

Article

An LED Array-Based 2D MIMO OCC System with Deep Learning for Mobile Environments

Oanh Giap, Huy Nguyen and Yeong Min Jang * 

Department of Electronics Engineering, Kookmin University, Seoul 02707, Republic of Korea;
oanh.giap@kookmin.ac.kr (O.G.); ngochuy@kookmin.ac.kr (H.N.)

* Correspondence: yjang@kookmin.ac.kr; Tel.: +82-2-910-5068

Abstract

Optical wireless communication (OWC) has emerged as a complementary technology to conventional radio frequency (RF)-based communication systems, particularly in scenarios requiring low electromagnetic interference, enhanced security, and efficient spectrum utilization. Within various OWC approaches, optical camera communication (OCC) has attracted increasing attention due to its ability to utilize commercially available image sensors as receivers. This paper presents a 2D multiple-input-multiple-output (MIMO) OCC system based on light-emitting diode (LED) arrays for reliable communication in mobile environments. The proposed system employs on-off keying (OOK) modulation, which supports both rolling shutter and global shutter cameras. To improve decoding reliability under mobility conditions, a deep learning-based decoding model is introduced to enhance LED state detection compared with conventional zero-crossing approaches. In addition, a sequence number-based synchronization is implemented to compensate for frame rate variation and packet missing in a real-time environment. Besides that, by applying YOLOv13 for light source detection and tracking, we can achieve 98% accuracy at 3 m/s velocity. Experimental results show reliable communication performance at transmission distances of up to 22 m under various mobility conditions. Furthermore, the proposed system is validated through real-time environmental data transmission using temperature and humidity sensors with 20 links. The results indicate that the proposed scheme provides stable and reliable OCC performance for mobility Internet of Things (IoT) applications.

Keywords: optical camera communication; LED array; multiple-input and multiple-output on-off keying; deep learning-based decoder



Academic Editor: Ephraim Suhir

Received: 4 June 2026

Revised: 22 June 2026

Accepted: 24 June 2026

Published: 1 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The IoT has become an indispensable part of modern life, with applications in smart homes, health monitoring, traffic control, smart grids, and smart factories [1]. Based on some standardized Internet protocols, massive devices can be easily integrated into IoT networks [2]. The IoT systems include not only electronic modules and micro-controller units but also living entities such as plants, animals, and humans. They are designed to enable seamless global connectivity, although this increasing automation may reduce direct human interaction [3]. As a key component of the Fourth Industrial Revolution, IoT relies heavily on wireless communication technologies, primarily those based on RF waves, such as LoRa, ZigBee, Wi-Fi, and Sigfox, to enable worldwide interconnection. However, a lot of concerns have been raised regarding the potential health effects of electromagnetic

RF radiation, particularly in sensitive environments such as schools and hospitals. To address these concerns, researchers are exploring safer and more effective alternatives to conventional RF-based communication systems.

OWC has emerged as a substitute for traditional RF-based systems by leveraging the visible light spectrum for data transmission. OWC technologies are primarily categorized into three methodologies: Visible Light Communication (VLC), Light Fidelity (Li-Fi), and optical camera communication (OCC). While VLC and Li-Fi rely on photodiodes to sense LED intensity variations, OCC utilizes image sensors—specifically rolling shutter or global shutter cameras—as receivers [4]. Compared with RF-based systems, OWC offers several advantages:

- Although RF systems are widely used, they are associated with health concerns and electromagnetic interference (EMI), whereas visible light poses no known health risks [5].
- OWC enables secure and efficient data transmission under line-of-sight conditions.
- The visible light spectrum offers bandwidth that is more than 1000 times greater than that of RF signals.

Motivated by these advantages, significant research efforts have been devoted to the development of OWC technologies. The IEEE 802.15.7-2011 standard [6] originally defined OWC protocols with a focus on VLC. The revised IEEE 802.15.7-2018 standard [7] expanded this framework to include: VLC (detailed in the 2011 standard) and OCC (using image sensors or cameras to receive data from light sources).

LEDs are well-suited for next-generation OWC systems due to advances in LED technology, including long operational lifetimes, high energy efficiency, and fast switching speeds [8,9]. Although RF-based systems remain widely used in communication and monitoring applications, their sensitivity to electromagnetic interference and potential biological effects pose significant challenges. In contrast, OWC systems do not generate electromagnetic interference, making them an attractive alternative. In VLC and Li-Fi systems, LEDs typically transmit data using OOK modulation, which is detected by photodiodes. The application of MIMO techniques has further improved data throughput in these systems. However, photodiode-based OWC systems are generally limited to short-range indoor communication due to signal degradation in outdoor environments. In contrast, OCC systems offer advantages in terms of transmission range, as they employ camera image sensors for reception and have demonstrated successful communication over much longer distances, reportedly up to 300 m [10]. The performance of OCC systems depends strongly on the type of image sensor used [11]. Rolling shutter-based systems rely on frame rate and rolling shutter speed, whereas global shutter cameras require the frame rate and packet rate to satisfy the Nyquist sampling requirement [12]. To achieve longer communication distances, parameters such as signal-to-noise ratio (SNR), camera focal length, and exposure time must be carefully optimized. Currently, Li-Fi systems achieve data transmission distances of up to approximately 10 m using photodiode-based receivers [13], whereas OCC continues to extend its communication range, making it a strong candidate for future optical wireless communication systems.

In OCC systems, MIMO techniques enable simultaneous data transmission from multiple light sources to a camera receiver, which do not work well in photodiodes. Within OCC frameworks, region-of-interest (RoI) signaling algorithms [14–16] are commonly employed to allow the receiver to identify and process multiple light sources that are simultaneously visible within the camera's field of view. Besides that, OCC can support better performance, by considering mobility environment, compared to VLC/Li-Fi systems.

A variety of modulation techniques for OCC have been published in the existing literature. One method exploits the rolling shutter effect to implement OOK modulation;

however, this approach is limited by a short operating range and reduced performance. To address low data rates, rolling orthogonal frequency-division multiplexing (rolling shutter-OFDM) was introduced in [17,18]. Similarly, the system proposed in [19] employs a MIMO scheme based on color intensity modulation using a high-speed global shutter camera operating at 330 frames per second (fps). In addition, global shutter cameras are expensive and not widely available, and the use of color channels introduces higher bit error rate (BER) and reduced range compared with OOK-based systems. The localization of IoT devices is one area of important work in IoT systems; Refs. [20,21] introduced the localization algorithms in integrated sensing and communication to improve communication performances. Our work showed reliable communication performance at transmission distances of up to 22 m under various mobility conditions. Furthermore, the proposed system is validated through real-time environmental data transmission using temperature and humidity sensors with 20 links. It makes sure that it is suitable for industrial IoT smart city applications by supporting any type of cameras in the market. Besides that, we can scale up to a hundred connectivity links in big area IoT.

This study presents a 2D-MIMO OCC system based on LED arrays for IoT applications. The proposed system focuses on improving communication reliability in practical mobility environments while maintaining compatibility with commercially available cameras. The main contributions of this work are summarized as follows:

- Suitable for almost all commercial cameras: The proposed system supports both rolling shutter and global shutter cameras through the combination of RoI detection and adaptive exposure control.
- Data synchronization: To address frame rate variations in real OCC system, a sequence number-based synchronization algorithm is proposed. This method enables the receiver to identify packet loss, duplication, and timing mismatches during the decoding process.
- Mobility support: The proposed system is evaluated under different transmission distances, camera types, and mobility conditions. By applying YOLOv13 for LED detection instead of conventional algorithms, we can support a high-mobility environment to improve OCC performance.
- Deep learning decoder: A deep learning-based decoding method is employed to improve OCC performance under various noise and mobility conditions. The proposed approach is evaluated and compared with conventional zero-crossing decoding methods.
- Multi-link support: By applying deep learning, we can support simultaneous multi-links at long distance and mobility environments with low bit error rate.

The remainder of this paper is organized as follows. The technical architecture and system design are detailed in Section 2. Evaluation of the system performance through experimental results is presented in Section 3. Finally, Section 4 provides a summary and concludes this study.

2. System Design

The main architecture of the proposed OCC scheme is based on the control of light intensity to enable data transmission. A key aspect of this design is the modulation scheme, which directly affects system throughput and link reliability. OOK stands as a prevalent modulation method favored for its straight-forward implementation. As an amplitude-shift keying (ASK), OOK encodes binary information through two distinct light levels: an 'ON' state for bit '1' and an 'OFF' state for bit '0'.

In this study, an LED matrix-based OCC scheme is proposed for IoT applications. The LED array is arranged in a spatial frame format that enables efficient data detection and

decoding using RoI algorithms. Each LED in the array functions as an independent light source and transmission channel, enabling simultaneous multi-channel data transmission. The system architecture of the proposed system is illustrated in Figure 1, which presents the OOK-based 2D MIMO transceiver architecture and highlights the interaction between the transmitter (LED array) and the receiver (camera). The following subsections describe the functional components and processing flow of the system, including data encoding, LED control, image acquisition, signal processing, and data decoding. The proposed architecture is designed for interoperability with commercial cameras and incorporates several technical enhancements.

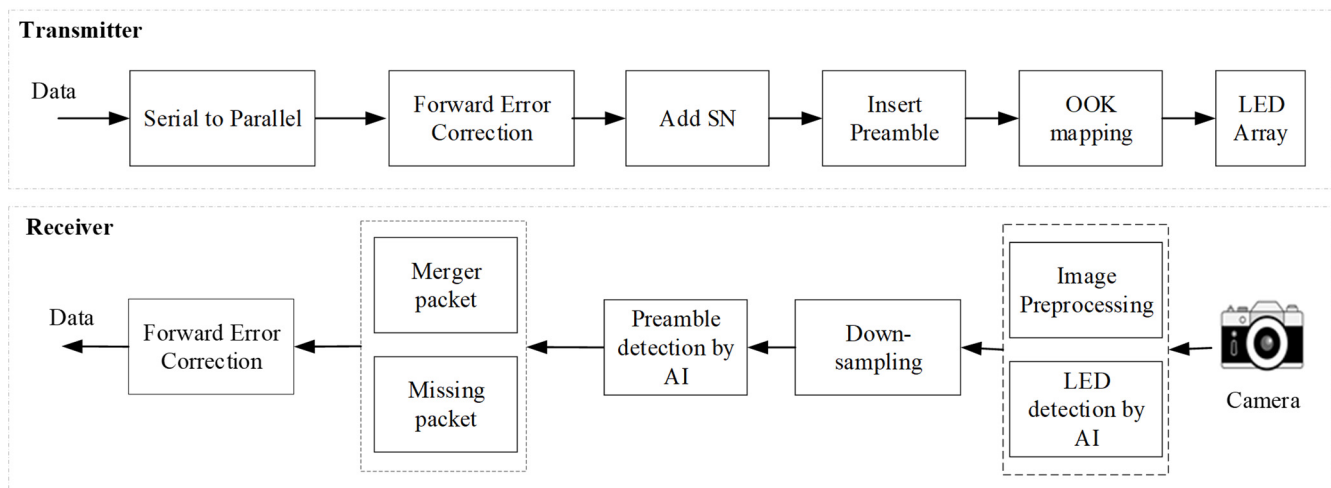


Figure 1. System architecture.

2.1. Channel Coding

Channel coding, or Forward Error Correction (FEC), is a fundamental element in digital systems, primarily used to reduce bit errors during transmission. This algorithm functions by appending parity information to the data stream at the source, allowing the receiver to reconstruct the original message despite the presence of distortion and noise. From the proposed architecture, we apply a coding scheme to the LED matrix's configuration, ensuring stable communication and enhanced data decoding.

2.2. Start-of-Frame and Sequence Number Insertion

To synchronize the sequence data frame and detect the received signal at the receiver, a synchronization part and a header frame are inserted into each data packet. Although the preamble does not carry information, it is necessary for reliable frame alignment. The preamble length is optimized to ensure accurate frame detection while minimizing overhead.

In this work, a 6-bit preamble Synchronization Field (SF = "011100") is employed, providing a clear reference for frame detection in each received image. Following the preamble, each data packet includes a sequence number (SN) to synchronize and detect the potential frame rate mismatches between the transmitter and the receiver. The SN enables identification of missing packets due to undersampling, elimination of redundant packets caused by oversampling, and reconstruction of large data packets from adjacent frames. The performance impact of the SN part is discussed in Section 4.

2.3. LED Detection

Instead of relying on conventional RoI-based detection methods, a deep learning-based LED detection using YOLOv13 algorithm is implemented to accurately detect LED arrays

in each received image and to maintain stable identification in a real-time environment. Frame structure of the proposed scheme includes two primary parts: (i) anchor LEDs, corresponding to the four corner LEDs that remain continuously illuminated and serve as geometric reference points, and (ii) data LEDs, which transmit information. The dataset of YOLOv13 includes 10,000 images captured under varying real-time conditions, including transmission distances from 2 to 22 m, different illumination levels, background light, camera parameters (exposure time, contracts, brightness, focal length, etc.), and motion blur. The images in dataset are split into training, validation, and test datasets with by following ratio: 60% dataset for training, 20% dataset for validation, and 20% for test model. Data augmentation techniques such as Gaussian blur, random brightness and contrast adjustments were applied to improve model robustness. During real-time operation, each camera frame is processed by the YOLOv13 model to generate bounding boxes, confidence scores, and class labels.

The Deep Simple Online and Realtime Tracking (Deep SORT) algorithm assigns a stable track ID to each detected LED, ensuring consistent channel association across frames. The RoI corresponding to each track ID is extracted and stored in a temporal buffer to form an intensity sequence over time. This sequence is subsequently passed to signal processing modules, including downsampling, deep learning-based decoder and preamble detection, and SN-based packet reconstruction. This approach differs from conventional RoI methods, which rely on static thresholding or manual calibration, by enabling adaptive and robust LED localization in dynamic environments.

The integration of YOLOv13 and Deep SORT provides several benefits:

- Improved RoI detection accuracy under varying lighting and background conditions;
- Supporting multi-LED and multi-user scenarios;
- Enhanced signal consistency, resulting in reduced BER prior to OOK demodulation.

Experimental parameters include image input resolution, confidence and Intersection over Union (IoU) thresholds, average inference latency per frame, and anchor LED detection accuracy.

2.4. Deep Learning Decoding

Communication performance degrades as transmission distance increases due to a reduction in SNR values, making it difficult to reliably distinguish between ON and OFF signal states. To address this limitation and extend the communication range, an enhanced signal detection approach is required in place of conventional methods. In wireless transmission, matched filtering is extensively employed as an optimal linear method for SNR maximization. Its widespread adoption algorithms form a balance between operational simplicity and its effectiveness against background noise. The IEEE Std 802.15.7a-2024 standard [18] recommends matched filtering for OCC systems to improve signal reliability. However, in mobility environments, we need other methods to apply in the OCC system to achieve good performance. Although deep learning [22–25] is the future for OCC decoding, this work presents a deep learning-based alternative that outperforms standard matched filtering in current applications. By applying DL decoding, we can achieve 98% accuracy at 3 m/s velocity, at transmission distances of up to 22 m under various mobility conditions. Furthermore, the proposed system is validated through real-time environmental data transmission with 20 links.

After LED detection, OCC signals are recovered from multiple light sources via a downsampling technique, whereby a 2D MIMO is generated by sampling the central intensity of each LED. The proposed system is evaluated on a dataset of 10,000 raw samples (including preamble and payload), captured using two camera types (rolling/global shutter cameras) at distances between 2 and 22 m under varying motion conditions. The signals in

dataset are split by the following ratio: 80% dataset for training, 15% dataset for validation, and 15% for test model. Figure 2 shows the experimental result of 2D OOK-MIMO signals with (a) high SNR and (b) low SNR. With long distance, the SNR values will reduce, making it difficult to recognize the threshold value at the receiver side. The deep learning decoder architecture, depicted in Figure 3, incorporates only two hidden layers to maintain a low model complexity and prevent overfitting during the training process. After preamble detection, the DL model accurately estimates threshold values for the 2D-MIMO signals, resulting in improved decoding performance compared with conventional techniques in mobile environments. Increasing to over five hidden layers in the DL model resulted in overfitting and reduced performance on the test dataset. Figure 3 illustrates the deep learning-based decoder model, which predicts threshold values instead of relying on conventional algorithms.

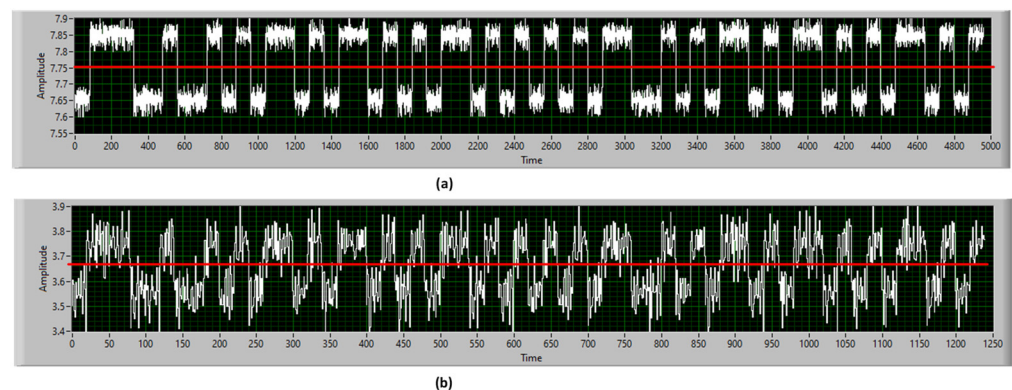


Figure 2. An experimental result of 2D OOK-MIMO signals with (a) high SNR and (b) low SNR.

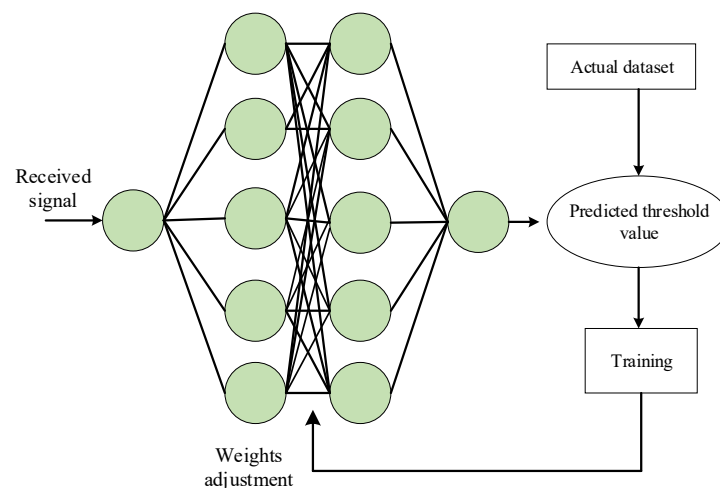


Figure 3. Structure of the deep learning decoder.

3. Implementation

3.1. Noise Modeling and Pixel- E_b/N_0

In Complementary Metal-Oxide-Semiconductor (CMOS) and Charge-Coupled Device (CCD) image sensors, pixel-level noise is commonly modeled as a Gaussian distribution. Following the stochastic framework established in [26], the noise component n is modeled as follows:

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

where s denotes the received pixel intensity, and the noise variance $\sigma^2(s)$ is defined by $\sigma^2(s) = s \times \alpha \times \alpha + \beta$, where α represents the amplitude difference between the mark

and space logic levels. The fitting coefficients α and β are system-specific parameters determined through experimental calibration, following the procedure described in [27].

To evaluate signal quality at the receiver, the pixel-level energy-to-noise ratio E_b/N_0 should be computed. Assuming a one-bit-per-symbol mapping, the estimated expression is given by

$$\text{Pixel} = \frac{E_b}{N_0} = \frac{E[s^2]}{E[n^2]} \approx \frac{a^2 \times \Delta}{a \times \alpha \times \Delta + \beta} \quad (2)$$

The parameters in Equation (2) are defined as follows:

- The parameters E_b and N_0 correspond to the energy associated with each bit and the noise's power spectral density.
- Δ is the ratio of the exposure time T_{exposure} to the bit T_{bit} , defined as $\Delta = \frac{T_{\text{exposure}}}{T_{\text{bit}}}$.

α and β are hardware-dependent constants determined through curve fitting during the initial system calibration phase.

3.2. Experimental Analysis of SNR Values

To evaluate the stability of the proposed system, an analysis of the SNR was processed under varying communication distances and camera exposure times. The primary objective of this analysis was to characterize the trade-off between the transmission range and BER by adjusting the camera's exposure settings.

Experimental Configuration and Setup:

- Transmitter Specifications: The transmitter employed an 8×8 LED matrix array powered by a 5 V supply.
- Receiver Architecture: The receiver consisted of a commercial rolling shutter camera.

The camera exposure time was incrementally varied from 100 μs to 400 μs to examine its effect on the pixel-level energy-to-noise ratio E_b/N_0 with distances ranging from 5 to 25 m. As transmission distance increases, the optical signal intensity captured by the CMOS sensor decreases. Increasing exposure time improves signal acquisition by accumulating more optical energy; however, this improvement must be balanced against potential issues such as sensor saturation and inter-symbol interference.

For an analysis of signal quality in a real-time environment, the SNR values were obtained at different communication ranges of 5, 10, 15, 20, and 25 m. The measurement procedure alternated the LED array states between "ON" (active transmission) and "OFF" (background reference) to separate signal power from ambient noise.

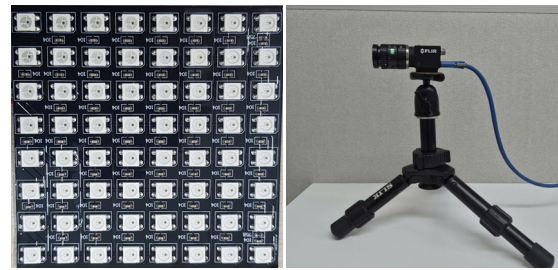
When the LEDs are in the "ON" status, the measured signal corresponds to the emitted power. In contrast, measurements obtained during the "OFF" state establish the baseline noise for each distance. The experimental SNR, expressed in decibels (dB), is calculated using the root-mean-square (RMS) formulation given by

$$\text{SNR}_{dB} = 20 \times \log \left(\frac{\sqrt{\frac{1}{n} \sum_{i=0}^{n-1} |A_i|^2}}{\sqrt{\frac{1}{n} \sum_{i=0}^{n-1} |B_i|^2}} \right) \quad (3)$$

where A_i is the discrete power sample extracted from the LED array, B_i denotes the recorded ambient noise power levels, n represents the cumulative count of samples gathered throughout the observation period.

As indicated in Equation (3), by increasing the camera exposure time, we can improve SNR values, but it reduces the bandwidth of the system. However, a relationship is observed between communication distance and signal quality. The measured SNR values decrease as the transmitter-to-receiver distance increases. Figure 4 illustrates the experimental setup

used to measure SNR, while Figure 5 presents the measured SNR values over distances from 5 to 25 m under different exposure times.



- | | |
|--------------------|-------------------------|
| LED matrix: | Camera: |
| • Model: 8×8 Array | • Type: Rolling Shutter |
| • Input: 5V DC | • Frame Rate: 60fps |

Figure 4. SNR measurement devices.

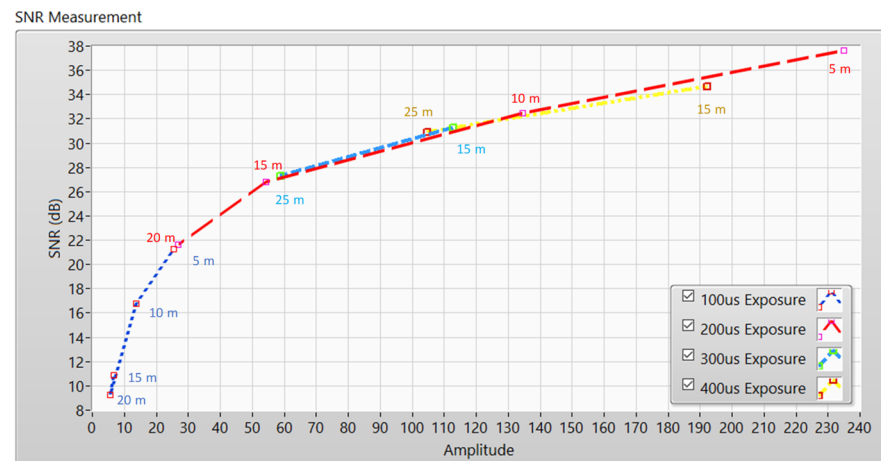


Figure 5. SNR measurement results.

3.3. Bit Error Rate (BER) Analysis for Optical OOK Modulation

The transformation of the received optical signal into its electrical equivalent, denoted as $r(t)$, can be mathematically formulated as follows within the OOK framework:

$$r(t) = I(t) + \sum_{i=-\infty}^{+\infty} I(t) \times a_i \times g(t - iT_{symbol}) + n(t) \quad (4)$$

Here $a_i \in \{0, 1\}$ denotes the i -th OOK symbol level, while $g(t)$ and T_{symbol} denote the rectangular pulse-shaping function and the symbol period, respectively. Under the assumption that additive white Gaussian noise (AWGN) is the primary channel impairment, the bit error rate (BER), or P_e , for OOK modulation is estimated by

$$P_e = \frac{1}{2} \operatorname{erfc} \left(\sqrt{\frac{E_b}{2\sigma_n^2}} \right) \quad (5)$$

where E_b represents the bit energy and σ_n^2 denotes the noise variance. Accordingly, the received signal under AWGN conditions can be characterized based on the transmitted logical state as follows:

$$r(t) = \begin{cases} n(t) & a_i = 0 \\ 2I(t) + n(t) & a_i = 1 \end{cases} \quad (6)$$

Figure 6 shows the bit error rate performance of OOK scheme considering the pixel E_b/N_0 as mentioned in Equations (4)–(6).

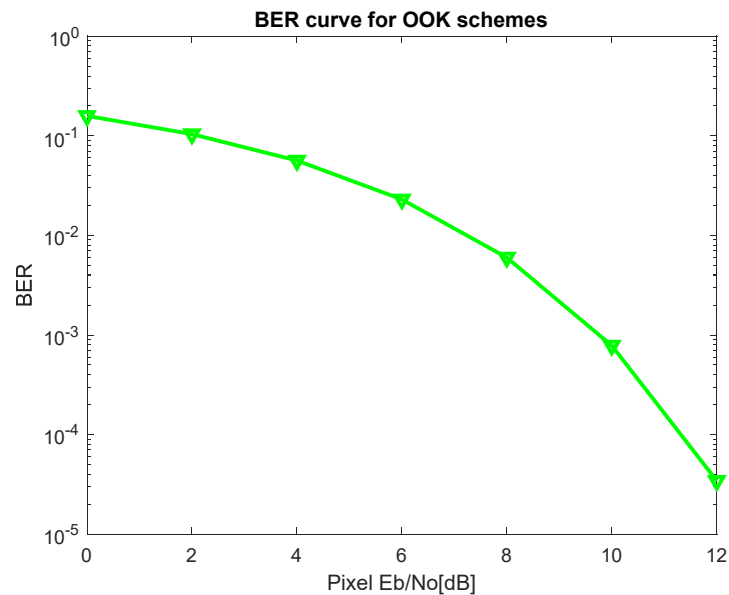


Figure 6. BER curve for the optical OOK modulation.

3.4. Proposed Methodology

Figure 7 presents the spatial arrangement of the signaling scheme, which incorporates an 8×8 LED matrix. Within this structure, the LEDs situated at the four outermost vertices work as reference anchors, ensuring reliable acquisition of spatial coordinates and corner detection. By leveraging the geometric coordinates of these reference points, the receiver applies a perspective transformation to accurately map the LED array and identify the positions of the internal LEDs.

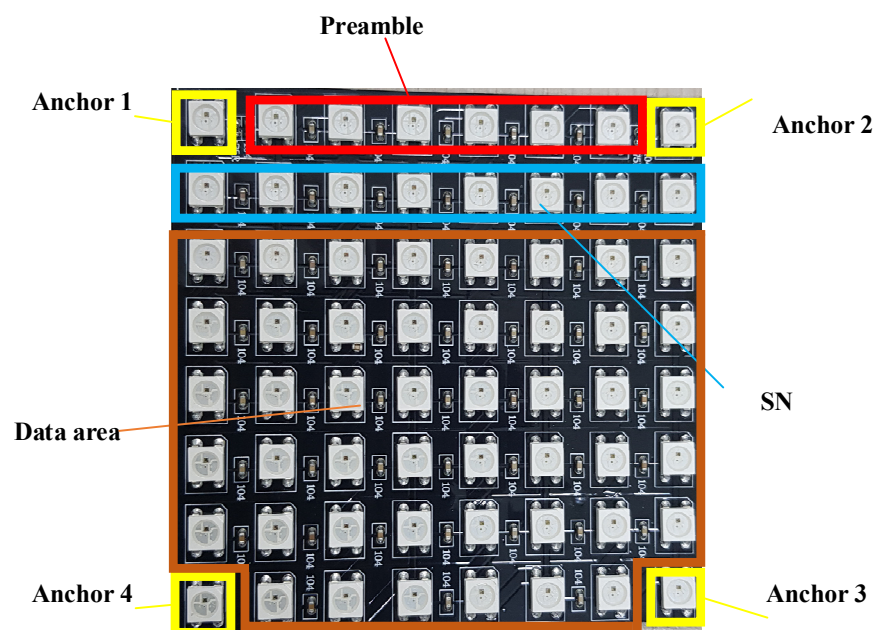


Figure 7. Spatial frame structure of the LED matrix.

Within this architecture, 46 LED units are dedicated exclusively to data payload transmission. As shown in Figure 7, this allocation ensures a fixed spatial frame structure,

allowing the RoI algorithm to reliably distinguish data-carrying LEDs from the background, even when the camera orientation changes.

The proposed structure includes four anchor LEDs positioned at each corner of the array. In addition, eight LEDs in the SN field are reserved as training signals to provide reference levels for the deep learning-based decoder, enabling more accurate classification of ON and OFF states compared to conventional zero-crossing. A preamble sequence is inserted at the beginning of each frame to enable reliable frame boundary detection. To further enhance signal reliability, a deep learning decoder is applied, improving the SNR and extending the achievable communication distance.

To improve system performance under mobility conditions, the proposed architecture incorporates DL-based algorithms for robust detection and real-time tracking multiple LED arrays. To address the challenges posed by temporal desynchronization and varying frame rates between the transmitter and the receiver, an SN is integrated into each data packet.

The synchronization algorithms should consider for two primary sampling conditions:

- **Oversampling:** When the camera's frame rate is higher than the transmission rate, SNs enable the receiver to detect and filter out redundant data packets.
- **Undersampling:** In cases where the sampling rate falls below the packet rate, the system utilizes SNs to pinpoint dropped packets and identify discontinuities in the data stream.

The packet rate is determined by the number of transmissions occurring within a specified time interval (e.g., 30 packets/s). Through SN values, the receiver maintains data integrity and accurately reconstructs the original data sequence despite temporal inconsistencies.

3.4.1. Oversampling

Oversampling occurs when the camera's frame rate is at least twice the transmission frequency of the LED. While this satisfies basic sampling requirements, it causes each packet to be recorded across several consecutive frames, creating redundancies that must be managed during data reconstruction. To address this issue, an SN is included in each data sub-packet (DS), ensuring that all duplicated DSs contain identical payloads and corresponding SN values. By referencing the SN, the receiver can identify and discard redundant packets caused by oversampling. Consequently, only packets with unique, incrementally increasing SN values, for example, n , $n + 1$, $n + 2$, are retained and correctly merged to reconstruct the transmitted data stream. By doing that, we can reduce the redundancy packets to merge the payloads.

3.4.2. Undersampling

Undersampling occurs when the frame rate of the camera is insufficient relative to the transmitter's data rate. In such cases, certain data sub-packets (DSs) are not captured, resulting in discontinuities in the received frame sequence that hinder accurate packet reconstruction. To mitigate this issue, the proposed architecture embeds an SN within each transmitted packet, enabling the receiver to detect and identify missing payloads. For example, if a DS associated with packet n is followed by a packet with an SN $n + 2$, the receiver identifies packet $n + 1$ as missing. This comparison of sequential SN values preserves communication integrity despite frame rate mismatches. The system's ability to detect lost payloads depends on the length of the SN. For instance, a 3-bit SN configuration allows the receiver to track and identify up to seven consecutive missing packets. When non-sequential SN values are detected, the system infers the loss of intermediate data segments, enabling appropriate recovery mechanisms to maintain overall data throughput.

By applying the SN, we can detect the missing packets in undersampling cases, and then we can notify the transmitter to send the missing packet again.

3.4.3. Implementation Results

To evaluate the reliability of the proposed system against frame rate fluctuations, the scheme was experimentally validated using various cameras to make sure that this scheme can adapt to any scenario. This combination of hardware platforms enables a comprehensive assessment of communication performance under varying resolutions. This approach ensures a more comprehensive dataset, ultimately leading to better-performing object detection algorithms. The YOLO algorithm is trained on an NVIDIA GeForce RTX 3050 with 30 epochs and it is trained with our dataset. The evaluation results of the training process are illustrated in Figure 8.

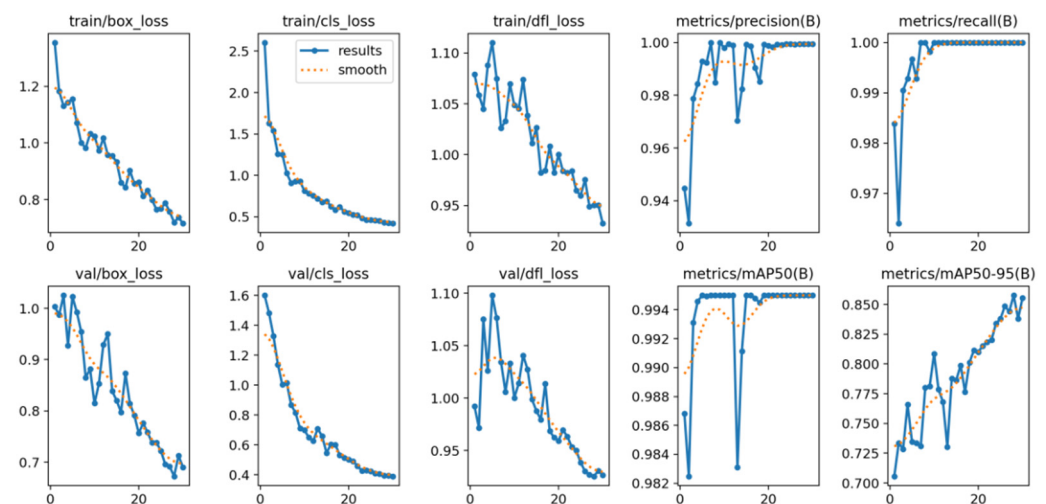


Figure 8. Object detector YOLOv13 training results.

A critical parameter of this architecture is the selection of an SN length. The SN parameters were optimized to ensure reliable synchronization and compatibility between the different frame rates of cameras. This design enables accurate data reconstruction within the asynchronous processing environment, even under nonideal sampling conditions. Using a 16×16 LED matrix, the system achieves uncoded data rates of up to 7.68 kbps and coded data rates of up to 5.76 kbps with good BER values. However, with 8×8 LED matrix, we only achieve uncoded data rates of up to 1.92 kbps and coded data rates of up to 1.44 kbps. In our work, we demonstrated the 16×16 LED matrix at 22 m distance with 20 links simultaneously, with a latency of lower than 30 milliseconds. To achieve longer range, we can use better focal length to improve the communication distance. To achieve good performance under extreme lighting conditions or high ambient noise, we should control camera parameters (exposure time, contracts, brightness, focal length, etc.).

The proposed scheme was implemented and evaluated using multiple camera devices, including a webcam, a CCTV camera and a rolling shutter camera, to investigate the effects of camera frame rate variation. An asynchronous processing incorporates non-synchronous decoding, packet merging, and missing-data detection. The SN length (3 bits) was determined based on the integrated system configuration and operational requirements. Figures 9 and 10 illustrate the quantized intensity profiles under varying exposure times and the integrated system configuration, respectively. The experimental setup using a rolling shutter camera (with exposure time of 100 μ s, resolution of 1920×1080 pixels) is shown in Figure 10. As shown in Figure 10, the 20 links LED arrays were allocated in panel with a distance of 10 cm between each LED array. The packet rate is 30 packets per second.

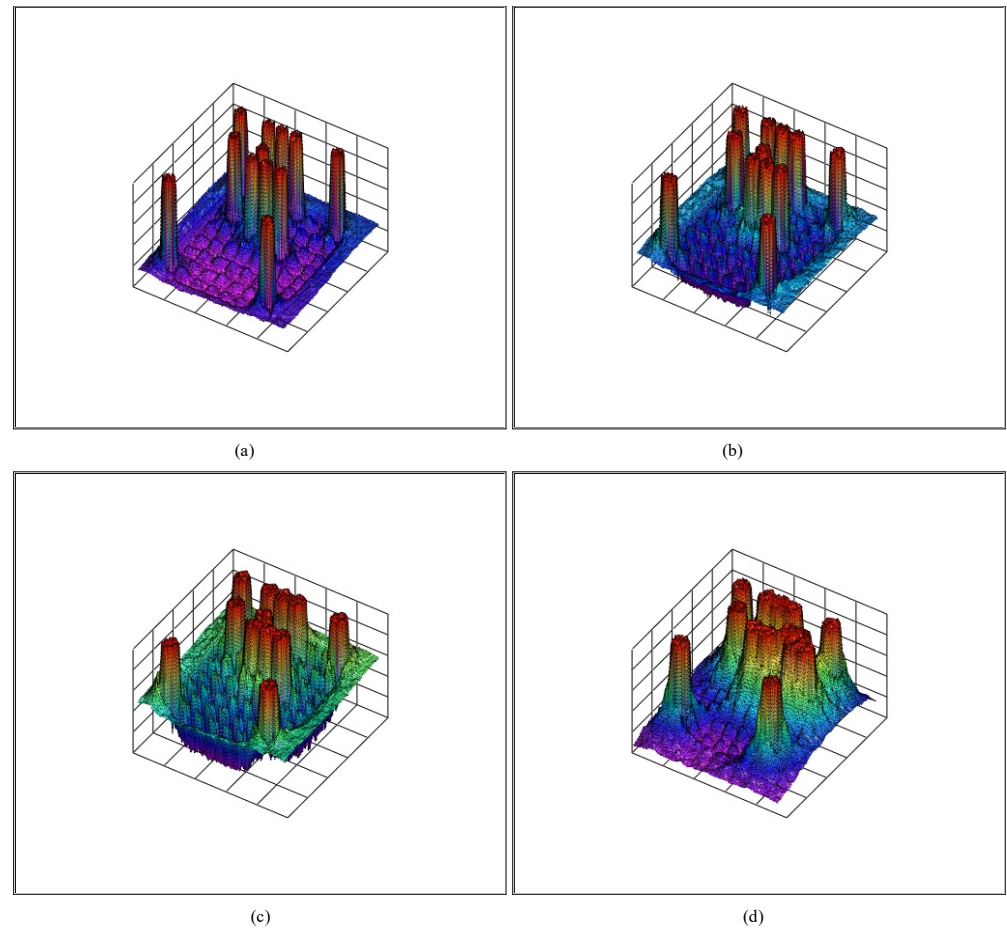


Figure 9. The quantized intensity profile of the LED array with exposure times (a) 100 μ s, (b) 200 μ s, (c) 300 μ s, (d) 400 μ s.



Figure 10. Demonstration setup.

The performance evaluations under various transmission distances are provided in Figure 11 with 20 links. Bit error rate (BER) performance was evaluated three times while maintaining constant transmission distance and noise conditions in Figure 12. At a distance of 10 m, the proposed deep learning-based decoding method achieved a BER of 10^{-3} , whereas the conventional zeros-crossing approach exhibited inferior performance, with a BER of approximately 10^{-2} , as shown in Figure 12. The effectiveness of the proposed

DL-based decoder in enhancing OCC performance under mobility and long-distance constraints is evident from these results. With long distance, the SNR value is reduced, and the receiver cannot define the threshold between ON status and OFF status of signals. As such, a deep learning decoder is helpful for long range communication. Besides that, YOLOv13 is a good candidate to detect and track objects in mobile environments. Since communication bandwidth is highly dependent on exposure settings, both the duration itself and the critical trade-off between exposure time and noise must be carefully optimized. In addition, the application of channel coding is recommended to further reduce BER or increase the communication range, thereby improving overall system performance. The implementation utilizes a 16×16 LED matrix with a smart camera supporting 20 links as well as an 8×8 LED matrix paired with a standard webcam supporting 3 links in a mobile environment. The proposed system is capable of operating over distances from 2 to 22 m, as further detailed in the Supplementary Materials.

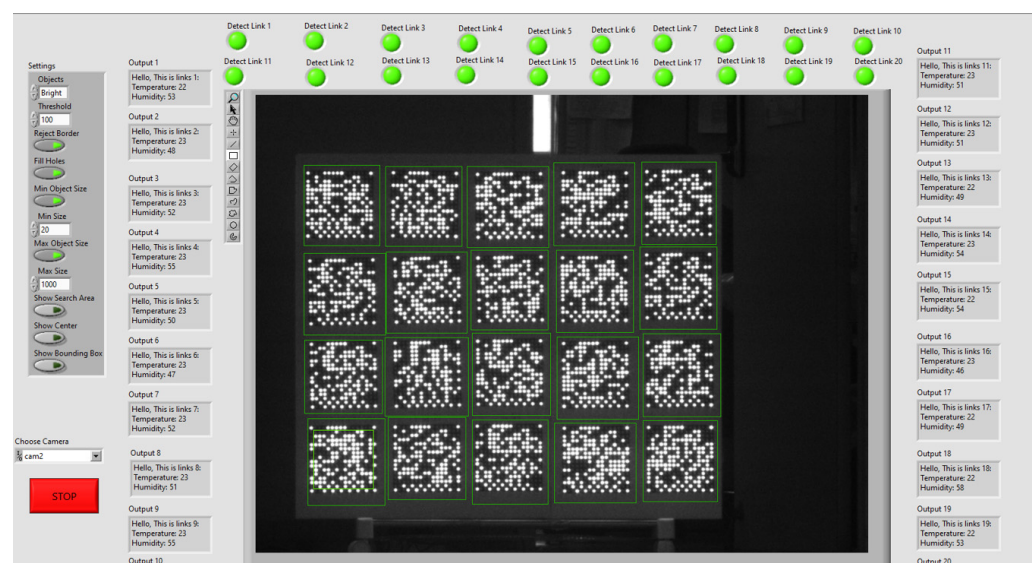


Figure 11. Experimental setup of the proposed system using an LED matrix.

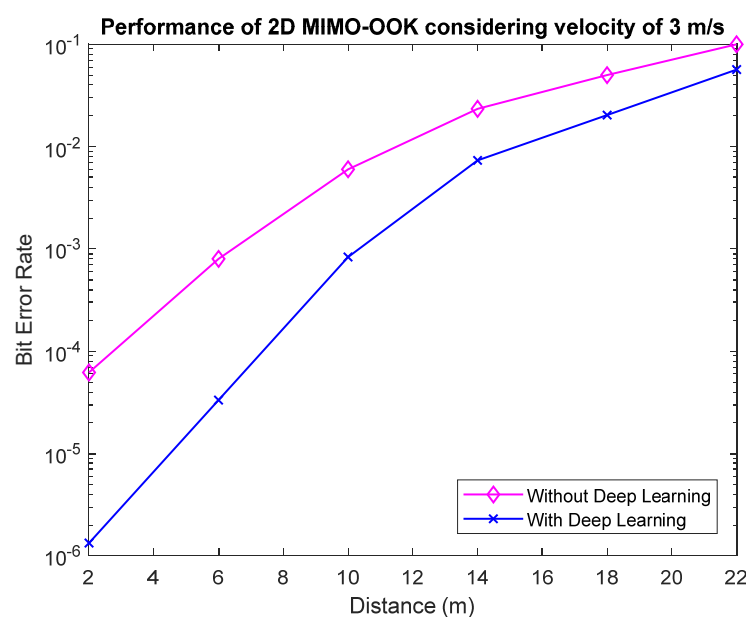


Figure 12. BER performance of the proposed scheme for a 16×16 LED matrix with different distances with a velocity of 3 m/s.

4. Conclusions

In this paper, we proposed an optical camera communication (OCC) system, which applies deep learning to light source detection and signal decoder. By applying YOLOv13, our system detects and tracks multiple LED matrixes under mobility conditions in real time, considering long-range environment. The experiments were conducted with two types of cameras (rolling shutter camera and global shutter camera), which make sure that our study is applicable for any camera, demonstrating high-speed OCC with low-cost hardware. In addition, the proposed asynchronous algorithm, based on sequence numbering for packet merging and missing packet detection, effectively addresses frame rate variation. Experimental results show that the system achieves good performance, supporting multi-links over up to 22 m of distance, thus making it suitable for multi-user IoT scenarios.

Supplementary Materials: The following supporting information can be downloaded at <http://zenodo.org/records/20888658>, accessed on 3 May 2026.

Author Contributions: Conceptualization, O.G. and H.N.; methodology, O.G. and H.N.; formal analysis, O.G. and H.N.; investigation, O.G. and H.N.; writing—original draft preparation, O.G. and H.N.; writing—review and editing, H.N.; supervision, Y.M.J.; funding acquisition, Y.M.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2022R1A2C1007884).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the datasets due to the project policy.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Elijah, O.; Rahman, T.A.; Orikumhi, I.; Leow, C.Y.; Hindia, M.N. An overview of IoT and data analytics in agriculture: Benefits and challenges. *IEEE Internet Things J.* **2018**, *5*, 3758–3773.
2. Wazid, M.; Das, A.K.; Odelu, V.; Kumar, N.; Conti, M.; Jo, M. Design of secure user authenticated key management protocol for generic IoT networks. *IEEE Internet Things J.* **2018**, *5*, 269–282.
3. Kim, J.H.; Lee, J.K.; Kim, H.G.; Kim, K.B.; Kim, H.R. Possible effects of radiofrequency electromagnetic field exposure on central nerve system. *Biomol. Ther.* **2019**, *27*, 265–275. [\[CrossRef\]](#)
4. Nguyen, H.; Utama, I.B.K.Y.; Jang, Y.M. Enabling Technologies and New Challenges in IEEE 802.15.7 Optical Camera Communications Standard. *IEEE Commun. Mag.* **2024**, *62*, 90–95.
5. Nikola, S.; Volker, J.; Yeong Min, J.; John, L.Q. An Overview on High-Speed Optical Wireless/Light Communications. Available online: <https://mentor.ieee.org/802.11/dcn/17/11-17-0962-02-00lc-an-overviewon-high-speed-optical-wireless-light-communications.pdf> (accessed on 1 February 2026).
6. *IEEE Std 802.15.7-2011*; IEEE Standard for Local and Metropolitan Area Networks—Part 15.7: Short-Range Wireless Optical Communication Using Visible Light. IEEE-SA: Piscataway, NJ, USA, 2011.
7. *IEEE Std 802.15.7-2018*; IEEE Standard for Local and Metropolitan area Networks—Part 15.7: Short-Range Optical Wireless Communications. IEEE-SA: Piscataway, NJ, USA, 2018.
8. Pathak, P.H.; Feng, X.; Hu, P.; Mohapatra, P. Visible light communication, networking, and sensing: A survey, potential and challenges. *IEEE Commun. Surv. Tutor.* **2015**, *17*, 2047–2077. [\[CrossRef\]](#)
9. Ong, Z.; Rachim, V.P.; Chung, W.Y. Novel electromagnetic-interference-free indoor environment monitoring system by mobile camera-image-sensor-based VLC. *IEEE Photonics J.* **2017**, *9*, 7907111.
10. Takano, H.; Nakahara, M.; Suzuoki, K.; Nakayama, Y.; Hisano, D. 300-Meter Long-Range Optical Camera Communication on RGB-LED-Equipped Drone and Object-Detecting Camera. *IEEE Access* **2022**, *10*, 55073–55080.
11. Hamidnejad, E.; Gholami, A. Developing a comprehensive model for underwater MIMO OCC system. *Opt. Express* **2023**, *31*, 31870–31883. [\[CrossRef\]](#) [\[PubMed\]](#)

12. Nguyen, H.; Al-Imran, J.; Jang, Y.M. Survey of next-generation optical wireless communication technologies for 6G and Beyond 6G. *ICT Express* **2025**, *11*, 576–589. [[CrossRef](#)]
13. Ayyash, M.; Elgala, H.; Khreishah, A.; Jungnickel, V.; Little, T.; Shao, S.; Rahaim, M.; Schulz, D.; Hilt, J.; Freund, R. Coexistence of WiFi and LiFi toward 5G: Concepts, opportunities, and challenges. *IEEE Commun. Mag.* **2016**, *54*, 64–71.
14. Yu, Q.; Wang, B.; Su, Y. Object Detection-Tracking Algorithm for Unmanned Surface Vehicles Based on a Radar-Photoelectric System. *IEEE Access* **2021**, *9*, 57529–57541.
15. Lin, H.; Si, J.; Abousleman, G.P. Region-of-interest detection and its application to image segmentation and compression. In *Proceedings of the 2007 International Conference on Integration of Knowledge Intensive Multi-Agent Systems*, Waltham, MA, USA, 30 April–3 May 2007.
16. Yan, C.; Chen, W.; Chen, P.C.Y.; Kendrick, A.S.; Wu, X. A new two-stage object detection network without RoI-Pooling. In *Proceedings of the 2018 Chinese Control and Decision Conference (CCDC)*, Shenyang, China, 9–11 June 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1680–1685.
17. Nguyen, H.; Thieu, M.D.; Nguyen, T.; Jang, Y.M. Rolling OFDM for image sensor based optical wireless communication. *IEEE Photon. J.* **2019**, *11*, 6500817. [[CrossRef](#)]
18. *IEEE Std 802.15.7a-2024*; IEEE Standard for Local and Metropolitan Area Networks—Part 15.7: Short-Range Optical Wireless Communications. IEEE-SA: Piscataway, NJ, USA, 2024.
19. Huang, W.; Tian, P.; Xu, Z. Design and implementation of a real-time CIM-MIMO optical camera communication system. *Opt. Express* **2016**, *24*, 24567–24579. [[PubMed](#)]
20. Jabeen, N.; Lei, H.; Muhammad, A.; Ali, A.; Khan, Z.U.; Pan, G. Localization in ISAC: A Review. *IEEE Internet Things J.* **2025**, *12*, 46526–46552. [[CrossRef](#)]
21. Aman, M.; Gang, Q.; Shang, Z.; Khan, Z.U.; Khan, M.S.; Ullah, I. Realization of RSSI Based, Three Major Components (Hx, Hy, Hz) of Magnetic Flux Created around the MI-TD Coil. In *Proceedings of the 2023 IEEE International Conference on Electrical, Automation and Computer Engineering (ICEACE)*, Changchun, China, 29–31 December 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1012–1017.
22. Lee, H.; Lee, S.H.; Quek, T.Q.S.; Lee, I. Deep Learning Framework for Wireless Systems: Applications to Optical Wireless Communications. *IEEE Commun. Mag.* **2019**, *57*, 35–41. [[CrossRef](#)]
23. Wu, H.; Chen, Z.; Liu, Z.; Geng, X.; Zhao, Y.; Liu, Z. CRS-Based Joint CFO and Channel Estimation Using Deep Learning in OFDM-Based Vehicular Communication Systems. *IEEE Trans. Wirel. Commun.* **2025**, *24*, 3882–3893.
24. Kong, M.; Pan, Y.; Zhou, H.; Yu, R.; Le, X.; Yuan, H.; Wang, R.; Yang, Q. Deep Learning-Based Acquisition Pointing and Tracking for Underwater Wireless Optical Communication. *IEEE Photonics Technol. Lett.* **2025**, *37*, 555–558.
25. Jia, B.; Ge, W.; Cheng, J.; Du, Z.; Wang, R.; Song, G.; Zhang, Y.; Cai, C.; Qin, S.; Xu, J. Deep Learning-Based Cascaded Light Source Detection for Link Alignment in Underwater Wireless Optical Communication. *IEEE Photonics J.* **2024**, *16*, 7801512. [[CrossRef](#)]
26. Li, J.; Wu, Y.; Zhang, Y.; Zhao, J.; Si, Y. Parameter Estimation of Poisson–Gaussian Signal-Dependent Noise from Single Image of CMOS/CCD Image Sensor Using Local Binary Cyclic Jumping. *Sensors* **2021**, *21*, 8330. [[PubMed](#)]
27. Roberts, R.D. Intel Proposal in IEEE 802.15.7r1. 2016. Available online: <https://mentor.ieee.org/802.15/dcn/16/15-16-0006-01-007a-intel-occproposal.pdf> (accessed on 21 February 2026).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.