

Received April 20, 2020, accepted May 25, 2020, date of publication June 1, 2020, date of current version June 11, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2998944

# Deep Learning for Optical Vehicular Communication

TUNG LAM PHAM<sup>1</sup>, MD. SHAHJALAL<sup>1</sup>, (Student Member, IEEE),  
VAN BUI<sup>1</sup>, AND YEONG MIN JANG<sup>1</sup>, (Member, IEEE)

Department of Electronics Engineering, Kookmin University, Seoul 02707, South Korea

Corresponding author: Yeong Min Jang (yjang@kookmin.ac.kr)

This work was supported in part by the Institute for Information and Communications Technology Promotion (IITP) Grant through the Korea Government (MSIT) under Grant 2017-0-00824, and in part by the Development of Intelligent and Hybrid OCC-LiFi Systems for Next Generation Optical Wireless Communications.

**ABSTRACT** In this paper, we consider the current status and technical issues involved in the use of optical camera communication (OCC)/visible light communication (VLC) technologies in vehicular communication systems. Hybrid spatial phase-shift keying was introduced in IEEE 802.15.7-2018 as the standard hybrid modulation scheme for vehicular OCC/VLC systems. We herein propose a functional communication system architecture for vehicular systems based on this hybrid waveform, and we also present state-of-the-art research work on an artificial intelligence (AI)-based vehicular OCC system. Every AI module within the proposed system architecture is discussed in detail. Finally, our experimental procedures and results are analyzed to evaluate the performance of the proposed system over a complex channel model in a vehicular environment. We effectively employed the popular You Only Look Once version 2 object detection algorithm for real-time region-of-interest tracking in city driving (at a vehicular velocity of around 30 km/h and highway night driving (at a vehicular velocity of > 60 km/h) scenarios. Moreover, our novel neural-network-based decoder and AI-based error correction proved effective in improving the data decoding accuracy, resulting in a best-case reduction of 2.2 and 9.0 dB, respectively, in the signal-to-noise ratio needed to achieve the desired bit error rate of  $10^{-4}$  in a vehicular OCC/VLC system.

**INDEX TERMS** Visible light communication (VLC), light-emitting diode (LED), image sensor (IS), deep learning, optical camera communication (OCC).

## I. INTRODUCTION

Intelligent transportation systems (ITSs) are currently a promising area of research, whose key features include autonomous vehicle management, traffic efficiency, road safety, and inter-vehicle and inter-passenger communication [1]. Due to the rapid development of mobile communication systems, radiofrequency (RF) spectrum (30-75 MHz) is not sufficient to support the requirements for high throughput [2]. Moreover, along with the density of vehicles, communications through tunnels and subway using RF based GPS signals are becoming a challenging issues for ITS. In this regard, visible light communication (VLC) can be used as a complementary, hybrid, or heterogeneous system alongside RF. VLC can provide vehicle-to-vehicle (V2V) and infrastructure-to-vehicle or vehicle-to-infrastructure communications using the existing lighting infrastructures inside tunnel or subway [3], [4]. This technology can support

high-throughput data transmission due to the broad spectrum of visible light (380-780 nm) and has great potential for development because of its low cost, widespread availability, and use of an unregulated radiation spectrum. Another important advantage of VLC is that it is safe for human health and can be used in RF-restricted areas [3]–[5]. VLC plays an important role in ITSs because of the widespread use of visible lighting infrastructures, such as vehicular front and rear light-emitting diodes (LEDs), traffic LEDs, traffic signage, and lamp post LEDs. Not only are these lighting infrastructures cost-effective, but also consume less energy. In VLC systems, LEDs are configured to transmit signals via light emission. LED driver circuits are programmed to modulate LEDs with input data according to a specified modulation scheme. Vehicular front and rear lights can, thus, be used to transmit ITS data, such as safety information, to other vehicles.

Optical camera communication (OCC) is different from VLC because of various types of receivers being used. PD is used for VLC system, whereas OCC utilizes image sensor

The associate editor coordinating the review of this manuscript and approving it for publication was Edith C.-H. Ngai<sup>1</sup>.

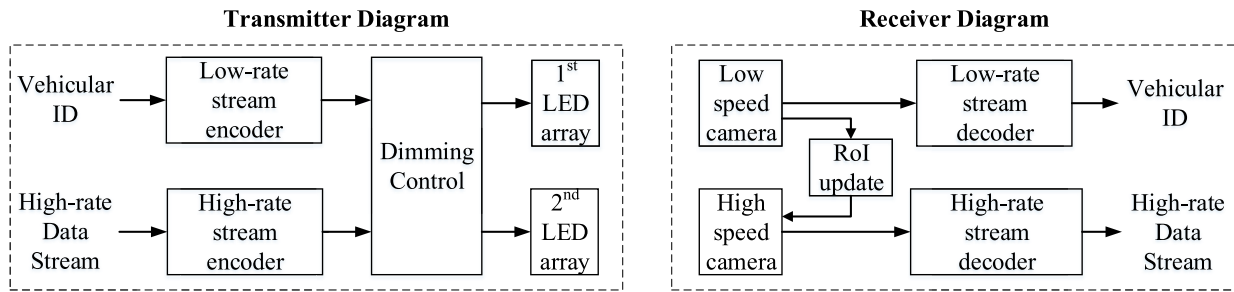


FIGURE 1. The reference vehicular OCC system architecture.

or camera as a receiver. Image sensor is beneficial compared to PD as it can perform spatial separation of light using its existing lens and can communicate with multiple sources. In such systems, a high-speed camera is used to receive lighting signals from both vehicle lights and street lighting infrastructures. To receive signals from either the front or the rear, cameras can be installed at both ends of the vehicle. Cameras or image sensor receiver systems can perform a number of tasks, including object tracking, data reception, distance measurement from pedestrians, and reading traffic signal statuses and LED signage. However, current low-speed cameras (with frame rates of up to 60 fps) limit the data rate to within the low bps to kbps range [6]–[8]. Undersampling-based modulation schemes have been proposed to achieve higher data rates in OCC systems, but they also limit the communication range [9], [10]. High-speed cameras can be installed to process data faster in high-mobility vehicles [11], [12].

#### A. STATE OF THE ART

In [12], an optical communication image sensor with DC-biased optical orthogonal frequency division multiplexing was employed in automotive VLC. This system achieved a data rate of 55 Mbps using an ultrahigh-speed camera, but it also incurred very high deployment costs. A novel image-sensor-based optical wireless communication system was proposed by Takai *et al.* [13], for which they claimed a data rate of 15 Mbps per pixel with 16.6 ms real-time LED detection for automotive applications. In [14], a novel VLC system that could communicate over a 130 m link between a traffic light and a vehicle was presented. The authors used a phototransistor for improved sensitivity and a transimpedance amplifier circuit to reduce noise. However, VLC-based automotive applications can be affected by parasitic light from sunlight and other outdoor light sources. These extraneous light sources can have a significant impact on the system's bit error rate (BER). Additionally, the phenomenon of blur, either motion blur or blurring due to rainy, foggy, or snowy weather conditions, as what typically occurs in optical vehicular communication systems, was analyzed in [15]. In such cases, the VLC or optical camera communication (OCC) receiver should perform region-of-interest (ROI) selection to

detect the target transmitter, in order to reduce the amount of incident parasitic light. A test bed was designed in [16] to evaluate infrastructure-to-vehicle VLC system performance over a 50 m link in the presence of bright sunlight. The test showed that bright sunlight affects the BER performance, increasing the error rate from  $10^{-7}$  to  $10^{-4}$ .

The recent related research works are summarized in Table 1. In our most recent paper, we provided a vehicular OCC system architecture [17], on which the developments mentioned in the next sections of this paper are based. Fig. 1 illustrates the architecture of our vehicular OCC system.

In the transmitter (Tx) of this system, vehicular headlights or tail lights can be used to transmit a hybrid low- and high-rate data stream. The technique of simultaneously sending two data streams in a single waveform is defined as an ROI signaling method [4], [17], [18]. Low-rate data streams often carry short data for vehicular identification, whereas high-rate data streams provide supporting information required to assist safe driving. The ROI signaling technique for OCC has recently been introduced as the PHY IV standard in IEEE 802.15.7-2018, updated by IEEE 802.15.7m OWC Task Group (TG). For this standard, Intel has proposed and implemented Twinkle VPPM and Kookmin University has introduced hybrid spatial phase-shift keying (HS-PSK) as a hybrid waveform for use in OCC systems.

For the receiver (Rx), Tx light source identification and high-rate data stream demodulation can be processed by a single camera, provided that it is reasonably time-slotted. The Rx camera first detects the ROI from among several natural and artificial light sources based on the low-rate identification signal, and then it focuses on the selected ROI for high-speed data decoding. This demands skillful programming of the camera control and image processing software. Additionally, in the real world, movement of vehicles is continuous and unpredictable, making a single-camera receiver ineffective and difficult to implement. An alternative approach is to use a dual-camera system, in which a low-frame-rate camera is used to continuously detect and track the Tx light source and a high-frame-rate camera focuses on decoding the data transmitted from the identified ROI at high speed [4], [18], [17].

TABLE 1. Related research work.

References	Used algorithm	Featuring metrics	Challenges
[1]	Proposed laser light based indirect diffused FSO for vehicular system. This system used reflected light to receive vehicular signal	This approach shows quite a high data rate communication over a long distance while the vehicle stays normal to the diffuse reflector	This system is quite vulnerable when car moves different directions
[3]	Made a survey of the VLC systems proposed for automotive applications.	Focused on current challenges and provided overview of up-to-date solutions.	The main challenge mentioned is rejecting the parasitic lights.
[11]	Provided channel modeling for moving vehicle of image sensor based VLC for automotive applications.	Described transmitter movement in the image plane and effect of illumination variation.	The readout processes for CMOS image sensors are a bottleneck
[12]	DC-biased optical-OFDM was employed in automotive VLC	55 Mbps data rate was achieved using ultrahigh-speed camera	Incurred very high deployment cost
[13]	Employed specialized communication pixel (CPx) and 1-bit flag image detection to obtain higher transmission rate in automotive OWC system.	Claimed a data rate of 15 Mbps per pixel with 16.6 ms real-time LED detection	LED blur effect during moving condition was not considered.
[14]	In a VLC system for vehicle, phototransistor and trans-impedance amplifier was used to improve sensitivity and reduce noise, respectively.	Achieved 130 m communication link between traffic light and vehicle	Need to employ noise mitigation technique and increase the data rate.
[15]	Multilayer Perceptron Neural Network-based Decoder for using in Vehicular OCC system	Blurred phenomenon which is model as the parameter for training the neural network-based decoder, results in better SER performance of data decoding compared with using traditional mathematics model for signal processing.	System complexity and overall systematic improvement have not been considered deeply.
[16]	Proposed an infrastructure-to-vehicle VLC system in the presence of sunlight.	Achieved 50 m communication link	Sunlight affects the error rate.
[17]	A standardized hybrid waveform for RoI-Signaling technology using in vehicular OCC system	The line-of-sight (LoS) link quality always guarantees the minimum pixel SNR requirement of 10 dB for the BER of $10^{-4}$ in vehicular OCC system. The RoI signaling technique could also deal with rotation problem, given that the number of deviation pixel rows between two LEDs must not exceed more than 34-pixel rows of the image sensor.	Different type of weather conditions (rainy, foggy, snowy, etc.) are not considered in performance analysis.
[19]	A vehicle to smartphone camera communication was tested using UFSOOK.	Achieved less BER and successful frame detection in a shorter communication range.	Error free communication distance is very low (< 0.5 m)

TABLE 1. (Continued.) Related research work.

References	Used algorithm	Featuring metrics	Challenges
[18]	Low rate S2-PSK scheme used in RoI signaling OCC system, with NN-based light detecting and XOR-decoding model are proposed	The system could work well with both global and rolling shutter camera, with no less than 20fps in frame rate. The NN-based light detecting and XOR-decoding model are also outperform other well-known method in term of accuracy and the complexity	The effect from various type of environments to the channel model in vehicular communication system are not put in concerned intensively
[20]	Proposed an intelligent transport system positioning scheme using two image sensor based VLC.	Measured the inter-vehicular distance considering low and high resolution camera	Distance error increases higher in case of low resolution camera.

Even though the space-time coding for high frame-rate Nyquist sampling optical system could be implemented in high speed camera and multi-LED array for optical MIMO concept [4], however, to the best of our knowledge, we have not found a research related to dimming support for those related systems. It may be acceptable in using traffic light at this time while the traffic light source is turned on and off at a time without any dimming requirement.

The other matter of oversampling technique is that the high speed camera used in Rx need to maintain the high frame rate even when detecting the Tx, which is computational cost-ineffective compare with using the dual camera in RoI-signaling OCC system. Due to those reason, HS-PSK, which is the hybrid waveform of S2-PSK for carrying the low-rate data stream and DS8-PSK for carrying the high-rate data stream is chosen as a suitable modulation scheme for implementing in vehicular-to-vehicular OCC system.

## B. DEEP LEARNING-BASED OPTICAL VEHICULAR COMMUNICATION: A POTENTIAL APPROACHING

Despite the enormous potential of OWC/OCC vehicular systems in the lucrative industrial market, there are several challenges that need to be overcome in the development of these technologies, as pointed out in [3]. One of the main causes of degradation in OWC/OCC vehicular systems is white noise from sources such as ambient light radiation from the Sun, or street lighting. Many studies have been performed to investigate the feasibility of OWC/OCC technologies in V2V/V2X systems (see, e.g., [18]–[20]), predominantly focusing on Gaussian white noise with different signal-to-noise ratios (SNRs). Another type of distortion that can reduce vehicular OWC/OCC system performance is blur, which can occur in any camera or image processing system. In an OWC/OCC vehicular scenario, vehicle vibration and mobility, weather conditions (e.g., rain, fog, and snow), and camera focusing issues are the main causes of blur in the received images. Recent developments in AI and deep-learning technologies have shed some light on

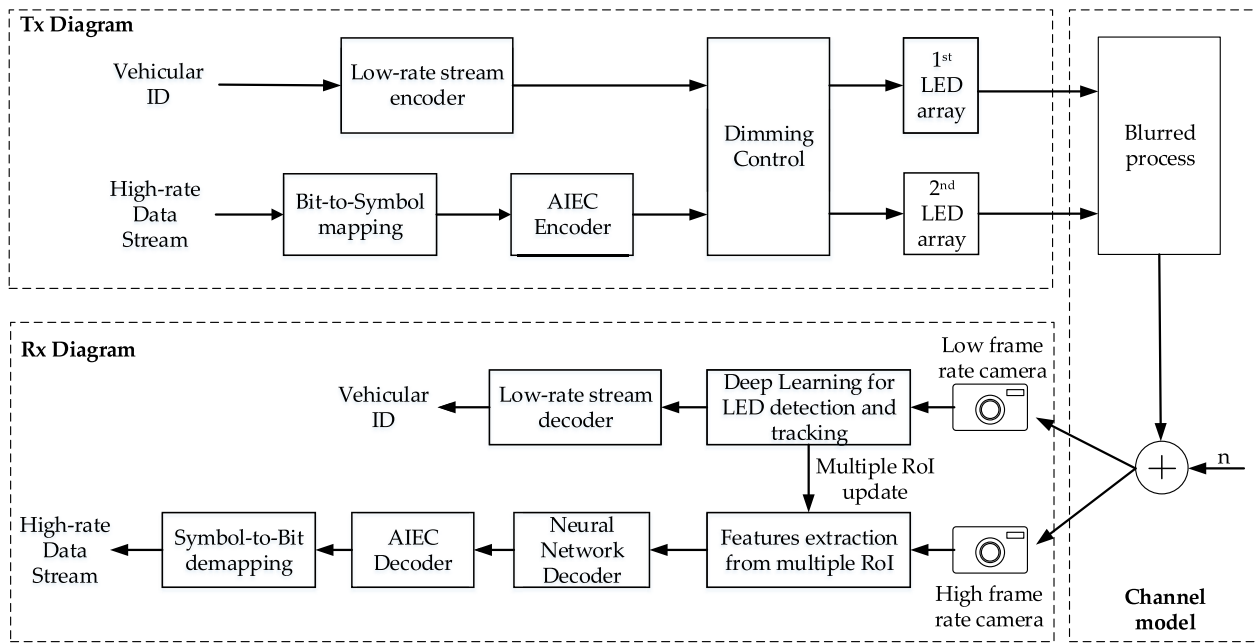
new approaches for dealing with such problems. A broad range of AI tools are now available, supporting a diverse range of applications. Fig. 2 illustrates the concept of a fully AI-equipped system.

Our primary aim in applying AI technology to an optical vehicular communication system is performance enhancement in two key tasks: (1) ROI detection and tracking and (2) communication functionality or data decoding. For the first task, it is recommended to use the YOLO platform for real-time multiple-light-source tracking before the low-frame-rate camera is able to extract all the ROI information from actual transmitters. Then, for further explicit high-rate data communication, a neural network decoder [15] and an AI-based error correction (AIEC) method [21] are proposed to enhance the accuracy of the received data, particularly when the signal is distorted by a complex channel model. Further explanation and experimental results will be provided in the following sections.

## C. OUR CONTRIBUTIONS

In the previous section, we presented the current challenges hindering the use of OCC technologies in vehicular communication systems. We will now move on to a detailed discussion of the uses of AI to improve these systems' performance. We have previously described the blocks from which our proposed system is formed in other recent papers. Although some of these areas of work are complete, with correspondingly consistent results [15], others have thus far only been introduced as concepts or early-stage research [21]. Moreover, as AI-based modules demand a high computational power, typically achieved using graphics processing units (GPUs) to accelerate their performance, it is difficult to be fully confident in the benefits of AI technology, particularly when applied to very specific applications or stand-alone modules.

With the above considerations in mind, this paper serves to present those previously documented technologies as a single proposed system that can be comprehensively evaluated to ensure that all basic performance requirements



**FIGURE 2.** The concept of optical vehicular communication system architecture with AI supporting.

are met. Comparing the performance enhancements provided by each AI-supported functional block provides a systematic approach for determining whether AI represents an excellent alternative. Furthermore, as some adjacent blocks function in similar ways, they can be combined to simplify the system. We also verify the ability to use the same set of deep-learning libraries for designing all the sections of the system, minimizing hardware costs.

The simulation environment used for the evaluation purposes was a combination of additive white Gaussian noise (AWGN) and blur phenomena, as described in [21]. As previously mentioned, although these are common issues in vehicular OCC systems, very little research has been performed to analyze the communication quality in noisy and blur-inducing environments. In summary, our main contributions in this paper are as follows:

- We briefly review our original concept of applying AI techniques to camera-based communication systems. The purpose of the proposed techniques is to mitigate the effects of the vehicular environment on the communication channel model. To concisely describe these effects, blur is added to the primary AWGN channel, which can be considered as representing the standard channel model of a vehicular OCC system.
- For the first time, we achieved a highly AI-integrated optical vehicular communication system by applying several trending technologies and algorithms, including LED detection and tracking method using the YOLO framework as well as a combination of semantic segmentation and fully-connected neural networks for feature extracting and data decoding process.

- To investigate the efficiency of the chosen parameters, several metrics for evaluating AI and communication systems will be analyzed. The corresponding experiments will then be presented, with detailed information provided on data preparation and annotation for model training.

The rest of the paper is organized as follows: In Section 3, we provide a detailed description of each enhanced receiver module, in terms of the newly defined channel model for vehicular environments. Finally, Section 4 describes our experimental procedures and results, demonstrating the benefits of deploying each proposed module in an optical vehicular communication system.

## II. PROPOSED SYSTEM DESIGN

### A. CHANNEL MODEL APPROXIMATION OF OPTICAL VEHICULAR COMMUNICATION

In the field of wireless communication, channel modeling and coding theory have been studied and developed over many decades. A lot of the current research is focused on estimating the channel model from real-world data using deep-learning technologies [22]–[24]. However, OCC is still a relatively new area of research; thus, to be able to commercialize OCC in vehicular wireless communication scenarios, intensive research into channel models for vehicular OCC systems is still required. Fig. 3 illustrates our idea of modeling the channel in the V2X communication system.

The most fundamental and unavoidable cause of degradation in any communication system is Gaussian white noise. In a vehicular OCC system, white noise sources include ambient light radiation from the Sun and street lighting. As the





FIGURE 3. Channel model in vehicular OCC system.

receiver side of the OCC system is heavily reliant on cameras and image sensors for its communication functionality, blur is also a significant concern, particularly in a vehicular environment. Sources of blur are varied, including motion blur, loss of camera focus, and weather conditions (e.g., fog, rain, or snow).

1) PIXEL NOISE MODELING AND SNR COMPUTATION

Pixel noise in charge-coupled device/complementary metal-oxide semiconductor (CCD/CMOS) cameras can be approximately modeled by

$$n \sim N(0, \sigma s^2), \tag{1}$$

where  $s$  is the pixel value,  $\sigma^2(s) = s.a.\alpha + \beta$  with  $a$  representing the mark and space amplitude, and  $\alpha$  and  $\beta$  representing the fitting parameters obtained experimentally. Here we used the values  $\alpha = 0.01529$  and  $\beta = 0.1973$ , which were estimated experimentally from [25]. Assuming that one symbol contains one bit, the pixel SNR (or  $E_b/N_0$ ) of an Rx camera can be estimated by:

$$Pixel \frac{E_b}{N_0} = \frac{E[s^2]}{E[n^2]} \approx \frac{a^2 \cdot \Delta}{a \cdot \alpha \cdot \Delta + \beta}, \tag{2}$$

where  $E_b$  is the bit energy,  $N_0$  is the noise density,  $s$  is the pixel value,  $a$  is the mark and space amplitude,  $\delta = T_{exposure}/T_{bit}$  is the ratio of the camera’s exposure time and bit interval, and  $\alpha$  and  $\beta$  are the fitting parameters.

Pixel SNR estimation has to be performed for each camera as there are no universal standard parameters for all cameras. Using Eq. (2), a theoretical estimation for a chosen camera with a shutter speed of 10 kHz and a transmitter with an optical clock rate of 1 kHz, giving  $\delta = 0.1$ , is shown in Fig. 4. Even when no Tx signal is transmitted, white noise is still present. When the received signal reaches the maximum value of an 8-bit analog-to-digital converter, the pixel SNR is approximately 40 dB.

To achieve a higher pixel SNR, we could increase the received signal strength by either increasing the transmitter power or using a longer camera exposure time. However, neither solution is appropriate in this application, not only because power saving is also of a high priority in wireless communication systems, but also because increasing the camera’s exposure time increases the probability of capturing

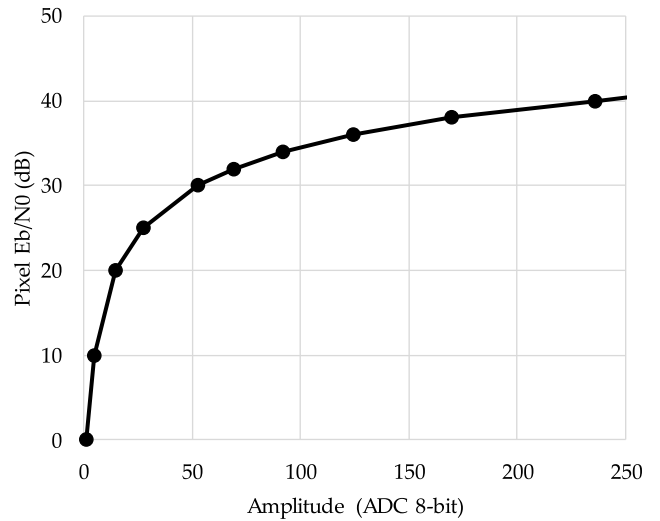


FIGURE 4. Calculation of the relationship between pixel SNR and BER.

fuzzy LED states, which occur when samples are obtained during pulse transitions [15].

Moreover, single-carrier modulation schemes such as S2-PSK and DS8-PSK require at least 10 dB of SNR to achieve a BER of  $10^{-4}$ . The SNR measurement for vehicular OCC in [17] also shows that the line-of-sight link quality always guarantees this minimum SNR requirement.

2) BLUR PROCESS – PRINCIPLE AND SIMULATION

Based on the findings presented in [26]–[28], a received image of LEDs, contaminated with blur and noise, can be obtained by the convolution of a blur kernel with a clear image of LEDs, followed by the addition of noise as follows:

$$y = h^*x + n, \tag{3}$$

where  $y$  is the captured image matrix,  $h$  is the blur kernel matrix,  $x$  is the original image matrix, and  $n$  is the noise matrix. Note that all of the matrices here are two-dimensional (2D), as the LEDs image will be converted to a grayscale format.

In order to blur an image with a size of  $r \times c$  pixels (rows  $\times$  columns), we use a 2D blur kernel matrix (the size of the kernel ( $b$ ) is crucial in determining the level of blur). All cells in a blur kernel will take equal values and sum to

one, so each cell will take a value of  $\frac{1}{b^2}$ . The output matrix of the convolution between the blur kernel and an LED image will have a full size of:  $(b + r - 1) \times (b + c - 1)$ . To obtain a blurred vision with the same size as the original, we crop the product matrix  $c$  at a row and column index equal to  $b/2$ .

Based on our research in [15], a sample image of an eight-LED group is used to analyze the effects of blur for different sizes of convolutional kernel. Simulation work also shows that the convolution-based mechanism results in blurring that flattens all LED intensities within an image.

## B. LED DETECTION AND TRACKING BASED ON YOLO FRAMEWORK

The evolution of YOLO, the most popular convolutional-neural-network-based object detection platform, has provided us with many customization options for our application [29]. Herein, we discuss the customization and training of different YOLO model (YOLOv2, YOLOv3 and their tiny model) for vehicular front or tail light detection. From [30], a comparison of state of the art YOLOv3 model with other detectors, as well as the comparison on other version of YOLO is provided in detailed. In our application of LED detection and tracking for vehicular OCC system, minimizing the processing time with reasonable trade back on accuracy is our priority. Therefore, the tiny model of YOLOv2 or YOLOv3 is considered to be modified for our application, with only 9 or 7 convolution layers make it is lightweight and could achieve 200fps detection performance. The number of class for detection is only one, and the corresponding number of filters in the final convolution layers is set as 30. To function within an OCC system, these light sources need to be adapted to produce high-frequency flickering in accordance with the modulation scheme. The transmitter could be a single LED [31]–[33], or multiple small LEDs [34]. When using the S2-PSK hybrid waveform and DS8-PSK for modulating two data streams in optical vehicular communication [34], the number of small LEDs in each group must be a multiple of eight. Hence, our dataset for training the YOLOv3 model had to be able to classify these two light source categories.

For the object tracking algorithm, in [35], the author proposed a method based on the Euclidean distance from the detected object in the current frame to  $n$  previous frames. In most cases, this algorithm performs well while satisfying the requirements for real-time object tracking with a high accuracy. However, the labels of multiple objects can be shuffled when object crossing occurs [36]. For more reliable tracking results, the ROI signaling technique is considered to be preferable to an image-processing-based solution, in which vehicle identity information is actively transmitted to the receiver. The use of ROI signaling with OCC has recently been introduced in IEEE 802.15.7m PHY IV for OCC modes. Intel's UFSOOK scheme [37] can deliver short identities of multiple-input multiple-output (MIMO) light sources. The aim of the S2-PSK scheme, proposed by Kookmin University [18], is to support vehicles in terms of tracking/identifying MIMO light sources. These two

modulation schemes support the detection and tracking of numerous communication data light sources in a flicker-free manner compatible with various types of low-frame-rate cameras, supporting communication over large distances among high levels of light noise on the road. However, with ROI signaling, the low frame rate camera needs to entirely decode the vehicular ID from the hybrid waveform before the high frame rate camera can obtain further information from the actual transmitter, which is suboptimal in terms of processing time. For example, by noting that an interval of a bit ( $T_{bit}$ ) is multiple times the waveform cycle ( $T$ ):

$$T_{bit} = NT, \quad (4)$$

In our practical system, a clock rate of no lower than 125 Hz is used to modulate the optical light source for mitigating against potential flicker outdoors [18]. Using a clock rate of 200 Hz and  $N = 20$  [17], a 10 bps link can be provided, which is feasible with a 20 fps camera at the receiver side. This means that if around 32 cars need to be labeled for simultaneous detection and tracking, the vehicle ID must at least be 5 bits. After using half-rate line coding [18] the transmitted codeword will be 10 bits, requiring 1 s to complete the transmission, during which period the high-frame-rate camera is completely idle. To minimize this idle time, the YOLO-based tracking method is used to temporarily detect and track all the connected OCC transmitters from the previous communication session. Taking advantage of these primitive ROIs, the high-frame-rate camera can continue decoding all transmitting high-rate data streams simultaneously until they are correctly labeled with each vehicular ID.

## C. HYBRID ROI-SIGNALING MODULATION SCHEME FOR MASSIVE-LEDs ARRAY

By using the HS-PSK modulation scheme in optical vehicular communication system, each LED group in the OCC vehicular system will be a combination of  $8n$  LEDs, and the number of LEDs will be the same for both the reference and the data LED group. The PHY PIB attributes for HS-PSK mode, as defined in IEEE 802.15.7-2018, only refer to a basic implementation using eight LEDs within each light source in `phyHSpskNumLightSources`, so the  $n$  number can be modulated and transmitted using a low-rate modulation scheme. Figure 5 provides an illustration of the sample LED arrangement, with  $n = 20$ . Our hybrid waveform implementation is summarized below:

- A low-rate modulation scheme for carrying low-rate data (vehicular ID, number of eight-LED groups) is generated by controlling the dimming level of two LED arrays. There may be a difference in the dimming level between the two LED arrays, but the dimming of all the LED groups within each array will be similar at every sampling point and can be considered as the dimming of the whole LED array. Using S2-PSK as the low-rate modulation scheme, the Tx can only transmit 1 bit for every sample taken by the low-frame-rate camera.

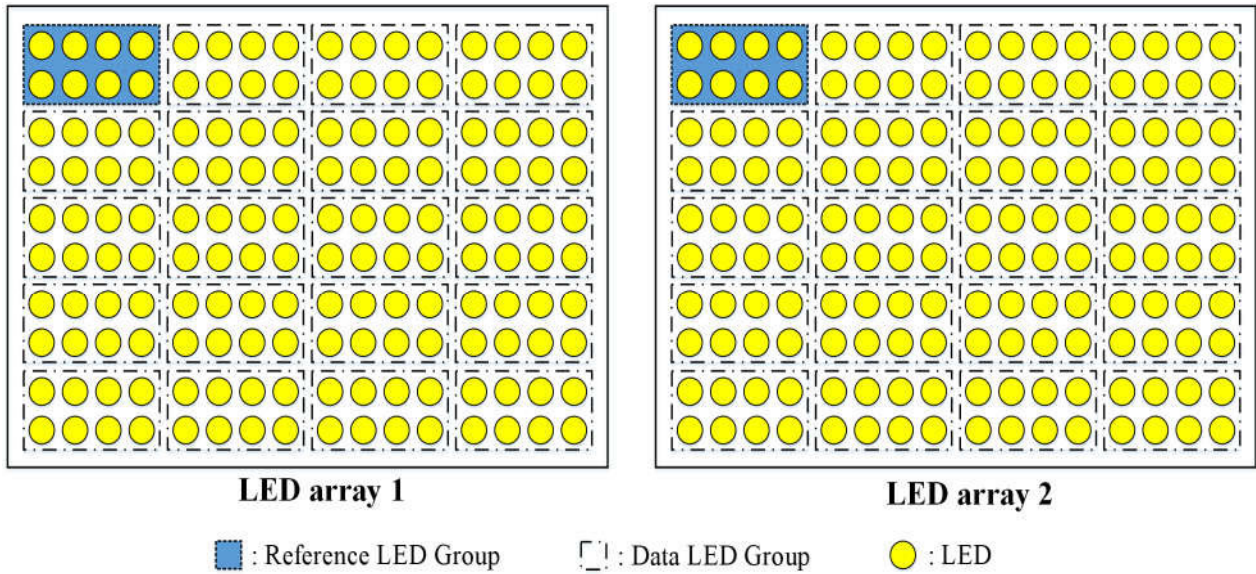


FIGURE 5. LED arrangement in vehicular light source for OCC functionality.

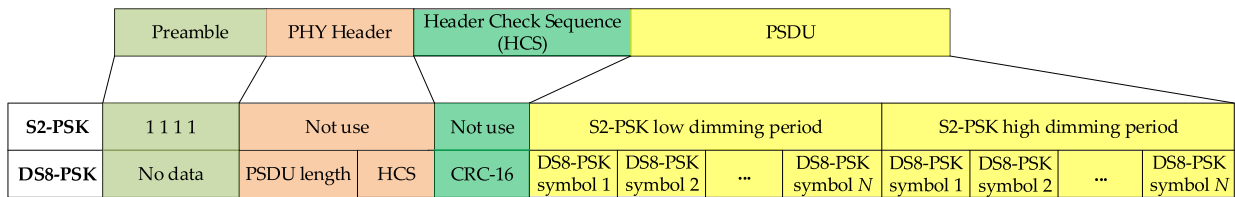


FIGURE 6. PPDU format using in HS-PSK.

- A high-rate modulation scheme for carrying high-speed data (e.g., vehicle speed, engine status) will be generated by every data LED group, with each group containing eight LEDs thanks to the use of DS8-PSK, within each LED array. In our example, each LED array contains one LED group for the reference phase and 19 LED groups for high-speed data encoding. At every sampling point of the high-frame-rate camera, each data LED group can transmit 3 bits per symbol.

The physical protocol data unit and physical layer service data unit (PSDU) using the HS-PSK format are shown in Figure 6. The clock of the dimming control for this high-rate communication is synchronized to the clock of the low-rate ROI signaling stream. In more detail, the HS-PSK PSDU field consists of multiple S2-PSK cycle times; each cycle is a subframe with a low-dimming period and a high-dimming period so that each period also consists of multiple DS8-PSK data symbols.

The presence of another HS-PSK preamble indicates the end of the HS-PSK PSDU. The configuration of the PSDU length is implemented via the PHY PIB attribute phyPsdLength and is announced by sending the updated PSDU length via the PHY header subfield.

Based on the basic knowledge of transmitting data format, Kookmin University already proposed a decoding method using a dual-camera receiver system for HS-PSK in [17]. A low-frame-rate camera, of either a global shutter or a rolling shutter type, is used for detecting the S2-PSK signal from the variations in transmitter dimming; whether a high-speed global shutter camera is used to receive DS8-PSK transmitting symbols. In both contexts, the camera’s frame rate should be higher than the modulation clock rate. A cross-check-XOR model [17] and a matched filter [34] are two effective decoding methods for the S2-PSK and DS8-PSK modulation schemes, respectively, which have been proven with the AWGN channel.

#### D. NEURAL NETWORK-BASED DECODER

A neural-network-based model for OCC signal decoding has already been described in [15]. In summary, data decoding from the light signal transmitted by vehicular LEDs can be performed by extracting the central intensity point of each LED within the group detected. As the YOLO-based detection and tracking block will form a boundary on every possible vehicular LED group, knowing the number of LEDs inside each LED group as well as the LED arrangement is



critical for extracting the intensity of every LED, as well as for binary or fuzzy logic state mapping.

Despite the effectiveness of those conventional decoding method over such common channel in wireless communication like AWGN, by considering the probability of fuzzy LED state sampling and the significant effects of blur in the vehicular communication channel, the neural network model could improve the data decoding accuracy and reduce the frame rate requirement for the high-speed camera, as per Nyquist's theorem [17]. Fig. 7 illustrates the neural-network-based decoder model architecture employing an iterative learning process with a real-world data feature map to adjust its parameters.

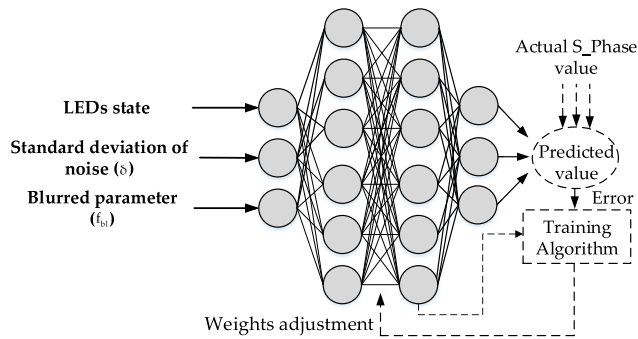


FIGURE 7. A neural network (NN)-based decoder for blurry and noisy.

The feature map extracted from the input data includes the following:

- LED State: From each image of the transmitter LED groups, we extract the intensity of central points of every LED in each group. The approaching method could be normally dividing the detected area based on the shape of LED arrays, which could be informed by using S2-PSK signage. Another approach could be based on semantic segmentation [38], [39] or instance segmentation [40], which could be built up from various of powerful object detection baselines. In our application, we need a real-time LED segmentation (at around 30 fps) solution with an acceptable mask accuracy, for which either Mask-YOLO [38], [41] or YOLACT [42] could be a suitable candidate. For those system does not have the computational power from GPU, intensity-threshold based segmentation represents a simpler method for bright object segmentation which can be deployed in CPU-only system in real time, and this technique has already been functioned as an IMAQ Count Objects block in LabVIEW's Vision Development Module (National Instruments, Austin, TX, USA).
- Standard Deviation of Noise: After measuring and estimating the SNR value for use in a camera with a specific distance and exposure time, as was done intensively in [17], the standard deviation corresponding to the presence of Gaussian white noise in a communication channel can also be calculated by:

$$\delta^2 = \frac{1}{N-1} \sum_{i=0}^{N-1} (\chi_i - \mu)^2, \quad (5)$$

where  $\chi_i$  is the intensity of the signal at the sampling time  $i$ ;  $\mu$  is the signal's mean;  $N$  is the number the of samples, and  $\delta$  is the standard deviation of the noise. The signal's mean can also be calculated by:

$$\mu = \frac{1}{N} \sum_{i=0}^{N-1} \chi_i, \quad (6)$$

- Blurred parameter ( $f_{bl}$ ): This is the novel feature of input data for the information of the dimming case and the ratio of the blur kernel in a specific area of the image:

$$f_{bl} = \text{dimming} \frac{\text{kernelarea}}{\text{imagearea}}, \quad (7)$$

Here, the dimming value ranges from 1 to 7,  $\text{kernelarea}$  is the number of pixels occupied by the blur kernel on the LED image, and  $\text{imagearea}$  is the total number of pixels in the LED image ( $200 \times 100$  pixels). It is worth noting that the kernel area in this equation is calculated on the central point of each LED in the LED group.

In Fig. 8, we consider the case in which the blur kernel size does not exceed the area of one LED in the image ( $50 \times 50$  in this case). Therefore, when calculating the post-blur intensity of the LED's central point, the whole area of the blur kernel will be within the cropped image of the LED group.

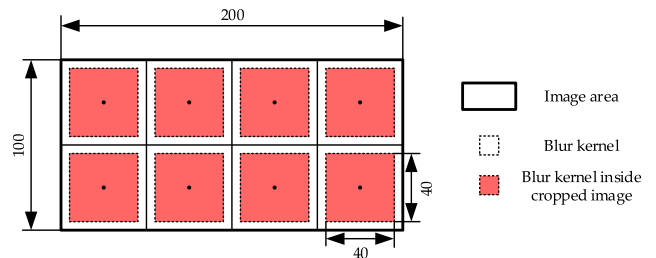


FIGURE 8. The calculation of the ratio of the blur kernel area in the image with kernel size =  $40 \times 40$ .

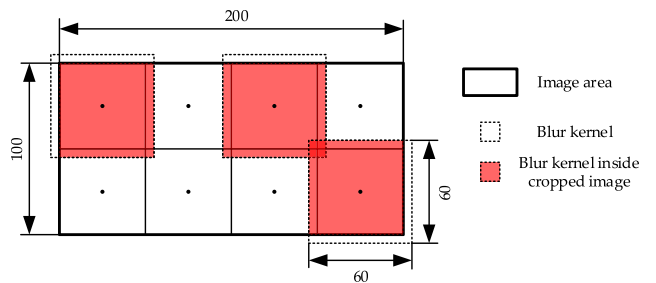


FIGURE 9. The calculation of the ratio of the blur kernel area in the image with kernel size =  $60 \times 60$ .

In Fig. 9, the size of the blur kernel ( $60 \times 60$ ) exceeds the area of one LED in the image ( $50 \times 50$ ). Therefore, the exact area of the blur kernel inside the cropped image for calculating the post-blur intensity of the LED's central point will be part of the blur kernel.

### E. AIEC

Current forward error correction (FEC) technologies are designed to reduce noise when transmitting data over

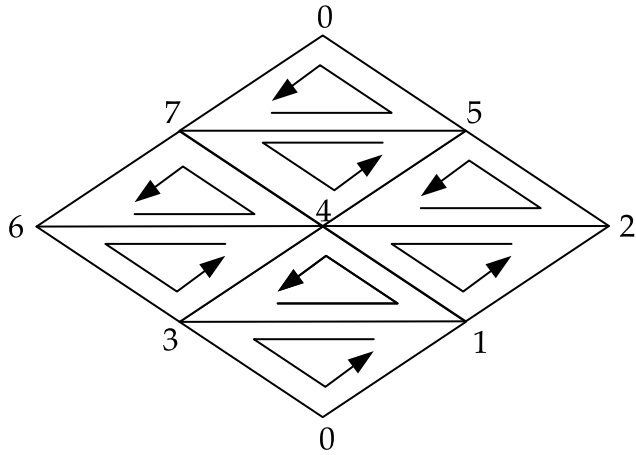


FIGURE 10. Rhombus diagram for encoding the 1/3 code rate AIEC (v1).

long distances. However, if we take into account the effects of blur on our communication channel, a model that can be directly trained and adapted by real-world data might offer a good solution. We herein propose an AIEC method that includes a novel encoding and decoding strategy for use in vehicular OCC applications. Symbols are encoded according to a predefined error correction encoding rule before transmitting. After neural network-based decoding process, these output symbols will be grouped as cluster of symbols before being passed through our pretrained AIEC decoder which is expected to output our original transmitted bitstream.

1) AIEC ENCODER

Traditionally, for error correction purposes, the data bitstream will be encoded using channel coding before being mapped into symbols [18]. Our proposed error correction technique is somewhat different. While it tries to transform each symbol or group of symbols in transmitting the data stream to a new form—a group of new symbols, following a predefined mapping table. The number of symbols in each group (original and transformed groups) forms the code rate used for AIEC encoding.

In this section, we present our recently designed AIEC encoding table for two code rates: 1/3, which maps each data symbol into a group of three symbols, and 1/5, which maps each data symbol into a group of five symbols. Note that the AIEC encoding method presented here is designed specifically for the DS8-PSK scheme. However, with some minor modifications, the encoding and decoding principles used for neural-network-based error correction could also be applied to other modulation schemes.

*a: CODE RATE 1/3*

We recently designed and tested two encoder versions with the same 1/3 code rate. The first was based on a rhombus diagram that allocates eight symbols in the DS8-PSK modulation scheme. The proposed rhombus diagram for three-symbol group mapping is illustrated in Fig. 10.

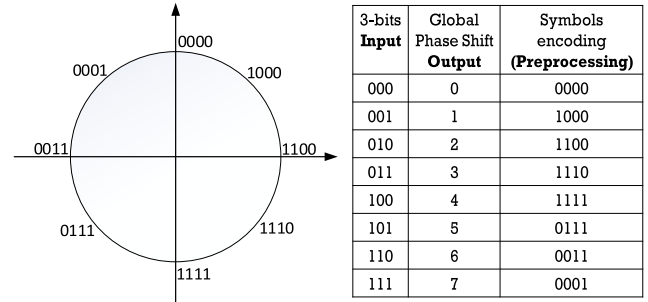


FIGURE 11. DS8-PSK constellation diagram for symbol-to-pseudo-bit mapping.

TABLE 2. AIEC encoding table for the 1/3 code rate (version 1).

Original symbol	0	1	2	3	4	5	6	7
Encoded 3-symbol group	075	476	634	457	542	314	130	124
	750	764	346	574	425	143	301	241
	507	647	463	745	254	431	013	412

The procedure for designing the encoding table is as follows:

- Mapping 3 symbols follow diagram: each 3-symbol sub-triangle will be served together as a symbol group, for example (3-7-6), (4-7-3).
- Packet be built anti-clockwise: (3-7-6); (7-6-3); (6-3-7)
- Totally we can build up to 24 different groups labeled from 0 to 23, as provided in Table 2.

The Hamming distance between two groups can be calculated by:

$$d = \sum_{i=1}^3 d_{Hamming}(S_i(A), S_i(B)), \tag{8}$$

where  $S_i(A)$  and  $S_i(B)$  represents  $i^{th}$  symbol of group  $A$  and  $B$  respectively. Using the encoding method in Table 2, we increase the minimum Hamming distance between symbols to 3. This enables our first proposed coding technique to deliver a significantly reduced symbol error rate (SER).

Our second encoder design aims to gradually build just enough eight mapping groups for eight symbols of DS8-PSK instead of 24 groups, which results in a significant increase in the Hamming distance between symbols. The design procedure for this approach is summarized as follows:

- Symbol-to-pseudo-bit mapping: To directly analyze the Hamming distance between symbols after AIEC encoding, eight symbols defined in the DS8-PSK modulation scheme are remapped into pseudo-4-bit form. These pseudo-4-bit groups are defined in accordance with the constellation diagram, as illustrated in Fig. 11.
- To form three-symbol groups for 1/3 code rate AIEC, after the symbol-to-pseudo-bit mapping process, each

TABLE 3. Mapping group design with  $s = 3$ .

3-symbol group	12 pseudo bits sequence
005	000000000111
050	000001110000
500	011100000000
111	100010001000

TABLE 4. Mapping group design with  $s = 9$ .

3-symbol group	12 pseudo bits sequence
441	111111110000
414	111110001111
144	100011111111
555	011101110111

symbol group will correspond to a 12-pseudo-bit sequence. To design an efficient encoding table, we are interested in the sum of the 12 bits in each sequence:

$$s = \sum_{i=1}^{12} b_i, \tag{9}$$

where  $s$  is the sum of the pseudo-bits in the sequence and  $b_i$  is the binary value (0 or 1) of the  $i^{th}$  pseudo-bit.

- The sum of bits will have a minimum value of 0 when all the pseudo-bits are 0, and a maximum value of 12 when all the pseudo-bits are 1. Here we choose the sum values of 3 and 9 for constructing the mapping table.
  - For  $s = 3$ , four sequences are chosen—005, 050, 500 and 111—as provided in Table 3. The aim is to maximize the Hamming distance while avoiding an even distance (2 or 4) between two corresponding symbols.
  - For  $s = 9$ , by inverting four sequences in Table 3, we obtain four sequences with  $s = 9$  in Table 4.
- In total, we construct eight different groups for the eight symbols in the DS8-PSK modulation scheme, as shown in Table 5.

By applying our second AIEC encoding method, as shown in Table 5, the minimum Hamming distance between groups will be raised to 6. The next section provides a detailed performance analysis of both of our encoding versions with a 1/3 code rate over the AWGN channel.

**b: CODE RATE 1/5**

For significantly more reliable data error correction, we can increase the number of transmitted symbols to five symbols for each original data symbol, AIEC with code rate 1/5. The design of this method also follows a unique rhombus diagram, as illustrated in Fig. 12.

TABLE 5. AIEC encoding table for the 1/3 code rate (version 2).

Original symbol	0	1	2	3	4	5	6	7
Encoded 3-symbol group	111	005	050	500	144	414	441	555

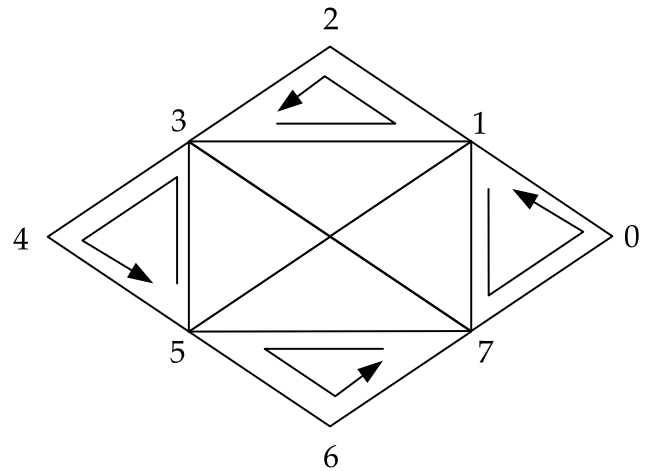


FIGURE 12. Rhombus diagram for AIEC encoding with code rate = 1/5.

The detailed procedure is as follows:

- Forming a group of three symbols: Four triangles will be chosen in the rhombus diagram to form three-symbol groups: (1-2-3), (3-4-5), (5-6-7), and (7-0-1). Working in an anti-clockwise direction, we can build up to 12 different three-symbol groups.
- Doubling the number of group - adding the fourth symbol either 0 or 4: By adding a fourth symbol, either 0 or 4, the mapping group increases from 12 to 24. As a trade-off, the minimum Hamming distance correspondingly decreases from 6 to 4.
- Increase the Hamming distance between groups: A fifth symbol can be added based on two trivalues: (0-3-6) and (1-4-7).

In total, we have constructed 24 different groups, labeled from 0 to 23, as showed in Table 6. The minimum Hamming distance among eight symbols is now increased to 8, and among 24 groups is 6.

**2) AIEC DECODER**

The AIEC decoder’s operating principle is also based on a fully connected MLP neural network model. The symbols output from the neural-network-based decoder described in Section 3.3 will be gathered to each group of three or five symbols, depending on the code rate used. Before feeding these symbols group by group to the AIEC decoder model, they must be passed through a data preprocessing process, which remaps every symbol to its corresponding pseudo-4-bit form, as shown in Fig. 11.

**TABLE 6.** AIEC encoding table for the 1/5 code rate.

Original symbol	Encoded 5-symbol group		
0	12300	23103	31206
1	34504	45307	53402
2	56700	67503	75606
3	70104	01707	17002
4	12344	23147	31242
5	34540	45343	53446
6	56744	67547	75642
7	70140	01743	17046

**TABLE 7.** Parameters set for the AIEC decoder model.

Code rate	1/3 (version 1)	1/3 (version 2)	1/5
Number of hidden layers	4	4	4
Number of neurons in layers	12-241-241-24	12-81-81-8	20-241-241-24
Activation function in hidden layers	Tanh-ReLU	Tanh-ReLU	Tanh-ReLU
Optimization algorithm	Adam Optimizer	Adam Optimizer	Adam Optimizer
Beta value for regularization	$10^{-4}$	$10^{-4}$	$10^{-4}$
Number of learning epochs	500	500	500

According to experimentation, to make the model rapidly converge and deliver better performance on the test set, all pseudo-bits of value zero are replaced by  $-1$ . The sample set of parameters (number of hidden layers, number of neurons, optimization algorithm, etc.) used for designing the AIEC decoder, as shown in Table 7, is compatible with both code rates proposed for AIEC encoding. Since the main task of AIEC is to detect and correct the data decoding error caused by the interference noise from communication channel, we propose a simple NN model to prevent the overfitting of the model. The number of the hidden layer is kept at four layers, robust enough to detect and classify the pattern effectively. With 5 hidden layers or more, the accuracy on test dataset decreases because of the overfitting. The number of neural is also picked as small as possible. Based on the requirement of the encoder, we propose three different NN that can effectively detect and correct the data decoding error caused by the interference noise. The version 2 model is the most robust AIEC model come up with the high code

rate efficiency, compare with other proposed AIEC models. The ADAM optimization also picked due to its effectiveness and high convergence speed on shallow classification model (2 hidden layers only) compare with other optimization. For three AIEC decoder models, the practical NN training efforts are 500 epochs. The accuracy of AIEC decreases significantly with higher learning effort, as the model becoming overfitting.

### III. EXPERIMENTS SET UP AND PERFORMANCE EVALUATION

#### A. LED DETECTION AND TRACKING BASED ON YOLO FRAMEWORK

To detect vehicular LED groups, we used a 3 min portion of a night driving video recorded on the highway from Daegu to Seoul [44] and extracted 400 frames for the training and validation datasets. The labeling procedure illustrated in Fig. 13 employs the labeling tool developed by Manivanan Murugavel in [45].

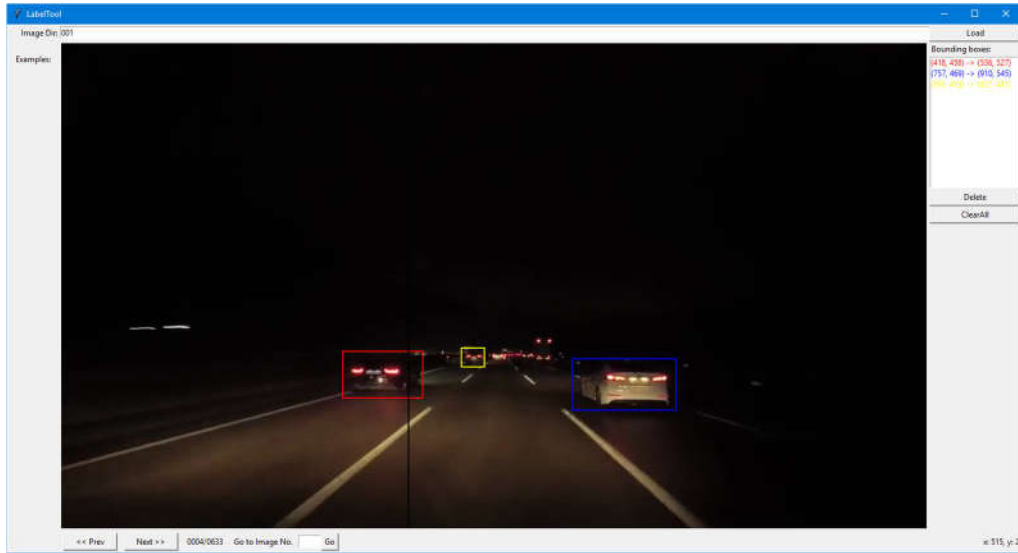
After 9,000 training epochs, the average loss of the YOLOv2 model was approximately 0.14. We also tested the performance of the trained model on two videos of night driving scenarios in Korea, which are provided in Fig. 14 and 15. Our test video for the highway night driving case can be found online at <https://www.youtube.com/watch?v=m0SGZHkukzk>, and the test video for city driving in Seoul at night can be found at <https://www.youtube.com/watch?v=sH7pRZPGNm4>.

The results demonstrated that real-time object detection could be achieved using the YOLOv2 model (in which the acquired processing frame rate varied from 40 to 50 *fps* in our experiment). For detecting those RoI with confidence score higher than 60%, our first video show that we could detect those RoI at a distance within 150 *m*, with moving speed over 60 *km/h* in highway. The second video shows a weaker RoI detecting ability (detection range is within 80 *m*, the RoI area is also not correctly fit with confidence score above 40%) in brighter background, since this lighting condition is not included in our training set. However, it prove that YOLOv2, which is based on CNN, can actually perform excellent for RoI detecting task in optical vehicular communication even at night time, given that it is trained with that type of background condition. Therefore, we could use recent version of YOLO framework as a solution for RoI detection in challenging weather conditions such as rainy, foggy, snowy, which are all currently presenting in Korea.

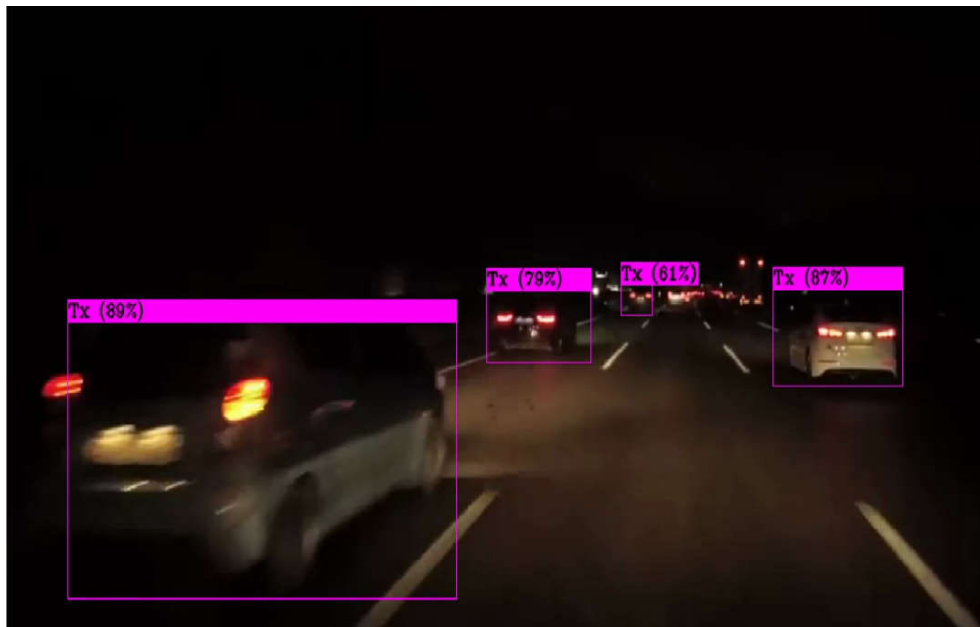
#### B. NEURAL NETWORK-BASED DECODER

As the input to the neural-network-based decoder, LED states can be extracted from the central point intensity of each LED area, or from the mean value of all intensities within each LED area, by comparison to a threshold (binary logic case) or mapping to a value between 0 and 1 (fuzzy logic case). In this paper, we consider instance segmentation as an efficient and practical method for extracting every LED area.





**FIGURE 13.** Label tool interface for ground truth bounding box labelling on training image dataset.



**FIGURE 14.** A test on using YOLOv2 for highway night driving with rain.

Decoder performance analysis was carried out under different blur and dimming conditions, with data in fuzzy logic form, simulated using LabVIEW.

1) LED SEGMENTATION WITHIN THE RoI AREA

After object detection phase, we analyzed the LED segmentation phase. The detected area which is defined within the rectangular bounding box will be extracted as the input of segmentation phase. Our LED segmentation algorithm was developed in LabVIEW, which uses the IMAQ Count

Objects block to segment bright objects according to a pre-defined intensity threshold value, as well as the minimum and maximum object size, to minimize unwanted detection of bright areas other than vehicular LEDs. Following detection, the LED area is marked from 0, so we need to sort all of the detected LED numberings prior to central intensity point extraction. Our strategy is to first detect the four corner LEDs of a pair of LED arrays, located at the top left, top right, bottom right, and bottom left of the array pair. These corner LEDs will consequently be closest to the top-left point, top-right point, bottom-right point, and bottom-left point of



FIGURE 15. A test on using YOLOv2 for city driving inside Seoul at night.

TABLE 8. Parameters setting for LED segmentation from background experiment.

Pixel intensity threshold (8 bit)	220
Minimum detected LED area (pixels)	400
Maximum detected LED area (pixel)	1600
Camera frame rate (fps)	30
Camera exposure time (ms)	0.488
LED diameter (mm)	10
Distance (m)	0.5 ~ 1

the ROI bounding box. For reference, the parameter settings in our experiment can be found in Table 8.

The process interface of this phase is illustrated in Figure 16. After extracting the locations of the four corner LEDs, other LEDs' central point coordinates can be interpolated from the coordinates of the four corner LEDs, and the LED state matrix can be obtained from those central intensity points to form the input to our proposed neural-network-based decoder.

The distance is varied from 0.5 to 1m in our experiment. In the real-world scenario, because we have not make test on LED segmentation process after RoI detection, so it is hard to conclude the boundary on achievable distance in real world scenario. However, since the real-time RoI detection and tracking task on high-mobility vehicles still be more challenging and require more computational cost over whole image, compare with the LED segmentation phase using

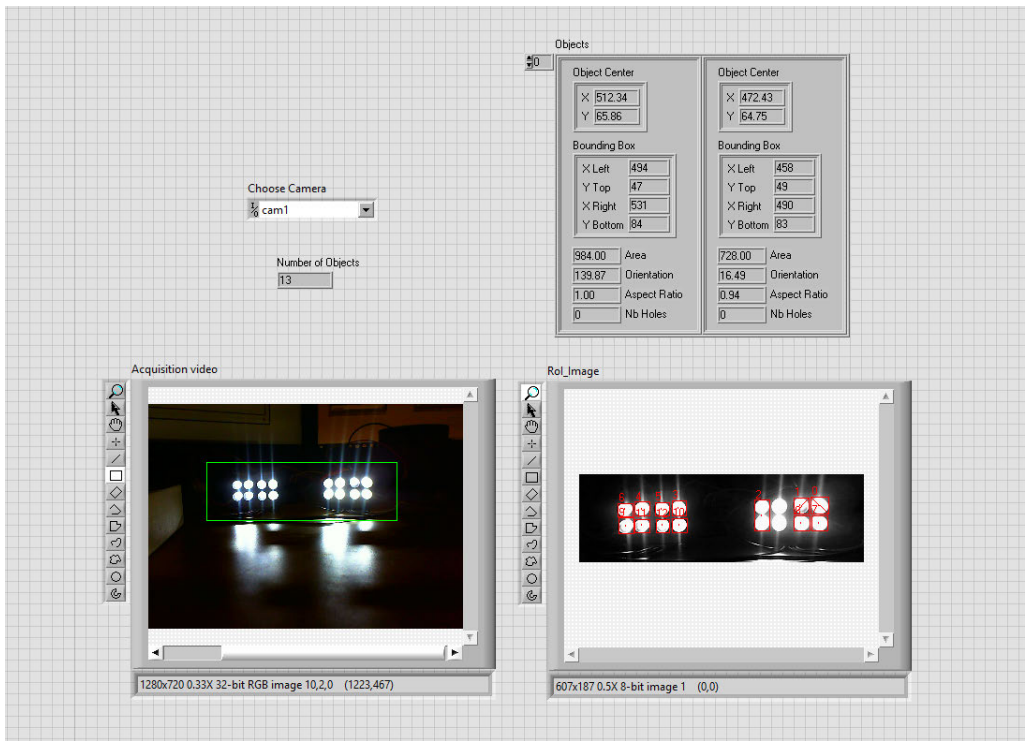
LabVIEW; therefore, it is also feasible to achieve the same mobility and distance in segmentation phase as in the object detection experiments. It is also noted that the RoI detection phase is handle on the frames which are sampled from low-frame rate rolling shutter camera, while segmentation and data decoding phase will be handle on the frames which are sampled from high-speed global shutter camera.

## 2) ANALYSIS OF THE EFFECTS OF BLUR AND DIMMING LEVEL ON THE DECODING PERFORMANCE OF THE NEURAL-NETWORK-BASED DECODER

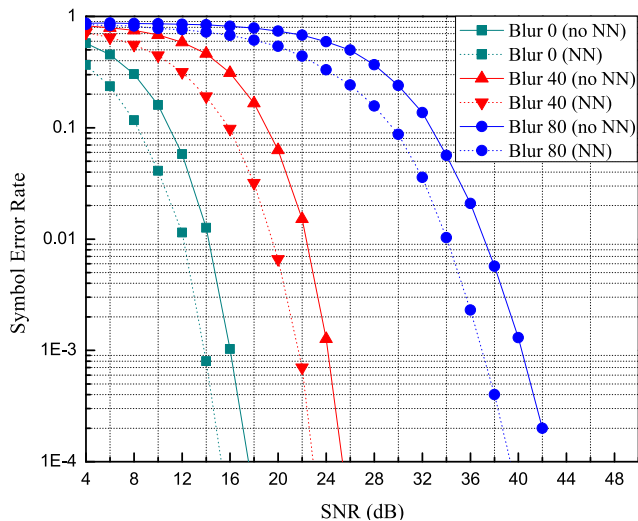
To analyze the effects of the ratio of the area occupied by the blur kernel over the total LED image area, several datasets with different blur kernel sizes were generated using LabVIEW to measure the SER performance. A comparison between our proposed neural-network-based decoder and a conventional matched filter model is shown in Fig. 17, with the dimming level set to 4/8 or 50% throughout our experiments.

In the nonblur case, using a neural-network-based decoder, the required SNR for  $10^{-4}$  SER level decreased by 2 dB compare with the conventional decoder. For the  $40 \times 40$  blur kernel case, the improvement was approximately 2.2 dB. However, with a blur kernel size of  $80 \times 80$ , the proposed scheme could achieve  $10^{-4}$  SER level with an SNR of 40 dB, whereas the matched filter could achieve only  $10^{-3}$  SER level for the same SNR condition.

In addition, the dimming level may also affect the blur in sampling images, the impact of which could be significant or trivial, depending on the blur kernel size. In our most recent paper, it was demonstrated that, without blur, using the matched filter for data decoding, performance lines corresponding to dimming from 2/8 to 6/8 were approximately

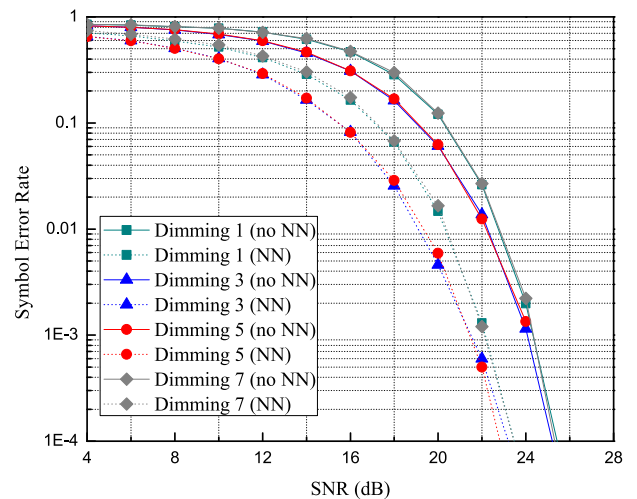


**FIGURE 16.** Bright object segmentation using threshold on pixel intensities method for LED central points localization within RoI.



**FIGURE 17.** The comparison of symbol error rate (SER) performance between an NN-based decoder and a matched filter on difference blur kernel size, dimming level is 50%.

similar, whereas dimming of 1/8 and 7/8 yielded the worst SER performance with similar performance lines [10]. When blurring occurred with an increase in the blur kernel size exceeding the LED area, the distinction between performance lines corresponding to each dimming level also increased. Fig. 18 and 19 compare the BER performance between an NN decoder and a matched filter for different dimming cases with a blur kernel size either contained by  $40 \times 40$  pixels

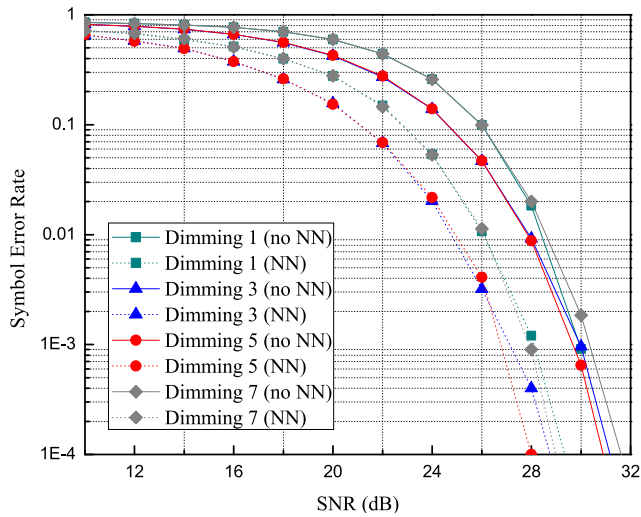


**FIGURE 18.** The comparison of symbol error rate (SER) performance between an NN-based decoder and a matched filter on different dimming cases with a blur kernel size of  $40 \times 40$ .

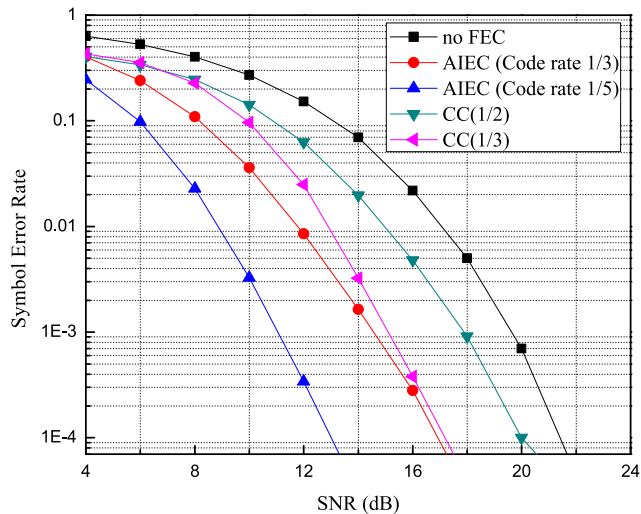
compared with  $50 \times 50$  pixels for each LED) or exceeding ( $60 \times 60$  pixels compared with  $50 \times 50$  pixels for each LED) the LED area.

### C. AIEC

Experiments were devised to demonstrate the efficiency of the new error correction scheme for two code rates: 1/3 and 1/5. For the 1/3 code rate, we used the second version



**FIGURE 19.** The comparison of symbol error rate (SER) performance between an NN-based decoder and a matched filter on different dimming cases with a blur kernel size of  $60 \times 60$ .



**FIGURE 20.** Performance comparison on using AIEC with code rate 1/3 and 1/5 to the method of using Convolution Code (with blur kernel size is  $40 \times 40$ ).

of the mapping table previously provided for SER analysis of the first encoding method in [11]. An SER performance comparison using the new version of AIEC versus using the convolution code for FEC is provided in Fig. 20.

Fig. 20 shows that, to achieve our desired SER performance of less than  $10^{-4}$  in an optical vehicular communication system, AIEC could significantly reduce the required SNR, when compared with the convolution code. AIEC with the 1/3 code rate is guaranteed to reduce the required SNR by approximately 5 dB, whereas the 1/5 code rate can deliver a 9 dB reduction in the required SNR, when compared with just using a neural-network-based decoder.

Fig. 20 shows that to achieve our desired SER performance, which is below  $10^{-4}$ , in an optical vehicular communication system, the AIEC could reduce significantly the requirement on SNR, compared with the use of the Convolution Code.

The AIEC with code rate 1/3 could guarantee to reduce the required SNR value by 5 dB approximately, while code rate 1/5 could give the ability to reduce the required SNR value by 9 dB, compare with just using a neural network-based decoder.

#### D. SYSTEM COMPLEXITY AND HARDWARE REQUIREMENTS

By taking consideration on the computational complexity when designing any AI-based system, we have also calculated the FLOPS of each proposed in-use neural network-based block in our implementation by using the method which is described in [46].

- Neural network-based decoder (18-81-81-8): 14.7 kFLOPs
- AIEC code rate 1/3 (version 1) and code rate 1/5: 24.1 kFLOPs
- AIEC code rate 1/3 (version 2): 14.3 kFLOPs

Though these feed forward neural network models are simple-designed and not require many computational power, so the pre-trained model for Neural Network-based decoder and AIEC could run faster on CPU also.

About using tiny YOLOv2/v3 pretrained model for setting up a demo on real time LED detection and tracking, our GPU specification is NVIDIA GTX 1080Ti, with CUDA version is 10.1. However, the lower generation of GPU with computing capability above 3.5 still could be used. The comparison on different object detection backbone about the accuracy and the amount of required FLOPs could be found in [47]. The tiny YOLOv3 could be found to achieve more than 200fps with 33.1 mAP over COCO dataset, with number of required FLOPs is approximately 5.5 billion. Therefore, it is feasible to be used in detecting and tracking the ROI in vehicular optical communication system.

#### IV. CONCLUSION

Our aim in this paper was to provide a comprehensive architecture for a vehicular OCC system with AI support. Based on a conventional ROI signaling OCC system architecture with two predefined tasks, different AI techniques were considered in terms of performance enhancement in the system under investigation to address the challenges of vehicular environments. For the ROI autodetection and tracking model, YOLOv2 was identified as a suitable solution for real-time applications, delivering an acceptable accuracy in boundary box detection, intended to cover two front/tail vehicle lights for decoding low-rate ROI signals. We also investigated the process of LED segmentation within each bounding box using the IMAQ Count Objects module within LabVIEW. The average intensity of the light of two vehicles was compared to a threshold value to determine whether each light source was of a low or a high-dimming level following the low-rate ROI signal. All LED center intensities within two vehicle lights were extracted to form the input to our neural-network-based decoder, which was used to decode a high-rate data stream. Moreover, we proposed a novel



end-to-end channel coding method and a neural-network-based decoding technique, which provided significant improvements in the SER performance compared to the use of the convolution code over the vehicular channel model. In the near future, the design of AI-based optical vehicular communication systems could benefit from the development of powerful computing hardware (e.g., the quantum computing hardware being developed by the Google AI team) and multiple megaframes-per-second cameras on the receiver side, paving the way for future commercialization of vehicular OCC technologies.

## REFERENCES

- [1] Y. Kaymak, S. Fathi-Kazerooni, and R. Rojas-Cessa, "Indirect diffused light free-space optical communications for vehicular networks," *IEEE Commun. Lett.*, vol. 23, no. 5, pp. 814–817, May 2019.
- [2] Z. Zhang, J. Dang, L. Wu, H. Wang, J. Xia, W. Lei, J. Wang, and X. You, "Optical mobile communications: Principles, implementation, and performance analysis," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 471–482, Jan. 2019.
- [3] A.-M. Cailean and M. Dimian, "Current challenges for visible light communications usage in vehicle applications: A survey," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2681–2703, 4th Quart., 2017.
- [4] T. Nguyen, A. Islam, T. Hossan, and Y. M. Jang, "Current status and performance analysis of optical camera communication technologies for 5G networks," *IEEE Access*, vol. 5, pp. 4574–4594, 2017.
- [5] M. Z. Chowdhury, M. T. Hossan, A. Islam, and Y. M. Jang, "A comparative survey of optical wireless technologies: Architectures and applications," *IEEE Access*, vol. 6, pp. 9819–9840, 2018.
- [6] M. Shahjalal, M. K. Hasan, M. Z. Chowdhury, and Y. M. Jang, "Smartphone camera-based optical wireless communication system: Requirements and implementation challenges," *Electronics*, vol. 8, no. 8, p. 913, Aug. 2019.
- [7] H.-W. Chen, S.-S. Wen, X.-L. Wang, M.-Z. Liang, M.-Y. Li, Q.-C. Li, and Y. Liu, "Color-shift keying for optical camera communication using a rolling shutter mode," *IEEE Photon. J.*, vol. 11, no. 2, pp. 1–8, Apr. 2019.
- [8] W. A. Cahyadi, Y. H. Kim, Y. H. Chung, and C.-J. Ahn, "Mobile phone camera-based indoor visible light communications with rotation compensation," *IEEE Photon. J.*, vol. 8, no. 2, pp. 1–8, Apr. 2016.
- [9] L. Shanguan, Z. Zhou, Z. Yang, K. Liu, Z. Li, X. Zhao, and Y. Liu, "Towards accurate object localization with smartphones," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 10, pp. 2731–2742, Oct. 2014.
- [10] V. P. Rachim and W.-Y. Chung, "Multilevel intensity-modulation for rolling shutter-based optical camera communication," *IEEE Photon. Technol. Lett.*, vol. 30, no. 10, pp. 903–906, May 15, 2018.
- [11] T. Yamazato, M. Kinoshita, S. Arai, E. Souke, T. Yendo, T. Fujii, K. Kamakura, and H. Okada, "Vehicle motion and pixel illumination modeling for image sensor based visible light communication," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 9, pp. 1793–1805, Sep. 2015.
- [12] Y. Goto, I. Takai, T. Yamazato, H. Okada, T. Fujii, S. Kawahito, S. Arai, T. Yendo, and K. Kamakura, "A new automotive VLC system using optical communication image sensor," *IEEE Photon. J.*, vol. 8, no. 3, pp. 1–17, Jun. 2016.
- [13] I. Takai, S. Ito, K. Yasutomi, K. Kagawa, M. Andoh, and S. Kawahito, "LED and CMOS image sensor based optical wireless communication system for automotive applications," *IEEE Photon. J.*, vol. 5, no. 5, Oct. 2013, Art. no. 6801418.
- [14] A.-M. Cailean, M. Dimian, and A. Done, "Enhanced design of visible light communication sensor for automotive applications: Experimental demonstration of a 130 meters link," in *Proc. Global LIFI Congr. (GLC)*, Paris, France, Feb. 2018, pp. 1–4.
- [15] T. L. Pham, H. Nguyen, T. Nguyen, and Y. M. Jang, "A novel neural network-based method for decoding and detecting of the DS8-PSK scheme in an OCC system," *Appl. Sci.*, vol. 9, no. 11, p. 2242, May 2019.
- [16] S.-A. Avatamanitei, A.-M. Cailean, E. Zadobrischi, A. Done, M. Dimian, and V. Popa, "Intensive testing of infrastructure-to-vehicle visible light communications in real outdoor scenario: Evaluation of a 50 meters link in direct sun exposure," in *Proc. Global LIFI Congr. (GLC)*, Paris, France, Jun. 2019, pp. 1–5.
- [17] M. D. Thieu, T. L. Pham, T. Nguyen, and Y. M. Jang, "Optical-Rol-Signaling for vehicular communications," *IEEE Access*, vol. 7, pp. 69873–69891, 2019.
- [18] T. Nguyen, A. Islam, and Y. Min Jang, "Region-of-interest signaling vehicular system using optical camera communications," *IEEE Photon. J.*, vol. 9, no. 1, pp. 1–20, Feb. 2017.
- [19] P. Ji, H.-M. Tsai, C. Wang, and F. Liu, "Vehicular visible light communications with LED taillight and rolling shutter camera," in *Proc. IEEE 79th Veh. Technol. Conf. (VTC Spring)*, May 2014, pp. 1–6.
- [20] V. T. B. Tram and M. Yoo, "Vehicle-to-vehicle distance estimation using a low-resolution camera based on visible light distance communications," *IEEE Access*, vol. 6, pp. 4521–4527, 2018.
- [21] T. L. Pham, T. Nguyen, M. D. Thieu, H. Nguyen, H. Nguyen, and Y. M. Jang, "An artificial intelligence-based error correction for optical camera communication," in *Proc. 11th Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Zagreb, Croatia, Jul. 2019, pp. 137–140.
- [22] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, Feb. 2018.
- [23] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8549–8560, Sep. 2018.
- [24] Y. Yang, F. Gao, X. Ma and S. Zhang, "Deep learning-based channel estimation for doubly selective fading channels," *IEEE Access*, vol. 7, pp. 36579–36589, 2019.
- [25] R. D. Roberts. (Jan. 17, 2016.). *Intel OCC Proposal*. Accessed: Sep. 9, 2019. [Online]. Available: <https://mentor.ieee.org/802.15/dcn/16/15-16-0006-01-007a-intel-occ-proposal.pdf>
- [26] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Process. Mag.*, vol. 13, no. 3, pp. 43–64, May 1996.
- [27] H. Zhang, D. Wipf, and Y. Zhang, "Multi-observation blind deconvolution with an adaptive sparse prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 8, pp. 1628–1643, Aug. 2014.
- [28] T.-C. Lin, L. Hou, H. Liu, Y. Li, and T.-K. Truong, "Reconstruction of single image from multiple blurry measured images," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2762–2776, Jun. 2018.
- [29] J. Hui. (Mar. 18, 2018). *Real-time Object Detection With YOLO, YOLOv2 and Now YOLOv3*. A Medium Corporation. Accessed: Sep. 9, 2019. [Online]. Available: [https://medium.com/@jonathan\\_hui/real-time-object-detection-with-yolo-yolov2-28b1b93e2088](https://medium.com/@jonathan_hui/real-time-object-detection-with-yolo-yolov2-28b1b93e2088)
- [30] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*. [Online]. Available: <http://arxiv.org/abs/1804.02767>
- [31] M. K. Hasan, M. Z. Chowdhury, M. Shahjalal, V. T. Nguyen, and Y. M. Jang, "Performance analysis and improvement of optical camera communication," *Appl. Sci.*, vol. 8, no. 12, p. 2527, Dec. 2018.
- [32] Y. M. Jang and T. Nguyen. (Jan. 21, 2016). *Kookmin PHY Sub-Proposal for ISC Using Temporal Scheme (C-OOK)*. Kookmin University. Accessed: Sep. 9, 2019. [Online]. Available: <https://mentor.ieee.org/802.15/dcn/16/15-16-0013-02-007a-kookmin-phy-sub-proposal-for-isc-using-temporal-scheme-c-ook.pptx>
- [33] Y. M. Jang, T. Nguyen, and H. C. Hyun. (Jan. 19, 2016). *Kookmin University PHY Sub-Proposal for ISC Using a Compatible M-FSK Scheme (CM-FSK)*. Kookmin University. Accessed: Sep. 9, 2019. [Online]. Available: <https://mentor.ieee.org/802.15/dcn/16/15-16-0014-01-007a-kookmin-university-phy-sub-proposal-for-isc-using-a-compatible-m-fsk-scheme-cm-fsk.pptx>
- [34] Y. M. Jang and T. Nguyen. (Jan. 21, 2016). *Kookmin University PHY Sub-Proposal for ISC Using Dimmable Spatial M-PSK (DSM-PSK)*. Kookmin University. Accessed: Sep. 9, 2019. [Online]. Available: <https://mentor.ieee.org/802.15/dcn/16/15-16-0015-02-007a-kookmin-university-phy-sub-proposal-for-isc-using-dimmable-spatial-m-psk-dsm-psk.pptx>
- [35] M. Murugavel. (Apr. 2019). *Object\_Tracking\_Ref\_Previous\_N\_Frame.py*. Accessed: Sep. 9, 2019. [Online]. Available: <https://gist.github.com/ManivannanMurugavel/85cab9e1549a0722cc06712ab92e2c02>
- [36] M. Murugavel. (Apr. 2019). *Object Tracker by Using YOLO*. Accessed: Sep. 9, 2019. [Online]. Available: [https://www.youtube.com/watch?v=N2voR\\_qi6Y](https://www.youtube.com/watch?v=N2voR_qi6Y)
- [37] R. D. Roberts, "A MIMO protocol for camera communications (CamCom) using undersampled frequency shift ON-OFF keying (UFSOOK)," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Atlanta, GA, USA, Dec. 2013, pp. 1052–1057.

- [38] H. Liu. *LightNet: Light-Weight Networks for Semantic Image Segmentation*. Accessed: Sep. 25, 2019. [Online]. Available: <https://ansleliu.github.io/LightNet.html>
- [39] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," May 2015. *arXiv:1505.04597*. [Online]. Available: <https://arxiv.org/abs/1505.04597>
- [40] H. Liu. *Mask-YOLO: Efficient Instance-level Segmentation Network Based on YOLO-V2*. Accessed: Sep. 25, 2019. [Online]. Available: <https://ansleliu.github.io/MaskYOLO.html>
- [41] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," 2017, *arXiv:1703.06870*. [Online]. Available: <https://arxiv.org/abs/1703.06870>
- [42] D. Bolya, C. Zhou, F. Xiao, and Y. Jae Lee, "YOLACT: Real-time instance segmentation," 2019, *arXiv:1904.02689*. [Online]. Available: <http://arxiv.org/abs/1904.02689>
- [43] J. Sun. *Mask-YOLO: A Multi-task Learning Architecture for Object Detection and Instance Segmentation*. Accessed: Sep. 9, 2019. [Online]. Available: <https://github.com/jianing-sun/Mask-YOLO>
- [44] RideScapes. (Mar. 23, 2019). *ASMR Highway Driving in the Rain at Night (No Talking, No Music)—Daegu to Seoul, Korea*. Accessed: Dec. 30, 2019. [Online]. Available: <https://www.youtube.com/watch?v=oqLiK-U0Jzs&t=3223s>
- [45] M. Murugavel. (Mar. 28, 2018). *How to Train Multiple Objects in YOLOv2 Using Your Own Dataset*. Accessed: Jan. 3, 2020. [Online]. Available: [https://medium.com/@manivannan\\_data/how-to-train-multiple-objects-in-yolov2-using-your-own-dataset-2b4fee898f17](https://medium.com/@manivannan_data/how-to-train-multiple-objects-in-yolov2-using-your-own-dataset-2b4fee898f17)
- [46] Anakin. (Jun. 10, 2018). *How to Train Multiple Objects in YOLOv2 Using Your Own Dataset*. GitHub. Accessed: Apr. 4, 2020. [Online]. Available: <https://github.com/anakin1028/Tensorflow-Examples/blob/master/notebooks/basic/Calculate%20FLOPs%20and%20Number%20of%20Model%20Parameters.ipynb>
- [47] J. Redmon and A. Farhadi. (Jun. 10, 2018). *ImageNet Classification*. Accessed: Apr. 4, 2020. [Online]. Available: <https://pjreddie.com/darknet/imagenet/#reference>



**VAN BUI** received the bachelor's degree in electronic engineering from the Danang University of Science and Technology (DUT), Vietnam, in 2015. He is currently pursuing the M.Sc. degree in electronics engineering with Kookmin University (KMU), South Korea. His research interests include AI applications and energy.



**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. He was with the Electronics and Telecommunications Research Institute (ETRI), from 1987 to 2000. Since 2002, he has been with the School of Electrical Engineering, Kookmin University, Seoul, South Korea,



where he has been the Director of the LED Convergence Research Center, since 2010, and the Internet of Energy Research Center, since 2018. He was a Visiting Professor with Cornell University, Ithaca, NY, USA, in 2008 and 2009. He has organized several conferences and workshops, such as the International Conference on Ubiquitous and Future Network (ICUFN), the International Conference on ICT Convergence (ICTC), the International Conference on Information Networking (ICOIN), the International Conference on Artificial Intelligence in Information and Communication (ICAIC), and the International Workshop on Optical Wireless Communication Networks. His research interests include 5G/6G mobile communications, internet of energy, AI platform, e-Health, public safety, optical wireless communications, optical camera communication, and the Internet of Things. He is currently a Life Member of KICS. He received the Young Science Award from the Korean Government, from 2003 to 2006, and the KICS Dr. Irwin Jacobs Award in 2017. He had served as the Founding Chair of the Korean Institute of Communications and Information Sciences (KICS) Technical Committee on Communication Networks, in 2007 and 2008. He had served as the Executive Director of KICS from 2006 to 2014, the Vice President of KICS, from 2014 to 2016, the Executive Vice President of KICS for 2018, and the President of KICS for 2019. He had served as the Chairman of the IEEE 802.15 Optical Camera Communications Study Group in 2014 and also the Chairman of the IEEE 802.15.7m Optical Wireless Communications Task Group. He is currently the Chairman of the IEEE 802.15 Vehicular Assistive Technology Interest Group. He had been the Steering Chair of the Multiscreen Service Forum, since 2011 and the Society Safety System Forum, since 2015. He is currently serving as the Co-Editor-in-Chief of *ICT Express (Elsevier)*.

**TUNG LAM PHAM** received the bachelor's degree in Control and Automation Engineering Technology from the Hanoi University of Science and Technology (HUST), Vietnam, in 2017, and the M.Sc. degree in electronics engineering from Kookmin University (KMU), South Korea, in 2020, and earned the Excellent Researcher Award. He is currently working as a full-time AI Software Developer at Gridone Co., Ltd., South Korea. His research interests include Optical Wireless Communications (OWC), Optical Camera Communications (OCC), Optical V2X Communications, Neural Networks, Image Processing and Artificial Intelligence (AI). He is currently working as a full-time AI software developer at Gridone Co, Ltd, Korea.



**MD. SHAHJALAL** (Student Member, IEEE) received the B.Sc. degree in electrical and electronic engineering (EEE) from the Khulna University of Engineering and Technology (KUET), Bangladesh, in May 2017, and the M.Sc. degree in electronics engineering from Kookmin University, South Korea, in August 2019, where he is currently pursuing the Ph.D. degree in electronics engineering. His research interests include optical wireless communications, wireless security, non-orthogonal multiple access, the Internet of Things, low-power wide-area networks, and 6G mobile communications. He received the Excellent Student Award for his M.Sc. degree.