Crosswalk Safety System for Pedestrian Protection*

Donggue Kim[0009−0007−5344−5360], Kwanho Kim[0009−0004−3094−0746], Kanghyun Jo[0000−0001−8317−6092]

School of Electrical Engineering
University of Ulsan, Ulsan, South Korea
kdk6859@gmail.com, aarony12@ulsan.ac.kr, acejo@ulsan.ac.kr

Abstract. Traffic accidents continue to be a prevalent issue annually, with some incidents lacking effective countermeasures. One such overlooked circumstance involves accidents caused by illegal right-turn vehicles. In South Korea, a comparison between 2022 and 2023 reveals a nearly twofold increase in accidents attributed to illegal right-turn vehicles, rising from 0.34 to 0.64. Despite this surge, there is a notable absence of surveillance systems or preventive measures to address this concern. To address this gap, this paper aims to enhance a previous study by incorporating a preventive component. Using 118,287 images from the COCO dataset for pedestrian and vehicle perception and an additional 36,808 images from the AI Hub's pedestrian crossing video dataset for crosswalk perception, this study can train model for detection. after training, this paper make a system for protecting pedestrian. For algorithm testing, the previous study utilized JTBC news videos for the research, while this paper tests algorithm in a Unity development environment. This approach demonstrates the potential efficacy of the proposed system in addressing the highlighted issues and contributing to overall pedestrian safety.

Keywords: Traffic accidents · Illegal Right-turn · Pedestrian.

1 INTRODUCTION

Computer vision is one of the prominently acclaimed fields leveraging artificial intelligence, leading to a surge in research and the development of associated systems. Autonomous driving, Unmanned Aerial Vehicles (UAVs) and Traffic Analysis are examples for using computer vision. It contributes to improving people's quality of life by being applied to many fields.
This paper presents a system developed using computer vision technology aimed at ensuring the safety of pedestrians crossing crosswalks. This research builds

* This results were supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(MOE)(2021RIS-003)

upon previous research, where the focus was on a surveillance system for deterring illegal right-turn.
However, it lacked the capability to prevent accidents to apply for real-world.
Interviews with driver involved in accident revealed instance where driver failed to notice pedestrian, particularly in incident involving large vehicle [1].
However, there is currently no preventative solution in the Korean traffic system. To address this, this paper introduces a system that triggers alerts when a vehicle approaches a certain zone while a pedestrian is on the crosswalk. Also, this system is made with previous system. The implementation of this system utilizes computer vision module called You Only Look Once (YOLO). It is a single-stage object detection model and represents widespread popularity due to its speed and accuracy.
Training data for crosswalk, pedestrian, and vehicle recognition is sourced from the Ai Hub dataset and the COCO dataset. Using this training data, this paper can get information about location of crosswalk, pedestrian and vehicle.
The ultimate goal of this paper is to develop a system that detects illegal right-turn vehicle and alerts drivers when pedestrians are on the crosswalk, with the overarching aim of reducing pedestrian accidents.

2 RELATED WORK

2.1 Detection for Illegal Right Turn Vehicles in the Korea Traffic System

This paper is a continuation of previous research [2], focusing on the development of a system utilizing the YOLOv5 computer vision technology module to detect illegal right-turn vehicles by analyzing the coordinates of observed detection boxes. In specific terms, as depicted in Fig. 1, the system identifies pedestrians on the crosswalk by analyzing coordinate values. Subsequently, it utilizes the measured coordinates of vehicle to distinguish between illegal right-turn vehicle and normal vehicle. This enables the system to recognize instances of illegal right-turn.
It is important to note that this paper is designed as a system applicable to the traffic laws of Korea [3]. because of this it is applicable to Korea's traffic environment. The datasets used for this research include the COCO dataset and the Ai Hub dataset. The main goal of this paper is to address the issue of directly arresting illegal right-turn vehicles by either the police or pedestrians, using the aforementioned system.

2.2 Video Surveillance for Road Traffic Monitoring

The paper [4] introduces a system for tracking a single car measured from multiple cameras, contributing to the advancement of camera systems in traffic-related scenarios. This paper employs Siamese network [5] and clustering techniques to present an algorithm for analyzing a single car measured by two cameras. This

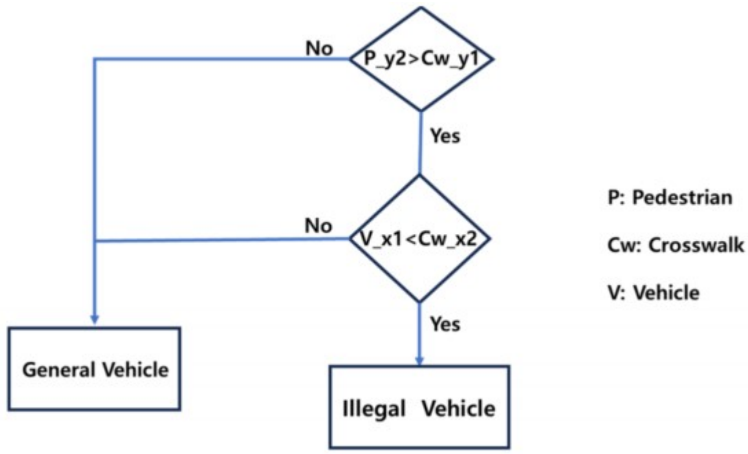

Fig. 1: Architecture of Algortithm ; Detection for Illegal Right Turn Vehicles in the Korea Traffic System

algorithm is called Re-ID(Re-Identification) system. When using two cameras, the algorithm considers the images to be the same if the identity distance is below a certain threshold.
However, The study further extends to utilizing six cameras but it was not work well and they mentioned about the reason.
This paper focuses on developing an algorithm that distinguishes identical car images measured by multiple cameras in traffic monitoring system and enables the tracking of these cars.

2.3 BoxCars: Improving Fine-Grained Recognition of Vehicles using 3D Bounding Boxes in Traffic Surveillance

The paper [6] presents an algorithm using 3D Bounding Boxes to address constraints from the perspective of traditional Bounding Boxes. The proposed CNN model in this paper is designed to generate 3D Bounding Boxes, and comparative evaluations with existing CNN models show improved performance results. The paper emphasizes that utilizing 3D Bounding Boxes allows for better consideration of the shape and perspective features of vehicles, enhancing performance in fine-grained vehicle recognition problems. It particularly highlights achieving high accuracy in traffic surveillance scenarios.
Additionally, the paper mentions undergoing Data Augmentation processes, such as color transformation and ImageDrop, to enhance the diversity of training data.

2.4 Where are the Blobs: Counting by Localization with Point Supervision

The paper [7] introduces the LC-FCN (Localization Counting Fully Convolutional Network), proposing a system for detecting the number of objects. LC-FCN is a fully-convolutional neural network architecture developed in this paper. Additionally, a novel loss function is introduced for counting and localization purposes. The paper suggests a detection-based method that accurately counts objects without the need for size or shape estimation.
Through experimental results, it demonstrates that this approach outperforms regression-based methods. Furthermore, the paper experimentally proves the superior performance of the proposed method compared to other latest model, utilizing datasets such as PASCAL VOC 2007, Trancos, and Penguins.
The paper also mentions plans for future research, exploring various FCN architectures and segmentation methods to achieve better performance.

3 Proposed ALGORITHM

3.1 YOLOv5

This paper utilized the YOLOv5 computer vision module to implement the proposed algorithm. YOLOv5 is a high-performance deep learning model designed for real-time object detection, representing the fifth version in the You Only Look Once series. This model delivers fast inference speeds and high accuracy, making it applicable in various computer vision applications.
In this paper, the smaller parameter model YOLOv5s was employed, and during training, input images of size 640 by 640 are utilized. This allowed for the recognition of crosswalks, pedestrians, and vehicles.

3.2 Detecting vehicles, pedestrian, crosswalks

This paper trains YOLOv5s using the COCO dataset for vehicle and pedestrian recognition. The weight obtained from this training is then used to carry out the process of vehicle and pedestrian recognition.
Similarly, for crosswalk recognition, the paper utilizes crosswalk samples provided by the AI Hub dataset for training the model and uses the generated weight file for the crosswalk recognition process.
Subsequently, the proposed algorithm is implemented by analyzing the coordinate values obtained from the generated bounding boxes. Coordinate-related information for each image can be verified through Fig. 2. The left coordinates for each image are represented as (x1, y1), and the right coordinates are represented as (x2, y2).

3.3 Algorithm for Pedestrian Protection

Utilizing the coordinates analyzed through YOLOv5, an algorithm was formulated as described earlier. The overall algorithmic process is depicted in Fig. 3.

Fig. 2: Coordinates information ; Left coordinate means (x1,y1) and Right coordinate means (x2,y2)

This algorithm is based on South Korean traffic laws [3] and is constructed upon previous research [2]. According to this legislation [3], the determination of illegal right-turn is based on whether a pedestrian is present on the crosswalk or not.
The first condition, as evident in Fig. 3, 'P_y2 > Cw_y1', represents the condition indicating the presence of a pedestrian on the crosswalk. The second condition, 'V_x1 < C_x2', signifies that the vehicle has crossed the crosswalk. The third condition, P_M, is a novel proposal in this paper.
P_M refers to instances where two conditions are simultaneously satisfied.
The first condition is when the x-coordinate of V_x1 is smaller than the sum of the distances of the cw coordinates and the center of the x-axis coordinates of the CW, and the second condition is when 'V_x1 > Cw_x2,' indicating that the vehicle has not yet crossed the crosswalk.
If the P_M condition is satisfied, it is classified as 'Danger Vehicle,' addressing the issue introduced in the introduction, where driver in vehicle failed to perceive pedestrians on the crosswalk.
As illustrated in Fig. 3, the remaining cases are classified as 'General Vehicle.'
In summary, by constructing conditions based on the coordinate information obtained through YOLOv5 and determining vehicle with these conditions, it can be classified as 'General Vehicle,' 'Illegal Vehicle,' or 'Danger Vehicle.'

4 EXPERIMENT

4.1 System Environment

The system environment employed RTX 3090 GPUs during training, featuring a system with 96GB of graphics memory and 188GB of RAM.
During training, this paper utilizes images with a size of 640 by 640, a batch size of 16, and train is run for 100 epochs. Additionally, a Cosine Learning Rate Scheduler [8] is applied with a learning rate set to 0.01. Stochastic Gradient Descent (SGD) [9] served as the optimization function.
In this setup, training is conducted for the recognition of vehicles, pedestrians, and crosswalks.

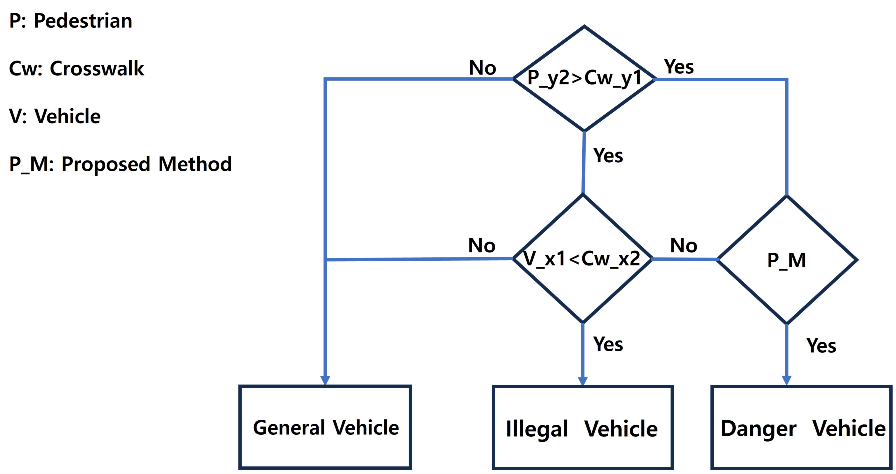

Fig. 3: Flowchart for Proposed Algorithm

4.2 COCO Dataset

The COCO dataset [10] consists of 118,287 samples for training and 5,000 samples for validation. There are approximately 80 class categories, including prominent ones such as pedestrian, bicycle, vehicle, and motorcycle.
In this paper, only the pedestrian and vehicle classes are utilized.

4.3 AI hub Dataset

The AI Hub dataset is a platform that hosts datasets from various fields. Among these, this paper utilizes a dataset [11] measured from the perspective of pedestrians, comprising a total of 3,608 samples for training. The class information includes red light, green light, and crosswalk.
In this paper, only the class information related to crosswalk is used.

4.4 Unity for making test dataset

Acquiring a realistic dataset for experimenting with the algorithm proposed in this paper posed challenges. Consequently, Unity [12] is employed to create a virtual environment, allowing for testing the algorithm presented in this paper. Unity serves as a game engine providing a development environment for 3D and 2D video games.
It is also a comprehensive authoring tool for interactive content creation, including 3D animation, architectural visualization, and Virtual Reality (VR).
Using Unity, a virtual environment is constructed to simulate traffic scenarios at the same point in time, enabling the testing of the proposed algorithm. Video captured through Unity is composed of frame width of 1180, frame height of 666

and 24 frames per second.
This paper converts these video into images and conducts the algorithm testing process on a total of 210 images.

4.5 Result of Algorithm

As mentioned earlier, as seen in Fig.3, this paper shows three cases. The first case is for General Vehicle, the second case is for Illegal Vehicle, and the last, the third case is for Danger Vehicle.
Instances like Fig.4 fall into the category of General Vehicle. As observed in the illustration, the recognition of crosswalk, pedestrian and vehicle can be confirmed.
Analyzing it from a coordinate perspective through the proposed algorithm, it is evident that the pedestrian has not yet reached the crosswalk, hence classified as a General Vehicle, as indicated by the 'No' in the first condition.
In the case of Fig. 5, Pedestrian is classified as on the crosswalk through the proposed algortihm. However, it is not a case where the vehicle's position is P_M, and the second condition is not satisfied, resulting in classification as a General Vehicle.

Fig. 4: Case 1 : Pedestrian is not on the crosswalk

Fig. 6 represents a scenario where the proposed algorithm identifies a pedestrian on the crosswalk.
In this paper, this process is assumed to be the movement of a vehicle proceeding straight without recognizing a pedestrian. Conducting experiments based on this assumption, Vehicle is classified as danger vehicle due to the P_M condition.
If this system is implemented, it could potentially address incidents caused by large vehicles proceeding straight without recognizing pedestrians [1].
Fig. 7 shows results related to the algorithm presented in a previous study [2].

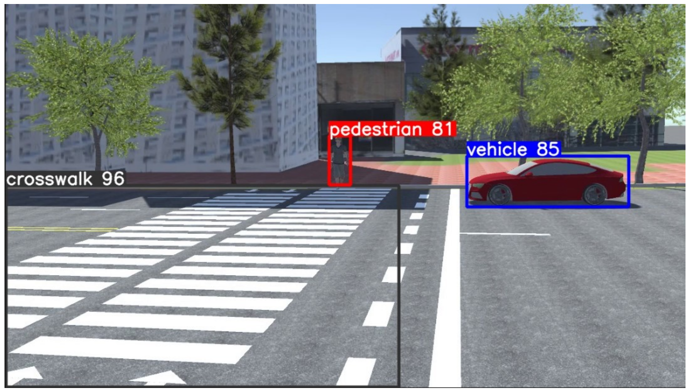

Fig. 5: Case 2 : Vehicle is classified as ‘General Vehicle’

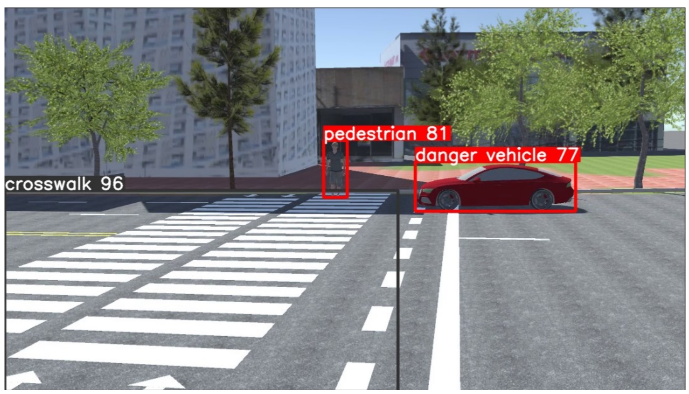

Fig. 6: Case 3 : Vehicle is classified as ‘Danger Vehicle’

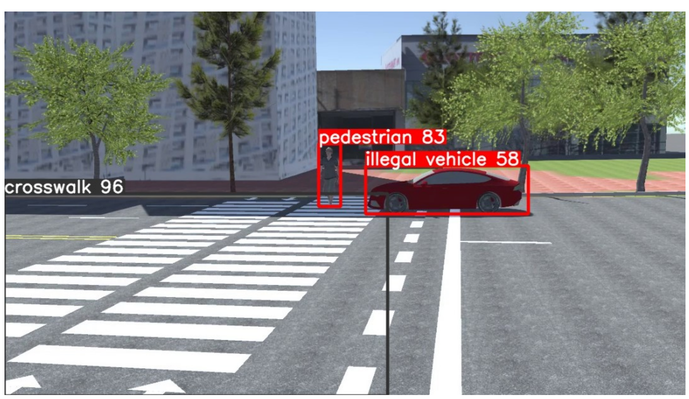

Fig. 7: Case 4 : Vehicle is classified as ‘Illegal Vehicle’

After identifying a pedestrian on the crosswalk through the proposed algorithm, the analysis of vehicle coordinates and crosswalk coordinates is conducted based on the second condition, resulting in classification as an 'illegal vehicle.'
The algorithm's outcomes demonstrate the simultaneous operation of an illegal right-turn vehicle monitoring system [2] and a system ensuring pedestrian safety when crossing the crosswalk. Based on these results, the goal of this paper is to reduce accidents involving pedestrians and vehicles on crosswalks.

5 CONCLUSION

Fig. 8: Detection results applied in a real-world; before the modification of the detection box

The proposed method in this paper utilizes YOLOv5 to analyze the coordinates of generated bounding boxes. Based on this information, it classifies Vehicle into 'General Vehicle,' 'Illegal Vehicle,' and 'Danger Vehicle' by analyzing the spatial relationships among pedestrian, vehicle, and crosswalk.
Through this approach, a system for monitoring illegal right-turn and ensuring pedestrian safety while crossing crosswalk is established, aiming to reduce accidents between vehicles and pedestrians on crosswalks. The experiments in this paper are conducted in a virtual environment using Unity, which is an ideal space. because of this, result of this paper is going well.
However, when applying the system presented in this paper to real-world scenarios, several issues are identified. To conduct experiments in a real-world environment, news data provided by the Korean comprehensive channel 'JTBC' is utilized [13]. Fig.8, Fig.9 and Fig.10 represent data obtained by real environment.
Comparing the three sets of data, it is evident that the size of the bounding box

Fig. 9: Detection results applied in a real-world; Modification of the detection box

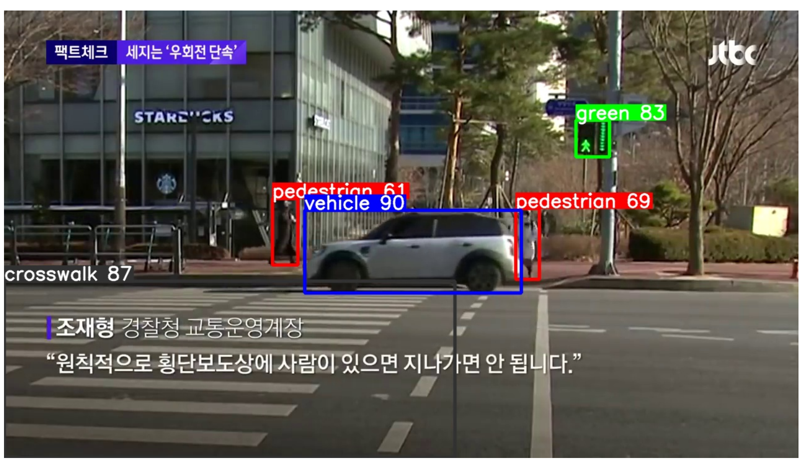

Fig. 10: Detection results applied in a real-world; Modification of the detection box, When vehicle on the crosswalk is gone

for the crosswalk is not fixed but changes. The algorithm proposed in this paper relies on coordinate analysis. However, if the coordinate values keep changing, as seen in the three datasets, the reliability of the algorithm may decrease.
Therefore, addressing the issue of dynamically changing bounding boxes according to the environment is a goal for future research.
Furthermore, the objective is to create a model that can operate the system in a real environment. Ultimately, the goal of this paper is to contribute to reducing accidents between pedestrians and vehicles on crosswalks through the proposed system.

Acknowledgment

This results were supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(MOE)(2021RIS-003)

References

1. The 'right turn' bus that ran away after hitting... "daughter, i'm scared of the crosswalk.", youtube, apr,28,2023. accessed on: Jan, 02, 2024, [https://www.youtube.com/watch?v=n-xrxrahtoy].
2. Donggue Kim and Kanghyun Jo. Detection for illegal right turn vehicles in the korea traffic system. *2023 International Workshop on Intelligent Systems (IWIS)*, pages 1–4, 2023.
3. Korea road traffic act,2022.(article 27,section 1).
4. Pol Albacar, Oscar Lorente, Eduard Mainou, and Ian Riera. Video surveillance for road traffic monitoring. *arXiv preprint arXiv:2105.04908*, 2021.
5. Xinlei Chen and Kaiming He. Exploring simple siamese representation learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 15750–15758, 2021.
6. Jakub Sochor, Jakub Špaňhel, and Adam Herout. Boxcars: Improving fine-grained recognition of vehicles using 3-d bounding boxes in traffic surveillance. *IEEE transactions on intelligent transportation systems*, 20(1):97–108, 2018.
7. Issam H Laradji, Negar Rostamzadeh, Pedro O Pinheiro, David Vazquez, and Mark Schmidt. Where are the blobs: Counting by localization with point supervision. In *Proceedings of the european conference on computer vision (ECCV)*, pages 547–562, 2018.
8. Ilya Loshchilov and Frank Hutter. Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983*, 2016.
9. Tsung-yi lin, michael maire, serge j. ruder, s. (2016). an overview of gradient descent optimization algorithms. arxiv, abs/1609.04747.
10. Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, pages 740–755. Springer, 2014.
11. Ai hub object detection dataset, footage of walking on the sidewalk ai hub, 2021. [online], https://open.selectstar.ai.

12. Sabir Hossain and Deok-Jin Lee. Autonomous-driving vehicle learning environments using unity real-time engine and end-to-end cnn approach. *로봇학회 논문지*, 14(2):122–130, 2019.
13. What if it's green light when you turn right? how will it change in 2022, youtube, jan,11,2022. accessed on: June, 23, 2023, https://www.youtube.com/watch?v=exhs4h4ycko.