Development of Automatic Multiple LED Detection for Hybrid OOK-OFDM Optical Camera Communication System

Eun Bi Shin, Yeong Min Jang
Department of Electronics Engineering, Kookmin University Seoul 02707, Korea
Email: seb0440@naver.com, yjang@kookmin.ac.kr

***Abstract*—Optical Camera Communication (OCC) has developed as a viable substitute for radio-frequency wireless systems by utilizing light-emitting diodes (LEDs) and image sensors for data transmission. To improve data throughput, robustness, and multi-user capability, recent OCC systems increasingly adopt multiple LED transmitters and hybrid modulation schemes such as Hybrid On–Off Keying and Orthogonal Frequency-Division Multiplexing (OOK–OFDM). However, accurately detecting and locating multiple LEDs in real time in a variety of settings is still a big problem, especially when the lighting is low or the LEDs are moving. This paper discusses the creation of a deep learning-based automatic multiple LED detection framework for a hybrid OOK–OFDM OCC system. A unique dataset is created with a global shutter camera that has controlled exposure to record realistic changes in distance, orientation, motion blur, and modulation patterns. A semi-supervised, model-assisted labeling strategy is used to quickly make high-quality annotations for both OOK and OFDM transmitters. An AI-based object detection model is trained on this dataset to accurately find and categorize multiple LED transmitters at the same time. The experimental results show that the training behavior is stable, the generalization is strong, and the detection performance is accurate even in difficult imaging conditions. The suggested method gives hybrid OCC systems a reliable perception front-end and lets them communicate with multiple LEDs in real time in changing environments.**

***Index Terms*—OOK, OFDM, OCC, Detection**

I. Introduction

Optical Camera Commuication (OCC) is a next-generation wireless technology that utilizes the light sources such as Light-Emitting Diode (LED) as data transmitting medium and paired with camera or image sensors to process captured images as received data. OCC is based on the wireless communication category of the Visible Light Communication (VLC) family. The distinction however, lies in the usage of camera instead of photodiode as receiver, normally used in the laser communication systems. Researches are developing OCC as alternative technology to the already highly congested radio frequency (RF) communication network. However, to establish connection between transmitter and receiver of OCC system the communication link must be line of sight (LoS). Other important consideration is the camera frame rate to support the effective data throughput of OCC system. Every captured image frames from the camera will define how much of data throughput can be extracted from them.

By using multiple LEDs in OCC systems, we can increase the overall data throughput, reduce the outage probability in case of view blockage, and multiple access for the users. Each LED can be defined as independent transmitter. This concept is derived from the Multiple-Input Multiple-Output (MIMO) method, which leverages the multiple nodes each of the transmitter and receiver side. Compared to single LED, the use of Multiple LEDs can set the data channel into parallel data streams which effectively increase the throughput into multi-fold of the LED counts. In this method the camera will detect each LED separately and count as different data stream. The second advantage of multi-LED is the mitigation of outage errors due to the view blockage. By utilizing different LED on the transmitting side, the camera can have more options to receive data compared to only one single LED as light source. The configuration of multi-LED can also provide the Multi-user or Multi-Access to the users. The separate LED can acts as single entity of OCC node that work entirely independent with other LEDs. The camera should be able to differentiate these multiple LED to be capable of the Multi-User feature [1]. Recently, the Hybrid On-Off Keying Orthogonal Frequency-Division Multiplexing (HOOK-OFDM) concept is explored by researchers to enhance the performance of current OCC version.

Incorporation of Artificial Intelligence (AI) algorithm is required to enable the multiple LED in OCC system. This automates the processing and keep the system works seamlessly in real-time. The real-world images captured by the camera has a large variant so it is difficult to solve the decoding process by just calculating the received images in direct approach decoding method. For the sake faster LED detection the AI is used on OCC system such as the Convolutional Neural Network (CNN) or the You Only Look Once (YOLO) AI models. Building dataset is essential for AI enhanced Automatic Multiple LED Detection for HOOK-OFDM OCC system. The dataset is utilized to train models to detect multiple LED setup simultaneously and find out the variations of many different kind of LED images captured by camera. The variance of image data can be due to distance, scale, viewing angle, motion blur, camera exposure, and lens focus [2].

The technical limitation of conventional method for Automatic Multiple LED Detection for HOOK-OFDM OCC system forces the designer to build robust and high performance AI model. The complexity of the decoding process in the hybrid modulation and multiple LED configuration requires the system to be AI enhanced. The image capture of OCC receiver must be able to process the single frame shot of the image of the multiple LED detected by the camera. This will be resource intensive in the conventional decoding system, therefore implementing the AI model will significantly enhance the performance. In the real-world scenario, the OCC system will have to be able to handle multiple LEDs that transmitting data simultaneously with highly variate size, distance, orientation, and brightness values. With the AI-based model for multi-LED detection, the system will be able to localize multiple LED simultaneously, separate closely arranged LEDs and flexibly increase effectiveness in dense OCC system.

II. Recent Works

Optical Camera Communication (OCC) leverages image sensors and the rolling shutter effect to enable low-cost optical wireless links but suffers from low frame rates and challenging channel conditions. Hybrid modulation schemes combining On-Off Keying (OOK) for low-rate identification and Rolling-Shutter OFDM (RS-OFDM) for high data throughput have been studied to enhance achievable rates and BER performance over conventional single-modulation schemes in OCC. However, in hybrid systems where OOK coexists with OFDM (or ACO-OFDM), fixed intensity thresholds become unreliable under low SNR—particularly over longer distances—due to interference between waveform components, causing a hybrid threshold conflict that undermines simple demodulation [3]. These limitations are amplified in multi-LED scenarios requiring real-time localization and modulation classification, motivating recent works that apply deep learning–based detection and classification to jointly localize LEDs, mitigate optical channel effects, and decode hybrid OOK–OFDM signals robustly [4].

Automatic detection and localization of LED transmitters is a foundational challenge in optical camera communication (OCC) systems, as accurate ROI extraction directly impacts communication reliability, throughput, and bit error ratio (BER). Traditional OCC detection methods are computationally lightweight but struggle under varying illumination, motion blur, and multiple simultaneous transmitters, motivating the adoption of more robust approaches. Recent research has explored the utilization of deep learning-based object detection models to address these limitations, leveraging convolutional neural networks (CNNs) to localize LEDs under complex backgrounds and dynamic environments.

Recent research has increasingly adopted deep learning-based object detection models to address limitations of traditional ROI extraction. Sun et al. [5] proposed an end-to-end LED detection and recognition architecture based on the YOLOv5 object detector to localize LED arrays and mitigate motion blur in vehicle-to-vehicle OCC, demonstrating real-time detection and improved LED status recognition accuracy in complex scenes. Similarly, Cheng et al. [6] developed a lightweight deep learning pipeline combining LED detection and segmentation specifically tailored for OCC, yielding higher detection accuracy and inference speed across various communication distances. Recent work has also explored novel detection and classification frameworks based on YOLOv8 have been shown to achieve high symbol recognition accuracies under color-based modulations, indicating the effectiveness of modern CNN detectors in OCC contexts [7].

Despite these advances, the limited availability of standardized LED-centric datasets remains a bottleneck for generalized model training in detection and deep learning research. The E-VLC dataset provides synchronized event camera and frame camera data for visible light communication and localization tasks across diverse environmental settings, offering a potential benchmark for LED detection and localization approaches [8].

These developments collectively underscore the trend towards integrating deep learning-based object detection, tracking, and classification in OCC systems to enable robust, real-time multiple LED detection suitable for hybrid modulation schemes such as OOK-OFDM. However, challenges remain in improving detection under varying lighting conditions, handling multiple simultaneous transmitters, and developing comprehensive datasets that support both localization and decoding tasks in diverse real-world environments.

III. LED Dataset Development

In this paper we establishes a specialized dataset designed for the robust detection of Optical Camera Communication (OCC) transmitters in high-speed, mobile environments. The development process is divided into a rigorous data collection phase, which simulates realistic channel conditions, and a semi-supervised labeling phase designed to maximize annotation efficiency for multi-class detection.

A. Data Collection

We captured data using a global shutter camera with the exposure locked to 22 µs. This low exposure removes background light, making the transmitters appear as bright shapes on a dark background. The dataset includes two types of transmitters: a 16x16 LED matrix for OOK modulation (OOK tx) and a rectangular LED with three extra corner LEDs for OFDM modulation (OFDM tx). To make sure the camera can always see the transmitters, even when they send ”off” signals, we used an anchor strategy. For OOK tx, we programmed the corner LEDs to stay on. For OFDM tx, the three extra LEDs act as permanent markers.

To ensure the model works well on a moving drone, we recorded continuous videos instead of taking single photos. Data collection process varied in several ways such as distance, camera angle, blinking patterns, and motion blur. The camera was maneuvered from close-range views to long-range views and subjected to angular tilts to introduce different perspective distortion. Simultaneously, the transmitters cycled through various modulation patterns. This approach ensures the model learns to generalize the spatial features of both the OOK tx and

OFDM tx classes, preventing overfitting to static, front-facing orientation

B. Labeling Method

A model-assisted active learning pipeline was employed to annotate the dataset efficiently, beginning with the manual curation of an initial reference subset. This subset was carefully balanced to include equal distributions of both transmitter classes OOK tx and OFDM tx. The annotation protocol strictly defined the bounding boxes to enclose the implied geometric area of the transmitters, for the OOK tx class, the box encompasses the 16x16 grid area defined by the corner anchors and for the OFDM tx class, the box encloses the rectangular main LED and its three auxiliary anchors. This consistent labeling strategy trains the model to recognize the complete physical extent of the transmitter even when the modulated data payload is not visible.

To maximize the utility of the limited reference subset, extensive data augmentation was applied during the preliminary training phase. Geometric transformations, including perspective warping and rotation, were heavily utilized to simulate the spatial orientation of a drone relative to the transmitter. Additionally, synthetic image composition techniques were employed to combine multiple transmitter instances into single frames, forcing the model to learn context-independent features. A lightweight object detection model was then trained on this augmented reference set to establish a baseline learner capable of generating preliminary bounding box predictions for both classes.

The final labeling stage utilized this baseline model to automatically predict annotations for the remaining unlabelled dataset. Rather than creating annotations from scratch, human verifiers only needed to validate the class classification (OOK tx, OFDM tx) and refine the bounding box dimensions. This semi-supervised workflow significantly reduced the manual labor required, ensuring high-quality ground truth labels while maintaining strict class consistency across the entire dataset.

IV. Detection Model Development

Over the years, numerous object detection model has been proposed. Mainly, the object detection model can be categorized into two groups: single-stage and double-stage detection. The double stage detection performs better than the single-stage model due to the additional steps when predicting the object class and location. However,the double-stage model has significantly slower computes compared to the single-stage model. The popular model of double-stage model are R-CNN and Fast R-CNN.

Due to the computation speed issue in double-stage model, the single-stage model is chosen to perform the LED detection. The YOLO model is the most popular single-stage model in the world. The YOLO has evolved since several years ago with the most recent one is the YOLO12, meaning the version 12 in official YOLO model family. Moreover, in most YOLO family, they have varying model size that usually can be classified into nano, medium, large, or extra-large depending on the number of layers in the YOLO model.

For this work, the LED object detection is performed in a resource-constrained device. Hence, the size of the model should be small enough to run in such device with suitable speed. Due to the varying options of YOLO models, we choose to use some of the most recent YOLO model where each model uses only the smallest size variant, the nano. In this work, we use YOLO version 5 to 12 by using the implementation from Ultralytics [9].

V. Experiment Results

Based on the developed system, we conduct a detailed assessment of our work using several quantitative metrics. These metrics are applied to rigorously validate the AI model's capability to detect the LED matrix. The evaluation is performed directly on diverse image data where the LED matrix appears under varying conditions, ensuring that the model's performance is reliable and consistent in realistic scenarios.

A. Dataset

In this work, we collected dataset of around 40,000 images of LED with three classes of tx, OOK tx, and OFDM tx. The 40,000 images are taken using the similar camera settings but taken in multiple different environment and distance. The composition of environment and distance are not proportional to add more variance in the data.

B. Model

Based on the Table I, it is clear that the model YOLOv9n and YOLOv11n are the fastest in inference time in terms of processing speed, meaning that the model is small enough and fit well with the requirements for the LED detection. The model YOLOv9n is better in the mAP@0.95 where the model has advantage of 0.001 than the YOLOv11n. Therefore, for the object detection model, the YOLOv9n is chosen as the main model to perform the task.

Based on Figure 1, it showing the overall metrics related to the training results of the proposed detection model. We can analyze that the LED matrix detection model's training losses for bounding box regression (box_loss), classification (cls_loss), and distribution focal loss (dfl_loss) decrease steadily and smoothly over 100 epochs. This indicates stable optimization without divergence or oscillation and suggests that the model is continually improving its fit to the training data. The smoothing curves closely follow the raw results, reinforcing that the overall trend is a consistent reduction in error across all three training loss components.

On the validation side, the box, classification, and dfl losses start higher than their training counterparts but also decrease over time, then plateau, which is typical of a model that is learning meaningful features rather than memorizing the data. The validation curves show more noise and spikes, especially for the classification loss, but the general downward trend and eventual stabilization imply that overfitting is limited and that the model generalizes reasonably well to unseen

TABLE I
EVALUATION OF VARIOUS YOLO MODEL TRAINING PERFORMANCE

Model name	Inference Time (ms)	mAP@0.50	mAP@0.95
YOLOv5n	1.0	0.829	0.533
YOLOv8n	1.1	0.829	0.534
YOLOv9n	0.7	0.829	0.532
YOLOv10n	1.3	0.829	0.534
YOLOv11n	0.7	0.829	0.531
YOLO12n	2.2	0.83	0.534

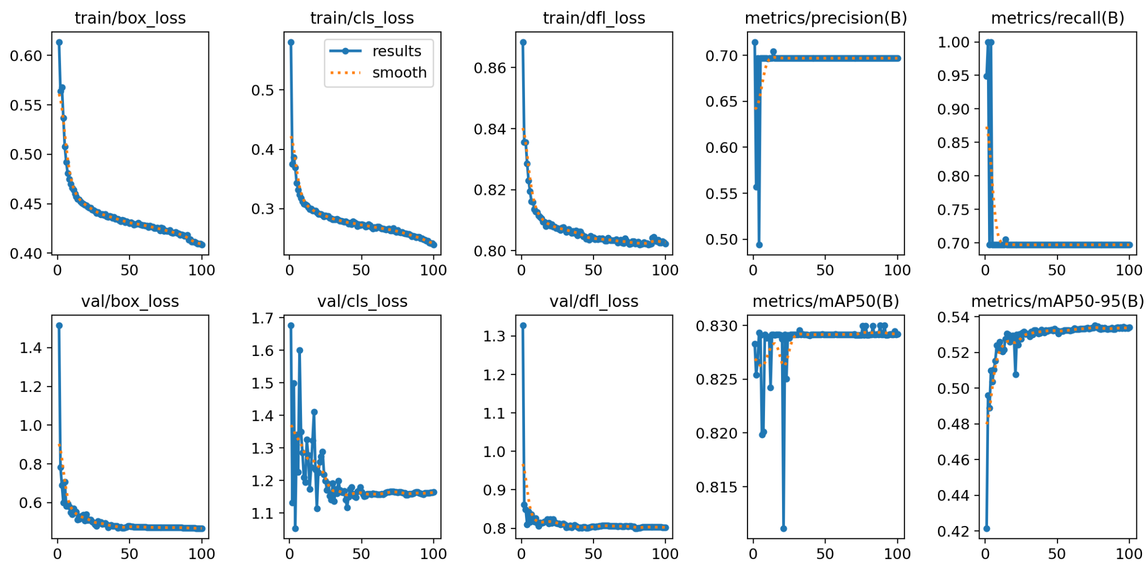

Fig. 1. Training result.

LED matrix samples. The narrowing gap between training and validation losses near the end of training further supports this interpretation.

The metric plots indicate that the model rapidly reaches high performance for LED matrix detection, with precision quickly climbing and stabilizing around a high value, while recall slightly drops from an initially inflated value to a more realistic but still strong level. The mAP@50 and mAP@50–95 metrics both increase throughout training and converge to relatively high plateaus, showing that the model is accurate across different IoU thresholds, not just at the easiest one. Overall, these curves suggest that the trained AI model can reliably localize and classify LEDs in the matrix with good balance between precision and recall and robust performance across varying localization tolerances.

As shown in Figure 2, its visualizes the ground-truth annotations for the LED matrix and associated objects across a sequence of low-light frames. Each frame contains bounding boxes and class label which is ofdm_tx, indicating the expected positions of the LED-based transmitters. This layout confirms that the dataset captures consistent spatial placement of the LEDs while still including background clutter and darkness, which makes the detection problem non-trivial.

As for the model's predictions are shown in Figure 3, the predicted bounding boxes and class labels for the same frames, now overlaid with confidence scores (e.g., tx 0.3). The predictions align closely with the ground-truth box locations, and the model repeatedly detects all visible LED elements across frames, demonstrating strong spatial consistency. Even with relatively modest confidence values, the repeated correct detections in every frame indicate that the model has learned a

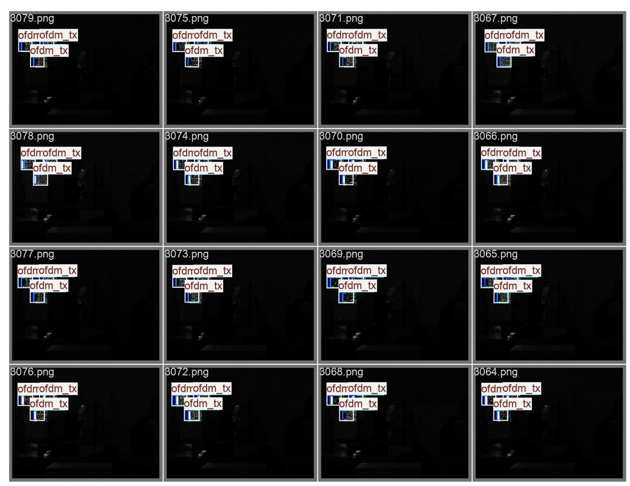

Fig. 2. Validation Labels Data.

Fig. 3. Validation Prediction Results.

robust representation of the LED transmitters under challenging illumination conditions.

Comparing labels and predictions side by side highlights that the model not only localizes the LED matrix accurately but also maintains stable classification performance over time. The dense clustering of blue predicted boxes around the true LED positions suggests minimal localization error and limited false positives in the surrounding dark background. This qualitative validation supports the quantitative metrics from training, showing that the LED matrix detection network generalizes well and is suitable as a perception front-end for communication or control tasks involving LED-based transmitters.

VI. Conclusion

This paper introduced an AI-driven automatic multiple LED detection framework for hybrid OOK–OFDM Optical Camera Communication systems. A specialized OCC dataset was created utilizing low-exposure global-shutter imaging to record authentic variations in distance, orientation, motion blur, and modulation patterns. A semi-supervised labeling strategy was implemented to effectively produce uniform annotations for both OOK and OFDM transmitters. Using this dataset, a deep learning-based object detection model was trained and tested. It showed stable convergence, strong generalization, and accurate localization and classification of several LED transmitters in difficult low-light and dynamic situations. The experimental findings validate that the proposed methodology delivers a resilient and scalable perception front-end for hybrid OCC systems, facilitating dependable multiple LED detection appropriate for real-time and mobility-sensitive communication contexts.

Acknowledgment

This work was supported by Korea Research Institute for defense Technology Planning and advancement (KRIT) grant funded by the Korea government (DAPA (Defense Acquisition Program Administration)) (KRIT-CT-23-041, LiDAR/RADAR Supported Edge AI-based Highly Reliable IR/UV FSO/OCC Specialized Research Laboratory, 2024)

References

[1] H. Herfandi, O. S. Sitanggang, M. R. A. Nasution, I. B. K. Y. Utama, M. M. Rahman, N. Huy, S. M. Ko, and Y. M. Jang, “Implementation of a multiple-transmitter rs-ofdm-based occ system with advanced bytetrack for mobility environments,” *IEEE Access*, vol. 13, no. 1, p. 119411, 2025.

[2] M. M. Faridh, H. Herfandi, O. S. Sitanggang, M. R. A. Nasution, H. Nguyen, and Y. M. Jang, “Implementation of csma/ca-based mac protocol for optical camera communication,” *IEEE Access*, vol. 13, no. 1, p. 14611, 2025.

[3] H. Nguyen, N. T. Le, D. T. A. Le, and Y. M. Jang, “Design and implementation of a hybrid optical camera communication system for indoor applications,” *Sensors*, vol. 24, no. 1, p. 300, 2024.

[4] H. Nguyen, V. L. Nguyen, D. H. Tran, and Y. M. Jang, “Rolling shutter ofdm scheme for optical camera communication considering mobility environment based on deep learning,” *Applied Sciences*, vol. 12, no. 16, p. 8269, 2022.

[5] X. Sun, W. Shi, Q. Cheng, W. Liu, Z. Wang, and J. Zhang, “An led detection and recognition method based on deep learning in vehicle optical camera communication,” *IEEE access*, vol. 9, pp. 80 897–80 905, 2021.

[6] Q. Cheng, H. Ma, and X. Sun, “Vehicle led detection and segmentation recognition based on deep learning for optical camera communication,” *Optoelectronics letters*, vol. 18, no. 8, pp. 508–512, 2022.

[7] F. V. V. Vera, L. Muñoz, F. Pérez, L. B. Alvarez, S. Montejo-Sánchez, V. M. Icaza, L. Rodríguez-López, and G. Saavedra, “High-accuracy deep learning-based detection and classification model in color-shift keying optical camera communication systems,” *Sensors*, vol. 25, no. 17, p. 5435, 2025.

[8] S. Shiba, Q. Kong, and N. Kobori, “E-vlc: A real-world dataset for event-based visible light communication and localization,” in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 4877–4886.

[9] G. Jocher, J. Qiu, and A. Chaurasia, “Ultralytics YOLO,” Jan. 2023. [Online]. Available: https://github.com/ultralytics/ultralytics