4th International Conference on Advances in Signal Processing and Artificial Intelligence (ASPAI' 2022), 19-21 October 2022, Corfu, Greece

Type of Presentation:
Oral: ☒ In-person: ☐
Poster: ☐ Virtual in Zoom: ☐
The same: ☐

Topic: Applied Artificial Intelligence

Drone radio signal detection with multi-timescale deep neural networks

C. Horn [1], M. Nyfeler [1] , Georg Müller [1] and C. Schuepbach [2]
[1] Zurich University of Applied Sciences, ICLS, Schloss 1, 8820 Wädenswil, Switzerland
[2] Armasuisse, Feuerwerkerstrasse 39, 3602 Thun, Switzerland
E-mail: matthias.nyfeler@zhaw.ch

Summary:

We develop a multi-time scale deep learning algorithm to detect drones from radio signals. While previous approaches focused on the analysis of high-frequency radio data alone we integrate signals from the higher timescale of the drone communication protocol in an end-to-end architecture. To this end, we develop a new meta-CNN layer, which generalizes the idea of the standard CNN (which slides a single, fully connected kernel along a higher level input) towards arbitrarily complex kernel models. As a result, our model is able to extend drone identification abilities significantly toward very small SNRs.

Keywords: drone signal detection, deep learning, multi timescale modeling.

1. Introduction

As consumer drones are becoming more popular and omnipresent, the need to detect them to ensure privacy and security increases. A promising detection technology for consumer drones is the detection of the radio signals transmitted. The detection range is in the order of magnitude of the flight range of the drone, the location accuracy is relatively good and the type of signal allows for conclusions to be drawn about the drone model.

Following previous work [1,2,3], we focus on deep learning-based approaches to separate drone signals from WIFI background and the distinction of different drone types. Our main contributions are the following:

- Introduction of a multi-timescale approach for drone signal detection
- Implementaion of a meta-CNN layer in PyTorch

2. Method

Typical deep learning architectures originally developed for image data continuously compress the input dimension by applying a sequence of convolution and pooling layers.

In the case of radio signals from drones, however, information is present at two distinct timescales: The high-frequency microseconds scale of the radio signal itself and a seconds-level timescale determined by the drone communication protocol. An architecture that trys to learn signals at all intermediate timescales would be very parameter inefficient here. Our approach makes use of three key performance drivers: The statistical benefit of repeated observations, the detection of temporal patterns, and the power of end-to-end learning.

2.1. Lower timescale model

Previous studies found that deep convolutional neural networks with many layers achieve high performance on radio data. Our radio signal model is inspired by the VGG architecture adapted to 1-dimensional input data with the convolution applied along the time dimension. We use 11 blocks, where each block contains 2 convolutional layers followed by a batch-normalization and a pooling layer. The number of filters is increased from 32 for the first block to 256 for the last block.

The model produces a classification output roughly every millisecond, while human evaluation happens at the seconds timescale. This offers the benefit that based on counting statistics alone we can achieve a much higher accuracy on the operator timescale compared to the model output timescale.

2.2. Meta-CNN

The meta-CNN approach offers the additional advantage that we can train our model over multiple timescales end-to-end. The full architecture of our approach is depicted in **Fig. 1**. A standard CNN takes a short fully connected layer and slides it over the input which enables it to identify the (lower level) patterns learned by it independently of their location on the higher-level input scale. Our meta-CNN layer generalizes this idea in that it allows to slide any arbitrarily complex kernel model along a higher level input. In our case, we use our radio signal network as the kernel.

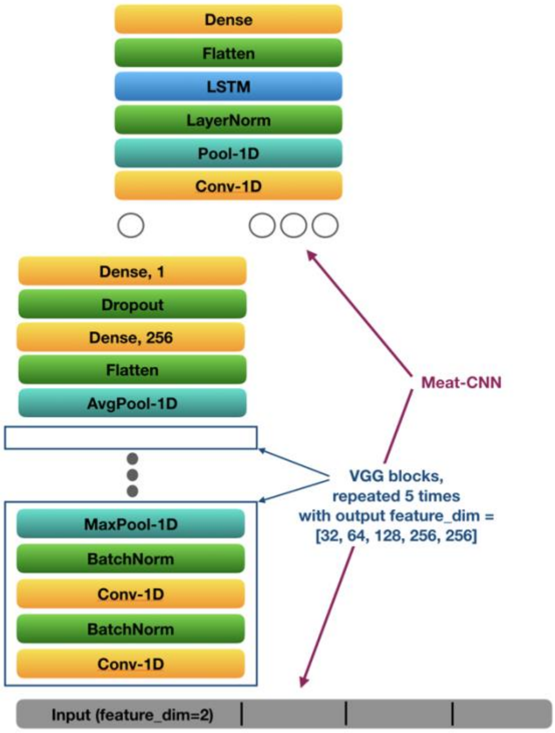

Fig. 1. Complete neural network architecture. The Meta-CNN shifts the input producing a single neuron output (denoted by ◯) for each step. The combined signal of all these neurons is then processed by the higher timescale (upper part) of the network.

Importantly, the meta-CNN allows for choosing a stride of less than the kernel size which increases the resolution compared to a simple meta learning approach.

2.3. Higher timescale model

For the detection of temporal patterns in the drone communication protocol, we use a 1-d convolution and pooling layer, followed by a layer normalization and an LSTM layer, and a fully connected classification layer.

2.4. Data preparation and model training

We recorded drone radio signal data (DJI Phantom pro 4) while a person was actively manipulating the remote control, as well as WiFi noise background. The signals are all at a carrier frequency of 2.44 GHz. They were with a USRP in an echo-free at a sampling rate of 56 MHz and then downsampled down to 14 MHz using an anti-aliasing filter, to investigate if detection is still possible at lower frequencies.

The signal have gaps with low power between bursts as shown in **Fig. 2** and perform frequency hopping, i.e. the bursts hop between different frequencies in the frequency band as can be seen in **Fig. 3**.

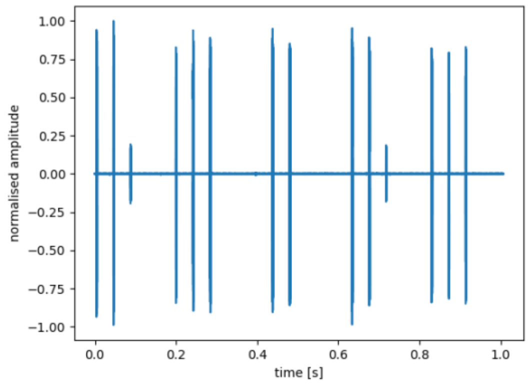

Fig. 2 The real part of a 1 second of a normalised I-Q-radio signals from a DJI phantom pro 4 system in the 2.44 GHz frequency band. The short signal bursts are clearly visible on which the lower timescale model was trained. The meta model then classifies the sequences of outputs from the lower timescale model.

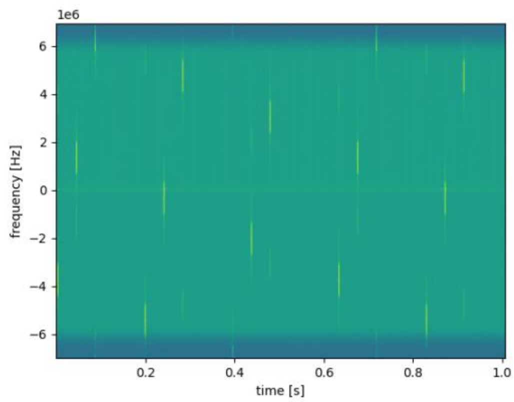

Fig. 3. Spectrogram of 1 seconds of drone signals from the DJI phantom pro 4 in the 2.44 GHz frequency band at a sampling rate of 14 MHz showing the short signal bursts and the frequency hopping

We remove input vectors with an average power of less than 0.001 of the maximum, to facilitate training. All input vectors were normalized individually. We construct noisy signal data by mixing the recorded signal with Gaussian noise and WiFi noise with SNR between -40, and 0 dB. Additionally, data augmentation is used by constructing multiple datasets from the same background data blended with individual Gaussian noise. We find that this approach increases the robustness of our results. For training our models we use a combination of datasets with SNRs of -30, -20, -10 and 0 dB. The training of the full model (lower-timescale model + meta-CNN + higher timescale model) happened end-to-end over all parameters.

3. Related Work

We compare our performance with previous results on the DeepSig dataset [4] and obtain very similar results as our “Radio Signal Model only” model. A comparison on the higher timescale is not possible, since the vectors in the dataset are not sequential signals.

4. Results

As a baseline for our higher timescale model we consider the case where we apply the meta-CNN but simply use a fixed (optimized) threshold on the total count of positive outputs from the lower-level model. Here we focus on the discrimination ability between drone signals and WiFi backgrounds, leaving the differentiation between drone types for later.

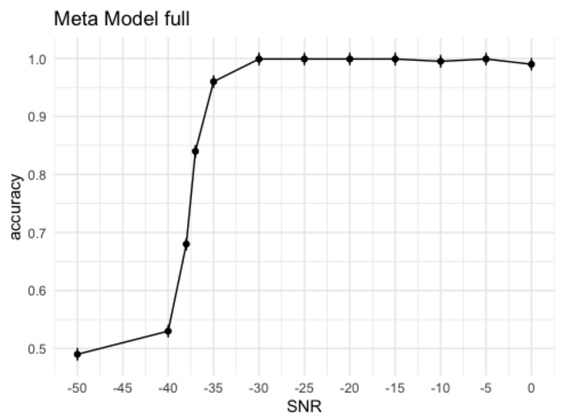

Fig. 4. Performance of the full model versus SNR.

Table 1. Accuracy comparison for different models and SNRs.

SNR [dB]	Lower timescale model only	Meta-CNN with fixed threshold	Full model
0	99%	99%	99%
-10	99%	99%	99%
-20	64%	99%	99%
-30	50%	98%	99%
-35	50%	62%	96%
-40	50%	50%	53%

The performance of the three model variants is shown in **Table 1**. While the radio signal model is useless at -30 dB, the meta-CNN baseline operates still at full performance. The dependence of the performance of our full model on SNR is shown in **Fig. 4.** We observe that the model is still effective even below -35 dB.

5. Conclusions

We developed a multi-timescale model for the detection of radio signals from drones. The novel neural network architecture we propose combines a 25-layer CNN-based kernel model for the identification of lower timescale signals with a LSTM-based model on the top to extract higher-level communication pattern. As our experiments show this enables a significant extension of drone identification ability down to SNR values below -30 dB. We note that the idea of a meta-CNN with a complex kernel model might be useful in general for other tasks where relevant information is present at different timescales.

6. Outlook

Further drone signals from different brands (DJI, Futaba, Taranis, Graupner, Turnigy) have been recorded at carrier frequencies of 2.4 GHz, 5.8 GHz and telemetry link at 0.869 GHz and the performance of the model on the new data is under current investigation. Furthermore a protype for live-testing in the field is under construction and will be the topic of a further publication.

Acknowledgements

We thank Selina Malacarne and Nicola Ramagnano from the University of Applied Sciences Rapperswil for recording the data and setting up the implementation.

References

[1]. T. J. O’Shea, T. Roy, and T. C. Clancy, Over-the-air deep learning based radio signal classifi- cation, IEEE Journal of Selected Topics in Signal Processing, vol. 12, pp. 168–179, 2018.

[2]. M. Ezuma, F. Erden, C. K. Anjinappa, O. Ozdemir, and I. Guvenc, “Micro-uav detection and classification from rf fingerprints using machine learning techniques,” in *2019 ieee aerospace conference*, 2019, pp. 1–13.

[3]. S. Scholl, Classification of radio signals and hf transmission modes with deep learning, 2019, https://arxiv.org/pdf/1906.04459.pdf

[4]. RF datasets for machine learning, https://www.deepsig.ai/datasets/ , 2018