

Offloading Deep Learning Empowered Image Segmentation from UAV to Edge Server

Hüseyin Enes İlhan^a, Sedat Ozer^b, Güneş Karabulut Kurt^{a,c}, and Hakan Ali Cirpan^a

^aDepartment of Electronic and Telecommunication Engineering, Istanbul Technical University, Istanbul Turkey

^bDepartment of Computer Science, Bilkent University, Ankara Turkey, (sedat@cs.bilkent.edu.tr)

^cPoly-Grames Research Center, Department of Electrical Engineering, Polytechnique Montréal, Montréal, Canada

Abstract—Image and video analysis in unmanned aerial vehicle (UAV) systems have been a recent interest in many applications since the images taken by UAV systems can provide useful information in many domains including maintenance, surveillance and entertainment. However, a constraint on UAVs is having limited battery power and recent developments in the artificial intelligence (AI) domain encourages many applications to run computationally heavy algorithms on the taken UAV images. Such applications drain the power from the on-board battery rapidly, while requiring strong computationally strong resources. An alternative to that approach is offloading heavy tasks such as object segmentation to a remote (edge) server and perform the heavy computation on that server. However, the effect of the communication system and the used channel introduce noise on the transferred data and the effect of the noise due to the use of such LTE communication system on pre-trained deep networks has not been previously studied in the literature. In this paper, we study one such scenario where the images taken by UAVs and (the same images) transferred to an edge server via an LTE communication system under different scenarios. In our case, the edge server runs an off-the-shelf pretrained deep learning algorithm to segment the transmitted image. We provide an analysis of the effect of the wireless channel and the communication system on the final segmentation of the transmitted image on such a scenario.

Index Terms—Computational offloading, image segmentation, edge computing, UAV image processing, deep learning.

I. INTRODUCTION

Unmanned Aerial Vehicles (UAVs) equipped with cameras have been used in many domains including entertainment, surveillance, rescuing and maintenance. Many of those domains have applications requiring object tracking, object detection or object segmentation algorithms. Today's state of the art object tracking, detection or object segmentation algorithms utilize deep learning-based techniques and those techniques typically require heavy computation, large memory and power resources [1]–[3]. However, one of the most typical bottlenecks in UAV usage is the limited battery power of UAVs and therefore, it is not desired to use computationally heavy applications on UAV boards. Such a bottleneck enforced researchers to come up with various offloading techniques as in [4], [5]. While offloading is an interesting topic in various UAV applications, when deep learning-based techniques are considered, the offloading might have some constraints since the channel might affect the transmitted image significantly. That is because of the performance of standard deep learning techniques is sensitive to certain types of noise. The computer vision literature has studied such effects in various object segmentation and detection algorithms as in [6], however,

it has not been widely studied how the noise introduced by the communication system and the channel effects the performance of such deep networks.

In this work, we aim to study the performance of a standard and pre-trained deep learning based object segmentation algorithm, when the image segmentation task is offloaded from a UAV to a remote (edge) server under different conditions. In particular, we study how the fading wireless channel attenuation (in SNR which is mainly caused by the distance between the transmitter and the receiver), UAV's velocity (Doppler shift) and the transferred image resolution effect on the performance of pre-trained deep learning based segmentation algorithm: MobileNetV2 based on DeeplabV3 [7].

Our contributions include: (i) introducing the first work that studies the effect of wireless (Long-Term-Evolution, LTE) communication system under different channel parameters on the offloading the deep learning based segmentation task from a UAV to a remote server; (ii) comparing the segmentation results before and after the offloading under different communication scenarios, while giving the same image sets as input.

II. RELATED WORK

While UAV offloading has been studied heavily in many domains, those works typically focused on the transmission side of the systems and avoided the effect of the offloading process on the final results obtained at the edge server. To our best knowledge, this is the first work that studies the effect of offloading on the performance of deep learning-based segmentation algorithms that run on the edge server-side. Therefore, relevant work is limited.

Most of the relevant work on the offloading side focuses on how a UAV can efficiently be used as a mobile access point (mainly from the energy efficiency perspective) as in [8]. As another example, the work in [5] focused on how computing optimal offloading conditions under different communication scenarios. In [9], the authors introduced a method to use a UAV as a mobile edge server (MES) for providing task offloading services. Their goal was maximizing the migration throughput of the tasks and they proposed using a semi-Markov decision process (SMDP) without using transition probabilities for the problem. In [10], localization of multiple UAVs' hovering heights in disaster or emergency situations for public safety

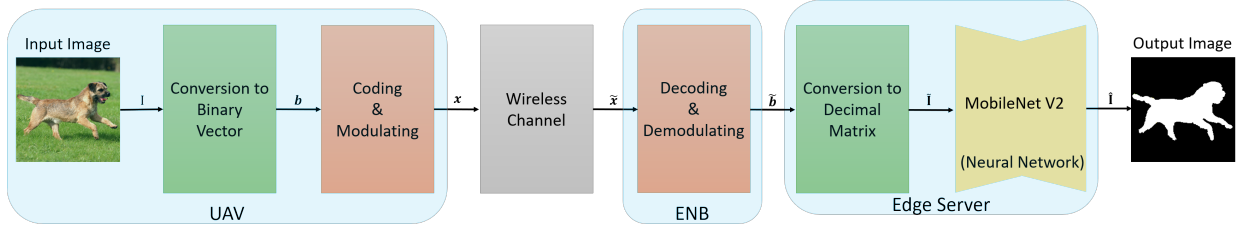


Fig. 1: Overview of our used system for deep learning-based image segmentation offloading from UAV to a remote server.

is studied, when the UAVs are used as mobile access points in 4G. The authors of [11] also studied the optimal height of UAVs when the UAVs are used as mobile edge servers to minimize the latency and the consumed power. In another study, when multiple UAVs are deployed and when multiple images are needed to be taken at different locations, the authors of [12] presented using Hungarian algorithm to assign multiple tasks to those multiple UAVs when the UAV size is the same as the task size. Then this work is extended to assign multiple tasks to multiple UAVs in [13] when the number of tasks is not matching the total number of available UAVs. The authors proposed using K-means and reinforcement learning algorithms for the problem considering energy efficiency. The problem of wireless charging of UAVs has also been studied in the literature as in [14] where the authors used microwave power transmission to charge UAVs.

The image and video analysis aspect of edge computing has also been studied intensively in the literature. For a detailed list of relevant work, please refer to [4]. However, there is no work that focuses on studying the effect of the wireless channel parameters and of the used communication system on the performance of offloading deep learning-based image segmentation task to a remote server.

Our work differs from the above-mentioned literature since by focusing on how the wireless communication parameters of LTE affect the performance of offloading the deep learning-based image segmentation task.

III. OBJECT SEGMENTATION WITH DEEP LEARNING

Deep learning has been an active research field in the AI world due to its significant contributions in many application domains [15]–[18]. One such application is semantic segmentation of images where the goal is semantically meaningful pixel-based extraction of the objects of interests. While many successful deep learning algorithms have been introduced in the literature, one problem that has been noticed, and also been an active research field in the AI community, is the sensitivity of deep learning-based algorithms to the noise as in [6]. There have been many solutions introduced for particular noise types as in [16] however, the problem of being sensitive to the noise is still an active research topic in AI. Another problem is the efficient computation of deep learning-based algorithms where the goal is obtaining deep architectures that are computationally efficient. MobileNetV2 [2] is one such algorithm that is designed for mobile platforms and it is paired with DeepLabV3 [7] model for segmentation. It can

perform multiple tasks such as segmentation, classification and object detection and its pretrained models are also available to be used on mobile platforms. MobileNetV2 is trained on PASCAL VOC 2012 [19] image segmentation dataset.

IV. SYSTEM OVERVIEW

Our system includes a mobile edge system equipped with a camera (UAV) and a wireless transceiver system for LTE, a wireless communication channel (LTE fading channel) and a remote server that is connected to a base station (eNB). The goal in our system is transmitting (offloading) the deep learning-based image segmentation task from UAV to the remote edge server. By doing so, our goal is freeing UAV's resources for energy efficiency as image segmentation (even with mobile deep algorithms) is a power-hungry task.

The overview of our system is given in Fig. 1. The image is taken by the UAV's camera is a colored (3-channel RGB) image where each pixel color is a discrete value and saved by 8-bits for each channel. The first step in our system is converting the taken image into a large binary vector. The next step focuses on coding and modulation to be transmitted over the wireless channel. The binary data streams \mathbf{b} are coded, modulated and mapped into LTE frame. The LTE frame \mathbf{x} is transmitted over the noisy fading channel and at the end of the channel $\hat{\mathbf{x}}$ is received on the eNB site. The received data is threaded by using channel estimation and equalization techniques to eliminate the channel noise and distortion effects. The corrected signal is then decoded, demodulated and remapped into binary data streams. The decoded binary stream $\hat{\mathbf{b}}$ is converted back to decimal-matrix format to form the received image. Once the received image is re-obtained on eNB, it is transmitted over a wired environment to the edge server where the pretrained deep learning algorithm is used. The noise of the wired environment is ignored.

LTE system uses turbo codes for channel coding to deal with the error correction and as we simulate a real LTE system, the error correction is also handled in our simulation by a turbo coding block with 1/3 Turbo code rate. We utilize a public image set from the FSS-1000 dataset [20] as the input and use MobileNetV2 as a common image segmentation network which is a well known segmentation algorithm that is commonly utilized on many mobile platforms in our simulations.

V. COMMUNICATION MODEL

The communication system of our study is based on the LTE physical uplink layer. Here, we model the fading channel

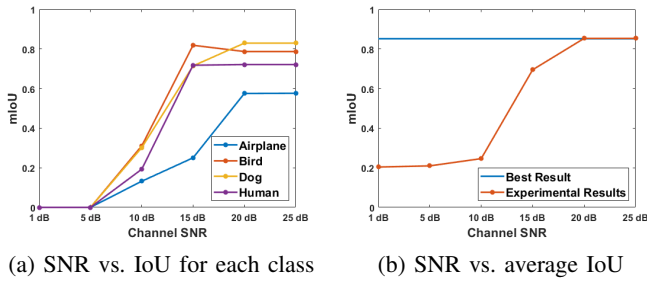


Fig. 2: SNR vs. IoU results.

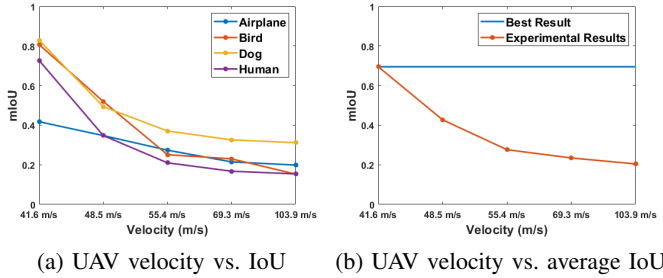


Fig. 3: UAV velocity vs. IoU results.

as Rayleigh distributed where the delay profile is set to Generalized Method of Exact Doppler Spread (GMEDS) [21], [22]. The attenuation of the transmitted signal due to the path loss of the channel [23] is defined as:

$$P_r[\text{dBm}] = P_t[\text{dBm}] + K[\text{dB}] - 10\gamma \log_{10}(d/d_0) - 20 \log_{10}|h_i|, \quad (1)$$

where h_i is the eNB antenna's height, P_r is the power of the received signal, P_t is the power of the transmitted signal, K is a constant depending on antenna characteristics and average channel attenuation, d is the distance of UAV to eNB, d_0 is the reference distance and γ is the constant for path loss exponent. The transmitted data rate is formulated as:

$$R = B \times \log_2(1 + \text{SNR}), \quad (2)$$

where B is the bandwidth and SNR is the signal to noise ratio. The SNR at the receiver is the ratio of transmitter power P_t and channel attenuation coefficient g that is affected by the path loss and fading over the noise ratio σ^2 . The transmitted data rate is affected by the SNR of the channel. Doppler shift is the frequency shift in the received signal and caused by the relative movement of the receiver or the transmitter. Doppler shift of a wireless communication system is formulated as [23]:

$$f_d = v \times \cos(\theta) \times f_c/c, \quad (3)$$

where f_d is the Doppler shift in Hz, v is the relative velocity, θ is the angle between transmitter & receiver, f_c is the carrier frequency of the transmitted signals and c is the speed of light.

Image binary data streams are the input signals of the communication subsystem. The binary stream \mathbf{b} is encoded with Turbo coding technique that is compliant with 3GPP Release 10. The Turbo encoded bit sequence is modulated with 256 QAM to transmit large image files in a challenging simulation situation. In 256-QAM, there are 256 symbol states and each symbol carries 8 bits. Each symbol is mapped into an LTE frame structure that consists of 100 resource blocks (RB) with the demodulation reference signal (DSR). Then each

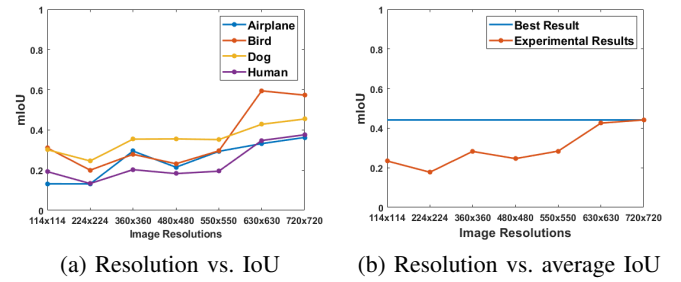


Fig. 4: Image Resolution vs. IoU results.

frame \mathbf{x} is modulated with the single-carrier frequency-division multiple access (SC-FDMA). SC-FDMA is the multiple access technique of the uplink LTE system. The main advantage of SC-FDMA is the lower peak-to-average power ratio (PAPR) that contributes to transmit power efficiently and reduce the cost of the power amplifier. The modulated frames are parsed through the noisy fading channel. We model the channel with the Rayleigh fading channel model where each individual of the channel matrix has a Rayleigh distribution, and the additive white Gaussian noise (AWGN) with varying SNR values is added. The output of the channel $\hat{\mathbf{x}}$, which is the noisy form of the input frames demodulated with SC-FDMA and the channel estimated with cubic interpolation with 12 averaged resource elements. Estimated channel is equalized with minimum mean square error (MMSE) to minimize both inter-symbol interference (ISI) and additive noise effect [24]. The equalized received binary data $\hat{\mathbf{b}}$ is the output signal of the communication layer. The data streams are then transmitted to the edge computing server for segmentation.

VI. EXPERIMENTS

We simulate our above-mentioned system using the LTE telecommunication uplink physical layer of Matlab LTE Toolbox. According to our simulation scenario, the images taken by a UAV are transmitted to the edge server through the LTE uplink physical layer and the bandwidth is fixed at 20MHz and $f_c 2.5\text{GHz}$. We study the effect of channel noise, UAV velocity and sending an image at different resolutions in our system on the segmentation results.

Channel is modeled by the Rayleigh fading to investigate a more challenging communication scenario that has no any Line-of-Sight (LoS) path component between the receiver and the transmitter. We also added Gaussian noise to the channel

TABLE I: Simulation parameters used in our experiments.

Carrier Frequency	2.5 GHz
Code Rate	1/3
Cyclic Prefix	Normal
Duplex Mode	FDD
Resource Block	100
Subcarrier Spacing	15 KHz
Band Width	20 MHz
Modulation	256 QAM
Channel Fading Type	Rayleigh
SNR (dB)	1, 5, 10, 15, 20, 25
Velocity (m/s)	41.6, 48.5, 55.4, 69.3, 103.9
Resolution	114 ² , 224 ² , 360 ² , 480 ² , 550 ² , 630 ² , 720 ²

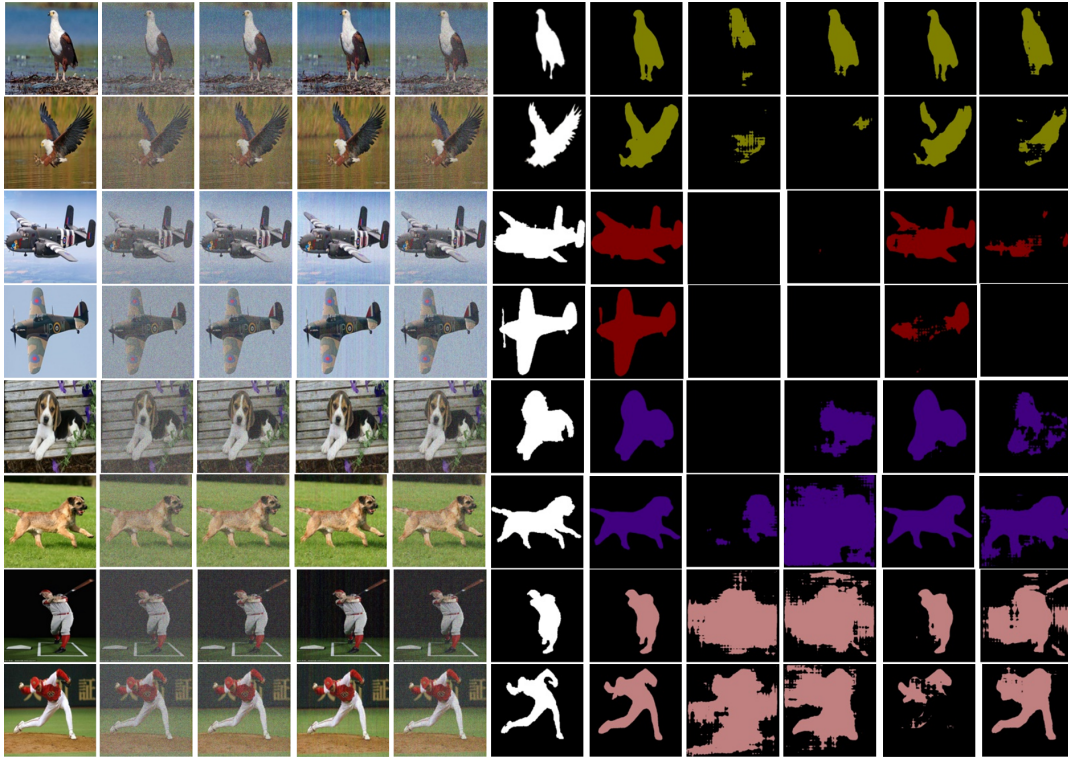


Fig. 5: Qualitative results obtained at different system parameters for sample images. The 1st column shows sample (input) images. The 2nd, 3rd and 4th columns show the results obtained at 5dB, 10dB and 15dB SNR values respectively (where UAV velocity is 41.6 m/s). The 5th column shows the result, when SNR=15dB but the UAV velocity is set to 55.4 m/s. The 6th column shows the ground truth segmentation, the 7th column shows the segmentation results of MobileNetv2, when there is no noise and no Doppler effect. The 8th, 9th, 10th and 11th columns show the offloaded segmentation results obtained for the images as shown in the 2nd, 3rd, 4th and 5th columns, respectively.

with varying SNR values. The velocity of the UAV is another key parameter that causes a frequency shift at the received signal. To measure the effect of that phenomenon, we also simulate that scenario by varying the Doppler shift frequency. Our simulation parameters are given in Table I.

A. Used metrics and the Dataset

In our experiments, we use Intersection Over Union (IoU) to measure the segmentation performance which is a common metric used in image segmentation tasks. it is defined as:
$$IoU = \frac{GroundTruth \cap ModelPrediction}{GroundTruth \cup ModelPrediction}$$
, where *GroundTruth* is the true segmentation and *ModelPrediction* is the predicted segmentation output. In this work, we use MobileNetV2 architecture as our pretrained segmentation algorithm to run on the server side [7]. We used a portion of the FSS-1000 dataset [20] as input which contains images with 224x224 pixel resolutions, varying from tiny objects to cartoon characters. Each class contains 10 images along with their segmentation ground truth data. We picked bird, human, airplane and dog classes to form a random input image set.

B. Results

Here, we study the effect of various parameters of a communication system on the offloading process of segmentation. We grouped those studies under three subjects: (i) effect of

noisy channel on segmentation; (ii) effect of UAV velocity on segmentation; (iii) resolution effect on segmentation IoU.

1) *Effect of noisy channel on segmentation:* We first studied the effect of the communication channel attenuation and its noise on the segmentation results. For that purpose, we simulated the communication channel as a *Rayleigh Fading Channel* model. AWGN noise is added to the channel at various SNR values. Those SNR values that are used in our experiments vary from 1dB to 25dB. To study how SNR affects the segmentation, we kept the transmitted image resolution fixed at 480x480 and we kept UAV's velocity at 41.6 m/s. The results are summarized in Fig. 2a for four object class separately (note that a segmentation algorithm may yield different IoU-based performance on different classes due to the class-imbalance problems and other class-based image properties). The averaged results are presented in Fig. 2b where the constant blue line represents the result obtained for those images when there is no any channel effect.

2) *Effect of UAV velocity on segmentation:* Next, we studied the effect of changing the UAV velocity on segmentation. We assumed that the UAV is in motion while offloading its captured images to the eNB. This effect is reflected in our communication system as Doppler Shift. The relation between the UAV velocity and the Doppler Shift is given in Eq. 3. We fixed SNR at 15dB and image resolution at 480x480. In

Fig. 3a, we summarized how the segmentation IoU changes at those velocities for each of the four class. Fig. 3b summarizes how the average IoU changes at each velocity for all of the four classes. The blue line represents the results obtained in the situation where there is no Doppler effect.

3) *Effect of image resolution on segmentation IoU:* The images are sent over the LTE noisy fading channel at different resolutions to observe the effect of the amount of the data transmitted over the noisy channel on the segmentation neural network. UAV velocity is fixed at 41.6 m/s and SNR is fixed at 15dB. The results are summarized in Fig. 4a for each class separately. The Fig. 4b for averaged classes results. Fig. 5 shows qualitative results for eight different images. It shows how the segmentation result changes for the same input at different parameters. As the figure suggests, the channel and Doppler shift can be significant at certain levels.

VII. CONCLUSION

In this paper, we studied the effect of offloading a deep learning-based image segmentation task from a UAV to a remote server over an LTE wireless communication system. To our best knowledge, this is the first work studying the effect of the noise introduced by a wireless communication system on a remote server where we run a pretrained deep learning-based object segmentation algorithm. In particular, we studied how the performance of the deep learning algorithm is affected by different SNR values (different distance values), at different UAV velocities and at different image resolutions, when there is only one UAV and one eNB exists in the system. Future work involves more complicated scenarios involving interference of having multiple UAVs and eNBs in a 5G system.

Our results demonstrate that when offloading process is concerned, the performance of pre-trained deep learning-based segmentation algorithms can be effected significantly by yielding low performance depending on different communication system parameters. That is due to the fact that the pretrained networks are typically trained on image sets that are taken by cameras directly. Therefore, the noise factor of a communication system is not considered by such deep networks. We provided a detailed analysis on that aspect to see at what rate the segmentation results can deteriorate due to the noise introduced by the used communication system.

We observed that at low-velocity values (smaller than 41.6 m/s), the Doppler effect remains insignificant on the segmentation results in a wireless LTE system. However, after that speed, the effect of Doppler shift becomes significant.

ACKNOWLEDGMENT

This paper has been produced benefiting from the 2232 International Fellowship for Outstanding Researchers Program of TÜBİTAK (Project No:118C356). However, the entire responsibility of the paper belongs to the owner of the paper. The financial support received from TÜBİTAK does not mean that the content of the publication is approved in a scientific sense by TÜBİTAK.

REFERENCES

- [1] B. M. Albaba and S. Ozer, "SyNet: An ensemble network for object detection in UAV images," in *International Conference on Pattern Recognition (ICPR2020)*, 2020.
- [2] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510–4520.
- [3] D. Gözen and S. Ozer, "Visual Object Tracking in Drone Images with Deep Reinforcement Learning," in *International Conference on Pattern Recognition (ICPR2020)*, 2020.
- [4] Q. Zhang, H. Sun, X. Wu, and H. Zhong, "Edge video analytics for public safety: A review," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1675–1696, 2019.
- [5] D. Callegaro and M. Levorato, "Optimal computation offloading in edge-assisted UAV systems," in *2018 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2018, pp. 1–6.
- [6] Z. Wu, S.-N. Lim, L. S. Davis, and T. Goldstein, "Making an invisibility cloak: Real world adversarial attacks on object detectors," in *European Conference on Computer Vision*. Springer, 2020, pp. 1–17.
- [7] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," *arXiv preprint arXiv:1706.05587*, 2017.
- [8] Y. Zeng and R. Zhang, "Energy-efficient UAV communication with trajectory optimization," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 3747–3760, 2017.
- [9] J. Li, Q. Liu, P. Wu, F. Shu, and S. Jin, "Task offloading for UAV-based mobile edge computing via deep reinforcement learning," in *IEEE/CIC International Conference on Communications in China (ICCC)*, 2018, pp. 798–802.
- [10] C. Ting, X. Yun, Z. Xiangmo, G. Tao, and X. Zhigang, "4G UAV communication system and hovering height optimization for public safety," in *IEEE International Conference on e-Health Networking, Applications and Services (Healthcom)*, 2017, pp. 1–6.
- [11] F. Costanzo, P. Di Lorenzo, and S. Barbarossa, "Dynamic resource optimization and altitude selection in UAV-based multi-access edge computing," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 4985–4989.
- [12] K. Kim and C. S. Hong, "Optimal task-UAV-edge matching for computation offloading in UAV assisted mobile edge computing," in *Asia-Pacific Network Operations and Management Symposium*, 2019, pp. 1–4.
- [13] K. Kim, Y. M. Park, and C. S. Hong, "Machine learning based edge-assisted UAV computation offloading for data analyzing," in *International Conference on Information Networking (ICOIN)*, 2020.
- [14] Y. Liu, M. Qiu, J. Hu, and H. Yu, "Incentive UAV enabled mobile edge computing based on microwave power transmission," *IEEE Access*, vol. 8, pp. 28 584–28 593, 2020.
- [15] Y. Yıldırım, S. Özer, and H. A. Çırpan, "Deep receiver design for multi-carrier waveforms using cnns," in *International Conference on Telecommunications and Signal Processing (TSP)*. IEEE, 2020.
- [16] A. Agarwal, M. Vatsa, R. Singh, and N. K. Ratha, "Noise is inside me! generating adversarial perturbations with noise derived from natural filters," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 774–775.
- [17] R. Valiente, M. Zaman, S. Ozer, and Y. P. Fallah, "Controlling steering angle for cooperative self-driving vehicles utilizing CNN and LSTM-based deep networks," in *IEEE Intelligent Vehicles Symposium (IV)*, 2019, pp. 2423–2428.
- [18] M. Turhan, E. Öztürk, and H. A. Çırpan, "Deep convolutional learning-aided detector for generalized frequency division multiplexing with index modulation," in *International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2019, pp. 1–6.
- [19] M. Everingham, S. A. Eslami, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge: A retrospective," *International Journal of Computer Vision*, vol. 111, no. 1, pp. 98–136, 2015.
- [20] X. Li, T. Wei, Y. P. Chen, Y.-W. Tai, and C.-K. Tang, "Fss-1000: A 1000-class dataset for few-shot segmentation," in *Conference on Computer Vision and Pattern Recognition*, 2020.
- [21] M. Patzold, C.-X. Wang, and B. O. Hogstad, "Two new sum-of-sinusoids-based methods for the efficient generation of multiple uncorrelated Rayleigh fading waveforms," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, pp. 3122–3131, 2009.
- [22] L. Polak and T. Kratochvil, "Simulation and measurement of the transmission distortions of the digital television DVB-T/H part 3: Transmission in fading channels," *Radioengineering*, vol. 19, no. 4, pp. 703–711, 2010.
- [23] A. Goldsmith, *Wireless communications*. Cambridge Univ. Press, 2005.
- [24] F. Conillera Vilar, "Implementation of zero forcing and MMSE equalization techniques in OFDM," Ph.D. dissertation, 2015. [Online]. Available: <http://hdl.handle.net/2117/78037>