IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 29 December 2023, accepted 10 January 2024, date of publication 17 January 2024, date of current version 25 January 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3355279

RESEARCH ARTICLE

Enhancing Optical Camera Communication Performance for Collaborative Communication Using Positioning Information

IDA BAGUS KRISHNA YOGA UTAMA[1], (Graduate Student Member, IEEE), ONES SANJERICO SITANGGANG[1], MUHAMMAD RANGGA AZIZ NASUTION[1], MD. IBNE JOHA[1], JAEJUN YOO[2], AND YEONG MIN JANG[1], (Member, IEEE)

[1]Department of Electronics Engineering, Kookmin University, Seoul 02707, South Korea
[2]Mobility UX Research Section, Electronics and Telecommunications Research Institute, Daejeon 34129, South Korea

Corresponding authors: Jaejun Yoo (jjryu@etri.re.kr) and Yeong Min Jang (yjang@kookmin.ac.kr)

This work was supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) Grant funded by the Ministry of Land, Infrastructure and Transport, under Grant RS-2022-00141819.

ABSTRACT Optical camera communication (OCC) has emerged as a promising alternative technology for radio frequency (RF)-based communication systems. However, existing OCC approaches only consider transmitting data through broadcasting, without any ability for point-to-point communication. In this paper, a new modulation scheme termed distance color-coded on-off keying (DCC-OOK) is proposed that uses distance information retrieved from relative localization. The proposed method allows OCC to perform point-to-point communication and facilitates bidirectional communication. An OCC system that uses multi-sensor fusion of a camera and LiDAR is considered, in which calibration of the sensors and LED array segmentation for performing relative localization between the OCC transmitter and the OCC receiver are explained. The results indicate that both point-to-point and bidirectional communication were achieved with the proposed OCC system. During implementation, the distance between transmitter and receiver was varied between 1.0 and 4.0 m, and the system demonstrated a maximum data rate of 38.4 kbps with a lowest BER of 0.03.

INDEX TERMS Optical camera communication, modulation, bidirectional, multi sensor fusion, relative localization, on off keying.

I. INTRODUCTION

To solve the limitations in RF communication, such as the limited spectrum and susceptibility to interference, optical wireless communication (OWC) has emerged as an alternative technology [1]. The advantages of OWC include a large available spectrum, resistance towards electromagnetic interference (EMI), license-free operation [2], good communication security, and high energy efficiency [3]. Optical camera communication (OCC) is a subsystem of OWC that employs an image sensor as the receiver and a light-emitting diode (LED) as the transmitter. The main advantage of OCC over other OWC subsystems (for example, LiFi) is its ability to use a variety of data carriers, including visible light, infrared (IR), and ultraviolet (UV). Additionally, OCC offers long-distance communication [4], resistance towards ambient light in outdoor applications, and the ability to capture data from different sources independently [5].

Although it appears to be comparable to other OWC approaches, such as LiFi or VLC, OCC is different. The key distinction is in the receiver approaches; other OWC techniques employ a photodiode as the receiver, but OCC employs a camera. To get the information, the OCC receiver processes a picture provided by the camera. As a result, because the image may contain undesirable items, the image processing approach is crucial in OCC to pick areas of the

The associate editor coordinating the review of this manuscript and approving it for publication was Barbara Masini.

image that contain the conveyed information. As a result of the receiver differences, the communication strategy in OCC, LiFi, or VLC differs.

Currently, collective intelligence is implemented in various fields [6]. In intelligent transportation systems (ITS), one application of collective intelligence is platooning, which significantly reduces fuel consumption and provides a smooth traffic flow [7]. In drone systems, collective intelligence is used to establish drone swarming when performing certain goals, such as formation control [8], espionage [9], or defense mechanisms [10]. Moreover, collective intelligence is widely used in robotics applications to create multi-agent systems, in which robots perform collaborative actions to accomplish a certain mission [11].

The application of collective intelligence in the aforementioned fields requires reliable communication between each entity in the system. Most research in collective intelligence proposes radio frequency (RF) as the communication technology. However, in some fields, using RF-based communication strategies can incur several critical disadvantages. For example, when vehicles use RF communication in ITS, the communication latency between vehicles is quite high because the data has to be transmitted to a nearby base station (BS) before reaching the destination [12]. This latency issue can be solved by using 5G communication, which allows low-latency data transmission. Nevertheless, the BS for 5G has a small coverage area, meaning numerous BSs need to be installed to cover the whole area in ITS, which incurs high installation costs. Moreover, if some areas are not covered by the BSs, the communication is inoperative, disabling vehicle collaboration and interrupting the collective intelligence between the vehicles.

Communication based on RF is also widely employed in drone systems to provide reliable communication between the drones. However, RF-based systems are prone to signal jamming [8] and spoofing [13], which could break the collective intelligence in the drone system and incur a higher risk of information leakage. In the field of robotics, there are certain applications where RF-based communication cannot be used, such as underwater situations.

To solve these issues, OCC has emerged as an alternative solution for replacing RF-based communications. For example, OCC has been widely proposed in ITS as a promising technology for vehicular communication because it allows for direct point-to-point communication between vehicles, resulting in significantly reduced latency [8]. In vehicular applications, the OCC receiver is mounted on the windshield of the vehicle to allow reading the transmitted data from the vehicle in front, while the transmitter is mounted on the tail lamp to send information to the following vehicle. The data is transmitted in a broadcast manner, allowing all the vehicles behind to receive the data.

In addition, OCC has emerged as a communication solution for drone systems. OCC improves drone system security, making the drones resistant to network jamming and spoofing. Several studies have applied OCC to drone systems. For example, Zhang et al. [8] proposed OCC for effectively controlling drone swarming. Long-range drone communication was also performed by Takano et al. [14], where long-distance OCC-based drone communication was achieved over a distance of 300 m. Importantly, it is easy to integrate OCC with existing ITS, drone, or robotics systems because most modern systems use image sensors for perception, meaning no changes are required in terms of hardware configuration when adopting OCC.

However, OCC suffers from several drawbacks compared to RF-based communication or other OWC methods. First, OCC has a relatively low throughput caused by the low frame-per-second (fps) rates of generally available cameras. This issue can be solved by using a high-fps camera to increase the throughput. Second, existing OCC systems transmit data in a broadcast manner, which only allows reception of the transmitted data provided that the receiver is located in the field of view of the transmitter. Moreover, although broadcast-manner transmission is beneficial in certain situations, when transmitting confidential data, a point-to-point approach is preferred to ensure information security. Third, because OCC mainly communicates in a broadcast manner, it is difficult to perform bidirectional communication using OCC.

In this paper, we propose a new modulation technique for OCC systems that enables point-to-point data transmission and easier bidirectional communication. We expect the proposed method to find applications in regular OCC systems that require point-to-point data transmission and bidirectional communication. However, we hope that the proposed method can be adopted in multi-agent robot systems, vehicular communication, or UAV swarms as its communication backbone in the future. By using the proposed method, the OCC system can securely transmit data to the target receiver directly without having to broadcast the information to all visible receivers, while also providing the ability to establish bidirectional communication. Currently, transmitting data to a target receiver can be performed by applying a preamble code at the beginning of the data packet, where each preamble code is unique to each OCC system [15].

However, assigning preamble codes to differentiate the receiver of the data packet is not practical in mobile environments because this would force the OCC receiver to continuously focus on the OCC transmitter and constantly receive and decode data packets from the source, even if the data packet is for another receiver. In other words, using the preamble method still requires the receiver to continuously decode frame-by-frame transmitted data packets even if the information is not for them, which wastes receiver resources.

Most modern vehicles and drones are equipped with range-measuring devices (such as LiDAR) that can provide distance information directly. In this work, we propose using relative distance information between the transmitter and the neighboring receiver by adding it to the OCC data modulation, where the relative distance information is coded

as a color and combined with existing on-off keying (OOK) modulation to control the LED. Accordingly, we propose a new modulation technique termed distance color-coded on-off keying (DCC-OOK) for mitigating broadcast issues in OCC, enabling point-to-point communication and providing easier bidirectional communication.

By localizing the relative location of the neighboring entities using a fusion of the camera and LiDAR, this enables the system to recognize the receiver easily while allowing other receivers in the field of view of the transmitter to automatically disregard the transmitted data and focus on other tasks rather than having to focus on receiving and decoding the transmitted data frame-by-frame. The proposed method exploits distance information extracted using the relative localization method to be coded as a color in the LED. When two OCC systems need to communicate with each other, they will be at the same distance from each other. Hence, they will have the same LED array color. When the LED array colors match, both OCC systems can start communicating. In contrast, if the LED colors do not match, the OCC systems can discard the information and focus on another task while waiting for a neighboring OCC system to present the same LED array color, requesting communication.

The overall contributions of this paper are as follows:

1) A method is provided to efficiently localize an LED as an OCC transmitter to obtain the relative distance by using multi-sensor fusion of a camera and LiDAR.
2) A new distance color-coded on-off keying (DCC-OOK) method is presented for enabling point-to-point communication in OCC.
3) Easier setup is achieved for bidirectional communication in OCC compared to existing methods.

The remainder of this paper is organized as follows: Section II highlights recent work on OCC modulation, while Section III provides an explanation of the method for performing relative localization used in this work. Section IV contains a detailed explanation of the proposed DCC-OOK modulation. Section V details the experimental results and subsequent analysis, and the paper is concluded in Section VI.

II. RECENT WORKS

Over the years, several modulation methods have been proposed for OCC. On-off keying (OOK) is one such modulation scheme, in which an ''on'' LED represents a ''1'' bit, while a ''0'' bit is represented by an ''off'' LED [16]. Several types of OOK schemes have been proposed by researchers. In [17], an asynchronous OOK was proposed to mitigate the synchronization problems present in traditional OOK. To improve the performance, Manchester coding and the camera rolling shutter effect were combined with OOK [18]. By combining these techniques, a higher data rate with a lower bit-error rate (BER) was achieved.

Instead of using the on-and-off states of LEDs, the frequency shift keying (FSK) scheme uses several LEDs that operate at different frequencies to transfer the data [19]. By transmitting data using FSK modulation, a higher OCC data rate can be achieved. However, based on [20], the communication distance is limited, with a maximum distance of less than 2 m. The emerging use of RGB LEDs and color cameras has facilitated a new modulation scheme termed color shift keying (CSK). The CSK scheme transmits data through a variation of colors emitted by an RGB LED [21]. The color camera captures these variations in color and decodes them accordingly. Although CSK is able to provide a higher data rate, it suffers from a short communication distance because the colors cannot be identified over longer distances [22].

To improve the communication distance, an undersampling modulation scheme has been proposed. Undersampled frequency shift on-off keying (UFSOOK) is a modulation technique that uses undersampling [23], which allows longer communication distances for OCC. Meanwhile, to improve the data rate of OCC, quadrature amplitude modulation (QAM) has proved effective [24]. By using QAM, several bits can be transmitted using a single LED, which significantly improves the data rate compared to other modulation schemes that only allow single-bit transmission by a single LED.

Recently, the usage of MIMO techniques has gained researcher attention. Instead of using a single LED, an array of LEDs is utilized as the transmitter. The MIMO technique enables a more efficient operation of OCC compared to using a single LED. In [25], a rolling-shutter-based MIMO using an LED array is introduced. By using the method, it promises flicker-free transmission, although the communication distance is short and does not support rotation detection. In [26], a color-intensity modulation using the MIMO technique is presented, where a high frame rate camera is employed. The work shows that LED colors with different intensities can be utilized to modulate data and successfully produce a quite high data rate of up to 126.72 kbps. However, the intensity level is difficult to detect by the camera, which limits the communication distance to 1.4 m and the BER value to up to 10^{-1}. The authors in [27] show OCC systems using the MIMO technique and OOK modulation. Their work shows promising results where the communication distance increased up to 20 m with a high data rate of 1.920 kbps using an 8×8 LED array. The proposed scheme is also not affected by camera types (global shutter and rolling shutter) used to perform the OCC systems. However, the authors did not present the possibility of performing bidirectional communication or point-to-point data transmission.

To date, many OCC modulation schemes have been proposed. However, all existing OCC modulation techniques only consider transmitting information in a broadcasting manner, where the data is transmitted to all other OCC systems that are in the field of view. However, broadcasting information places OCC at a disadvantage because data security is vulnerable due to the inability to select the destination. Bidirectional communication using OCC is also

difficult to perform when the data is broadcast. In this paper, a new OCC modulation technique that considers point-to-point communication using OCC is proposed, in which the system can select the data destination based on relative localization information. The proposed scheme also allows easier implementation of bidirectional OCC because the system is able to select the communication partner.

In the proposed method, a relative localization technique is employed, where a multi-sensor fusion of LiDAR and a camera is used for localization. There is no other modulation scheme that uses relative localization information to be added to the modulation schemes, making this work the first to combine relative localization with a modulation scheme for OCC systems. The relative localization commonly used in robotics systems, especially in robot swarm applications, provides accurate positioning to avoid collisions or cooperate between entities [28]. The relative localization is also utilized in the autonomous vehicle for cooperative perception to widen the sensing coverage among vehicles [29]. Numerous works also utilize relative localization in the UAV for navigation guidance and formation control [30], [31]. From the aforementioned works, the relative localization method is done by utilizing multi-sensor fusion of cameras, LiDAR, UWB, and other types of sensors. Therefore, multi-sensor fusion-based relative localization is proven to provide accurate localization performance.

The proposed method uses a fusion of LiDAR and cameras to perform the relative localization. To join the sensor data, a calibration technique, often called point cloud registration, is required. Generally, the methods for calibrating the LiDAR and camera can be grouped into two categories: traditional approaches and learning-based approaches. In traditional approaches, the transformation parameters are retrieved by using handcrafted features such as FPFH [32], SHOT [33], or Tensor [34]. The traditional approaches suffer from difficulty creating a handcrafted feature, which requires an expert to calculate it. However, the traditional approaches win in terms of processing time, where the computational complexity is very low and yields a fast system response time. Meanwhile, the learning-based approaches utilize deep neural networks to automatically generate the features or perform direct end-to-end learning to generate the transformation [35], [36], [37], [38]. Unfortunately, deep neural networks dramatically increase computational complexity, which results in the impossibility of applying them to time-sensitive applications such as modulation schemes. Therefore, the traditional approaches are better for modulation schemes due to their fast processing times.

III. MULTI-SENSOR FUSION FOR RELATIVE LOCALIZATION

To establish the DCC-OOK method, it is important to determine the precise localization of the OCC transmitter. In this work, the fusion of a camera and LiDAR is employed to estimate the relative localization between entities. Since the LED size of the OCC transmitter is quite small, a certain technique needs to be performed to localize the OCC transmitter precisely. First, the camera and LiDAR need to be calibrated to unify their field of view and correctly project the 3-D point cloud generated by the LiDAR to the 2-D images generated by the camera. Then, a segmentation method is used to localize the estimated location of the OCC transmitter in the 2-D images. After determining the location of the OCC transmitter, point cloud matching is performed based on the location of the OCC transmitter to recover its relative distance information. Fig. 1 summarizes the steps to perform the relative localization as a flowchart. The details of each method are discussed in the following subsection.

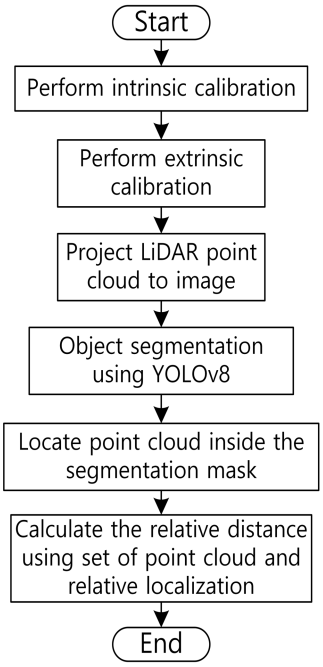
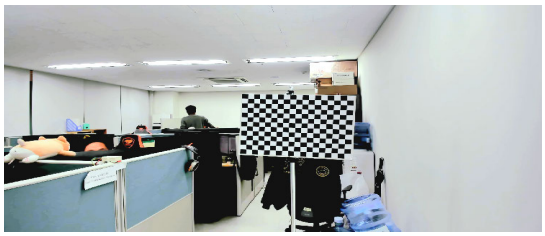


FIGURE 1. Flowchart of the relative localization method.

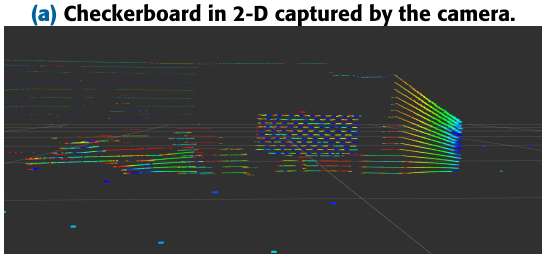
A. INTRINSIC AND EXTRINSIC SENSOR CALIBRATION

The camera used in this work is a pinhole model that creates multiple distortions in the produced images, meaning it required initial calibration to fix the distortion problem. This process is termed camera intrinsic calibration, where the aim is to recover the camera intrinsic parameters K_i and distortion coefficient β from the images and then generate rectified images. In this work, the camera intrinsic calibration module from OpenCV [39] was used to perform the calibration.

Next, a rigid-body transformation that defines the relative pose between the camera and LiDAR is calculated, which is also known as sensor extrinsic calibration. The aim is to determine the transformation vector of six parameters $(\theta = t_x, t_y, t_z, r_x, r_y, r_z)$, where the parameters represent the relative position of LiDAR in the reference frame of the camera. To find these parameters, reference points from the 2-D image and 3-D point cloud need to be extracted.



(a) Checkerboard in 2-D captured by the camera.



(b) Checkerboard in a 3-D image captured by LiDAR.

FIGURE 2. Captured checkerboard image.

The target-based calibration technique was adopted, where a checkboard is used as the reference point to calibrate the sensors. Fig. 2 displays an image of the checkerboard in 2-D and 3-D formates.

The translation and rotation parameters are achieved by applying the random sampling consensus (RANSAC) to the sampled 2-D and 3-D points. Initially, several reference points from the 2-D and 3-D images are selected, and the selected points must be of the same object presented in both images. After selecting the set of 2-D and 3-D points, RANSAC is employed to remove any outliers while keeping the inliers for calculating the translation and rotation vectors. Then, the translation and rotation vectors are calculated using the set of inlier points.

The generated translation and rotation vectors are estimations of the parameters. A Levenberg-Marquardt algorithm is employed to refine the calibration parameters to achieve lower rotation and translation errors between the 2-D and 3-D points. The equation for the Levenberg-Marquardt algorithm has a function to minimize the projection errors, which is represented as follows:

$$\hat{X} = Argmin_x F(x) \tag{1}$$

where x refers to the parameter vectors resulting from RANSAC, $F(x)$ refers to the errors between the 2-D and 3-D points, and $\hat{X}$ refers to the optimized parameter vectors that achieve the lowest errors.

The overall steps for performing the extrinsic calibration method are depicted in Algorithm 1.

Algorithm 1 Steps to Perform the Camera-LiDAR Extrinsic Calibration

Input: F_{cam}: frame from camera stream, F_{LiDAR}: frame from LiDAR stream

Output: Refined calibration parameters $\theta = (t_x, t_y, t_z, r_x, r_y, r_z)$

From the camera stream, get the recent frame F_{cam} and F_{LiDAR}
Make sure that F_{cam} and F_{LiDAR} are taken at the same timestamp
Select reference points from F_{cam} and store them in $img_{\text{point\_sets}}$
Select reference points from F_{LiDAR} and store them in $pcl_{\text{point\_sets}}$
Load the RANSAC model
while err $> TargetError_{\text{RANSAC}}$ **do**
    Find the inliers from $img_{\text{point\_sets}}$ and $pcl_{\text{point\_sets}}$
    Calculate the calibration parameters using the inliers
    Measure the projection error
end while
Load the calibration parameters from RANSAC
Load the Levenberg-Marquardt model
while err $> TargetError_{\text{Levenberg-Marquardt}}$ **do**
    Calculate the calibration parameters
    Measure errors
    Update the parameters
end while

B. OBJECT SEGMENTATION

Usually, object detection is employed to find an object's location in an image. However, using object detection only outputs a bounding box that is not sufficiently tight to fit the object shape. Hence, object segmentation is utilized because it can produce a pixel resolution mask of the target object. Typically, object segmentation can be grouped into three categories: semantic, instance, and panoptic.

In semantic segmentation, only the group of an object is detected, without any ability to distinguish different objects in the same categories. Instance segmentation is able to detect the object class as well as assign a unique identifier to each detected object. Panoptic segmentation is a combined version of semantic and instance segmentation. In this work, semantic segmentation was selected because it can differentiate between objects that belong to the same class. Although panoptic segmentation can also achieve this, it was not selected due to its heavy computational burden.

Instance segmentation was developed using the YOLOv8 model, which was initially developed and perfected by Ultralytics [40]. YOLOv8 is the latest version of the YOLO (You Only Look Once) series, which is noted for providing good performance in object detection tasks. YOLOv8 was selected because it has a good balance between detection performance and computational complexity. Moreover, YOLOv8 can generate an acceptable performance level within a short period. This was beneficial in this work because OCC requires a system with a high frame speed to improve its communication quality. The YOLOv8 model architecture is displayed in Figure 3.

YOLOv8 is available in several versions, including YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x. Although all variants use the same architecture,

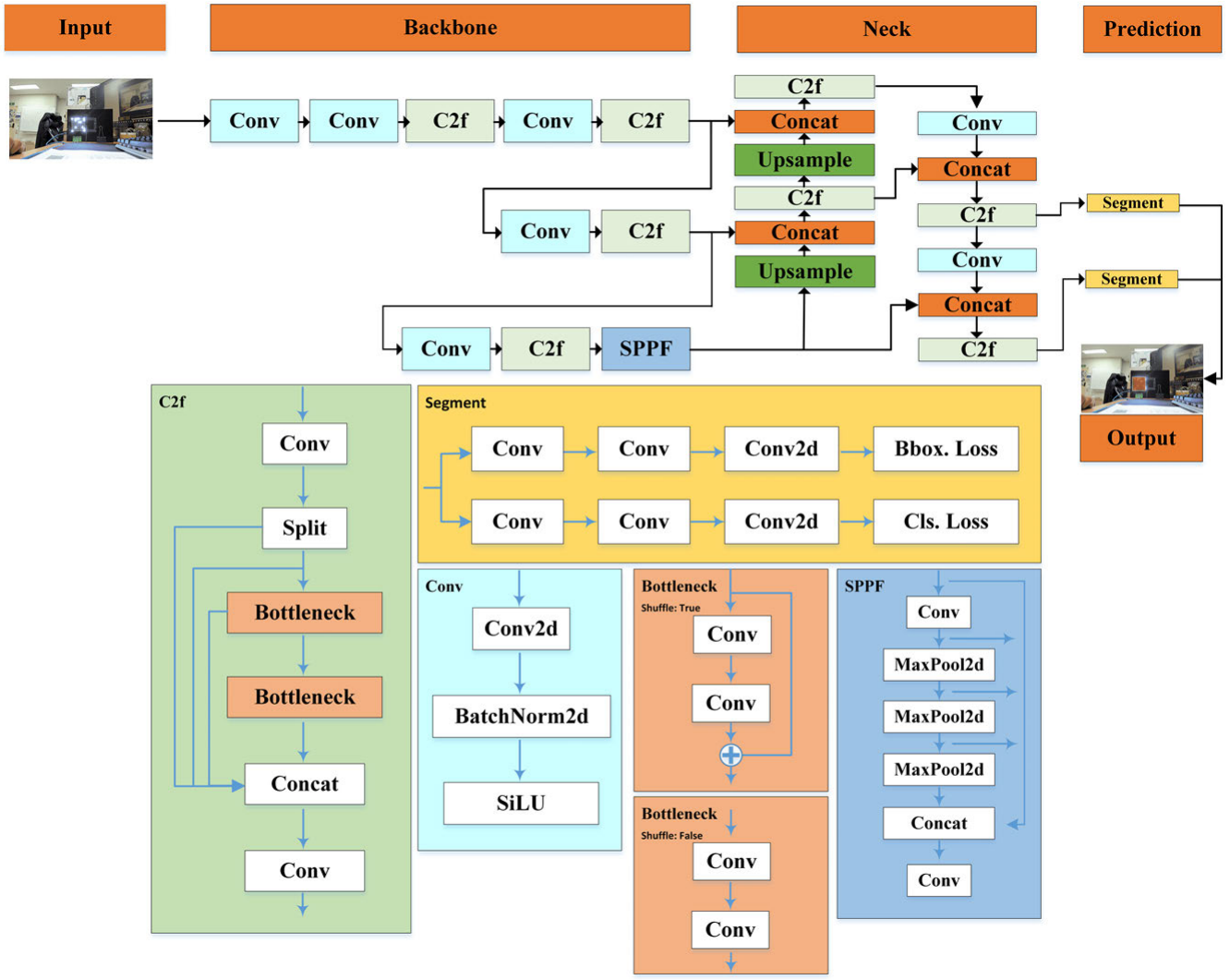


FIGURE 3. YOLOv8 architecture for segmentation task.

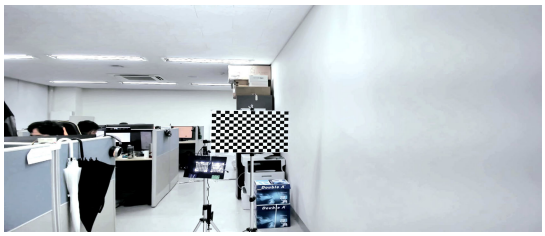
they are different in terms of the depth multiplier, width multiplier, and ratio, which determine the size of the calculated image. YOLOv8s was selected as the instance segmentation model in this work, as it has FLOPs of 28.6 with an input image size of 640 pixels.

As displayed in Figure 3, the overall format of YOLOv8 consists of four layers: input, backbone, neck, and prediction. The input layer receives the input image, which is resized to 640 pixels. Then, important features from the input image are extracted using the backbone layer, which uses the CSPDarknet53 architecture. The next layer is a C2f module, which replaces the neck layer used in previous YOLO models. To generate the segmentation mask, two segmentation heads are employed and positioned after the neck layer. The segmentation heads consist of detection modules and a prediction layer to generate a prediction of the segmentation mask of an object.

In this work, the objective was to detect the location of an LED array that was acting as the transmitter for OCC. However, there was no open dataset that provided the data of LED arrays with corresponding segmentation labels. Hence, we developed an image dataset for training the YOLOv8s segmentation model.

The image dataset was developed by ourselves and contains images of an 8×8 LED array from various angles, distances, and environments (indoors and outdoors). We collected the dataset using the LED array and camera, specifically the Logitech Brio camera. To collect the dataset, the LED array is installed on a tripod and controlled to show a random pattern with various colors as determined in the modulation schemes.

The camera is placed in front of the LED array at a distance between 1 and 6 m, where one image is taken from each distance. Additionally, the camera was also placed at a random, various angle from the LED array



(a) Indoor environment.

(b) Outdoor Environment.

FIGURE 4. LED array image dataset for training the segmentation model.

to improve segmentation performance when detecting the image at a skew angle. Moreover, the image collection from various distances and angles is performed in two environments: indoors and outdoors. In both indoor and outdoor environments, the LED array image is taken with a complex background where the settings of distance and angle are the same. Finally, we collected a total of 750 images of 8×8 LED arrays.

After collecting the images, each was labeled manually using the labeling software Computer Vision Annotation Tool (CVAT) [41] to provide the segmentation mask labels. Finally, after obtaining the labels for each image, the dataset was ready to be trained for the YOLOv8s model. An example of captured images in indoor and outdoor environments for the dataset is presented in Figure 4.

C. RELATIVE LOCALIZATION MEASUREMENT

After performing the sensor calibration steps, projected images that displayed the LiDAR points on the 2-D images were gathered. Then, from the segmentation, the estimated pixel location of the object could be determined. Since the LiDAR point cloud data was already in 3-D position format (x, y, and z with units in meters), the distance of the object could be instantly recovered.

In the mask of the detected object, more than one point cloud resided inside the mask. Hence, to obtain the relative distance of the object, the average point cloud data in the x, y, and z axes needed to be calculated, as demonstrated in Equations 2, 3, and 4. Finally, the relative distances could be calculated using Equation 5, where it refers to the Pythagorean theorem to find distance in a 3-D space [42] with the difference between two points is the average point cloud data in x, y, and z axes.

$$X_{avg} = \Sigma X / n_X \tag{2}$$

$$Y_{avg} = \Sigma Y / n_Y \tag{3}$$

$$Z_{avg} = \Sigma Z / n_Z \tag{4}$$

$$d_{relative} = \sqrt{X_{avg}^2 + Y_{avg}^2 + Z_{avg}^2} \tag{5}$$

where X_{avg}, Y_{avg}, Z_{avg} is the average position of the object in X, Y, and Z axis, respectively. ΣX, ΣY, and ΣZ is the total position of the points in X, Y, and Z axis. The total number of points in X, Y, and Z axis represented by n_X, n_Y, n_Z, respectively. The relative distance between objects is expressed by $d_{relative}$.

IV. PROPOSED DCC-OOK MODULATION SCHEME

DCC-OOK modulation is a scheme that uses distance information between the transmitter and receiver to enable point-to-point data transmission in OCC, instead of broadcasting the data, which is widely proposed in existing modulation schemes. The idea originates from a simple observation in OCC systems, where the distance between the transmitter and receiver is always the same when observed relatively from the transmitter or receiver-side. Hence, to ensure correct point-to-point transmission in OCC, the distance information can be leveraged.

In DCC-OOK, the distance between the transmitter and receiver is color-coded, where the color coding is set manually. The color code for the distance used in this work is displayed in Figure 5. Based on this color code, the LED array of the transmitter will display the corresponding color. Hence, if the receiver is located 1 m away from the transmitter, the LED array of the transmitter will turn blue. Then, the receiver who also know that the distance to the transmitter is 1 m, will receive and decode the data only when the LED array of the transmitter light up blue color. If the LED array of the transmitter is any other color, the receiver will not receive and decode the data because that information would be for another receiver positioned at a different distance relative to the transmitter.

Figure 6 displays a block diagram that explains the detailed steps of the proposed DCC-OOK modulation working principle. Initially, the transmitter needs to localize the receiver by using the fusing of the camera and LiDAR. The localization process is initiated by performing receiver detection, where the LED array of the receiver is detected by using the segmentation method. Then, relative localization is performed to retrieve the distance between the transmitter and the receiver.

Simultaneously, the data stream from the sensor that is to be transmitted is then encoded as the packet payload. The distance information is also encoded with the packet payload and the data stream. Embedding the distance information in the packet payload helps the receiver verify that the data is received from the correct transmitter. The data encoding functions to convert the decimal number from the sensor reading and relative localization into a binary number.

The sequence number (SN) is inserted into the data packet to provide each data packet with an identifier for easier packet merging when transmitting large data packets. The SN is also important for detecting missing packets that could occur

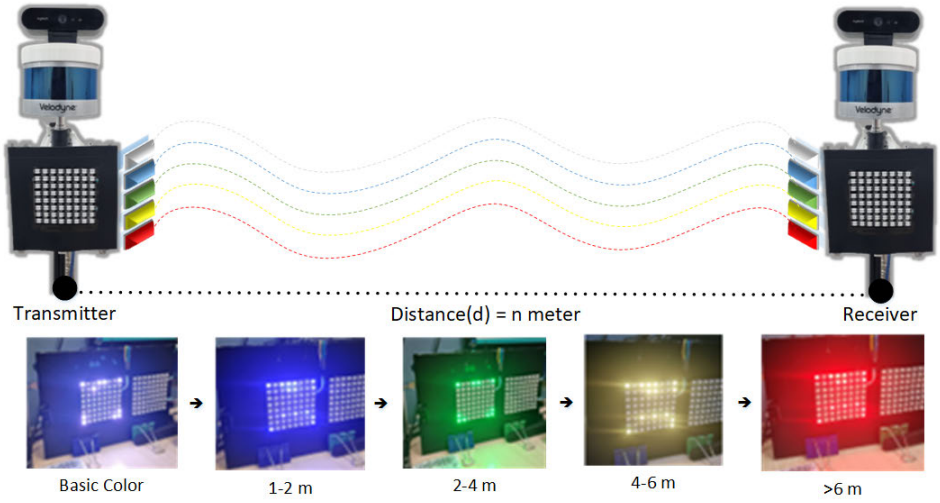


FIGURE 5. DCC-OOK color coding used in this work.

during data transmission. Then, the preamble is added to the data packet, where the preamble is the first part of the packet frame. The preamble acts as a sign for the start of the frame to help in synchronizing the data transmission. In this work, the preamble was modified so that the bit values corresponded to the distance results from relative localization.

Moreover, the preamble is also color-coded, where the LED array displays a different color according to the distance from the relative localization. Different distances between the transmitter and receiver will cause the LED array to display different colors, which are already coded. The data packet is then complete and can be mapped into the LED array. The mapping of the data packet to the LED array uses the OOK principle, where "0" bits turn off the LED inside the LED array, and "1" bits turn on the LED inside the LED array (hence the term OOK mapping). Thereafter, OOK mapping is used to control the LED array that emits the signal to the receiver. In this work, the OOK mapping is inspired by [27], where a similar 8×8 LED array is utilized.

In the receiver, the LED array of the transmitter is detected using a camera. This detection uses the segmentation method to localize the LED array in the images. Then, relative localization is performed to measure the distance to the transmitter. Preamble detection is then executed to verify that the detected preamble is the same as the color code of the distance from the relative localization. If the color code from relative localization is different than the detected preamble, this means that the data is not for that receiver. Hence, data transmission is halted. If the detected preamble has the same color code as the distance from relative localization, the process continues. When the detected preamble and color code from the distance from relative localization match, this means that the data is meant for the receiver and has been successfully transmitted to the correct receiver.

After the preamble is successfully detected, the transmitted data is decoded to recover the original data. In this stage, data verification is performed again to verify that the data is correctly transmitted to the receiver. This verification can be conducted by comparing the distance information decoded from the data with the distance measured from the relative localization, where both should be the same value. If the data decoder is passed, then the decoded data is checked for any missing bits during the transmission. Simultaneously, the decoded data is sorted based on the SN to create a complete big packet. Finally, the decoded data stream can be recovered and passed to further processes in the receiver.

A. BIDIRECTIONAL OCC USING DCC-OOK

Existing OCC systems widely use modulation schemes that utilize the rolling-shutter effect. By utilizing the rolling shutter effect of a camera, it is proven that it can increase the data rate compared to other modulation schemes. Some examples of modulation schemes that utilize the rolling-shutter effect are Rolling OFDM [43], DCO-OFDM [44], 2D-OFDM [45], and FSK [46]. The rolling-shutter effect of a camera generates a striped-pattern monochrome image where every white or black stripe represents the bit "1" or bit "0". However, this scheme has several disadvantages, such as:

1) It is difficult to detect the stripe correctly since it correlates with the transmitter power and the distance between the receiver and transmitter.
2) Short-distance communication when using a wide field of view camera.
3) Need to use a narrow field of view camera when communicating over long distances to enable focus only on the transmitter.

Based on the disadvantages of rolling-shutter effect-based modulation schemes, it is difficult to use them in OCC

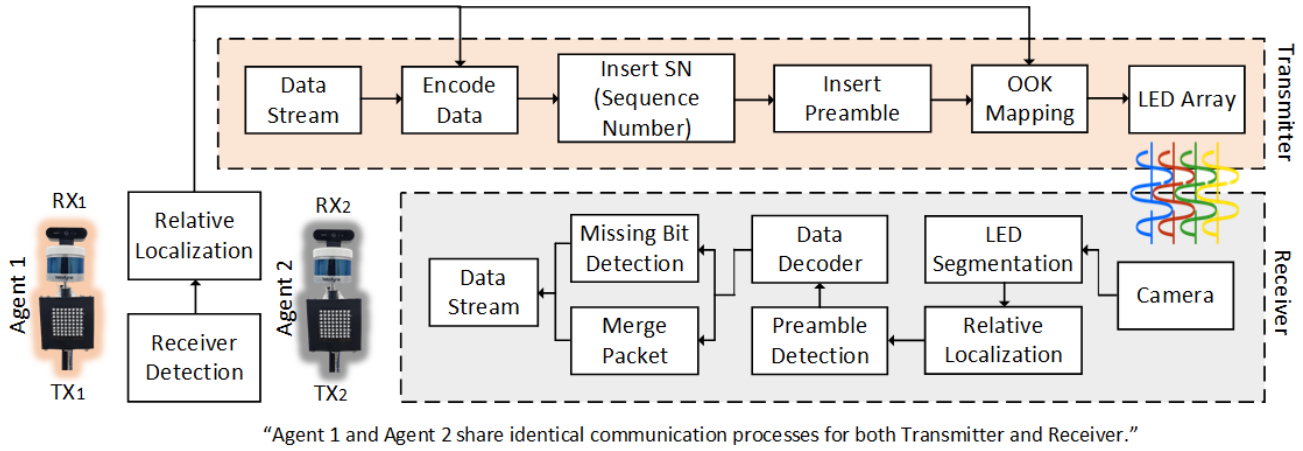


FIGURE 6. Block diagram of the proposed DCC-OOK modulation for OCC.

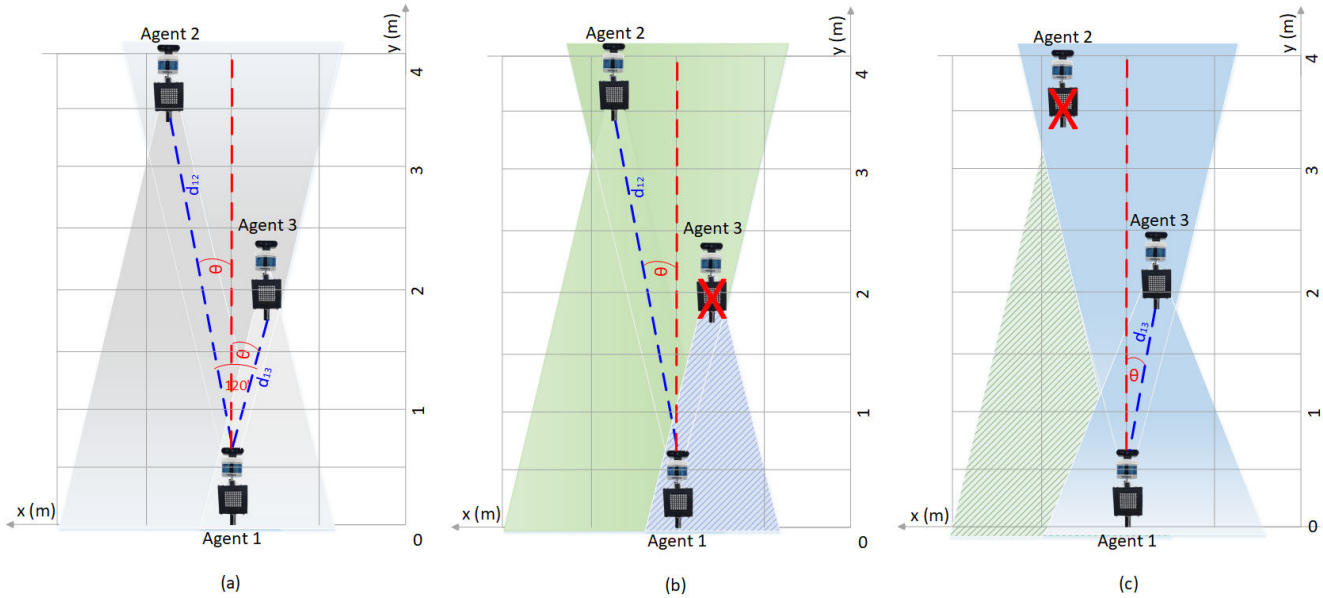


FIGURE 7. Scenario of bidirectional OCC using the DCC-OOK.

systems for bidirectional communication. In long-distance bidirectional communication, the receiver is only able to focus on one transmitter to retrieve the striped-pattern image, which means the receiver cannot see another transmitter. Additionally, when using these existing systems, the transmitter only broadcasts data by controlling the LED. Accordingly, the receiver does not know whether the transmitted data is for another receiver. Moreover, the receiver cannot acknowledge the transmitter because it cannot be identified due to the data only being broadcast. This means the OCC systems cannot perform bidirectional communication and multi-agent communication since the receiver is only tied to one transmitter, and the transmitter broadcasts the information without knowing the receiver. Therefore, we propose DCC-OOK as a solution for bidirectional communication in OCC. Using DCC-OOK, the LED array codes the relative distance between the transmitter and receiver as a color, allowing for their identification.

The transmitter sends data to the receiver by illuminating an LED array with a color based on the measured distance to the receiver. The receiver will only receive data from a transmitter that illuminates the LED array with a color that matches the distance between the receiver and transmitter. Hence, DCC-OOK enables a virtual link between the transmitter and receiver, and bidirectional communication can be established.

Figure 7. explains the steps to establish bidirectional communication in an OCC system using DCC-OOK. First, as demonstrated in Figure 7a, each systems agent uses its perception system of a camera and LiDAR to scan the

surrounding area and measure the relative distance to the surrounding agents. Hence, each agent has a list of the surrounding agents along with the relative distances to each agent. In this initial step, the LED array will be white, meaning "initial setup, not ready for communication."

As depicted in Figure 7b, Agent 1 needs to communicate with Agent 2, which is located around 3 m away relative to Agent 1. Therefore, Agent 1 transmits data with a preamble that is green based on the color code. Agent 3 receives the transmitted data from Agent 1. However, since the distance between Agent 3 and Agent 1 is less than 2 m, Agent 3 instantly realizes that the green preamble data is not for them. Hence, the data is instantly discarded by the Agent 3.

Meanwhile, Agent 2 receives the transmitted data from Agent 1 and knows that the distance to Agent 1 is approximately 3 m, which is coded green. Hence, the color code is a match, and Agent 2 starts to decode the data. If Agent 2 needs to acknowledge Agent 1, it simply transmits data in which the preamble will be green. Since Agent 2 is also green in its data preamble, Agent 1 instantly realizes that Agent 2 needs to communicate. Hence, Agent 3 and Agent 1 can establish bidirectional communication provided both transmit green preambles according to the color code.

A similar situation is explained in Figure 7c, where Agent 1 needs to transmit data to Agent 3. Based on the relative localization, the distance is less than 2 m and the preamble should be blue. Hence, Agent 1 displays blue on its LED array. Agent 3 retrieves the blue preamble and realizes that the data is for them. Agent 3 then decodes the data for further processing.

When establishing bidirectional communication between Agents 1 and 3, both Agents have the same blue preamble. When both Agents display blue, this means that both need to exchange data and communicate with each other. As such, based on Figure 6, if Agent 1 needs to communicate with Agent 2, it should transmit data with a green preamble. Meanwhile, if Agent 1 needs to transmit data to Agent 3, it should display a blue preamble. If the agents need to communicate bidirectionally, both agents should use the same preamble color.

V. EXPERIMENT RESULTS AND ANALYSIS

A. EXPERIMENT SETUP

To perform the experiment, two sets of agent setups were installed. The component configurations in each agent were set identically to ensure that all agents had the same characteristics. The components used in each agent are described in Table 1.

As displayed in Figure 6, the sensors for OCC communication were installed on top of a tripod. The camera was mounted on the top of the LiDAR and acted as the receiver for the OCC and for recognizing the other setup. Then, an 8×8 LED array was located on the bottom of the LiDAR as the transmitter for OCC. The LED array was controlled by an Arduino UNO board connected to an NVIDIA Jetson AGX Orin board. The camera, LiDAR, and Arduino board were connected to the NVIDIA Jetson AGX Orin board as the main processor of the system. The field of view of the LiDAR was limited to 120Â° to match the field of view of the camera.

TABLE 1. Components installed in each agent.

Component Name	Type
LiDAR	Velodyne VLP-168
Camera	Logitech Brio
LED	8×8 LED array
Processor	NVIDIA Jetson AGX Orin
Humidity sensor	AM2301
Temperature sensor	AM2302
Vibration sensor	Crowtail sensor
LED controller	Arduino UNO

However, LiDAR can produce extensive point cloud data that could slow the computation time. Hence, a distance filter was applied to limit the maximum object distance to 8 m, meaning it filtered out the point cloud of an object that was located more than 8 m from the sensor. To coordinate the components to work together, a robot operating system (ROS) was employed. All the software required in the experiments was written in Python and configured as a ROS node. After configuring the hardware and software for each agent, both agents were positioned next to each other at a minimum distance of 1 m and a maximum of 4 m, as displayed in Figure 5. Evaluation of OCC performance started after both agents were in the correct position.

B. CALIBRATION RESULTS

The approximate estimated parameters of the translation and rotation vectors generated by RANSAC were refined using the Levenberg-Marquardt model to reduce the errors. Based on the experiment, the errors were measured using root mean square error (RMSE). The best RMSE achieved was 2.354 with parameter values of 0.208, -0.001, -0.251, 0.941, -0.702, and 0.534 for t_x, t_y, t_z, r_x, r_y, r_z, respectively.

Based on the achieved best parameter values, these were used when projecting the 3-D data to the 2-D images, and the results are displayed in Figure 8. It is evident that the point cloud data were tightly matched with the checkerboard because all the point cloud data were correctly positioned inside the checkerboard.

C. SEGMENTATION RESULTS

To train the YOLOv8 for instance segmentation, the collected LED array dataset was split into 75% and 25% for training and validation, respectively. The dataset was also split proportionally for indoor and outdoor images to create balanced information for the model to learn. Additionally, the labels were normalized following the YOLOv8 model requirements. Pre-trained weights for the YOLOv8 model were used to speed up the training process, and these pre-trained weights were trained using the COCO dataset beforehand. Then, the pre-trained weights were loaded into

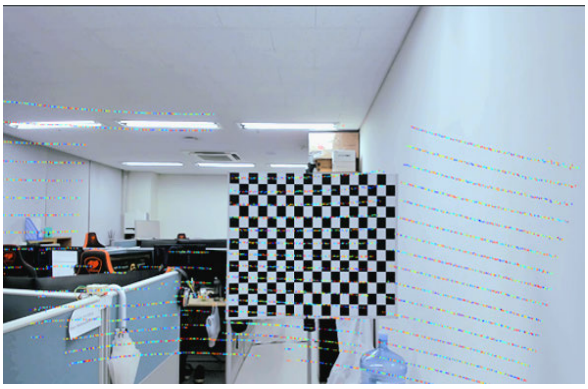


FIGURE 8. Extrinsic calibration results between the camera and LiDAR.

TABLE 2. Settings used for the YOLOv8 segmentation model training.

Hyperparameters	
Learning rate	0.01
Momentum	0.937
Weight decay	0.0005
Batch size	1
Epoch	100
Hardware	
CPU	Intel i7 12th gen
GPU	RTX 3060Ti
RAM	16GB
Software	
Framework	PyTorch and ONNX
OS	Ubuntu 20.04

TABLE 3. Training results of YOLOv8 segmentation for LED array.

mAP@0.50	0.91
Precision	0.95
Recall	0.86

the YOLOv8 model, and transfer learning was performed using the LED array dataset. In the training process, the model was trained using hyperparameters, hardware, and software, as displayed in Table 2.

From the training process, the results of the trained YOLOv8 model with its training performance are listed in Table 3. At 0.50 IoU, the model achieved up to 0.91 mAP, which indicated that the model successfully predicted most of the mask precisely. Moreover, the precision and recall performances were also stable, achieving values of 0.95 and 0.86, respectively. Overall, the training results exhibited a promising performance in terms of the YOLOv8 model and demonstrated that the model was ready for real-time inference.

The trained weight of the YOLOv8 model was used for the real-time inference. From the test results, it was evident that the processing time was 0.1 ms for one frame, including generating the mask location in the image. The experiment tested the model performance when detecting the LED array from distances of 1 to 4 m. From the experimental results, the model was able to detect the LED array seamlessly within

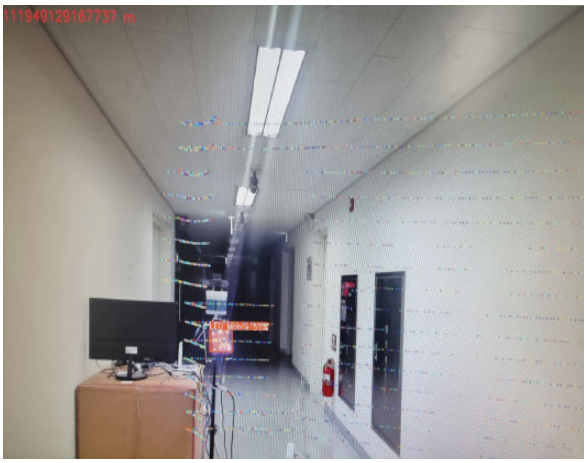


FIGURE 9. LED array detection results from the YOLOv8 model.

TABLE 4. Relative localization results.

Actual Distance (m)	Estimation Results (m)	Error (%)
1.0	1.0354	3.54
1.5	1.544	2.93
2.0	2.1904	9.52
2.5	2.7295	9.18
3.0	3.2850	9.5
3.5	3.5591	1.68
4.0	4.1530	3.82

distances of 1 to 3 m. However, starting from 3 to 4 m, the model had difficulty detecting the location of the LED array. With the camera, the LED became small and the model had difficulties in recognizing the LED array. However, although difficult, the model was still able to detect the LED array successfully with a lower confidence level. The detection results from the experiment are displayed in Figure 9.

D. RELATIVE DISTANCE LOCALIZATION RESULTS

The performance of relative distance localization was tested by performing eight measurement tests at different distances. The test was performed with a distance from 1 to 4 m in increments of 0.5 m. At each distance, the agent was tasked with localizing another agent that was located in front. The optimum calibration settings were applied, and the measurement results are listed in Table 4.

As demonstrated in Table 4, the performance of the proposed system had an error range of 1.68% to 9.52% with measuring distances from 1 to 4 m. Based on the implementation, the system experienced difficulties when detecting the LED array as the size was quite small, meaning the point cloud data inside the masking points were too few. Errors were also caused by calibration problems. Sometimes, the point cloud in the object mask was not the actual point cloud of the object, resulting in distance measurement errors.

Figure 10 displays the implementation of relative localization between the OCC systems. It is evident that the segmentation was still able to detect and segment the LED

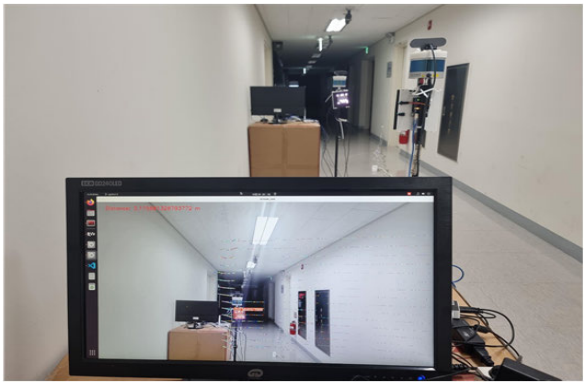


FIGURE 10. Relative localization implementation.

TABLE 5. Transmitter and receiver parameters.

Transmitter Side	
Preamble bit	4 bits
SN bit	4 bits
Anchor bit	52 bits
Anchor bit	4 bits
LED type	5 V
Receiver Side	
Camera	global shutter camera
Camera frame rate	15 fps
Capture resolution	1920 × 1080

array even though the LED array was small due to the distance. Moreover, the localization results were satisfactory considering the LED array size, indicating that calibration was successful.

E. BIDIRECTIONAL OCC USING DCC-OOK RESULTS

Table 5 lists the parameters used for creating the transmitter and receiver. The same settings were employed for all agents in the system. On the transmitter side, each data packet consisted of 64 bits according to the size of the employed LED array, 4 bits were reserved for the anchor to determine the corner of the data, and a further 4 bits were reserved for the preamble and SN bits. Then, most of the packet was filled with the payload, which uses 52 bits. The payload consisted of distance information resulting from relative localization and data from the sensor reading.

To demonstrate the feasibility of the OCC system for bidirectional communication, an experiment was conducted, as depicted in Figure 11. The implementation was successfully performed using an 8 × 8 LED array as the transmitter and camera as the receiver. As illustrated in Figure 11, the bidirectional OCC simultaneously collected data from both systems. Four types of sensors (as described in Table 1) were employed to transmit data simultaneously to each receiver, and the readings from the communication partner were displayed on individual monitors corresponding to the decoded results from each transmitter.

Based on the parameters presented in Table 5, each OCC system needed to transmit 64 bits of data per frame, of which

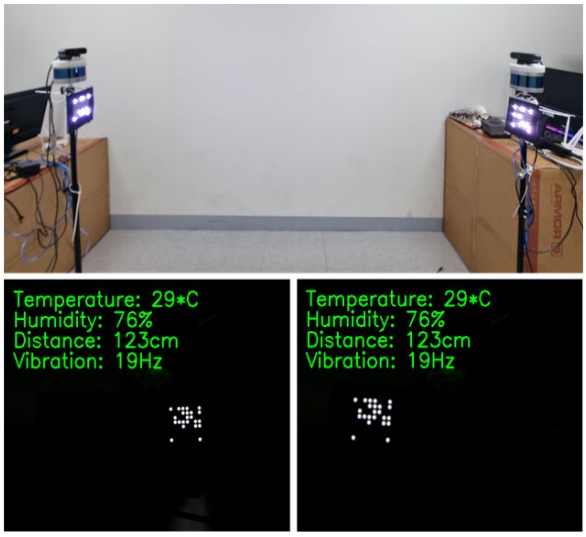


FIGURE 11. Real-time sensor data collection using bidirectional OCC.

TABLE 6. Data rate and BER performance analysis in DCC-OOK OCC systems.

Color and Wavelength	Distance	Data Rate	BER
Blue, 460–490	1 m	38.4 kbps	0.03
	1.5 m	19.2 kbps	0.283
	2 m	9.6 kbps	0.348
Green, 520-550	2.5 m	4.8 kbps	0.68
	3 m	2.4 kbps	0.742
	3.5 m	1.2 kbps	0.909
Yellow, 610-620	≥ 4 m	0.0 kbps	1

the payload data was only 52 bits. We used an image resolution of 1920 × 1080 pixels with a global sensor type operating at 15 fps to capture the LED matrix's flashing results. On the receiver side, the image was captured at a resolution of 1920 × 1080. The overall system fps was only 15 Hz because LiDAR can be maximally operated at 15 fps. Hence, the fps of the camera should match the LiDAR's fps to prevent object misdetection.

As displayed in Table 6, the performance of the proposed system exhibited a data rate range between 0.0 and 38.4 kbps, with a BER ranging from 1 to 0.03. The BER measurements were conducted by focusing on the color and distance characteristics in bidirectional OCC communication. High data rates and low BER levels could be achieved with optimal placement of the transmitter and receiver in the communication system, which reduced signal attenuation and distortion when transmitting light through the medium.

Furthermore, precise detection of the LED array by the receiver resulted in clear images, allowing each bit sent from the transceiver to be accurately read by the receiver. However, as the communication distance increased, there was a significant decrease in data rate and BER due to signal attenuation with increasing distance and the difficulty in distinguishing individual bits from the LED array.

TABLE 7. Comparison of proposed method with existing modulation schemes.

	Teli et al [25]	Huang et al [26]	Nguyen et al [27]	Proposed Method
Modulation	OOK	Color intensity	OOK	DCC-OOK
Camera type	Rolling shutter	Global shutter	Both types	Both types
Distance	Short (1.4 m)	Short (1.4 m)	Long (20 m)	Medium (3.5 m)
point-to-point transmission	Not mentioned	Not mentioned	Not mentioned	Yes
Bidirectional communication	Not mentioned	Not mentioned	Not mentioned	Yes
Max. data rate (kbps)	13.44	126.72	7.680	38.4
BER	Not mentioned	$> 10^{-2}$	$< 10^{-4}$	$> 10^{-2}$

Consequently, the camera used was no longer capable of detecting every bit available on the LED array. Moreover, at longer distances, OCC could not work due to the LED array not being clearly visible, causing a loss of data. Although the data rate and BER performance of DCC-OOK modulation were not significantly higher than other existing modulation schemes, DCC-OOK allowed for easier bidirectional communication in the OCC system.

Table 7 shows the comparison of the proposed method with the existing method. Since the proposed method also utilizes LED arrays as transmitters, several existing modulations with LED arrays as transmitters are chosen to provide a more fair comparison. The table clearly shows that the proposed DCC-OOK modulation is the only one equipped with point-to-point transmission and bidirectional communication. Then, the proposed method is able to use any camera type, not limited to certain camera types.

The communication distance of the proposed method is in the medium range, better than some schemes that only support up to 1.4 m. Moreover, the data rate of the proposed OCC is worse than the data rate of color-intensity-based modulation. However, although the data rate is the fastest, the color-intensity modulation can only perform short-distance communication. Finally, the BER of the proposed method is worse than the work by Nguyen et al. The high BER value indicates significant data loss during transmission, which is caused by blurry images resulting from the tight positioning of the LEDs in the array, making it difficult for the camera to differentiate each LED. A possible solution is to utilize a better LED array with wider-spaced LEDs and a higher camera resolution.

VI. CONCLUSION

In this paper, a new modulation scheme termed DCC-OOK was proposed to solve the current issues in OCC of the transmitter only broadcasting data and sending it to a specific receiver is not possible. By using the proposed method, the transmitter knew the receiver's location and the receiver also knew which transmitter had sent the data.

To establish DCC-OOK, a relative localization scheme using multi-sensor fusion of a camera and LiDAR was employed to precisely measure the distances between the OCC systems. Sensor extrinsic calibration using RANSAC and Levenberg-Marquardt was used to project the 3-D data into the 2-D images.

A YOLOv8-based segmentation model was also employed to locate the LED array of the transmitter in the images. The distance between the OCC systems was acquired by calculating the point cloud inside the object mask. To establish the DCC-OOK, the distance was inserted to encode the data and modify the data preamble to ensure that the LED arrays illuminated in different colors based on the distance.

From the implementation results, the proposed DCC-OOK system demonstrated that bidirectional communication between OCC systems was possible by the transmitter only communicating with a receiver that had the same color code. The implemented bidirectional communication can achieve data rate up to 38.4 kbps and BER of 0.03.

The positioning information between OCC systems was an important aspect of correctly transmitting data to the target receiver. Using DCC-OOK was also beneficial because the transmitter was able to select the data receiver, which was automatically aware due to the color code emitted by the transmitter. In future studies, the relative localization performance will be improved and additional OCC systems will be employed to implement multi-agent communication in OCC.

REFERENCES

[1] M. K. Hasan, M. O. Ali, M. H. Rahman, M. Z. Chowdhury, and Y. M. Jang, "Optical camera communication in vehicular applications: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 6260–6281, Jul. 2022.

[2] Z. Zeng, S. Fu, H. Zhang, Y. Dong, and J. Cheng, "A survey of underwater optical wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 204–238, 1st Quart., 2017.

[3] W. Liu, J. Ding, J. Zheng, X. Chen, and I. Chih-Lin, "Relay-assisted technology in optical wireless communications: A survey," *IEEE Access*, vol. 8, pp. 194384–194409, 2020.

[4] T. Nguyen, A. Islam, and Y. M. Jang, "Region-of-interest signaling vehicular system using optical camera communications," *IEEE Photon. J.*, vol. 9, no. 1, pp. 1–20, Feb. 2017.

[5] K. Dong, X. Ke, and H. Li, "Camera-based channel modeling and symbol error rate analysis of CSK modulation for outdoor OCC systems," *IEEE Access*, vol. 10, pp. 50254–50264, 2022.

[6] F. He, Y. Pan, Q. Lin, X. Miao, and Z. Chen, "Collective intelligence: A taxonomy and survey," *IEEE Access*, vol. 7, pp. 170213–170225, 2019.

[7] T. L. Pham, M. Shahjalal, V. Bui, and Y. M. Jang, "Deep learning for optical vehicular communication," *IEEE Access*, vol. 8, pp. 102691–102708, 2020.

[8] X. Zhang, G. Klevering, and L. Xiao, "PoseFly: On-site pose parsing of swarming drones via 4-in-1 optical camera communication," in *Proc. IEEE 24th Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, Jun. 2023, pp. 67–76.

[9] A. Pansare, N. Sabu, H. Kushwaha, V. Srivastava, N. Thakur, K. Jamgaonkar, and M. Z. Faiz, "Drone detection using YOLO and SSD a comparative study," in *Proc. Int. Conf. Signal Inf. Process. (IConSIP)*, Aug. 2022, pp. 1–6.

[10] J. A. Ricardo, L. Giacomossi, J. F. S. Trentin, J. F. B. Brancalion, M. R. O. A. Maximo, and D. A. Santos, "Cooperative threat engagement using drone swarms," *IEEE Access*, vol. 11, pp. 9529–9546, 2023.

[11] T. Liu, X. Sun, C. Hu, Q. Fu, and S. Yue, "A multiple pheromone communication system for swarm intelligence," *IEEE Access*, vol. 9, pp. 148721–148737, 2021.

[12] X. Zhang, G. Klevering, K. Wijewardena, and L. Xiao, "Demo: Integrated on-site localization and optical camera communication for drones," in *Proc. IEEE 24th Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, Jun. 2023, pp. 334–336.

[13] N. Q. Pham, K. Mekonnen, E. Tangdiongga, A. Mefleh, and T. Koonen, "Efficient mobility management for indoor optical wireless communication system," *IEEE Photon. Technol. Lett.*, vol. 33, no. 17, pp. 939–942, Sep. 1, 2021.

[14] H. Takano, M. Nakahara, K. Suzuoki, Y. Nakayama, and D. Hisano, "300-meter long-range optical camera communication on RGB-LED-equipped drone and object-detecting camera," *IEEE Access*, vol. 10, pp. 55073–55080, 2022.

[15] H. Wu, Y.-C. Chen, G. Xue, Y. Jiang, M. Wang, S. Qian, J. Yu, and P.-Y. Chen, "OnionCode: Enabling multi-priority coding in LED-based optical camera communications," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, May 2022, pp. 260–269.

[16] O. S. Sitanggang, V. L. Nguyen, H. Nguyen, R. F. Pamungkas, M. M. Faridh, and Y. M. Jang, "Design and implementation of a 2D MIMO OCC system based on deep learning," *Sensors*, vol. 23, no. 17, p. 7637, Sep. 2023.

[17] T. Nguyen, N. T. Le, and Y. M. Jang, "Asynchronous scheme for unidirectional optical camera communications (OCC)," in *Proc. 6th Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Jul. 2014, pp. 48–51.

[18] C. Danakis, M. Afgani, G. Povey, I. Underwood, and H. Haas, "Using a CMOS camera sensor for visible light communication," in *Proc. IEEE GLOBECOM Workshops*, Dec. 2012, pp. 1244–1248.

[19] N. Rajagopal, P. Lazik, and A. Rowe, "Visual light landmarks for mobile devices," in *Proc. 13th Int. Symp. Inf. Process. Sensor Netw. (IPSN)*, Apr. 2014, pp. 249–260.

[20] W.-C. Wang, C.-W. Chow, L.-Y. Wei, Y. Liu, and C.-H. Yeh, "Long distance non-line-of-sight (NLOS) visible light signal detection based on rolling-shutter-patterning of mobile-phone camera," *Opt. Exp.*, vol. 25, no. 9, p. 10103, 2017.

[21] C.-W. Chow, R.-J. Shiu, Y.-C. Liu, Y. Liu, and C.-H. Yeh, "Non-flickering 100 m RGB visible light communication transmission based on a CMOS image sensor," *Opt. Exp.*, vol. 26, no. 6, p. 7079, 2018.

[22] P. Chavez-Burbano, J. Rabadan, V. Guerra, and R. Perez-Jimenez, "Flickering-free distance-independent modulation scheme for OCC," *Electronics*, vol. 10, no. 9, p. 1103, May 2021.

[23] R. D. Roberts, "Undersampled frequency shift ON-OFF keying (UFSOOK) for camera communications (CamCom)," in *Proc. 22nd Wireless Opt. Commun. Conf.*, May 2013, pp. 645–648.

[24] P. Luo, M. Zhang, Z. Ghassemlooy, H. Le Minh, H.-M. Tsai, X. Tang, and D. Han, "Experimental demonstration of a 1024-QAM optical camera communication system," *IEEE Photon. Technol. Lett.*, vol. 28, no. 2, pp. 139–142, Jan. 15, 2016.

[25] S. R. Teli, V. Matus, S. Zvanovec, R. Perez-Jimenez, S. Vitek, and Z. Ghassemlooy, "Optical camera communications for IoT-rolling-shutter based MIMO scheme with grouped LED array transmitter," *Sensors*, vol. 20, no. 12, p. 3361, Jun. 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/12/3361

[26] W. Huang, P. Tian, and Z. Xu, "Design and implementation of a real-time CIM-MIMO optical camera communication system," *Opt. Exp.*, vol. 24, no. 21, pp. 24567–24579, Oct. 2016. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-24-21-24567

[27] H. Nguyen, V. Nguyen, C. Nguyen, V. Bui, and Y. Jang, "Design and implementation of 2D MIMO-based optical camera communication using a light-emitting diode array for long-range monitoring system," *Sensors*, vol. 21, no. 9, p. 3023, Apr. 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/9/3023

[28] S. Chen, D. Yin, and Y. Niu, "A survey of robot swarms' relative localization method," *Sensors*, vol. 22, no. 12, p. 4424, Jun. 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/12/4424

[29] J. Dong, Q. Chen, D. Qu, H. Lu, A. Ganlath, Q. Yang, S. Chen, and S. Labi, "LiDAR-based cooperative relative localization," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2023, pp. 1–8.

[30] V. Pritzl, M. Vrba, P. Štěpán, and M. Saska, "Fusion of visual-inertial odometry with LiDAR relative localization for cooperative guidance of a micro-scale aerial vehicle," 2023, *arXiv:2306.17544*.

[31] V. Pritzl, M. Vrba, P. Štepán, and M. Saska, "Cooperative navigation and guidance of a micro-scale aerial vehicle by an accompanying UAV using 3D LiDAR relative localization," in *Proc. Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, Jun. 2022, pp. 526–535.

[32] R. B. Rusu, N. Blodow, and M. Beetz, "Fast point feature histograms (FPFH) for 3D registration," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2009, pp. 3212–3217.

[33] S. Salti, F. Tombari, and L. Di Stefano, "SHOT: Unique signatures of histograms for surface and texture description," *Comput. Vis. Image Understand.*, vol. 125, pp. 251–264, Aug. 2014.

[34] A. S. Mian, M. Bennamoun, and R. A. Owens, "A novel representation and feature matching algorithm for automatic pairwise registration of range images," *Int. J. Comput. Vis.*, vol. 66, no. 1, pp. 19–40, Jan. 2006, doi: 10.1007/s11263-005-3221-0.

[35] H. Deng, T. Birdal, and S. Ilic, "3D local features for direct pairwise registration," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 3239–3248.

[36] W. Lu, G. Wan, Y. Zhou, X. Fu, P. Yuan, and S. Song, "DeepVCP: An end-to-end deep neural network for point cloud registration," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 12–21.

[37] Z. J. Yew and G. H. Lee, "3DFeat-Net: Weakly supervised local 3D features for point cloud registration," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds. Cham, Switzerland: Springer, 2018, pp. 630–646.

[38] E. Arnold, S. Mozaffari, and M. Dianati, "Fast and robust registration of partially overlapping point clouds," *IEEE Robot. Autom. Lett.*, vol. 7, no. 2, pp. 1502–1509, Apr. 2022.

[39] A. Kirillov, K. He, R. Girshick, C. Rother, and P. Dollár, "Panoptic segmentation," 2018, *arXiv:1801.00868*.

[40] A. Dumitriu, F. Tatui, F. Miron, R. T. Ionescu, and R. Timofte, "Rip current segmentation: A novel benchmark and YOLOv8 baseline results," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2023, pp. 1261–1271.

[41] B. Sekachev et al., Aug. 2020, "opencv/cvat: v1.1.0," *Zenodo*, doi: 10.5281/zenodo.4009388.

[42] G. B. Thomas and R. L. Finney, *Calculus and Analytic Geometry*, 9th ed. London, U.K.: Pearson, 2010. [Online]. Available: https://api.semanticscholar.org/CorpusID:238556465

[43] H. Nguyen, M. D. Thieu, T. Nguyen, and Y. M. Jang, "Rolling OFDM for image sensor based optical wireless communication," *IEEE Photon. J.*, vol. 11, no. 4, pp. 1–17, Aug. 2019.

[44] J.-K. Lain, Z.-D. Yang, and T.-W. Xu, "Experimental DCO-OFDM optical camera communication systems with a commercial smartphone camera," *IEEE Photon. J.*, vol. 11, no. 6, pp. 1–13, Dec. 2019.

[45] T. Nguyen, M. D. Thieu, and Y. M. Jang, "2D-OFDM for optical camera communication: Principle and implementation," *IEEE Access*, vol. 7, pp. 29405–29424, 2019.

[46] N. Van Thang, N. T. Le, T. L. Vu, M. D. Thieu, and Y. M. Jang, "An implementation of binary frequency shift on-off keying modulation for optical camera communication," in *Proc. 10th Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Jul. 2018, pp. 121–125.

IDA BAGUS KRISHNA YOGA UTAMA (Graduate Student Member, IEEE) received the B.Eng. degree from the Department of Electrical Engineering, Universitas Indonesia, Depok, Indonesia, and the M.Sc. degree from the Department of Electronics Engineering, Kookmin University, Seoul, South Korea. He is currently pursuing the Ph.D. degree with the Wireless Communication and Artificial Intelligence Laboratory, Kookmin University. His current research interests include pattern recognition, optical wireless communication, sensor fusion, the Intelligent Internet of Things, and machine learning.

ONES SANJERICO SITANGGANG received the B.Eng. degree from the Department of Electrical Engineering, Universitas Indonesia, Depok, Indonesia, in 2018. He is currently pursuing the M.Sc. degree with the Wireless Communication and Artificial Intelligence Laboratory, Kookmin University, Seoul, South Korea. He contributed to the International Telecommunication Union (ITU), Telecommunication Standardization Sector, specifically in topic Q6/5, focusing on the procedure of optimization platform framework for cloud computing. His contribution also addresses areas, such as energy-efficient cloud computing, energy monitoring, and resource allocation. His research interests include optical wireless communication, optical camera communication (OCC), the Intelligent Internet of Things, and machine learning.

MUHAMMAD RANGGA AZIZ NASUTION received the B.Comp.Sc. degree from the Department of Informatics, Universitas Amikom Yogyakarta, Yogyakarta, Indonesia. He is currently pursuing the M.Sc. degree with the Wireless Communication and Artificial Intelligence Laboratory, Kookmin University, Seoul, South Korea. His research interests include optical camera communication, machine learning, and localization.

MD. IBNE JOHA received the B.Sc. degree from the Department of Electrical and Electronic Engineering, Khulna University of Engineering and Technology, Khulna, Bangladesh, in 2023. He is currently pursuing the M.Sc. degree with the Department of Electronics Engineering, Kookmin University, Seoul, South Korea. His current research interests include the Internet of Things (IoT), deep learning, and energy.

JAEJUN YOO received the M.S. and Ph.D. degrees from the School of Computing, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea. He is currently a Principal Researcher with the Electronics and Telecommunications Research Institute, Daejeon. His current research interests include positioning, databases, information retrieval, geographic information systems, intelligent transportation systems, location-based services, and machine learning.

YEONG MIN JANG (Member, IEEE) received the B.E. and M.E. degrees in electronics engineering from Kyungpook National University, South Korea, in 1985 and 1987, respectively, and the Ph.D. degree in computer science from the University of Massachusetts, USA, in 1999. From 1987 to 2000, he was with the Electronics and Telecommunications Research Institute, South Korea. Since 2002, he has been with the School of Electrical Engineering, Kookmin University, Seoul, South Korea, where he was the Director of the Ubiquitous IT Convergence Center from 2005 to 2010, the Director of the AI Mobility Research Institute from 2021 to 2023, and the Director of the Internet of Energy Research Center from 2018 to 2023. Since 2010, he has been the Director of the LED Convergence Research Center, Kookmin University. He has coauthored more than 500 technical papers and holds more than 140 patents. His research interests include the Internet of Energy, IoT platform, AI platform, cloud platform, sensor fusion, optical camera communication, 5G/6G mobile communications, and Internet of Things. Dr. Jang is a Fellow of the Korean Institute of Communications and Information Sciences (KICS), South Korea. From 2006 to 2014, he was an Executive Director with KICS. In 2019, he was the President of KICS. He was the recipient of the Young Scientist Award from the Korean Government from 2003 to 2006 and KICS Dr. Irwin Jacobs Award in 2018. From 2007 and 2008, he was also the Founding Chair of the KICS Technical Committee on Communication Networks. he was/has been the Organizing Chair of the MultiScreen Service Forum from 2011 to 2019, Society Safety System Forum from 2015 to 2021, and the ESG Convergence Forum, since 2022. He was also the Chair of the IEEE 802.15 Optical Camera Communications Study Group in 2014 and the IEEE 802.15.7m Optical Wireless Communications TG from 2015 to 2019, and successfully published IEEE 802.15.7-2018 and ISO 22738:2020 standard. He has been the Chair of IEEE 802.15.7a Higher Rate and Longer Range OCC TG since 2020 and the Chair of IEEE 802.15 IG Next Generation OCC since 2024. He has organized several conferences and workshops, such as the International Conference on Ubiquitous and Future Networks from 2009 to 2017, International Conference on ICT Convergence from 2010 to 2016, IEEE OWC Workshops in 2015 and 2022, International Conference on Information Networking in 2015, and International Conference on Artificial Intelligence in Information and Communication since 2019. He is also the Editor-in-Chief of ICT Express (indexed by SCIE).