", fires are on the decline [1]. However, with the influx of the population into the city, urbanization is accelerated due to the dense phenomenon, and as facilities using energy such as electricity and oil increase, human life and property damage due to fire are increasing. In 2020, the number of fires decreased by 11.9% compared to 2011, but the amount of property damage increased by 134%, human damage increased by 22.6%, and deaths increased by 38.8%.

Therefore, a fire monitoring system capable of quickly detecting a fire is required in order to reduce economic loss and human damage caused by a fire [2]. CCTV is convenient to monitor and it can respond quickly in the case of a fire by using intelligent video analysis technology [3].

In this study, in order to develop a fire detection model that is faster and more accurate than existing artificial intelligence models, we use a mixture of the YOLOv5s model and DeepSORT, which complements and extends SORT, one of the most widely used frameworks for object tracking [4-6]. In particular, the YOLOv5s

Manuscript received: January 28, 2023 / revised: March 3, 2023 / accepted: March 14, 2023
Corresponding Author: leesnag64@honam.ac.kr
Tel: +82-62-940-5285, Fax: +82-62-940-5285
Associate Professor, Department of Computer Engineering, Honam University, Korea

Copyright©2023 by The International Promotion Agency of Culture Technology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0)

model is capable of high detection accuracy, lightweight characteristics, and fast detection speed.

DeepSORT has the advantage of improving accuracy, but it also has the disadvantage of reducing the number of frames per second by requiring additional computing calculations. Therefore, by combining the advantages of the two models and applying them to fire detection, the false detection rate of fire is reduced and object tracking is accurately and quickly detected [7-8].

2. MIXED NEURAL NETWORK MODEL DESIGN OF YOLOV5S AND DEEPSORT

Figure 1 shows the learning process of the YOLOv5s and DeepSORT Mixed model proposed in this study.

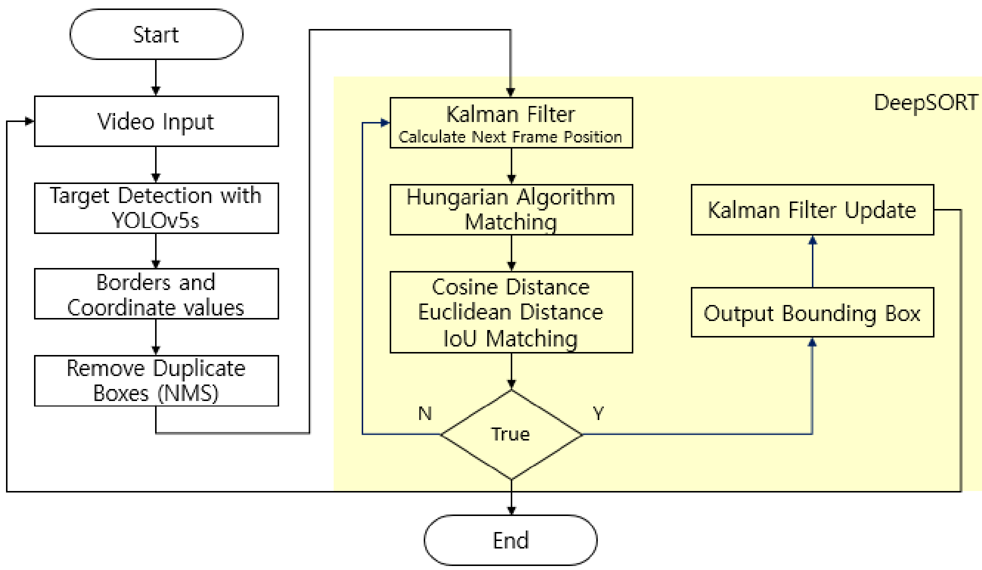

Figure 1. Structure of YOLOv5s and DeepSORT Mixed Model

The fire detection implementation sequence proceeds in four stages. ① The video is processed into video frames using OpenCV, checkboxes are extracted with the YOLOv5s model, overlapping boxes are removed with the NMS algorithm, and the final inspection result is obtained. ② The next frame position and state are predicted through the Kalman filter algorithm, and compared with the check box provided by the YOLOv5s model. The prediction result is determined as an expected frame with high accuracy. ③ It optimizes several targets between the preceding and following frames through a Hungarian algorithm. After obtaining the target trajectory of the video, the Cosine Distance is calculated using the Re-identification(ReID) algorithm, and the ID is reduced by measuring the distance of the target features. ④ After updating the output result and the Kalman filter, restart the target can.

3. IMPLEMENTATION OF MIXED NEURAL NETWORK MODEL OF YOLOV5S AND DEEPSORT

3.1 Dataset

Since there are relatively few data sets available on the Internet and the type of scene is single, 2,059 fire images were collected for various data sets. Since the amount of data is relatively small, the training data set is 70%, the validation data set is 15%, and the test data set is divided into 15%.

Table 1. Dataset configuration

Dataset	Image Number	%
Train Dataset	1,441	70
Validation Dataset	309	15
Test Dataset	309	15

The Computer Vision Annotation Tool (CVAT) labeling tool to be used in this paper is developed by Intel, used to label data in computer vision algorithms, and is a free, open, web-based image and video annotation tool. I used it because it was sourced.

At this time, uncertain labeling was minimized to show high accuracy, and as shown in Figure 2, the bounding box displaying the flame area, which is the labeling condition of this study, proceeds to accurately display the shape of the flame.

Figure 2. Flame zone labeling

3.2 Learning of the Model

The size of the image input for the YOLOv5 model test was 640*640, the size was 8, the epochs were 300, SGD optimization was used, and the change in Train Loss during the test was recorded in WandB software.

Table 2. Train Loss of the Model

Step	0	100	200	300
Train Loss (%)	0.08533	0.0211	0.01556	0.01389

As shown in Table 2, when comparing 0 Steps and 100 Step, the loss rate decreased from 0.08533% to 0.0211%, and from 0.0211% to 0.01556% for 100 to 200 Steps. And 200 ~ 300 Step decreased from 0.01556% to 0.01389%.

Figure 3 shows the loss rate graph according to learning of the YOLOv5s model.

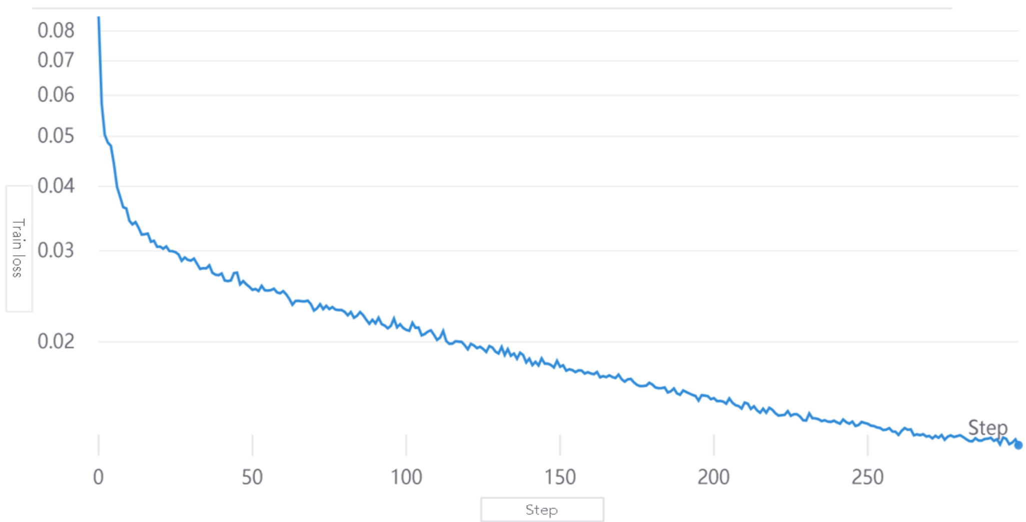

Figure 3. Flame zone labeling

3.3 Result of model test

Table 3 shows the flame detection test using the YOLOv5s model and the YOLOv5s_DeepSORT mixed model proposed in this paper. Accuracy and frames per second (FPS) were measured.

The accuracy of the YOLOv5s model was 96.3% and the number of frames per second was 30. However, the YOLOv5s_DeepSORT hybrid model proposed in this paper has an accuracy of 97.2% and the number of frames per second is 30, and the accuracy is 0.9% higher than that of one model, YOLOv5s.

Therefore, as a result of detecting a fire with the YOLOv5s_DeepSORT mixed model proposed in this paper, the accuracy is 0.9% higher than that of the existing model YOLOv5s and the error can be lowered, enabling tracking and detection of small to large flames.

Table 3. Test results of the Model

Model	Accuracy	FPS
YOLOv5s	96.3%	30
YOLOv5s_DeepSORT	97.2%	30

4. CONCLUSION

In order to reduce economic losses and human casualties caused by current fires, a system capable of detecting and tracking fires that can quickly detect fires in buildings, forests, island areas, and other environments is required. Therefore, in this paper, YOLOv5s model with high detection accuracy, lightweight characteristics, fast detection speed at the same time, and DeepSORT, which is most commonly used for object tracking, was implemented using a mixed model.

Accuracy and FPS were measured to compare the flame detection test results using the YOLOv5s model and the YOLOv5s and DeepSORT mixed model proposed in this study. As a result, the accuracy of the YOLOv5s model was 96.3% and the FPS was 30. The accuracy of the YOLOv5s and DeepSORT mixed model proposed in this paper was 97.2% and the FPS was 30. The accuracy of the proposed YOLOv5s and DeepSORT mixed model was 0.9% higher than that of YOLOv5s.

As a result of detecting fire with the YOLOv5s and DeepSORT mixed model proposed in this study, the accuracy is 0.9% higher than that of the existing model YOLOv5s, and the error can be reduced, and it can track and detect small to large flames.

REFERENCES

[1] National Fire Agency, Fire Statistical Yearbook 2020 (11-1661000-000011-10), pp. 21-22, 2020.

[2] S.H. Lee, "A Study on Fruit Quality Identification Using YOLO V2 Algorithm," *International Journal of Advanced Culture Technology*, Vol. 9, No. 1, pp. 190-195, 2021. DOI: https://doi.org/10.17703/IJACT.2021.9.1.190

[3] J. S. Shin, “Integrated Safety System based on IoT,” *Journal of The Korea Society of Computer and Information*, Vol. 24, No. 10, pp. 159-165, 2019. DOI: https://doi.org/10.9708/jksci.2019.24.10.159

[4] Nicolai Wojke, Alex Bewley, and Dietrich Paulus, "Simple online and real-time tracking with a deep association metric," *IEEE International Conference on Image Processing (ICIP),* Beijing, China, pp. 3645-3649, 2017.

[5] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple online and realtime tracking," IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, pp. 3464-3468, 2016.

[6] S. J. Yang, I. H. Jung, D. H. Kang and H. B. Baek, “Real-Time Multi-Object Tracking using Mixture of SORT and DeepSORT,” *Journal of KIIT*, Vol. 19, No. 10, pp. 1-9, 2021. DOI: http://dx.doi.org/10.14801/jkiit.2021.19.10

[7] H.T. Kim, S.H. Lee, “Detecting Jaywalking Using the YOLOv5 Model,” *International Journal of Advanced Culture Technology (IJACT),* Vol. 10, No. 2, pp. 300-306, 2022.

[8] G.S. Lee, S.H. Lee, “Prediction of changes in fine dust concentration using LSTM model,” *The International Journal of Advanced Smart Convergence*, Vol. 11, No. 2, 30-37, 2022. DOI: https://doi.org/10.7236/IJASC.2022.11.2.30