An Improved YOLO model for Large Object Detection for Black Ice Detection

Mst Ayesha Khatun[1], Mohtasin Golam[1], Adinda Riztia Putri[1], Philip Tobianto Daely[1], Jeong-Ho Lee[2], Young-Rea Cho[2], Jung-Hyeon Kim[1], Dong-Seong Kim[1], and Jae Min Lee[1]
[1]Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea
[2]Seokyung Industry, Jeju, South Korea
(20226138, trizkim, ljmpaul)@kumoh.ac.kr

***Abstract*—This paper proposes a model based on YOLOv5 to detect black ice on the road surfaces.As black ice is responsible for the frequent traffic accidents in winter,it is very urgent to identify and handle them correctly.To solve the problem of recognition of black ice detection, the real-time object detection system YOLOv5 is used and addressed to improve the YOLOv5 architecture. To train the neural network it used the data set consisting of approximately 1307 images. Comparing with others, it has achieved over 95% in accuracy.**

***Index Terms*—Black ice,YOLOv5, deep learning.**

I. Introduction

A particular kind of thin ice layer that develops on pavements and roadways is known as black ice. It is easily produced in damp places with low temperatures, particularly in the shadow [1]. Due to its extreme transparency, which allows for the ability to see through it the dark hue of the road and the pavement it develops on, black ice gets its name. It is translucent because it is so thin and mixes in with road pavements, making it practically hard to notice. According to research by the Korea Expressway Corporation, on black ice, a roadster's surface slides 14 times more than it would on dry pavement and six times more than it would on snow-covered roads [2]. Even in the period of development for autonomous vehicles (AVs), black ice may be conceived of as a potential accident component, hence it is anticipated that technology that can identify it in advance and avoid accidents would be needed [3].

Thousands of incidents, fatalities, and injuries have been caused by black ice worldwide. Numerous accidents occur each year, often with severe repercussions. In a sequence of black ice incidents on the Korean highway between Sangju and Yeongcheon in the winter of 2019, 44 automobiles were damaged and seven people died. According to data from the Ministry of the Interior and Safety of Korea, 5,042 vehicle accidents brought on by ice roads and frost resulted in 9,420 fatalities between 2015 and 2019 [2]. These issues have made it imperative to create technologies for detecting and preventing traffic accidents caused by black ice as well as protecting the security of winter passengers. Since the advancement of deep learning and computer vision, object detection has gained significant prominence in a variety of fields. The goal of this research is to find a solution by developing a deep learning-based black ice detection algorithm concentrating YOLOv5.

One-stage and two-stage object detectors are the two main categories into which object detection algorithms are currently categorized. The R-CNN (region convolutional neural network) [4]family, which includes the image segmentation tools Fast R-CNN, Faster R-CNN, R-FCN (region fully convolutional networks), and Mask R-CNN, are the most typical two-stage object detectors. The most prominent one-stage object detectors are the single-shot multi-box detectors (SSD), and the YOLO (you only look once) series, including YOLOX [5]. YOLOv5, a one-stage object detector that relies on anchors, is the foundation of the enhanced network structure shown in this research. The model for YOLOv5 is compressed and tailored into four sizes: s, m, l, and x, which is one of its advantages. It is possible to use it for various applications, and the detection accuracy is much increased when compared to other networks.

II. Proposed Methodology

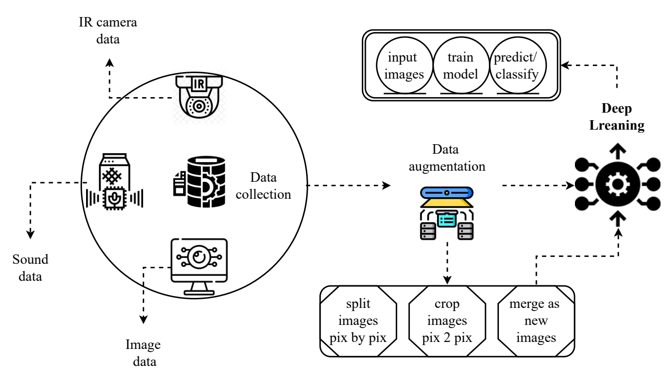

Fig. 1. Architecture of the proposed model

The architecture of the proposed model is illustrated in Fig. 1. Two distinct datasets, including acoustic and picture data, will be combined with the proposed model. However, in this study, the emphasis was on image-based black ice detection, and to enable this, an improved YOLO architecture was employed. This section provides a detailed explanation of the improved YOLO architecture. Even though it would take an extensive explanation to go through the entire workflow from data collection to preprocessing, it will be covered in a subsequent article, leaving this one to focus merely on model improvement. The procedure for gathering data, including the acquisition of data from acoustic sensors and infrared cameras (IR), as well as how the two separate datasets may be utilized to detect black ice, will also be covered in greater detail in a subsequent article.

The model is being improved using YOLO v5 starting from the feature extraction for large object detection model with bottleneckCSP functioning as the baseline for feature extraction. A sequence of 181 convolution replacement pooling layers and 3 * 3 layers are built to create the bottleneckCSP network. The authors of this research [6] introduced the SPP-Net (spatial pyramid pooling network) structure, which fixes the size of the input images and can be tailored to any network configuration. In YOLO v5, the SPP structure is made up of three branches: 5 * 5, 9 * 9, and 13 * 13. In order to transfer the feature data to the network's following layer, three branches will be merged into one branch.

The intermediate layer of the architecture for multi-scale feature fusion contains elements including feature pyramid networks (FPN) and path aggregation networks (PAN), which include four connection layers, four convolution layers, and five CSP layers respectively. In this network model, there are five feature fusions and a path aggregation network (PANet) [7] structure that is an optimization of a feature pyramid network (FPN) structure. P3, P4, P5, P6, and P7 are the labels assigned to the initial data that was taken from the network, and the modules P6 and P7 are added to the network. A new feature fusion layer was added to create P6 and P7, which make up the structure. This structure is known as the multi-scale feature cross-layer fusion network (M-FCFN), and it is a multi-scale, multi-feature, multi-fusion, and multi-box detector. The M-FCFN is also known as the M-FCFN detector.

The feature maps are P3, P4, P5, P6, and P7 following twice-upsampling and five feature fusion. The associated scale sizes are (80 * 80), (40 * 40), (20 * 20), (10 * 10), and (5 * 5). For output in detection, the feature maps from these five different scales are employed. Each of those five scales is designed with a distinct anchor size in order to fulfill the goal of anticipating objects of various sizes. These five scales correlate to the length of grids. Now, in order to create a feature map with a dimension of (10 * 10), the model combines the feature data from P5 with the feature data from the SPP layer. Furthermore, another feature map, P7, measuring (5 * 5), is added to the P6 structure. The experiment was repeated with the anchor sizes modified for P6 and P7 to obtain the results. Given the availability of the auto-anchor mechanism, training takes priority over initial anchor size, and the anchor sizes of P3, P4, and P5 are identical to those of YOLOv5.

A. Performance Evaluation

The simulation outcomes of black ice detection are discussed in this section. Mean average precision (mAP), accuracy, sensitivity (recall), and F1-score are the performance measures used to assess the model's performance. To evaluate the model, 1307 photos were obtained from a public dataset and divided into 70%, 20%, and 10% for training, testing, and validation, respectively. Additionally, the dataset was divided into three classes (dry, snowy, and wet) to train the object detector and then supervise its learning via bounding box annotations. Furthermore, the frame per second (FPS) was also considered as an valuable impact features on the performance results and the same FPS is considered for the proposed model along with existing YOLOv5 model. The performance results are illustrated in the table below.

TABLE I
PERFORMANCE RESULT OF THE PROPOSED MODEL

Models	Metrics			
	mAP	Recall (%)	F1-score (%)	FPS
Proposed Model	95.4	91.7	93.1	168
YOLO v5	94.6	90.2	91.3	168

III. Conclusion and Future Work

This article presents the proposed model for black ice detection, using YOLOv5 model as its foundation. The model architecture is improved with feature extraction along with bottleneck CSP at the initial stage. From the comparison with YOLOv5, the proposed model produce better performance in regards to accuracy, precision, and recall. In the future work, other of machine/deep learning models will be investigated.

IV. Acknowledgment

This work was supported by Priority Research Centers Program through the NRF (2018R1A6A1A03024003), the Grand Information Technology Research Center support Program supervised by the IITP (IITP-2022-2020-0-01612), and Project for Industry-University-Research Institute Platform Cooperation R&D funded Korea Ministry of SMEs and Startups in 2022.(S3310783).

References

[1] S. Kim, H. J. Oh, S. J. Han, H. S. Ko, Y. Shin, and D. H. Shin, "Development of black-ice removal system with latent heat thermal energy storage and solar thermal collectors," *Energy*, vol. 244, p. 122721, 2022.

[2] Y. Xu, D. Yao, X. Ren, and Y. Dai, "Intelligent black ice detection and alert system using thermal imaging camera and drone," in *2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys)*. IEEE, 2021, pp. 2328–2331.

[3] H. Lee, M. Kang, J. Song, and K. Hwang, "The detection of black ice accidents for preventative automated vehicles using convolutional neural networks," *Electronics*, vol. 9, no. 12, p. 2178, 2020.

[4] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-based convolutional networks for accurate object detection and segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 1, pp. 142–158, 2015.

[5] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, "Yolox: Exceeding yolo series in 2021," *arXiv preprint arXiv:2107.08430*, 2021.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 9, pp. 1904–1916, 2015.

[7] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8759–8768.