An Explainable Multi-Task Learning Approach for RF-based UAV Surveillance Systems

Rubina Akter*, Van-Sang Doan†, Ahmad Zainudin‡, and Dong-Seong Kim*
*‡Networked Systems Laboratory, *Department of IT Convergence Engineering,
‡Department of Electronic Engineering †Department of Communication and Radar
*‡Kumoh National Institute of Technology, Gumi, South Korea,
†Vietnam Naval Academy, Nhatrang, Vietnam
{rubinaakter2836, zai, dskim*‡}@kumoh.ac.kr, {doansang.g1†}@gmail.com

***Abstract*—Unmanned aerial vehicles or drones are ubiquitous among people, which can lead to technological, security, and community safety issues that must be addressed, monitored, and avoided. Intelligence services are always on the search for potential technology and intelligent systems that can identify drones. A potential drone surveillance system must be capable of detecting, localizing, identifying, recognizing the modes, and combating unauthorized drones. In this paper, we introduce a Multi-Task Learning (MTL) neural network for drone detection, identification, and drone mode detection using Radio Frequency (RF) signals. Due to the semantic abstraction of the drone RF signals, a single-task learning method can not fully meet the demands of the current anti-drone system. Moreover, executing each of the tasks, such as drone detection, type identification, and activity recognition, individually takes longer time, which is not applicable in a real-time drone surveillance system. Therefore, this paper proposes an MTL approach leveraging convolution layers to perform three tasks in parallel. A cross-entropy loss function used as the objective function optimization to improve the accuracy of the multiple tasks. The empirical results shows that the proposed MTL model achieve a better recognition accuracy compared to the existing solutions.**

***Index Terms*—Convolution neural network, drone detection and classification, multi-task learning neural network, radio frequency signal.**

I. Introduction

In the drone industry, there has been significant technological advancement [1]. Drones are increasingly being outfitted with cutting-edge technologies and sensors like GPS, LIDAR, radar, and vision sensors. Nowadays, drones are being used for a variety of purposes, including cinematography, agriculture, surveillance, and leisure activities, thanks to the advancements in drone technology. Aside from these advantageous functions, drones are also being utilized for illicit purposes, posing a threat to public safety [2]. Violations of civil security, drug trafficking, firearm smuggling, carrying explosives materials and breaching security-sensitive locations such as airports and nuclear power facilities are among the criminal actions.

A number of counter-unmanned aircraft systems have been developed. It has been proposed to disable drone attacks, which are hard and soft interception, the two main types of interception (a solution that is either kinetic or non-kinetic). Surveillance a drone with a trained bird of prey, a web gun [3], a laser beam, and a weapon are among the kinetic options. On the other hand, the non-kinetic options comprise GPS spoofing and RF jamming to confuse a drone's tracking system. The presence of a drone should be identified and categorised ahead of time, regardless of the method chosen for any situation. Automatically detecting and identifying a drone is a difficult operation. Radar detection [4], vision detection, acoustic detection, and RF fingerprint-based detection [5] are some of the most commonly used technologies for detecting and classifying drones. Different studies also proposed hybrid technologies for drone detection and classification.

The back-scattered RF signal is used in aerial radar surveillance to detect and classify drones. Due to the narrow radar cross-section area, the traditional radar system fails to detect mini-drones. To solve this challenge, researchers used a multistatic radar [6] or a Frequency Modulated Continuous Wave (FMCW) radar to identify and categorize a quadcopter or multi-rotor UAV's micro-Doppler signature. Vision-based detection covers both visual and thermal detection. Researchers presented numerous drone detection methods utilizing this technology in [7], [8]. The detection of the vision-based method is comparatively accurate, but it requires a direct line of sight (LOS) between the drone and the camera, and its performance is heavily reliant on sunshine and weather conditions such as dust, rain, fog, and clouds. In addition, the likeness of a bird to a drone makes it more difficult for a video detector to detect. The acoustic-based drone surveillance system uses microphones to detect the presence of flying drones by monitoring their sound [9]. A Hidden Markov Model is proposed to analyze drone sounds and identify the flying drones [10]. Acoustic detection works well in quiet or low-noise environments. However, in noisy environments, such as urban or industrial areas, or near seashores, performance suffers.

RF sensing [11] is one of the most viable technologies for detecting the presence of a drone [12]. Commercial drones communicate with their ground control station using RF signals for flight control and navigation, live video transmission, and telemetry data transfer. An RF fingerprint drone detection system can potentially detect a drone by analyzing the transmission frequency spectrum. In the literature, a few RF-based drone detection algorithms have been proposed. The detection was carried out in [13], by determining the data

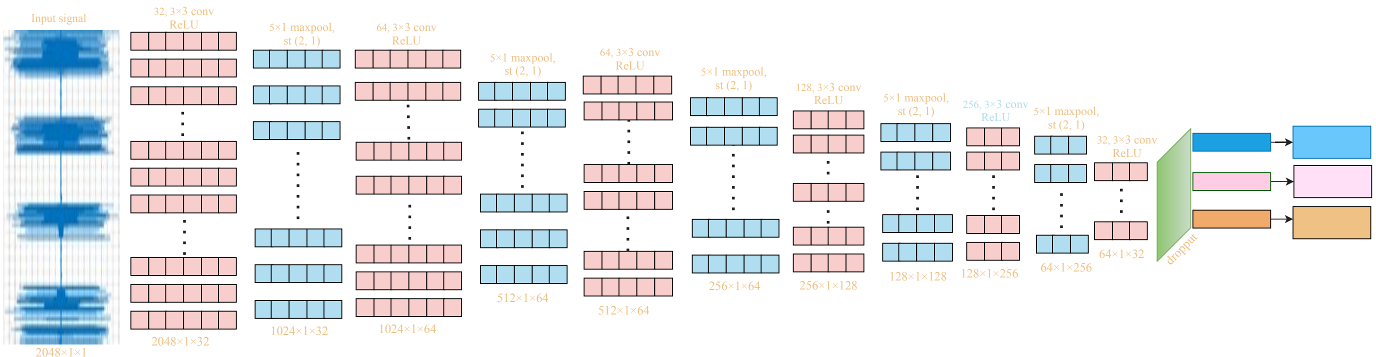

Fig. 1. Training window of the proposed Multi-task neural network.

packet duration of a drone's communication channel. These detection approaches are ineffective because an application communicating with an access point at the same packet transfer rate or packet length as a drone can readily spoof the detector. However, most of the drone surveillance papers discussed drone detection and classification by separately learning the model [14]. Moreover, the existing literature did not detect the drone's flying mode or activities, which is a crucial part of the drone surveillance system. Motivated by these aforementioned issues, in this literature, we intend to create a drone surveillance framework combining both drone detection, type classification and activity recognition.

II. Proposed Multi-Tasking Neural Network

Representation learning is a critical subject in the area of machine learning [15]. In recent years, there has been a growing interest in nonlinear feature transformation from different tasks utilizing several layers of deep networks [16]. Multi-task learning is a learning technique that requires the utilization of other related tasks to boost the generalization performance of the learning tasks. It is extensively used in transfer learning, particularly in the field of natural language processing. A multi-task recurrent neural network model-based collaborative joint training technique was discussed in [17]. This study illustrated that the multi-task technique could enhance the efficiency of both instant voice and speaker identification activities when compared to single-task systems. More multi-task learning models have been discussed in [18], [19]. However, nowadays, multi-task learning is the new research trend in the field of deep learning, which is widely acknowledged and offers more benefits in terms of accuracy and computational complexity compared to the single task learning. Therefore, these study adopts multi-task neural network to detect, classify, and recognize activities of the drone in parallel using a single network model which targets to solve the problem of the existing drone detection models.

Fig. 1. illustrates the architecture of the multitask learning network, which consists of six convolution layers. Each of the convolution layers is followed by a Rectified Linear Unit (ReLU) and a maxpool layer to extract the unique features and capture long-time dependencies from the input signals. For multi-tasking processing, three fully connected (fc) layers are employed in parallel, where the corresponding hidden layers share the extracted features to perform the specific task. This study uses the droneRF dataset, which is composed of three categories of datasets, such as drone RF and interference signals, three types of drone signals, and drone mode signals. To train the network, we configured these different categories dataset in the same dimension as $2048 \times 1 \times 1$. This input signal is processed through (3×3) convolution layer with a 32 channel. In general, the convolution layer processes the activation signals of the previous layers as $F^{l-1} \epsilon \mathbb{R}^{x \times y_h^{l-1} \times y_w^{l-1} \times y_C^{l-1}}$. Here, F, and x presents the feature map and total signal data at $l-1$ index layer. The kernels of the convolution layer are presented as $K^{[l]} \epsilon \mathbb{R}^{k^{[l]} \times k^{[l]} \times y_C^{l-1} \times y_C^{l}}$, where $k^{[l]} \times k^{[l]} \times y_C^{l-1}$ is the dimension of the kernels and y_C^l is the total number of kernel. The convolution layer conducts convolution operations between the input signals and each filter, adds bias, and applies ReLU as the non-linearity. The functions of the ReLU layer can be shown as:

$$y(F) = \begin{cases} F & \text{if } F \geq 0 \\ 0 & \text{if } F < 0 \end{cases}, \tag{1}$$

The ReLU layer conducts a threshold operation to scale the output of the convolution layer. Moreover, it also helps the network to prevent the vanishing gradient problem and offer better learning convergence during training. The output of the ReLU layer is received by a maxpool layer. The maxpool layer is organised as 5×1 pool size with the stride of $(2, 1)$. This configuration of the maxpool layer reduces the input dimension by half and offers reduced computational expense to the successive layers. Following the same design structure, we have applied six convolutions, ReLU, and the maxpool layers. Then we added a dropout layer with a 20% dropout rate to address the issue of overfitting.

As the proposed networks perform three tasks instead of training a separate network for each task; hence it has three output layers and share the processed features of the preceding blocks. Therefore, the output feature maps from the dropout layer is fed to the three fully-connected layers. Each of the fully-connected layers has a softmax layer to generate the class probabilities of each task. In our network, the first fully-connected layer is to classify the drone signal and the background signal. The second and third output layers work for the drone's signal identification and mode detection, respectively.

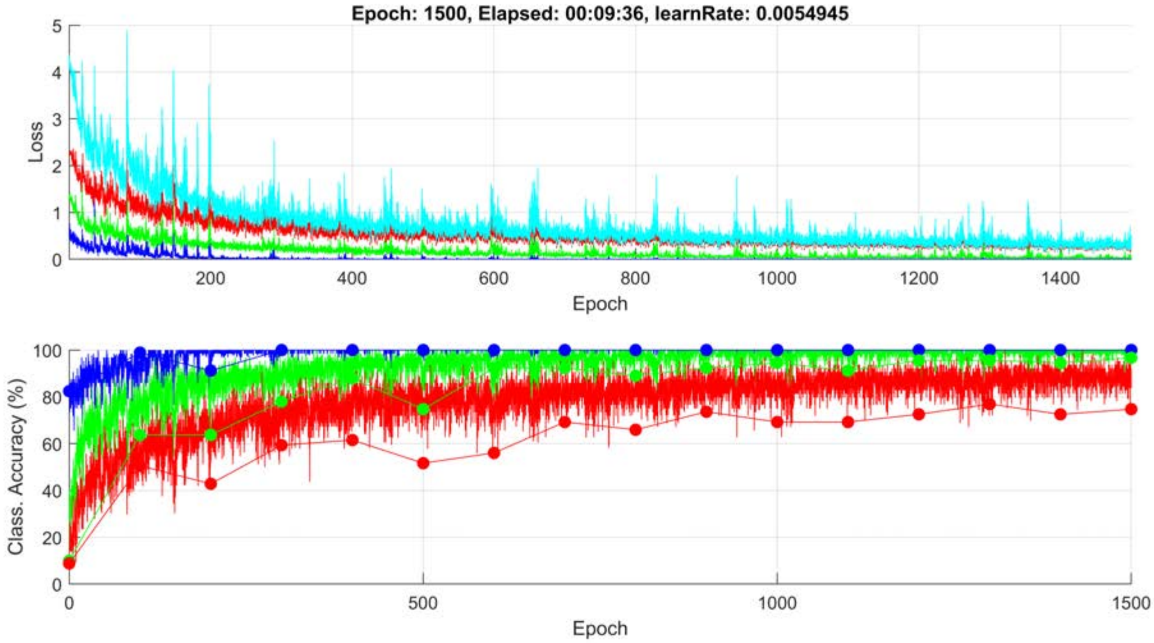

Fig. 2. Training window of the proposed Multi-task neural network.

The output of the fully-connected layers can be shown as:

$$y_1 = f\left(F_1.W_1 + b_1\right), \quad (2)$$

$$y_2 = f\left(F_2.W_2 + b_2\right), \quad (3)$$

$$y_3 = f\left(F_3.W_3 + b_3\right). \quad (4)$$

here, y_1, y_2, and y_3 is the output of the first task, second task, and third task fully-connected layers, respectively.

III. Experimental Results

A. Dataset Description

The dataset used in this study is an open source dataset generated by [20]. The dataset was collected from three types of drones, Bebop, AR, and Phantom. During data gathering, all drones use WiFi, operated at 2.4 GHz frequency band. They employed two RF receivers to collect the entire 2.4 GHz bandwidth. The parameter settings to record data is given in Table I. The RFdataset consists of raw RF signals, which were collected in four functioning modes, such as on and connected, hovering, flying without video recording, and flying with video recording. There are a total of 227 segments, each of which is made up of two equal-sized portions, each of which contains one million samples, for a total of 454 record files. Table II shows the total number of segments and samples for the recordings in the created drone RF database.

B. Result Analysis

In this section, we analyze the efficiency of the proposed multi-task CNN model that was presented in Section II. The proposed network performs three tasks, such as drone detection (two classes), type identification (four classes), and activity recognition (ten classes) in a single training consideration. In this case, we create our own custom training window as shown in Fig. 2. To calculate the difference between the probability distributions, the training option configured the cross-

TABLE I. System specification and parameter description of RFdataset generation.

Parameter Description	**System1**	**System2**
USRP-2943	RIO0	RIO0
Active channel	RX2	RX2
Radio frequency band	Low	High
Carrier frequency (MHz)	2422	2462
IQ rate (MHz)	40	40
Number of samples per segment	10^7	10^7
Gain	30	30

TABLE II. Detail description of the developed drone RF database including the number of raw RF samples and segments for each class.

Level	Class	Segments	Samples
1	Drone	186	3720×10^6
	No Drone	41	820×10^6
2	Bebop	84	1680×10^6
	AR	81	1620×10^6
	Phantom	21	420×10^6
	No Drone	41	820×10^6
3	Bebop mode 1	21	420×10^6
	Bebop mode 2	21	420×10^6
	Bebop mode 3	21	420×10^6
	Bebop mode 4	21	420×10^6
	AR mode 1	21	420×10^6
	AR mode 2	21	420×10^6
	AR mode 3	21	420×10^6
	AR mode 4	18	420×10^6
	Phantom mode 1	21	420×10^6
	No drone	41	420×10^6

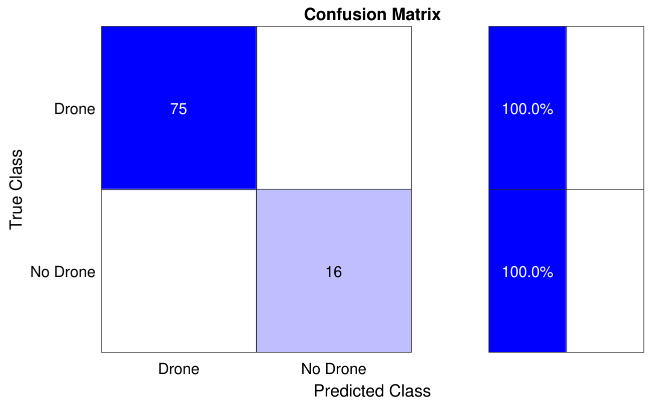

Fig. 3. Confusion matrix of drone detection.

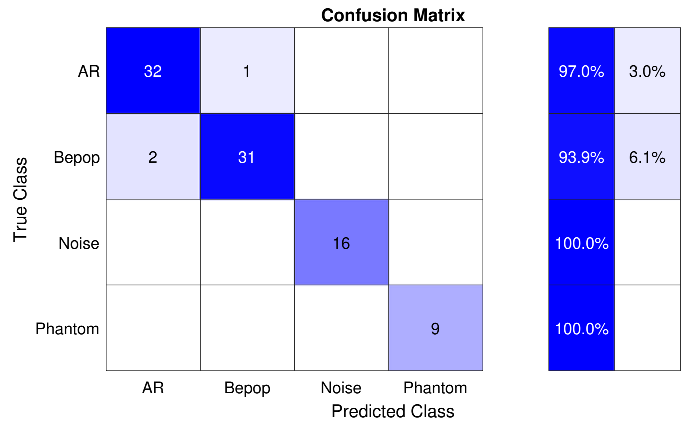

Fig. 4. Confusion matrix of drone type classification.

entropy loss function. The network training is carried out using 1500 epochs. Fig. 2. shows the loss and the accuracy of the multiple tasks, where blue, light green, and red colors indicate drone detection, type identification, and activity recognition, respectively. To clearly demonstrate the results, the confusion matrix of each task is shown in Fig. 3, Fig. 4, and Fig. 5. In the case of separating the drone signal from the interfering signals, the proposed model achieved a 100% classification accuracy. For drone type identification, the multi-task model achieved 96.70% accuracy rate. As shown in Fig. 4, the model gets a little confused to uniquely identify the AR drone and the bepop drone.

Lastly, Fig. 5. shows the confusion matrix of the drone mode detection of the three types of drones. Notably, the experiment considered four types of modes, such as "on and connected", "hovering", ”flying with video recording", and ”flying without video recording". Due to the large number of classes and similar features, the proposed model achieved 74.72% recognition accuracy, which is less compared to the other tasks. However, to evaluate the efficiency of the proposed model, we also compared the performance with three other deep learning models, named as, Existing-CNN [14], InceptionNet [21], and ResNet [22]. As shown in Fig. 6., almost all models performed well for drone detection (approximately 100% accuracy except for Existing-CNN (98.80%)). The accuracy of the InceptionNet, Existing-CNN, ResNet and the proposed model for drone type identification and mode recognition is 91.47%, 93.87%, 94.8%, 96.7%, and 62.38%, 68.12%, 68.13%, and 74.72%, respectively. From this analysis, it is revealed that the proposed multi-task model outperformed the existing CNN models.

IV. Conclusion

Given the rising number of instances of drone misuse, new technology is needed to assist in the enforcement of laws and regulations. In this research, we implement a MTL neural network for the detection, identification, and activity recognition of drones sensing RF signals. The proposed approach was trained and evaluated using a benchmark dataset, with the experimental findings indicating that employing RF signals in conjunction with CNN for drone detection and identification is successful and feasible. The proposed MTL model is composed of six convolution layers mounted sequentially with an increasing number of filters to extract the most disentangled features that contribute to the improvement of the model accuracy. Furthermore, the results of the studies demonstrated that CNN outperformed other existing CNN techniques reported in the literature. In future work, we will include

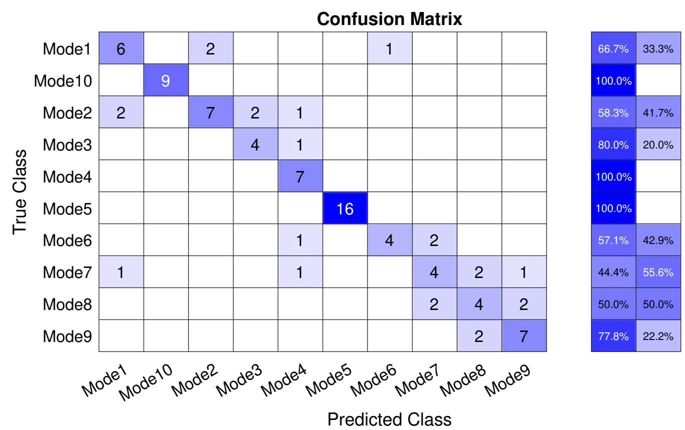

Fig. 5. Confusion matrix of drone mode detection.

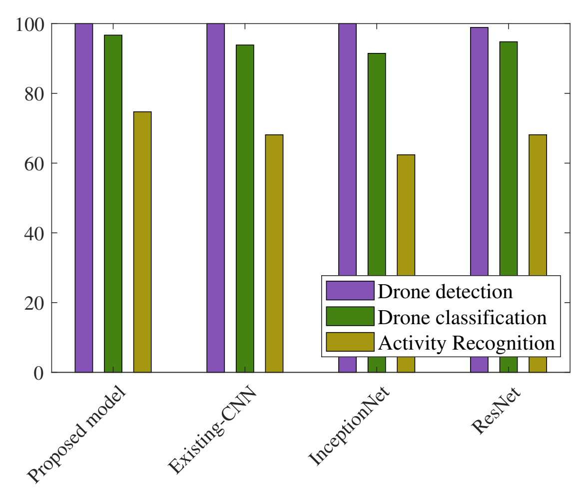

Fig. 6. Performance comparison between the multi-task and the existing CNN models.

the computation complexity analysis and other performance metrics measurement, which will act as the evidence of the potential anti-drone system in the real-time environment.

V. Acknowledgement

This research was supported by the MSIT (Ministry of Science and ICT), and Priority Research Centers Program, Korea, under the Grand Information Technology Research Center support program (IITP-2022-2020-0-01612) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation) and the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology"(2018R1A6A1A03024003), respectively.

References

[1] R. Akter, V.-S. Doan, J.-M. Lee, and D.-S. Kim, "CNN-SSDI: Convolution neural network inspired surveillance system for UAVs detection and identification," *Computer Networks*, vol. 201, p. 108519, 2021.

[2] B. Taha and A. Shoufan, "Machine learning-based drone detection and classification: State-of-the-art in research," *IEEE access*, vol. 7, pp. 138 669–138 682, 2019.

[3] D. Sathyamoorthy, "A review of security threats of unmanned aerial vehicles and mitigation steps," *J. Def. Secur*, vol. 6, no. 1, pp. 81–97, 2015.

[4] R. Akter, M. Golam, J.-M. Lee, and D.-S. Kim, "Doppler Radar-based Real-Time Drone Surveillance System Using Convolution Neural Network," in *2021 International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, 2021, pp. 474–476.

[5] R. Akter, V.-S. Doan, G. B. Tunze, J.-M. Lee, and D.-S. Kim, "RF-based UAV surveillance system: A sequential convolution neural networks approach," in *2020 International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, 2020, pp. 555–558.

[6] F. Fioranelli, M. Ritchie, H. Griffiths, and H. Borrion, "Classification of loaded/unloaded micro-drones using multistatic radar," *Electronics Letters*, vol. 51, no. 22, pp. 1813–1815, 2015.

[7] R. Stolkin, D. Rees, M. Talha, and I. Florescu, "Bayesian fusion of thermal and visible spectra camera data for mean shift tracking with rapid background adaptation," in *SENSORS, 2012 IEEE*. IEEE, 2012, pp. 1–4.

[8] S. O. Ajakwe, V. U. Ihekoronye, R. Akter, D.-S. Kim, and J. M. Lee, "Adaptive Drone Identification and Neutralization Scheme for Real-Time Military Tactical Operations," in *2022 International Conference on Information Networking (ICOIN)*. IEEE, 2022, pp. 380–384.

[9] A. Bernardini, F. Mangiatordi, E. Pallotti, and L. Capodiferro, "Drone detection by acoustic signature identification," *Electronic Imaging*, vol. 2017, no. 10, pp. 60–64, 2017.

[10] M. Nijim and N. Mantrawadi, "Drone classification and identification system by phenome analysis using data mining techniques," in *2016 IEEE Symposium on Technologies for Homeland Security (HST)*. IEEE, 2016, pp. 1–5.

[11] R. Akter, M. Golam, V.-S. Doan, J.-M. Lee, and D.-S. Kim, "IoMT-Net: Blockchain Integrated Unauthorized UAV Localization Using Lightweight Convolution Neural Network for Internet of Military Things," *IEEE Internet of Things Journal*, 2022.

[12] R. Akter, V.-S. Doan, T. Huynh-The, and D.-S. Kim, "RFDoA-Net: An efficient convnet for RF-based DoA estimation in UAV surveillance systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 11, pp. 12 209–12 214, 2021.

[13] P. Kosolyudhthasarn, V. Visoottiviseth, D. Fall, and S. Kashihara, "Drone detection and identification by using packet length signature," in *2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE)*. IEEE, 2018, pp. 1–6.

[14] M. S. Allahham, T. Khattab, and A. Mohamed, "Deep learning for RF-based drone detection and identification: A multi-channel 1-D convolutional neural networks approach," in *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*. IEEE, 2020, pp. 112–117.

[15] M. Golam, R. Akter, J.-M. Lee, and D.-S. Kim, "A long short-term memory-based solar irradiance prediction scheme using meteorological data," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2021.

[16] A. Maurer, M. Pontil, and B. Romera-Paredes, "The benefit of multitask representation learning," *Journal of Machine Learning Research*, vol. 17, no. 81, pp. 1–32, 2016.

[17] Z. Tang, L. Li, D. Wang, and R. Vipperla, "Collaborative joint training with multitask recurrent model for speech and speaker recognition," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 3, pp. 493–504, 2016.

[18] S. Liu, E. Johns, and A. J. Davison, "End-to-end multi-task learning with attention," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 1871–1880.

[19] X. Yang, H. Sun, X. Sun, M. Yan, Z. Guo, and K. Fu, "Position detection and direction prediction for arbitrary-oriented ships via multitask rotation region convolutional neural network," *IEEE Access*, vol. 6, pp. 50 839–50 849, 2018.

[20] M. F. Al-Sa'd, A. Al-Ali, A. Mohamed, T. Khattab, and A. Erbad, "RF-based drone detection and identification using deep learning approaches: An initiative towards a large open source drone database," *Future Generation Computer Systems*, vol. 100, pp. 86–97, 2019.

[21] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.