Deep Learning-based Anti-Drone System using Radar Technology

Ann Janeth Garcia, Jae Min Lee, Dong-Seong Kim,
Networked Systems Laboratory, Department of IT Convergence Engineering,
Kumoh National Institute of Technology, Gumi, South Korea.
{ajg.garcia, ljmpaul, dskim}@kumoh.ac.kr

***Abstract*—This paper proposes a drone detection system with the use of a radar device-based detection scheme. The radar used to collect the data uses a frequency modulated continuous wave (FMCW) on an 8.75 GHz based frequency band with a BW$_{max}$ of 500 MHz. The Deep Neural Network (DNN) is varied with different number of filters and the preference that will give the most accurate result is selected and is compared to different machine learning algorithms such as ResNet-18, SqueezeNet and Support Vector Machine (SVM).**

***Index Terms*—Deep neural network, drone detection, radar.**

I. Introduction

In the last decades, numerous surveillance technologies have been studied for drone detection because of the great threats that has been posed by drones. In [1], a radar-based system is used to secure an area against approaching unwanted drones by tracking and jamming the signal that is used by the controller. The authors' goal in this paper is to develop a drone-detection system that will provide long-term surveillance. Automatic classification techniques with quality labelled data is necessary to improve the efficiency of this system. This system is based on the database, Real Doppler RAD-DAR (Radar with Digital Array Receiver).

Neural Network (NN) is a computing system composed of many simple processing units working simultaneously or in parallel to understand experiential knowledge from a dataset. It has its input layer, output layer and hidden layers that are present in between. The complexity of the models depends on the number of hidden layers in every layer, as well as the number of nodes. NNs are commonly used in pattern recognition applications. Deep NN (DNN) is a supervised neural network that has numerous hidden layers within the input and output layer. This is applied in processing high-dimensional data and in learning progressively complicated models however, it has increased training difficulties and requires more computing resources. In every layer's neurons, it trains a feature representation centered on the former layer's output called feature hierarchy. This makes DNNs efficient in handling significantly large high-dimensional data sets. It provides much improved performance with other machine learning algorithms because of multiple-level feature representation learning.

RAD-DAR Dataset is used, and it is a quality labelled database that has been produced after an extensive controlled trial test campaign. This novel dataset is created for the sole purpose of the paper [2] by Roldan, et. al. It has more than 17,000 samples of cars, people and drones that is obtained in real outdoor scenarios. The radar that is used to capture the data is a constant radar system created by the Microwave and Radar Group called RAD-DAR (Radar-Digital Array Receiver). It operates on a frequency band that is centered at 8.75 GHz with a BW$_{max}$ of 500 MHz FMCW.

After a digital signal processing, a 4092x512 matrix for every single scene is acquired. The distance cells are in rows and the Doppler frequencies are in columns, all in *dBm* unit. These matrices have been reduced resulting in 11x61 matrices.

II. Network Structure

The input for our network is an 11x61 Doppler matrix. The parameters such as stride, it is the step size with which the filter moves. The stride in the pooling layer is set to (2,2) and (1,1) for the convolutional layer. Since there are 3 series of convolutional and pooling layer that is paralleled to each other the filter size is set to (3,3), (5,5) and (7,7). The filter size determines how many neighbor information a neuron can see when processing a current layer. When the filter size is (3,3), each neuron can see a total of 8 neighbor information around it.

Connected to the input layer are the hidden layers that is also called the feature detection layers. These do one of the three types of operations on the data, which is pooling, convolution, and/or activation layer.

- As shown in Fig. 1, convolution layer is chosen because this sets the input signal across a set of different convolutional filters where each filter activates some specific features from the frames.
- MaxPooling layer as the pooling technique is selected and is connected after the convolution layer because this simplifies the output by doing a non-linear down sampling which will decrease the number of parameters or features that the network wants to learn.
- The activation layer that is used is a Rectified Linear Unit or ReLU. This layer allows rapid and efficient training by plotting the negative values to zero and retaining positive values.
- The combination layer used is an addition layer. This layer combines all the output of the activation layer into one single output and then will be the input of the next mid-block our output block.

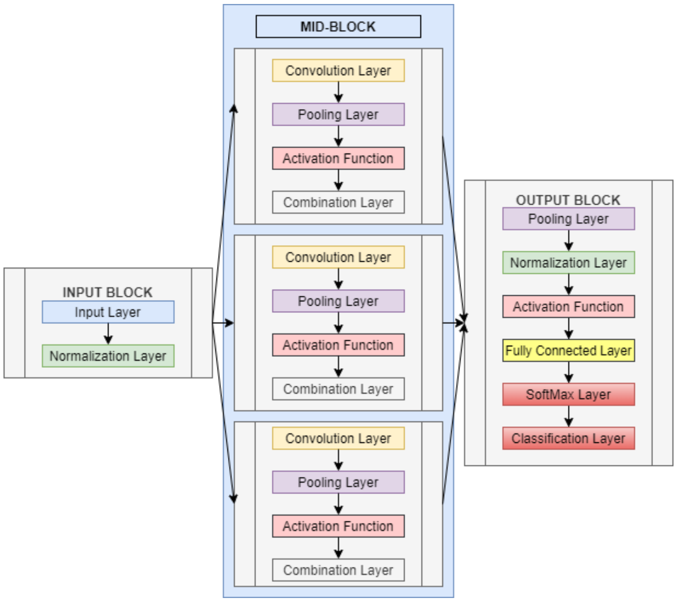

Fig. 1. Proposed network structure design implemented in MATLAB using Deep Network designer with the input layer, output layer and hidden layers in between.

The input block is composed of the input layer and normalization layer, this is where the signal enters. Mid-block is the feature detection layer that is composed of the convolution, pooling, and an activation function which is ReLU. The last layer in the mid-block is the combination layer, this layer combines the outputs of the activation layers into one single output.

Lastly, the output block is the last layer that comes next after the hidden layer. It composes of the classification layers of this network. The third to the last layer is a fully connected layer (FC) that gives an output of a vector of k-dimensions where k is the number of output classes that the network will be able to detect. This vector includes the probabilities for each class of any frame that is being categorized, and the final layer in the network is a SoftMax layer that provides the classification output that is being classified where in this case, it's either a drone, person or car.

III. Results

The number of filters that is set in the input layer is varied from 16, 24, 32 and 48, and then compared and checked which number of filters will give the most accurate result.

TABLE I
VARYING THE NUMBER OF FILTERS OF THE PROPOSED NETWORK

No. of filters	Accuracy	Time consumption	Learnable Parameters Parameters
16	95.85%	0.0032 s	3,715
24	96.00%	0.0032 s	5,571
32	96.40%	0.0033 s	7,427
48	96.54%	0.0034 s	11,139

Table 1 shows that a 32-filter network will give a higher accuracy than other numbers of filters and in Table 2, the results obtained using the RAD-DAR database with the proposed network is compared with ResNet-18 and SqueezeNet and Support Vector Machine (SVM). These are pretrained convolutional neural networks with a million images from ImageNet database and are readily available in MATLAB. The input and output blocks of these networks were varied so that it will match the database used.

TABLE II
COMPARISON OF RESULTS WITH OTHER NETWORKS

Network	Accuracy	Time consumption
Proposed network	96.54%	0.0034 s
ResNet-18	93.47%	0.0059 s
SqueezeNet	92.44%	0.0064 s
SVM	95.34%	0.0065 s

IV. Conclusion

The proposed network that has 48 filters gave the highest accuracy and it also exceeds the accuracy of the ResNet-18, SqueezeNet and SVM. Although the accuracy of these networks as shown in Table 2, shows not that much difference, the proposed network would still be the best choice of network to use because of the time consumption. This is the time consumed in processing one frame by the network. Therefore, a lower processing time indicates how fast the network is.

The proposed network is a promising model that can be applied for the radar-based anti-drone system. Despite outperforming other considered models, the network should be optimized prior to implementation and validated in more scenarios and for future works, varying the pooling technique and combination layer can be done to check if it will further improve the system.

Acknowledgement

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program (IITP-2020-2020-0-01612) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation).

References

[1] Multerer, T., Ganis, A., Prechtel, U., Miralles, E., Meusling, A., Mietzner, J., ... & Ziegler, V. (2017, October). Low-cost jamming system against small drones using a 3D MIMO radar based tracking. In 2017 European Radar Conference (EURAD) (pp. 299-302). IEEE.

[2] Roldan, I., del-Blanco, C. R., de Quevedo, Á. D., Urzaiz, F. I., Menoyo, J. G., López, A. A., ... & García, N. (2020). DopplerNet: a convolutional neural network for recognising targets in real scenarios using a persistent range–Doppler radar. IET Radar, Sonar & Navigation, 14(4), 593-600.