Horizon 2020
European Union Funding
for Research & Innovation

RF Fingerprinting Database

RadioSpin Project

H2020-FETPROACT09-2020

Author: Alice Mizrahi, THALES

Deliverable D5.2
Work Package 5

Version: Final, 31/03/2022

Project number: 101017098

Horizon 2020
European Union Funding
for Research & Innovation

Deliverable No. 5.2

Dissemination Level		
PU	Public	x
PP	Restricted to other program participants (incl. the Commission Services)	
RE	Restricted to a group specified by consortium (incl. the Commission Services)	
CO	Confidential, only for members of the consortium (incl. the Commission Services)	

Nature		
R	Report	x
DEM	Demonstrator	
DATA	Data sets, microdata, etc.	
ORDP	Open Research Data Pilot	
ETH	Ethics	

Deliverable Details			
Due date:	31/03/2022		
Submission date:			
Involved participants:	Thales, CNRS		
WP no. and title:	WP5 Evaluation and benchmarks		
WP leader:	UBT SRL		
Task no. and title:	Task T5.3 – Demonstrator evaluation on RF fingerprinting application		
Version/Date	31/03/2022		
Keywords:	Dataset, RF fingerprinting, drones, neural network		

Summary

We have developed a database of 400 RF signals in order to evaluate the demonstrator on an RF Fingerprinting task. The chosen task is to identify commercial drones and their radio-controllers out of four classes. The signals were adapted from real data recorded from the devices in an anechoic chamber. The resulting signals are suitable for the RadioSpin demonstrator and will be used for neural networks learning and inference.

Contents

1. Goal of the deliverable

This deliverable is part of Task 5.3 – Demonstrator evaluation on RF fingerprinting application.

In the Grant Agreement, this task is described as follows: *This task aims at evaluating the RadioSpin demonstrator capabilities in terms of training and performance before scaling-up to the actual use case developed in WP6. Thales will first constitute a sample dataset, based on the one collected in T6.2, and suitable for the demonstrator (size of the network, frequencies of the oscillators). Using a learning algorithm chosen by simulations in WP6, from the three algorithms developed in WP4, Thales will train the demonstrator on a smaller RF signal classification task. Then we will evaluate the performance, speed and energy consumption or the demonstrator, which lays the ground for estimating the scaling capabilities for this application. CNRS will provide assistance on the usage of the demonstrator.*

The goal of deliverable D5.2 RF Fingerprinting Database is to present the database that we have constituted. This database has been custom made for the evaluation of the demonstrator. The requirements are the following:

- The database should compose a meaningful task, linked to a real-life application and deriving from real data.
- The database should be compatible with the demonstrator in terms of functional capability. Because the demonstrator will have a restricted size, we need the task to be small enough to be solvable by the neural network implemented on the demonstrator.
- The database should be compatible with the demonstrators in terms of system parameters. This means that the amplitudes and frequencies of the RF signals must match the characteristics of the devices in the demonstrator.

2. Description of the database

2.1. Chosen task: identification of commercial drones

The Radio Spectrum is full of different radiofrequency (RF) signals. It is critical for the good functioning of the internet of things that smart connected objects are able to identify the origin of these RF signals, i.e. to perform RF fingerprinting.
We choose to focus on RF signals emitted by commercial drones and their radio-controllers.

We use the data collected by Basak et al. in "Drone classification from RF fingerprints using deep residual nets" (IEEE COMSNETS conference, 2021[1]). This database has been identified and investigated in WP6. Basak et al. collected signals in an anechoic chamber, using a universal software radio peripheral (USRP X310) placed seven meters apart from the devices (as shown in Figure 1). The signals were all in the 2.4 GHz ISM band and the whole 100 MHz band was received instantaneously using a receiving sampling rate of 100 MSps (i.e. the system down-converted the signal frequencies to the 0-100 MHz band to sample them correctly).

[1] https://ieeexplore.ieee.org/document/9352891

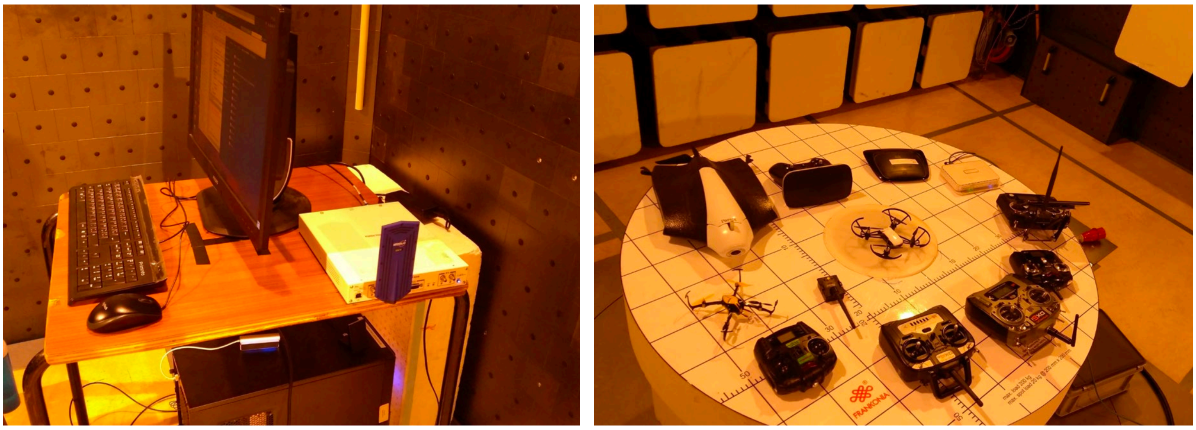

Figure 1. Left: the USRP recording system. Right: the drones and radio-controllers. Both photos are taken in the anechoic chamber and reproduced from Basak et al., 2021.

Basak et al. have recorded signals from ten devices. We would like to be capable of solving the chosen task with one tile of the demonstrator, i.e. a 8 neurons by 8 outputs network. In order to fit the size of the demonstrator, we restrict the database to four devices (two drones and two radio-controllers), presented in Figure 2. The task is to identify from a received RF signal which of the four devices emitted it.

Spektrum DX4e

Parrot Disco

Q205

Tello

Figure 2: The four chosen devices for the RF fingerprinting database.

2.2. Adaptation of the signals to the demonstrator

Each recording of a signal is a spectrogram of size 256 x 256 (i.e. 256 frequency bins and 256 time frames). These spectrograms are not suitable for our demonstrator so we make the following transformations, schematized in Figure 3:

- We are not interested in the temporal evolution of the signals because the demonstrator will not have a state memory or feedback loop. In a first phase, the implemented neural network will be feed forward. In consequence, we transform the spectrograms into simple spectra (power density versus frequency) by averaging over the time dimension.
- The input frequency dimension (256) is too large for our 8 input neurons tile. In consequence we shrink the signals to 8 input frequencies by averaging the 8 sequences of 32 consecutive frequencies into 1 central frequency each.

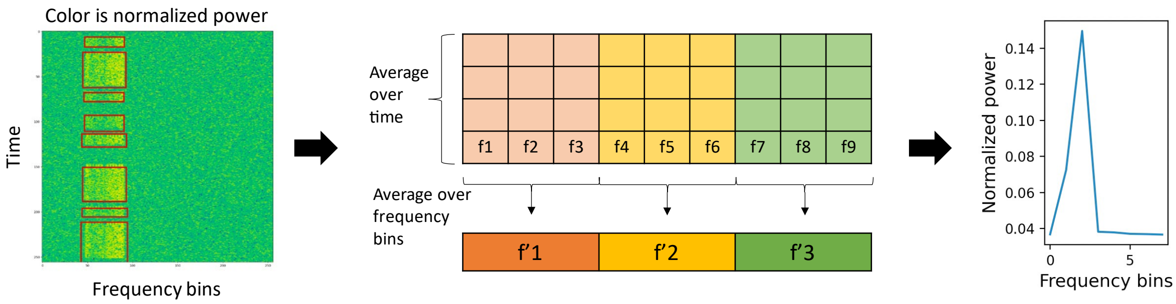

Figure 3: Schematic of the process to transform the spectrograms (left) into spectra suitable for the demonstrator (right). Here we show an example of signal corresponding to the "Tello" drone.

Figure 4 presents examples of resulting signals for each of the four classes.

In the experiment of Basak et al., the signals frequencies where down converted to the 0 – 100 MHz range so that they could be sampled and digitized. However, when applied to RadioSpin technology, the signals would not require digitization: the spintronics devices are able to process analogue RF signals because of their intrinsic RF dynamics. In consequence, the lowest frequency bin can be set to an arbitrary value. In a future application, this value would correspond to the chosen down-conversion frequency.

Regarding the frequency range, we will need to adapt it so that the inputs cover the whole range of the devices in the demonstrator. In consequence, we set the input frequencies to unlabeled bins that can be labeled according to the demonstrator's physical properties. The down side is that this is a departure from real data. However, this adaptation is required to evaluate correctly the demonstrator. The simulations in WP6 will be able to process real data.

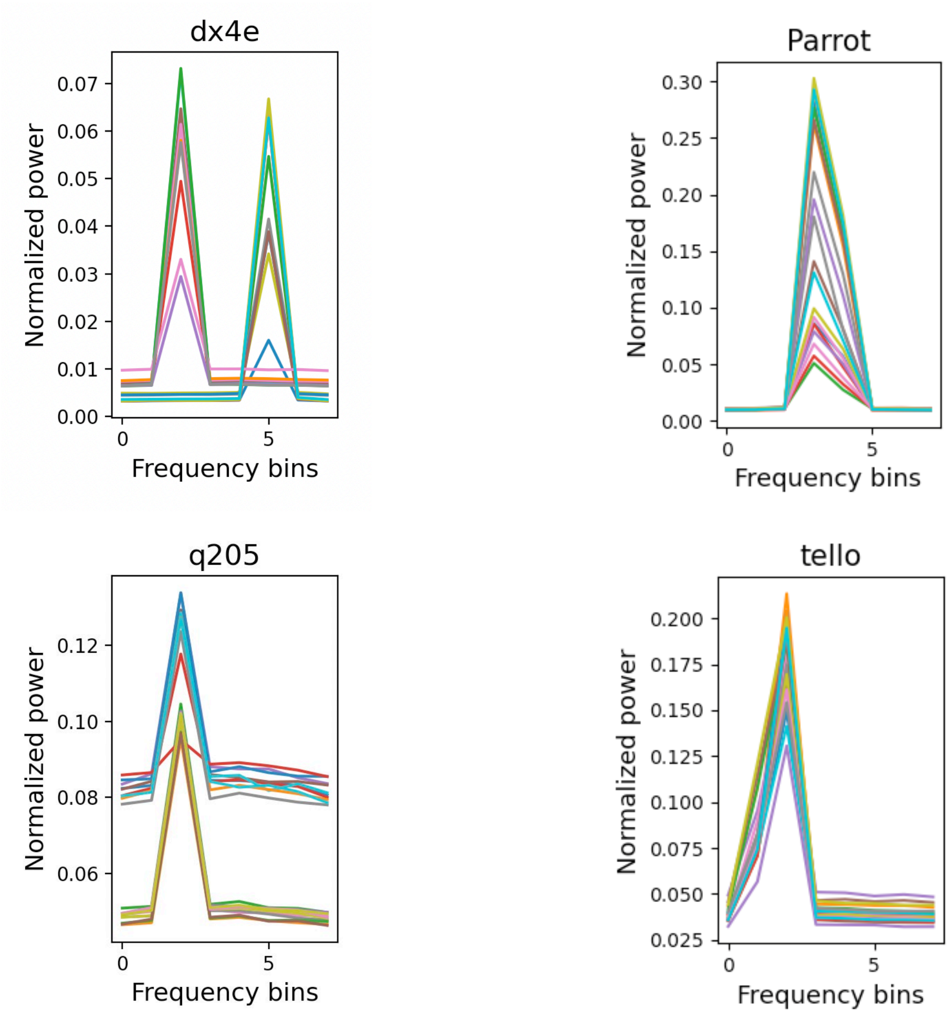

Figure 4: Examples of 20 signals for each of the four classes. Each color is one example.

3. How to use the database

The database is composed of 400 signals, 100 signals for each class. Contrary to the initial dataset proposed by Basak et al., we chose to not separate "test" from "train" signals upfront. Indeed, the choice of the algorithms and network models to evaluate the demonstrator will influence which proportion of "test" versus "train" signals is optimal. In consequence, we leave the database flexible.

Each signal is a vector of 8 elements. To each signal corresponds a target, which is the number of the class it belongs to.

Device	Target
DX4e	0
Parrot	1
Q205	2
Tello	3

The database is stored in an h5 file, a format adapted to databases. Inside the file there are two datasets: the signals (‘Signals’) and the targets (‘Targets’). The data can then easily be extracted and fed to a machine learning data-loader (from the Pytorch library for instance).

The database file is ***RadioSpin_D52_RF_fingerprinting.h5***

4. REFERENCES

[1] S. Basak, S. Rajendran, S. Pollin and B. Scheers, "Drone classification from RF fingerprints using deep residual nets," *2021 International Conference on COMmunication Systems & NETworkS (COMSNETS)*, 2021, pp. 548-555,
doi: 10.1109/COMSNETS51098.2021.9352891.

5. ACKNOWLEDGEMENT

This deliverable is part of the RadioSpin project. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101017098.