Impact of Source Coding on Downstream AI Applications

Nazmul Islam and Seokjoo Shin*
Dept. of Computer Engineering, Chosun University
Gwangju, 61452 South Korea
nazmul@chosun.kr, *sjshin@chosun.ac.kr (Corresponding author)

***Abstract*— The use of Artificial Intelligence (AI) in Internet of Things (IoT) ecosystem has been growing exponentially, enabling various Computer Vision (CV) applications. These applications must handle large image data demanding reliable communication systems that retain image quality for downstream Deep Learning (DL) tasks. Existing communication systems, such as Orthogonal Frequency Division Multiplexing (OFDM), promise improvements in data rate, spectral efficiency, and mitigation of multipath fading; however, these systems often distort the received images due to complex channel environments and impairments from various physical layer (PHY) blocks. Source Coding is one such PHY block, which aims for compression savings at the expense of image quality. Therefore, in this study, we evaluate the performance of a DL model for downstream image recognition tasks, where images are transmitted over communication systems utilizing various source coding schemes over complex channels. Experimental analysis shows that Variable-Length Coding (VLC) retains superior image quality, which results in over 95% DL model accuracy throughout the experiment.**

Keywords— Source Coding, OFDM System, Deep Learning.

I. INTRODUCTION

Image communication system is at the center of Internet of Things (IoT) ecosystem to enable Computer Vision (CV) applications such as traffic sign recognition, event monitoring, medical image analysis, and surveillance systems. With the recent advancements in Artificial Intelligence (AI), many of these applications are AI-driven which demands for robust communication systems ensuring high data rates and bandwidth efficiency while retaining image quality. This is essential for reliable transmission of vast amounts of image data generated by end devices in real-time, and effectively carrying out downstream Deep Learning (DL)-based IoT applications. The emergence of Fifth Generation (5G) communication systems aims to cater to these requirements, with potential advancements in network capacity, spectral efficiency and infrastructure evolution compared to previous generations [1]. One such innovation is the Orthogonal Frequency Division Multiplexing (OFDM) system, which is recognized to mitigate multipath fading, and provide high data rate with high order modulation and parallel transmission schemes [2]. Since its inception, it has become popular and has been adapted in many practical wireless communication systems and standards such as WiMAX, IEEE 802.11, Wi-Fi, and LTE. However, in OFDM-based image communication systems, there are multiple factors that can degrade signal quality, such as complex channel models, data transformation methods, modulation techniques, the number of antennas, as well as interferences such as Inter-symbol Interferences (ISI) and Inter-carrier Interferences (ICI). These factors can introduce heavy distortions in the recovered images, and severely hinder DL performance for downstream IoT applications [2]. Table 1 of [3] provides a comprehensive summary of several studies that have investigated the implementation of various OFDM communication systems, and evaluated the quality of recovered images under the influence of these systems.

The performance of DL models under the influence of various communication impairments needs to be well-studied to improve existing systems or develop new ones in the field of wireless communication and network. Several studies have evaluated the performance of DL models on images recovered from various OFDM-based communication systems considering factors such as channel correction [4], modulation schemes, and channel models. The authors in [2] conducted an extensive analysis of OFDM-based image communication systems and the DL performance on recovered images considering complex channel models, various channel estimation techniques, and high-order modulation schemes. Additionally, they evaluated the performance of multiple DL models and utilized various augmentation techniques to achieve better inference accuracy. However, the authors have considered only Variable-Length Coding (VLC) for source coding and do not evaluate other schemes. Therefore, in this study we consider OFDM-based image communication systems using various source coding schemes over different channel models. Specifically, we have considered VLC and Fixed-Length Coding (FLC) for source coding of image data. These coding schemes are often utilized in practical systems to enhance transmission quality and efficiency. For the channel model, AWGN and Rayleigh fading channel are used to replicate real world scenario. Performance of DL models on images transmitted over these systems are evaluated.

The rest of the paper is organized as follows. Section II presents the system model of the study; section III presents the results and discussion. Finally the conclusion and future work.

II. SYSTEM MODEL

A. Physical layer (PHY) of OFDM Communciation system

Orthogonal Frequency Division Multiplexing (OFDM) is a digital modulation technique used in numerous high-speed communication systems. It converts serial high-speed data stream into multiple parallel lower-rate data streams, which improves spectral efficiency, and minimize multipath interferences [1]. In OFDM, bitstream are first coded using a specific source coding, which may include data compression, error correction, and interleaving. The encoded data are then mapped onto subcarriers and represented on In-phase and Quadrature (IQ) plane based on modulation techniques such as M-ary Quadrature Amplitude Modulation (M-QAM). Data is then converted from serial to parallel symbols, and subsequently pilot symbols are inserted in the data stream for channel correction at the receiver. Comb-type pilot insertion with linear interpolation is used in the experiment, which is a common pilot insertion technique, particularly in fast fading channels [1]. These parallel symbols then undergo the Inverse Discrete Fourier Transform (IDFT) to convert from frequency-domain to time-domain OFDM symbols.

Adjacent symbols and multipath carriers can overlap during transmission due to channel time spread, causing ISI and ICI. To mitigate this, a guard interval, often called the cyclic prefix (CP), is added to the signal. After adding the CP,

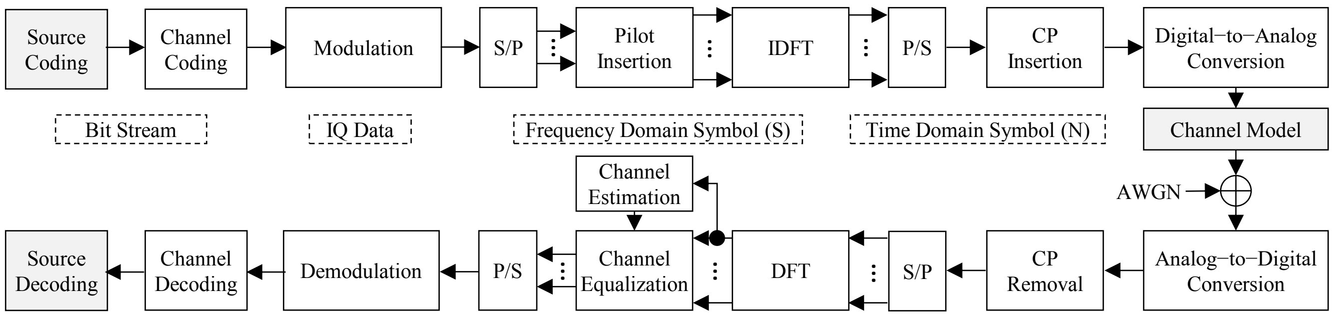

Fig. 1. Illustration of PHY in legacy OFDM.

the signal is then amplified to radio frequency (RF) and wirelessly transmitted. During wireless transmission, the signal experiences multi-path effects, fading, and variations in delays, phase shifts, and signal strengths. Different channel models can be used to simulate these effects. For the experiment, both simple AWGN channel and complex Rayleigh fading channel is considered for evaluation. At the receiving end, the signal is converted back to IQ samples. The OFDM symbols are recovered, and the CP is removed. The symbols are then converted to the frequency domain using DFT. In the frequency domain, symbols are adjusted using channel correction, which includes both channel estimation and equalization to recover the original IQ data. During channel estimation the received pilots are compared with the previously known ones to find the channel impulse response. This process counteracts signal attenuation by estimating the channel through the least squares (LS) method and linear interpolation. This interpolation method uses two neighboring pilots to predict the signal response. The estimated channel then recovers the modulated data using zero-forcing channel equalization. The parallel data then gets serialized, demodulated into soft bits (log-likelihood ratios), and turned into a binary stream by the decoder. To recover the transmitted image, the binary data is reverted to decimal and resized to match the original image dimensions. Fig. 1. illustrates the PHY of the OFDM system.

To summarize the parameters in the experiment, FLC and VLC are utilized for source coding. The modulation technique used is 16-QAM (rectangular constellation) with a symbol length of 4. The data stream for the system comprises 64 subcarriers, 16 CP, and 4 pilots (using comb-type pilot insertion). IDFT and DFT are used for the transformation techniques. LS with linear interpolation is employed for channel estimation, and zero-forcing is applied for channel equalization. For the channel model, AWGN and a Rayleigh fading channel are considered for wireless transmission.

B. Source Coding for OFDM System

Source coding is the process of assigning binary sequences, known as *codewords*, to elements of a specific *alphabet* comprised of symbols or *letters*. This collection of codewords is termed as *code* [5]. FLC is a source coding scheme where every symbol from the source alphabet gets an equal length of bits. For example, in an RGB image with pixel values from 0 to 255, each pixel value has a fixed length of 8 bits (as $2^8 = 256$ possible values). While FLC coding is simple and straightforward, it is not efficient. Lower pixel values are assigned with fixed length even when they have shorter codeword compared to higher pixel values [1], [5]. Additionally, redundant bits are added to the codewords with smaller code lengths to make them same as the assigned fixed length. Errors in these redundant bits, especially in the most significant ones, from the communication channel can drastically alter pixel values and heavily distort the recovered images. In practice, image formats usually incorporate advanced techniques, such as compression and VLC, to minimize the data size, especially for intricate or large images. Standards like JPEG combine transformations, quantization, and VLC techniques (Huffman coding) to attain notable compression ratios. Techniques like Huffman coding are often preferred due to their ability to exploit varying symbol probabilities for better compression. There are challenges such as synchronization issues that persist in Huffman coding as well. Furthermore, the probabilities of occurrence for the source symbols or their block lengths should be known which is impractical in real world scenarios and can leads to less-than-optimal compression [1], [5]. In such cases, transmitting the Huffman tree alongside the compressed data might be necessary leading to additional overhead.

Images can be compressed using either lossless or lossy methods. While lossless compression aims to preserve image quality, lossy compression ensures higher compression savings [6]. However, increased compression savings have adverse effects on DL performance [6]. Therefore, instead of compression savings, this study mainly focuses on evaluating the image quality recovered from OFDM systems using FLC and VLC, and how it affects the performance of DL models. In the experiment, for FLC, the corresponding decimal values of each pixel are coded to a fixed 8-bit length codeword. On the other hand, for VLC, the decimal values are coded as variable length according to the length of their respective codewords. The information about the length of each codeword for every pixel is transmitted as side information to the receiver for successful decoding. Although transmitting side information may increase the overall system overhead, it avoids the synchronization problems commonly experienced in VLC coding, such as the standard Huffman coding.

III. Result and Discussion

This section discusses the simulation analysis of the experiment. Fig. 2. presents sample images recovered from various systems for visual analysis. Based on visual inspection, VLC retains better image quality compared to FLC, especially over the Rayleigh fading channel. Using FLC over Rayleigh fading channel, the information in the Region of Interest (ROI) is invisible at lower E_b/N_0 regions. Following section quantifying the performance of communication system and image quality in terms of Bit Error Rate (BER) and Peak Signal to Noise Ratio (PSNR), respectively. The BER measures the number of error bits due to system impairments, and PSNR measures quality of received image.

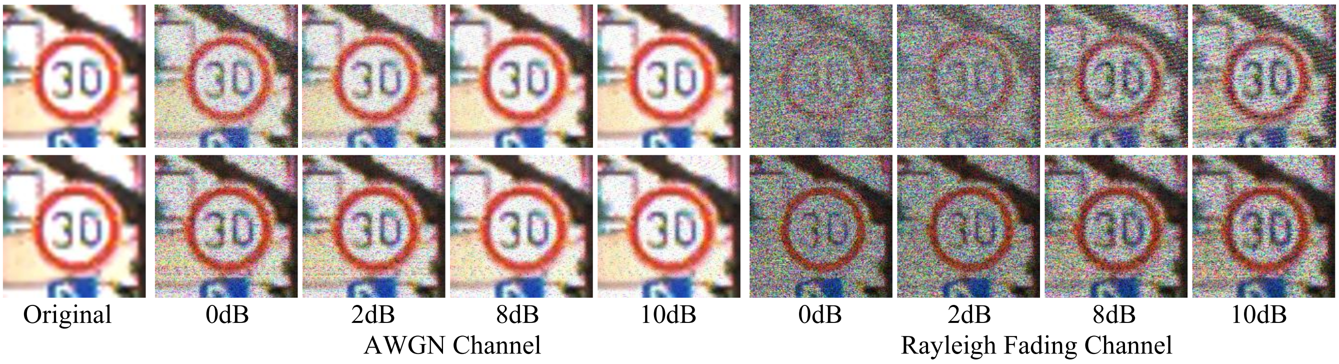

Fig. 2. Sample images from communication systems at various E_b/N_0 using FLC (top) and VLC (bottom) over the channels.

A. Image Quality Analysis

The BER analysis (on a logarithmic scale) for the proposed OFDM-based image communication system over both AWGN and Rayleigh fading channels is shown on Fig. 3. (a). Over the AWGN channel, there is a significant drop in BER as E_b/N_0 increases. At an E_b/N_0 of 10dB, the system achieved the lowest BER of 0.03. Throughout the experiment, FCL and VLC achieved identical BER. In contrast, for the system over the Rayleigh fading channel, the BER does not decrease as significantly at higher E_b/N_0 values compared to the AWGN channel. The lowest BER recorded was 0.32 at E_b/N_0 10dB. Similar to AWGN channel, both FLC and VLC yield identical BER values throughout the experiment. The reason for that could be multifold. Firstly, BER is the rate of number of bits in error, which remain consistent regardless of the number of bits transmitted using the different source coding scheme. Secondly, although FLC and VLC inherently differ in terms of average code length, their efficiencies is comparable over the given channel conditions, which did not particularly favor one coding scheme over the other. Lastly, the inherent characteristics of the AWGN and Rayleigh fading channels did not differentiate for the two different coding techniques, resulting in identical values.

The quality of the recovered images in terms of PSNR from the proposed systems over AWGN and Rayleigh fading channels is presented in Fig. 3. (b). Images recovered from the AWGN channel exhibit high quality, which further improves with an increase in E_b/N_0. Using VLC, as the E_b/N_0 increased, the PSNR improved by 41%, reaching a peak of 28.78dB at an E_b/N_0 of 10dB. Similar trend is observed with system using FLC, where the PSNR improved by 67%, attaining a maximum of 24.62dB at the same E_b/N_0. However, there is a significant difference between the image qualities from systems using FLC and VLC. Though the BER remain consistent between the two source coding schemes as observed previously, the error in the bits can drastically change a pixel value depending on its codeword position as discussed in in Section II(B), resulting in difference in image quality. Throughout the experiment, VLC consistently achieved better image quality compared to FLC, with improvement of highest 39% and lowest 17% observed at 0 dB and 10dB, respectively. Images from the communications system over a Rayleigh fading channel were heavily distorted with noise, which was observed in the PSNR analysis as well. At E_b/N_0 10dB, the highest PSNR of 18dB and 11dB was observed by the system using VLC and FLC, respectively. There were no significant improvement in image quality even with the increase in E_b/N_0. Throughout the experiment, the PSNR value improved only by 1dB and 2dB for VLC and FLC, respectively. Similar to the AWGN channel, there was a significant difference in image quality between FLC and VLC schemes. VLC achieved the highest improvement of 98% at E_b/N_0 0dB and lowest 64% at E_b/N_0 10dB, compared to FLC.

B. DL-based Downstream Application

The primary objective of this study is to evaluate the performance of a DL model on reconstructed images from systems using FLC and VLC over the two channels. For the experiment we consider DL-based traffic sign recognition (TSR) as the downstream IoT application deployed in an Intelligent Transport System (ITS) environment. Given the power and computational constraints of IoT devices, AI tasks are often offloaded to edge or cloud computing platforms. While cloud computing meets high computational and storage needs, primarily for training and testing DL models, it can be costly for larger capacities. As a solution, edge computing can bring cloud resources closer to the devices, allowing the DL

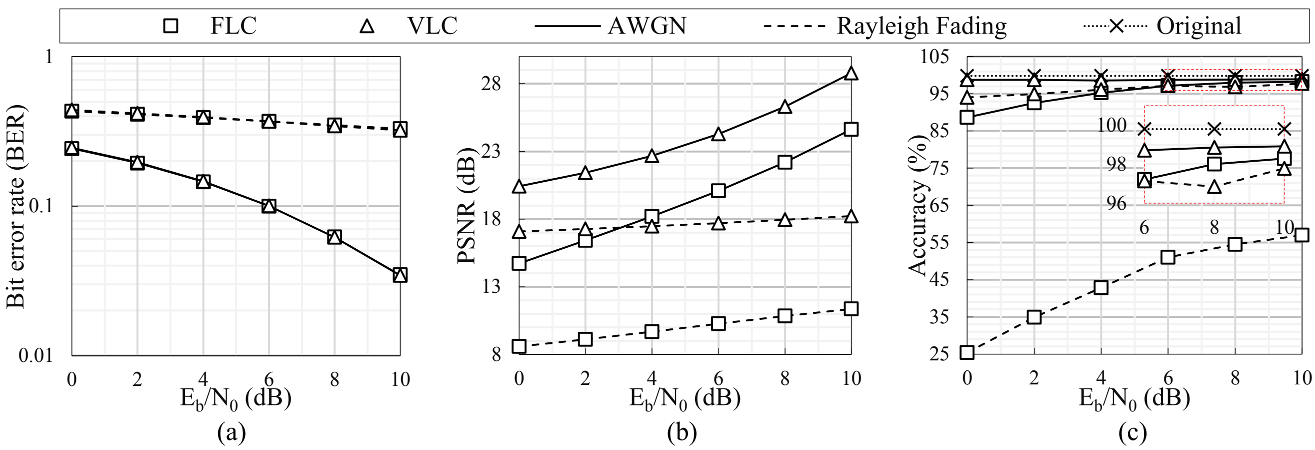

Fig. 3. Performance analysis in terms of (a) BER, (b) PSNR, and (c) DL model accuracy.

model to serve as an inference engine in an edge-cloud collaborative framework, which enables real-time CV applications such as the TSR in ITS. In this scenario, the images are initially transmitted to an edge server through OFDM-based image communication systems, exposing them to noise and distortion. The edge server processes these distorted images using a model trained on a public dataset in a cloud server. The cloud center offers multiple task-specific models to provide Machine Learning as a Service (MLaaS) [7]. Specifically, the edge server requests a model for TSR, which had been initially trained and tested on pristine traffic sign images from local database in cloud server. The DL model and training parameters are mentioned in Table. I.

The performance of the DL model in terms of accuracy (correctly classifying the reconstructed images) from communication systems using VLC and FLC over the different channels is shown in Fig. 3. (c). The accuracy of the DL model corresponds to the mean value calculated across all inferencing images for a specific E_b/N_0. From the graph, we can observe that at lower E_b/N_0 region (0dB-4dB), the model performance on images were distinct across the different communication systems. However, in higher E_b/N_0 regions (6dB-10dB), the DL model achieve high accuracy ranging between 97%-99% on images from different systems, demonstrating high DL model generalizability. An exception to this trend is the DL performance on images from communication system using FLC over the Rayleigh fading channel. This system has achieved poor DL performance, with the highest accuracy of only 57% at E_b/N_0 10dB. Such accuracy is not suitable for downstream DL applications. Additionally, while the BER and PSNR analysis of the communication systems indicates that the AWGN channel outperforms the Rayleigh fading channel significantly, it is not the case for the DL model accuracy. Specifically, the DL performance on images from systems using VLC over the Rayleigh fading channel outperforms system using FLC over the AWGN channel through E_b/N_0 0dB to 6dB. This indicates that for downstream DL applications, a system utilizing VLC scheme can achieve high accuracy, even on complex and noisy channels. The reason for that is VLC has high image quality retention compared to FLC, as mentioned in Section. II(B), even over the noisy channel. Fig. 4. shows the Grad-CAM visualization of model predictions for images retrieved from systems using FLC and VLC over the Rayleigh fading channel. Two incorrect predictions on images from system using FLC at E_b/N_0 0dB and 2dB show the heat map outside the ROI, indicating that the model is unable to extract essential features for accurate TSR due to the noise on the images.

TABLE I. Simulation parameters for the DL Application

Model Parameters	Model	EfficientNetV2-B0 [8]
	Model Input Size:	224×224
	Top Layers:	Average Pooling, 1280 Flatten, 1280 Fully Connected, 256 * Fully Connected, 7 ⁺
	Trainable parameter:	6,188,439
	Non-trainable parameter:	60,608
Training Parameters	Optimizer:	Stochastic Gradient Descent
	Learning Rate (LR):	0.01 to 0.0001 using Reduce LR
	Loss:	Categorical Cross-Entropy
	Accuracy:	Training: 99.9%, Validation: 99.9%
	Dataset:	GTSRB [9], Classes: 7
	Dataset Split:	Training: 10220 images – 80% Validation: 1400 images – 10% Testing: 1400 images – 10% ⭡

* ReLU Activation ⁺ Softmax Activation ⭡ Transmitted per E_b/N_0

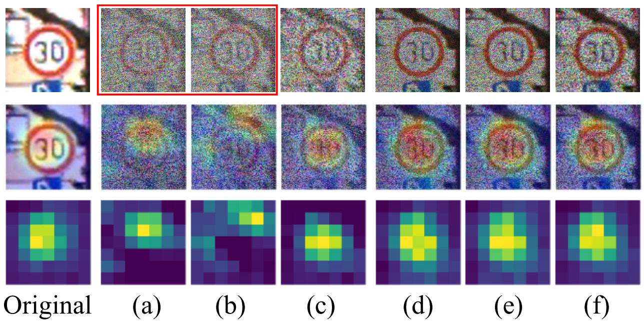

Fig. 4. Grad-CAM visualization on sample image from system over Rayleigh fading channel at $E_b/N_0 \in \{0,2,4\}$ using FLC (a-c) and VLC (d-f). Two images in the red box are incorrectly predicted; rest are correctly predicted.

IV. Conclusion and Future work

In this study, we evaluated the quality fo recovered images and the DL performance on these images from various OFDM-based communication systems. Specifically, we examined different source coding schemes and channel models for the PHY of the OFDM-based image communication system. The primary objctive was to assess the quality of the recovered images using various source codings and determine their suitability for downstream DL tasks. Analysis have shown that VLC outperformed FLC in terms of both image quality and DL performance. While VLC require the transmission of side information, that may lead to additional overhead, it retains superior image quality and proves more suitable for downstream DL applications. As a future direction, we aim to explore different DL algorithms in the PHY of the OFDM system to enhance the reliability and robustness of downstream DL applications.

Acknowledgement

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government. (MSIT) (RS-2023-00278294).

Reference

[1] J. G. Proakis, *Digital communications*. McGraw-Hill, Higher Edu, 2008.
[2] N. Islam and S. Shin, “Robust Deep Learning Models for OFDM-Based Image Communication Systems in Intelligent Transportation Systems (ITS) for Smart Cities,” *Electronics*, vol. 12, no. 11, p. 2425, 2023.
[3] L. Kansal, S. Berra, M. Mounir, R. Miglani, R. Dinis, and K. Rabie, “Performance Analysis of Massive MIMO-OFDM System Incorporated with Various Transforms for Image Communication in 5G Systems,” *Electronics*, vol. 11, no. 4, p. 621, 2022.
[4] I. Ahmad, N. Islam, and S. Shin, “Performance Analysis of Cloud-based Deep Learning Models on Images Recovered without Channel Correction in OFDM System,” in *2022 27th Asia Pacific Conf. on Com. (APCC)*, Jeju Island, Korea, Republic of: IEEE, 2022, pp. 225–259.
[5] K. Sayood, *Introduction to data compression*. Morgan Kaufmann, 2017.
[6] I. Ahmad and S. Shin, “A novel hybrid image encryption–compression scheme by combining chaos theory and number theory,” *Signal Process. Image Commun.*, vol. 98, p. 116418, Oct. 2021, doi: 10.1016/j.image.2021.116418.
[7] M. Ribeiro, K. Grolinger, and M. A. M. Capretz, “MLaaS: Machine Learning as a Service,” in *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, Miami, FL, USA: IEEE, Dec. 2015, pp. 896–902. doi: 10.1109/ICMLA.2015.152.
[8] M. Tan and Q. Le, “Efficientnetv2: Smaller models and faster training,” presented at the International conference on machine learning, PMLR, 2021, pp. 10096–10106.
[9] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, “The German traffic sign recognition benchmark: a multi-class classification competition,” presented at the The 2011 international joint conference on neural networks, IEEE, 2011, pp. 1453–1460.