

# Improved CNN-based Path Planning So an Autonomous UAV Can Climb Stairs By using a LiDAR Sensor

Yeon Ji Choi<sup>1</sup>, Tariq Rahim<sup>2</sup>, and Soo Young Shin<sup>3\*</sup>

Department of IT Convergence Engineering, Kumoh National Institute of Technology / Gumi, Korea  
yzygzy@kumoh.ac.kr, tariqrahim@ieee.org, wdragon@kumoh.ac.kr

\* Corresponding Author: Soo Young Shin

Received April 7, 2021; Revised May 25, 2021; Accepted July 9, 2021; Published October 30, 2021

\* Regular Paper

\* Extended from a Conference: Preliminary results of this paper were presented at the ICEIC Winter 2021. This paper has been accepted by the editorial board through the regular reviewing process that confirms the original contribution.

**Abstract:** Unmanned aerial vehicles (UAVs) have tremendous potential in civil and public areas. These are especially beneficial in applications where human lives are threatened. Autonomous navigation in unknown environments is a challenging issue for UAVs where decision-based navigation is required. In this paper, a deep learning (DL) approach is presented that aids autonomous navigation for UAVs in completely unknown, GPS-denied indoor environments. The UAV is equipped with a monocular camera and a light detection and ranging (LiDAR) sensor to determine each next maneuver and distance calculation, respectively. For deeper feature extraction, a version of You Only Look Once (YOLOv3-tiny) is improved by adding a convolution layer with different filter sizes. The process is observed as an exercise where the DL model classifies the targeted image as stairs or not stairs. We created our dataset considering the indoor scenario for specific implementation. Comprehensive experimental results are compared with YOLOv3-tiny, indicating better performance in terms of accuracy, recall, F1-score, precision, and maneuvering movements.

**Keywords:** UAVs, CNN, Path planning, Stair climbing, LiDAR sensor

## 1. Introduction

UAV use is growing in areas such as scientific research, rescue missions, commerce, and agriculture. Originally, UAVs were developed to be managed by an on-the-ground pilot via remote-control communication [1]. Recently, UAVs have been moving closer to navigating with unusual degrees of autonomy. Most UAVs employ global navigation satellite system technology and inertial sensors to determine their geospatial positioning. It is necessary to overcome factors such as GPS signal error, narrow passageways, and transparent glass for stable-flight UAVs in indoor environments [2]. Studies in image-based stair-recognition for robots [3] and of techniques for ground robots [4] are ongoing; however, there is a lack of such research with UAVs. An abundance of techniques, varying from learning-based to non-learning-based, have been suggested to resolve UAV navigation dilemmas. The most popular non-learning-based method is sensing and avoidance, which prevents accidents by steering vehicles

in a reverse orientation and navigating by path planning [5, 6]. Another type of non-learning-based technique takes advantage of simultaneous localization and mapping (SLAM). The inspiration is that, after creating a map of the surroundings by utilizing SLAM, navigation is accomplished by path planning [7, 8]. The work in [7] combines GraphSLAM [9] with an online path planning module in a proposal-approving UAV to determine obstacle-free trajectories in foliage. A general characteristic of non-learning-based approaches is that they demand precise path planning, which may result in unanticipated failures when environments are extremely dynamic and complicated. To address this matter, machine learning (ML) methods such as imitation learning and reinforcement learning (RL) have been explored [10-12]. For example, a model-based RL approach called TEXPLORE [12] was presented, which is a high-level control system for navigation of a UAV within a grid map having no barriers. And an imitation learning-based controller utilizing a small set of human displays was

presented that obtains reliable performance in forested areas [10].

Therefore, this paper proposes a convolutional neural network (CNN)-based system based on real-time stair recognition that can fly a UAV without colliding with stairs, and that obtains distance information between walls or stairs through 2D light detection and ranging (LiDAR) with a camera mounted on the UAV. In addition, algorithms were designed for systems that recognize stairs, avoid collisions, and maneuver themselves, which is one of the obstacles to an autonomous flight process, and flight experiments were carried out after the actual UAV was implemented.

Deep learning (DL), which is a subcategory of machine learning, acts like the human brain, and is therefore known as artificial intelligence (AI). Many applications of machine learning have been proposed, with different signals representing data such as music signals [13], 2D signals or images [14], and video signals [15]. CNNs are used for various purposes, such as classification, detection, and pattern recognition, especially in health [16], drone applications [17], and autonomous driving systems. Recently, You Only Look Once (YOLO) was introduced for real-time detection of objects, with each version improving the mean average precision (mAP) per frame per second [18].

In this work, we attempted for the first time to use the YOLOv3-tiny model, and improved the model further by adding a convolution layer to extract deep features for the detection of stairs. This DL detection model was used in a classification problem to determine each next maneuver.

The rest of this paper is organized as follows. Section 2 details related work, while Section 3 explains the proposed scheme. Section 4 summarizes the experimental results and the analysis. Section 5 provides concluding statements and suggests the scope of future work.

## 2. Related Work

Previously, a 3D map of the local area was developed for autonomous UAV navigation. In some cases, these methods were used to map exact quadcopters [19, 20]. However, these methods are based on a smart control scheme, thereby restricting their use to laboratory settings [21-23]. The map is learned through other manual route methods, and quadcopters travel the same path [24]. For most outdoor flights (where precision is not as high as indoors), a GPS-based posing projection is used.

Most applications use scale sensors, such as infrared sensors, RGB-D (red, green, blue depth) sensors, or laser range sensors [25]. A single ultrasonic sensor was used in [26] as an automated navigation device with an infrared sensor. The condition evaluation method of the LiDAR and inertial measurement unit (IMU) was advanced to work independently in uncertain conditions that are denied by a GPS [27]. Range sensors have limitations, being heavy and high in power consumption.

The simultaneous localization and mapping (SLAM) technique uses separate optical sensors to create a 3D image [21-23] from every UAV position on the map. A 3D

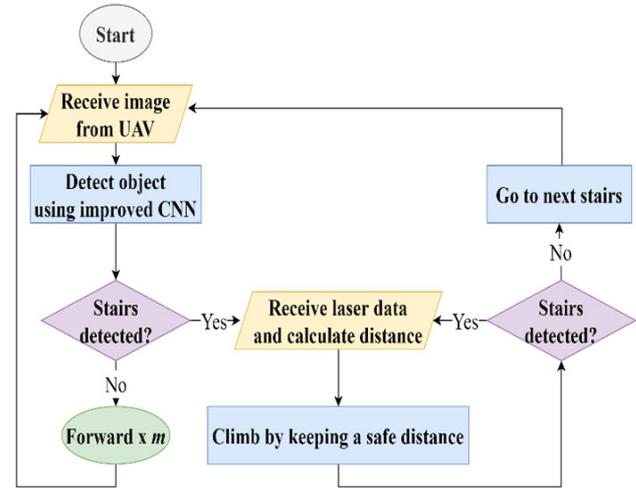


Fig. 1. Flowchart for the proposed implementation.

map of an unknown indoor scenario was used for the SLAM laser range finder [25]. The SLAM technique [29, 31] offers single-camera indoor navigation. SLAM is highly complicated when it comes to regenerating the 3D map region, requiring precise measurements and extensive resources because additional sensors are needed.

SLAM can also set contact delays during real-time navigation. The studies in [31] and [32] addressed these issues. SLAM is primarily a practical system, and its output with indoor materials (such as walls/roofs) is not considered good, because its differential intensity is very weak. The entire corridor comprises partitions, roofs, and floors, and SLAM technologies cannot attain the desired navigational quality.

## 3. The Proposed Scheme

This section discusses the system configuration for UAV recognition of stairs, the deep learning model using YOLOv3-tiny, and the improved YOLOv3-tiny for detecting stairs.

### 3.1 System Configuration

The proposed system was designed based on recognizing stairs with a camera mounted on the UAV for indoor environments and on distances measured via the 2D LiDAR sensor attached to the UAV's side. Fig. 1 shows the flowchart for the entire system. The connections and communications between the parts are both wired and wireless, as shown in Fig. 2. In particular, communications among the ground control station, the UAV, and the onboard PC is via Wi-Fi/LTE. Meanwhile, the wired connection is only used for the sensor.

The system's actual implementation uses a Parrot Bebop 2 drone, which is suitable for narrow passageways and convenient for load sensors. The UAV is equipped with an RPLiDAR S1 laser scanner, which rotates 360° and can measure distances up to 40m with a lightweight, mainboard Jetson TX2 embedded computing device (Auvidia J120 carrier board) as shown in Fig. 3(c). The

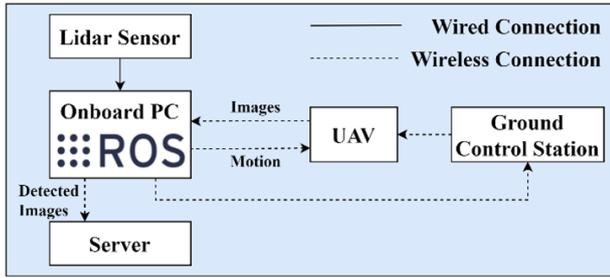


Fig. 2. Network connections and the architecture of the proposed system.

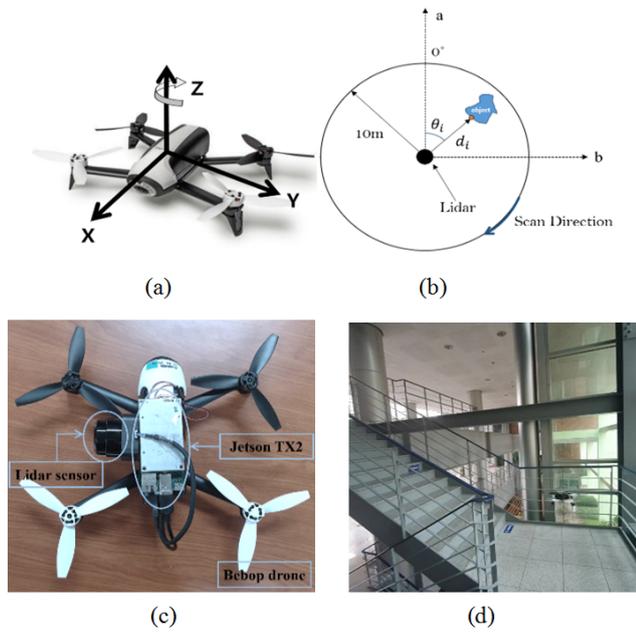


Fig. 3. System configuration: (a) UAV movement axes; (b) illustration of the RPLiDAR S1 scanning process; (c) the 2D-LiDAR sensor and the Jetson-TX2 onboard PC attached to the UAV; (d) the test environment.

Table 1. Experiment Parameters.

Device	Model name	Company
Lidar sensor	RPLiDAR S1	Slamtec
UAV	Bebop drone 2	Parrot
Onboard PC	Jetson TX2	Nvidia
Carrier board	Auvideo J120	Auvideo
GCS	ThinkPad T580	Lenovo
LTE modem	LTE USB Stick	Huawei

Lenovo ThinkPad T580 is used as a ground control system (GCS), and the equipment required for the experiment is listed in Table 1. All algorithms are implemented in Python, and the Robot Operating System (ROS) was used as middleware (software that can run multiple different programs together) in a kinetic version.

The LiDAR sensor uses distances measured along 360 points, as shown in Fig. 3(b). The distance data obtained by the LiDAR sensor were  $0^\circ$  to the floor,  $90^\circ$  to the front, and  $180^\circ$  to the ceiling, based on the direction of progress for the UAV. In the polar coordination system, each of the

### Algorithm 1. Stair-climbing algorithm.

#### Input :

Stairs detected  $\longrightarrow$  Data from DL-model

Images as input

#### Initialization :

1: Check stairs' distance  $\longrightarrow$  LiDAR sensor data

2: SET:  $r$  in meters  $\longrightarrow$  Threshold for maneuver

#### Decision:

for distance

if distance value  $a$  or  $b < r$   $\longrightarrow$  Maneuver UAV

up  $z$  meters

else

no movement

if distance value  $c < r$   $\longrightarrow$  Maneuver UAV

down  $z$  meters

if stairs not detected

go up next stairs

else

forward  $x$  meters

end if

end if

end if

end for

Output: Go up stairs with minimum space consumption.

raw laser points is defined as  $\{(d_i, \theta_i); 0 \leq i \leq 359\}$ , where  $d_i$  is the distance from the UAV center to the object, and  $\theta_i$  is the relative angle of measurement. The information obtained by the LiDAR is stored as a vector  $(d_i, \theta_i)$ , and the stored data are checked to convert the values of the infinity scan.

## 3.2 Stair-climbing System

Algorithm 1 is used by the UAV to climb stairs. When steps are recognized by the camera, the algorithm starts. If the distance between the UAV and the stairs is longer than  $r$  meters, a straight start is performed on the  $x$ -axis, or a rising maneuver on the  $z$ -axis, to avoid collisions if the distance is less than  $r$  m. At this instant, if a staircase is not recognized, the stair climb mission is determined as complete, and recognition for climbing the next step commences.

## 3.3 Deep Learning Model for Detection of Stairs

In this study, a DL approach is implemented for detecting stairs, which the drone uses to make decisions intelligently in order to follow the stairs and determine the next maneuver. In this work, we improved the YOLOv3-tiny default model. The backbone of YOLO is *darknet*, where the YOLOv3-tiny default model uses six max-pooling and seven convolution layers. We modified it by adding one more layer. Instead of the softmax function, and where multi-class classification and detection is an

issue, regression is employed to solve the multi-class detection and classification problem [33].

The proposed model starts by dividing the stair-image input into a  $G \times G$  grid in the training stage. A bounding box is used as a tool for labeling five features—width  $w$ , height  $h$ , vertical height  $v$ , horizontal height  $u$ —as shown in Fig. 4, and confidence score  $C$ , which represents the presence of stairs within the bounding box, and hence, represents the accuracy.

In the proposed YOLOv3-tiny method, we attempt to make the model computationally inexpensive, along with implementing it to extract more semantic features. Max-pooling is used after each convolution layer to reduce the computational complexity and improve image feature extraction. Fig. 6 shows the network architecture for both the default and the improved YOLOv3-tiny models. The loss function is obtained as an end-to-end network, and can be expressed as follows [33]:

$$loss = \sum_{i=0}^{S^g} iouErr + coordErr + clsErr \quad (1)$$

where  $iouErr$ ,  $coordErr$ , and  $clsErr$  indicate the IOU error, coordinates error, and classification error, respectively. We used a rectified linear unit (ReLU) as an activation function to achieve sparsity and reduce vanishing gradient issues [25]. Table 2 details the training configuration employed for both YOLOv3-tiny and the proposed

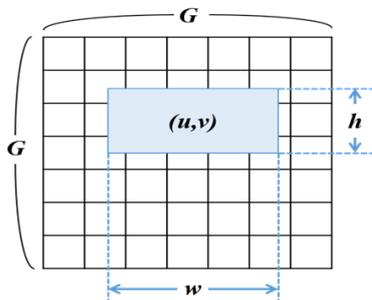


Fig. 4. Definition of the bounding box.

improved YOLOv3-tiny model.

### 3.4 ROS

The nodes that are separated and managed by the master are shown in Fig. 5. In addition, the topic node continuously communicates the results processed by the publisher node, and makes them available to other nodes by subscription. The system proposed in this paper is largely a UAV status message, a  $scan$  value obtained from the LiDAR, and a visual message obtained from the UAV camera. When running *darknet* on the ROS, the messages required from the published messages are subscribed. Among them, a message containing information on the bounding box is received through the *darknet\_ros* node. When the proposed DL model detects a staircase, a

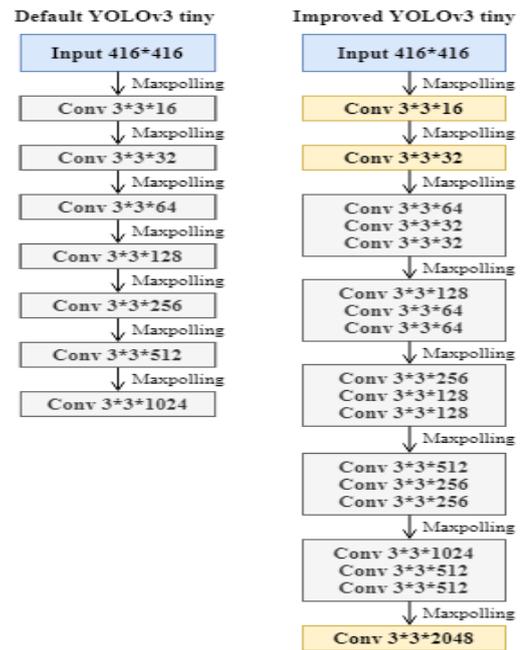


Fig. 6. YOLO models: the default YOLOv3-tiny and the improved YOLOv3-tiny.

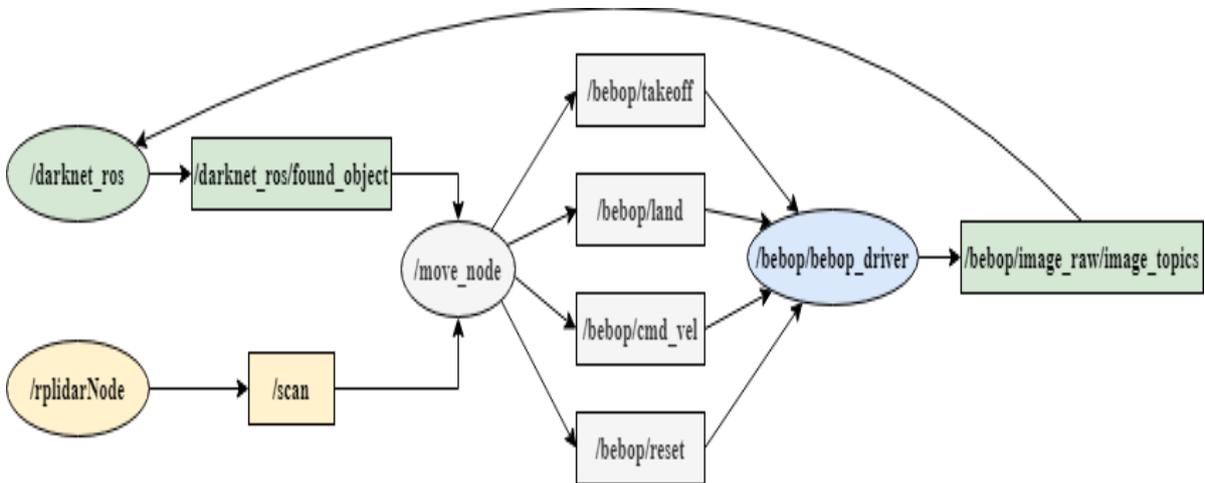
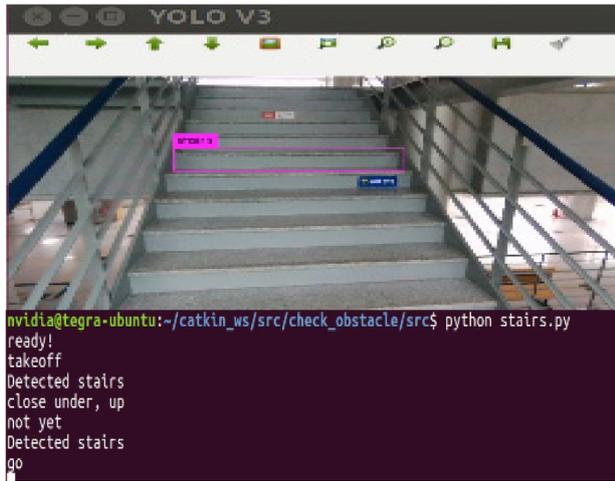
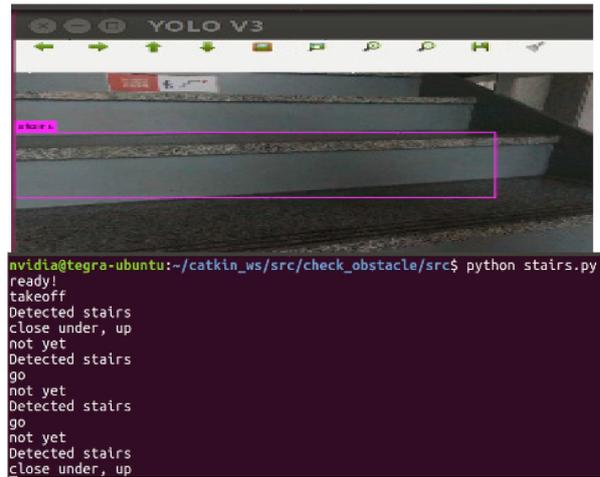


Fig. 5. ROS node graph.

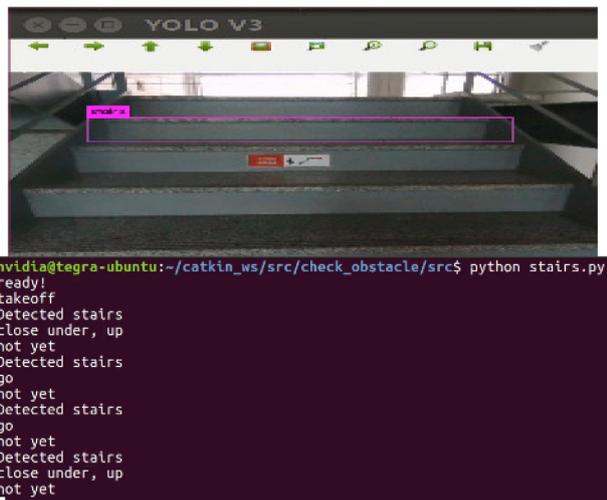




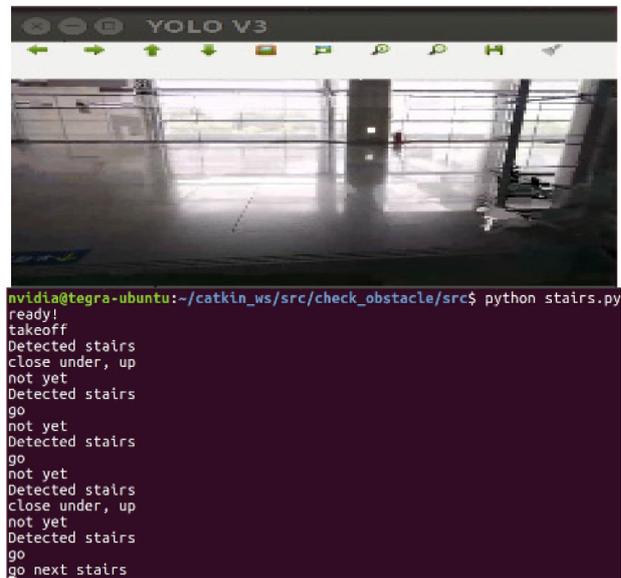
(a)



(b)



(c)



(d)

Fig. 9. GCS screen commands and screen shots from the UAV's built-in camera for (a) forward movement; (b) upward movement; (c) hovering; (d) going to the next stair.

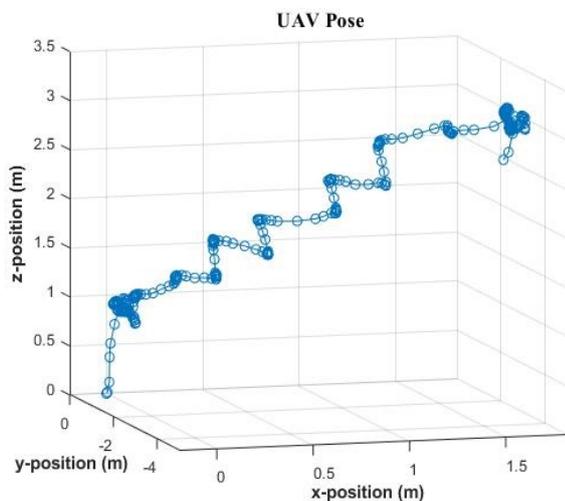


Fig. 10. Trajectory of the UAV.

Table 4. Performance Time of the Proposed Stair-climbing Scheme.

No.	Takeoff	Landing
1	0:06.35	1:00.91
2	0:05.22	0:57.16
3	0:05.78	1:07.18
Average	0:05.78	1:01.75

### 5. Conclusion

In this study, we designed, implemented, and experimented with a system in which a UAV recognizes and climbs stairs, which are obstacles often encountered during indoor flight. The system was implemented through a CNN-based imaging process for real-time stair

recognition and by using LiDAR-based distance measurements. The accuracy derived from stair recognition was 92.06%, and the actual test results showed that stair climbing was carried out without collisions.

Future research would require more efficient algorithms to climb various types of stairs. Moreover, the proposed system can be combined with SLAM navigation to expand studies to systems that can autonomously fly through multiple floors.

## Acknowledgement

This paper was supported by the National University Development Project in 2020.

## References

- [1] P. R. Prasad, et al., "Monocular vision aided autonomous UAV navigation in indoor corridor environments." *IEEE Transactions on Sustainable Computing*, Vol. 4, No. 1, pp. 96-108, 2018. [Article \(CrossRef Link\)](#)
- [2] Y. Lu, et al., "A survey on vision-based UAV navigation." *Geo-spatial information science*, Vol. 21, No. 1, pp. 21-32, 2018. [Article \(CrossRef Link\)](#)
- [3] M. Ilyas, et al., "Design of sTetro: A Modular, Reconfigurable, and Autonomous Staircase Cleaning Robot," *Journal of Sensors*, Vol. 2018, 16 pages. Jul. 2018. [Article \(CrossRef Link\)](#)
- [4] X. Gao, et al., "Dynamics and stability analysis on stairs climbing of wheel-track mobile robot," *International Journal of Advanced Robotic Systems*, Vol. 14, No. 4, pp. 1729881417720783, 2017. [Article \(CrossRef Link\)](#)
- [5] J. Israelsen, et al., "Automatic collision avoidance for manually tele-operated unmanned aerial vehicles." In 2014 IEEE International Conference on Robotics and Automation (ICRA), pp. 6638-6643, 2014. [Article \(CrossRef Link\)](#)
- [6] L. Chnibo, et al., "UAV position estimation and collision avoidance using the extended Kalman filter." *IEEE Transactions on Vehicular Technology*, Vol. 62, No. 6, pp. 2749-2762, 2013. [Article \(CrossRef Link\)](#)
- [7] J. Cui, et al., "Autonomous navigation of UAV in foliage environment." *Journal of intelligent & robotic systems*, Vol. 84, No. 1 pp. 259-276, 2016. [Article \(CrossRef Link\)](#)
- [8] Z. Huizhong, et al., "StructSLAM: Visual SLAM with building structure lines." *IEEE Transactions on Vehicular Technology*, Vol. 64, No. 4 pp. 1364-1375, 2015. [Article \(CrossRef Link\)](#)
- [9] A. E. Oguz, et al., "On the consistency analysis of A-SLAM for UAV navigation. Proc. SPIE 9084, Unmanned Systems Technology XVI, Vol. 9084, pp. 90840R, June. 2014. [Article \(CrossRef Link\)](#)
- [10] S. Ross, et al., "Learning monocular reactive uav control in cluttered natural environments." In 2013 IEEE international conference on robotics and automation, pp. 1765-1772, 2013. [Article \(CrossRef Link\)](#)
- [11] A. Fraust, et al., "Automated aerial suspended cargo delivery through reinforcement learning." *Artificial Intelligence*, Vol. 247, pp. 381-398, 2017. [Article \(CrossRef Link\)](#)
- [12] N. Imanberdiyev, et al., "Autonomous navigation of UAV by using real-time model-based reinforcement learning." In 2016 14th international conference on control, automation, robotics and vision (ICARCV), pp. 1-6, 2016. [Article \(CrossRef Link\)](#)
- [13] B. L. Sturm, et al., "Machine learning research that matters for music creation: A case study," *Journal of New Music Research*, Vol. 48, No.1, pp. 36-55, 2019. [Article \(CrossRef Link\)](#)
- [14] J. Raharjo, et al., "Cholesterol level measurement through iris image using gray level co-occurrence matrix and linear regression," *ARNP Journal of Engineering and Applied Sciences*, Vol. 14, No. 21, pp. 3757-3763, Nov. 2019. [Article \(CrossRef Link\)](#)
- [15] Y. Zhang, et al., "Machine learning based video coding optimizations: A survey." *Information Sciences*, Vol. 506, pp.395-423, Jan. 2020. [Article \(CrossRef Link\)](#)
- [16] M. Heidari, et al., "Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms," *International journal of medical informatics*, Vol. 144, pp. 104284, Sep. 2020. [Article \(CrossRef Link\)](#)
- [17] S. A. Hassan, et al., "Real-time uav detection based on deep learning network," In 2019 International Conference on Information and Communication Technology Convergence, pp. 630-632, Oct. 2019. [Article \(CrossRef Link\)](#)
- [18] J. Redmon, et al., "You only look once: Unified, real-time object detection," In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788, 2016. [Article \(CrossRef Link\)](#)
- [19] D. Mellinger, et al., "Minimum snap trajectory generation and control for quadrotors," In 2011 IEEE international conference on robotics and automation, pp. 2520-2525, 2011. [Article \(CrossRef Link\)](#)
- [20] D. Mellinger, et al., "Trajectory generation and control for precise aggressive maneuvers with quadrotors," *The International Journal of Robotics Research*, Vol. 31, No. 5, pp. 664-674, Jan. 2012. [Article \(CrossRef Link\)](#)
- [21] P. Checchin, et al., "Radar scan matching slam using the fourier-mellin transform," In *Field and Service Robotics*, Vol. 62, pp. 151-161, 2010. [Article \(CrossRef Link\)](#)
- [22] J. Engel, et al., "LSD-SLAM: Large-scale direct monocular SLAM," In *European conference on computer vision*, Vol. 8690, pp. 834-849, 2014. [Article \(CrossRef Link\)](#)
- [23] C. Mei, et al., "RSLAM: A system for large-scale mapping in constant-time using stereo," *International journal of computer vision*, Vol. 94, No.2, pp. 198-214, Jun. 2011. [Article \(CrossRef Link\)](#)
- [24] M. Müller, et al., "Quadcopter ball juggling," in

- 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 5113–5120, Sep. 2011. [Article \(CrossRef Link\)](#)
- [25] A. S. Huang, et al., “Visual odometry and mapping for autonomous flight using an RGB-D camera,” *Robotics Research*. Vol. 100, pp. 235–252, Aug. 2011. [Article \(CrossRef Link\)](#)
- [26] J. F. Roberts, et al., “Quadrotor using minimal sensing for autonomous indoor flight,” In *European Micro Air Vehicle Conference and Flight Competition (EMAV2007)*, Sep. 2007. [Article \(CrossRef Link\)](#)
- [27] A. Bry, et al., “State estimation for aggressive flight in GPS-denied environments using onboard sensing,” In *2012 IEEE International Conference on Robotics and Automation*, pp. 1-8, May, 2012. [Article \(CrossRef Link\)](#)
- [28] A. Bachrach, et al., “Autonomous flight in unknown indoor environments,” *International Journal of Micro Air Vehicles*, Vol. 1, No. 4, pp. 217-228, Dec. 2009. [Article \(CrossRef Link\)](#)
- [29] M. Achtelik, et al., "Onboard IMU and monocular vision based control for MAVs in unknown in-and outdoor environments." *2011 IEEE International Conference on Robotics and Automation*, pp. 3056-3063, 2011. [Article \(CrossRef Link\)](#)
- [30] M. Blösch, et al., "Vision based MAV navigation in unknown and unstructured environments." *2010 IEEE International Conference on Robotics and Automation*, pp. 21-28, 2010. [Article \(CrossRef Link\)](#)
- [31] G. Nützi, et al., "Fusion of IMU and vision for absolute scale estimation in monocular SLAM." *Journal of intelligent & robotic systems*, Vol. 61, No. 1, pp. 287-299, Nov. 2011. [Article \(CrossRef Link\)](#)
- [32] S. Weiss, et al., “Versatile distributed pose estimation and sensor self-calibration for an autonomous MAV,” In *2012 IEEE International Conference on Robotics and Automation*, pp. 31-38, 2012. [Article \(CrossRef Link\)](#)
- [33] T. Rahim, et al., “A Deep Convolutional Neural Network for the Detection of Polyps in Colonoscopy Images,” *Biomedical Signal Processing and Control* 68 (2021): 102654.



**Yeonji Choi** received her BSc in Electrical Engineering in 2019 and received her MSc from the Department of IT Convergence Engineering at Kumoh National Institute of Technology (KIT) Gumi, South Korea, in 2021. Currently, she is working as graduate research assistant at the Wireless and Emerging Network System (WENS) Lab in the Department of IT Convergence Engineering, Kumoh National Institute of Technology (KIT), Gumi, South Korea. Her major research interests include intelligent control and systems, Unmanned Aerial Vehicles, and wireless communications.



**Tariq Rahim** is a PhD student in the Wireless and Emerging Network System Laboratory (WENS Lab) of the Department of IT Convergence Engineering, Kumoh National Institute of Technology, Republic of Korea. He completed his master's degree in Information and Communication Engineering from Beijing Institute of Technology, PRC, in 2017. His research interests include image and video processing and quality of experience for high-resolution videos.



**Soo Young Shin** received his BSc, MSc, and PhD in Electrical Engineering and Computer Science from Seoul National University, Korea, in 1999, 2001, and 2006, respectively. He was a visiting scholar for the FUN Lab at the University of Washington, U.S.A., from July 2006 to June 2007. After three years working in the WiMAX Design Lab of Samsung Electronics, he is now an associate professor for the School of Electronics at Kumoh National Institute of Technology, joining the institute in September 2010. His research interests include wireless LANs, WPANs, WBANs, wireless mesh networks, sensor networks, coexistence among wireless networks, industrial and military networks, cognitive radio networks, and next-generation mobile wireless broadband networks.