Machine Learning Based Prediction of Frequency Hopping Spread Spectrum Signals

Pascal Thiele[1], Laura Bernadó[1], David Löschenbrand[1], Benjamin Rainer[1],
Christoph Sulzbachner[1], Maria Leitner[2], Thomas Zemen[1]

[1]Center for Digital Safety & Security, AIT Austrian Institute of Technology GmbH, Austria
[2]Faculty of Computer Science, University of Vienna, Austria
Email: pascal.thiele@ait.ac.at

***Abstract*—In an world shifting towards wireless communications, the already scarce electromagnetic spectrum within the unlicensed bands is becoming increasingly crowded. All wireless devices operating in those bands need to co-exist without interfering with each other. Frequency hopping spread spectrum (FHSS) is a communication technique especially resilient to interference due to its constant change of the carrier frequency and its narrowband transmission bandwidth. Furthermore, it produces minimal interference to other signals in the same frequency band using wider bandwidth. However, interference can also be harmful even for FHSS transmissions as a result of the loaded ISM bands. Intelligent spectrum sensing techniques can contribute to a more efficient spectral usage. In this paper, we propose a supervised learning algorithm which predicts the future time-frequency location of a FHSS signal. We design a convolutional neural network which is trained on a dataset, obtained from measurements of two FHSS sources. Based only on a small observation window of 50 ms, it predicts the signal appearance of the following 25 ms in a time-frequency representation. To show that we can accurately predict the signal, we introduce a special score measure. The mean score of about 0.9 with small standard deviation demonstrates the high fidelity prediction of the signal's evolution.**

***Index Terms*—FHSS, frequency hopping, supervised learning, CNN, pattern prediction**

I. Introduction

Technological advances enable the use of machines to carry out human activities, reducing risks and costs for people, society and environment. Unmanned vehicles are among one of the most proliferating application fields, with a notable momentum in the aerial set-up with usage in search and rescue operations, goods delivery or environmental monitoring [1], just to name a few. These unmanned robots, though human controlled, often communicate using a command and control (C2) wireless link in the unlicensed 2.4 GHz ISM frequency band. The control commands for unmanned aerial vehicles (UAVs) require little payload and, given the crowded electromagnetic spectrum at this frequency band, it is reasonable that UAVs adopt a modulation for the C2 link that can be easily carried in a narrowband signal without suffering the effect of interference.

Frequency hopping spread spectrum (FHSS) is a wireless communication scheme characterized by a rapidly changing (hopping) carrier frequency using narrowband signals within a wider frequency band. It is well established in the 2.4 GHz ISM band and adopted by the IEEE 802.15 Standard (Bluetooth) and the C2 links of UAVs, among others. In FHSS, the signal operates on a predefined set of carrier frequencies $\{f_i\}$, the respective frequencies are called channels. The order in which these frequencies are addressed yields the channel sequence $(f_i)_{i=1}^{K}$, where K is some integer. In the case of UAVs, the channel sequence length K is reasonably small, implying a periodic radio signal, since the carrier reiterates again through $(f_i)_{i=1}^{K}$ upon completion. This, and the fact that some UAVs use a certain special type of pattern [2], should make it possible to predict the UAVs signal evolution.

The biggest advantages of FHSS transmission are the resilience to eavesdropping, to jamming, and to the degeneration of the signal due to unwanted interference. Furthermore, the risk of a communication failure in the C2 link can cause the loss of control of a UAV and its fall on an unpredicted (and maybe populated) area. Therefore, it is crucial to guarantee an electromagnetic collision free C2 link. In order to improve the resilience of FHSS signals, it is necessary to provide the UAVs themselves with a method that allows them to sense the surrounding spectrum, identify other FHSS signals and, if possible, negotiate a new frequency pattern for the C2 link [3], [4].

Extensive literature can be found on the detection of radio controlled UAVs, as well as on UAV type identification, based on their respective RF signatures [5]–[7]. Only few literature investigates the possibilities to identify such FHSS patterns, or generally of patterns of FHSS based communications, to allow for precise prediction of the signal evolution. The found approaches require usually a long scanning of the signal before any prediction is possible [8], [9]. The work described in [10] made use of approximately 5 hopping times to identify the FHSS pattern, which is already a great improvement compared to the monitoring of channel activities. We propose a similar, but novel approach in which we require only a small and fixed observation time to predict the near future evolution of the FHSS signal. To do so, we use supervised machine learning (ML) with a convolutional neural network (CNN) architecture to obtain the signal prediction in a time-frequency representation. Since we use a time-frequency representation,

we allow for the possibility to actively avoid the signal, or in contrast, address specifically this signal. To achieve our goal we proceed as follows: First, we measure the radio signals of two different FHSS sources, see Sec. II. Second, we present the dataset construction and the performed preprocessing, the data labeling and the annotation of the signal regions, see Sec. III. Third, we outline the training process of the model in Sec. IV. Fourth and finally, we evaluate in Sec. V, the model performance on the test set. The *scientific contribution* of this work can be summarized as follows:

- We use a time-frequency representation of an FHSS signal to train a CNN designed to predict the near future evolution of an FHSS signal, only based on a small observation window.
- We present a metric assessing the signal prediction quality. This metric focuses on the informative regions, i.e regions occupied by the target signal.
- We present the ability of the CNN to recognize two different FHSS patterns, and predict their future evolution.

II. Measurements

The two measured FHSS sources were in both cases two commercially available UAVs, having FHSS as C2C link. These signals were measured in multiple measurement runs. The specifc UAVs were:

- A quadcopter DJI Phantom 2330D
- A fixed wing Skywalker EVE-2000 UAV with FrSky control unit.

We recorded about 5 s of data per measurement run in the lab with our AIT software-defined-radio testbed [11]. The measurement bandwidth was set to 80 MHz at a center frequency of 2.44 GHz. The repetition time of the measurement snapshots was 125 µs, and 512 I/Q samples were obtained in each snapshot. In the lab there is occasional WiFi coverage and therefore, it is expected, and desired, to be recorded during measurements.

Each UAV was measured multiple times in different operating modes as in [12]. The data for the Phantom UAV was acquired in five measurement runs, where as the for the Skywalker UAV, we performed three measurement runs. The discrepancy in the number of measured runs comes from the fact that Phantom and Skywalker have a different number of operating modes. We were unable to identify any difference in the measurements per signal source, i.e. no (if any existing) effect of the operating modes. We therefore don't make any distinction between these individual measurement runs. Additionally, two measurement runs with RF background signals of the lab were recorded. This data will also be added to the training- and test set for two reasons: first, it allows the supervised learner to not wrongly predict signals if there are none present, and second, it can give a baseline for the model performance [2].

III. Dataset Construction

Since we aim for a time-frequency signal prediction, we process our measurements to be represented with a spectrogram. Therefore, we used the fast Fourier transform (FFT) with 512 samples, on which a Hanning windowed was applied, to generate a spectrogramic representation of the data. To train (and test) our supervised model, we fix the dimensionality of the feature space, i.e. the dimension of the input to the ML model, as well as those of the target space, i.e. the resulting output dimension of the ML model. This ultimately boils down to the FFT size and the desired observation and prediction time windows.

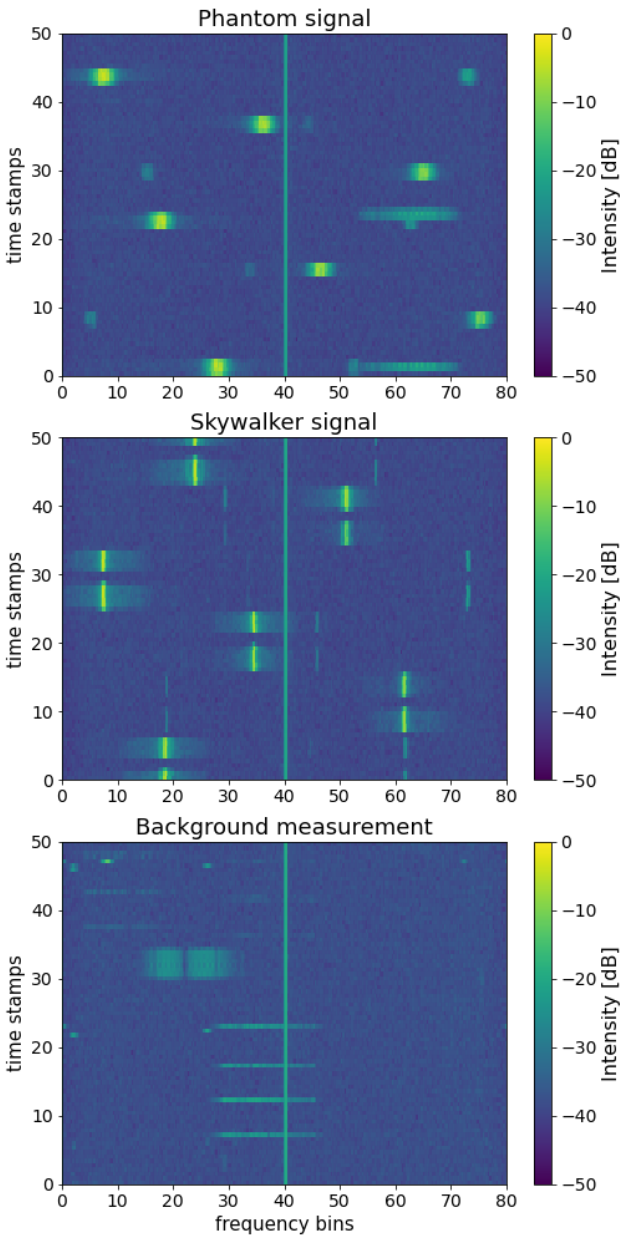

Fig. 1. Extract of the recorded UAVs. From up to down: Phantom signal, Skywalker signal, background measurement. Note the sporadic interferences and note also the difference in the FHSS signal appearance of, e.g. double hop vs single hop type pattern.

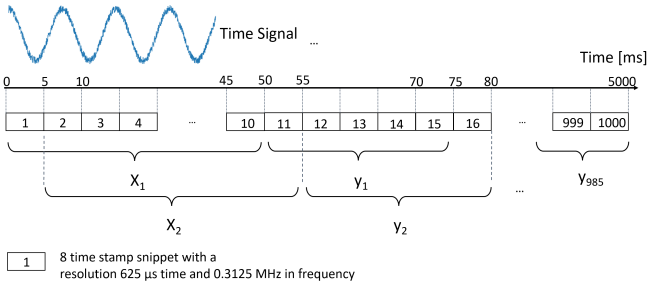

Fig. 2. Generation of X-y samples from measurement run.

A. Spectrogram construction

We choose an observation window of 50 ms to predict the following 25 ms of the FHSS signals. The chosen time lengths of both input sample X and target samples y are based on the observed time scales of different FHSS signals, cf. [2] such that one observes around three signal hops in the output sample y. An ad hoc choice was made for the observation window length, we chose the double of the prediction time, this should guarantee enough information to predict the signals' evolution.

Therefore, 75 ms worth of measurement data form a X-y sample. This process was repeated every 5 ms, on each measurement run, resulting in 985 ML samples per measurement run. After data compression, theses samples have 80 and 40 time stamps x 256 frequency bins for the X and y samples respectively, with resolutions of 625 s and 0.3125 MHz. The X-y sample generation procedure is captured in Fig. 2

The samples were source wise normalized to the maximum recorded energy. Consequently, each time-frequency pixel takes a value in $[0, 1]$.

B. Annotation

Since we are only interested in the signal's time-frequency localization, and not in its exact appearance, we propose here an annotation procedure to avoid redundant information in the prediction outcome. This has the additional advantage to speed up the training process [2].

The annotated y samples are therefore the actual used labels, to which we refer from now on as y. The annotation procedure involves a signal detection part followed by a hop rectification.

The signal detection consists of a two step thresholding scheme. The signal detection scheme requires two user defined input thresholds δ_1, δ_2 and computes y time stamp-wise. The scheme can be describe as

$$y_{i,j} = \begin{cases} 1 & \text{if } s_{i,j} > \delta_1 \text{median}(s_i) \wedge s_{i,j} > \delta_2 \max(s_i) \\ 0 & \text{else} \end{cases}, \tag{1}$$

where $s_{i,j} \in \mathbb{R}$ are the entries of the to be annotated spectrogram, with i indexing the time stamps and j the frequency bins. With the right parameter choices, this scheme shows to be robust in detecting signals while not detecting any noise. Table I summarizes the chosen parameters δ_1, δ_2.

The choice of δ_1 depends on heuristics and on the underlying data, it was designed to filter the noise background. The second threshold, δ_2 was designed to identify the FHSS signal. The choice $\delta_2 = 0.5$ defines the common full width at half maximum definition. For the Skywalker UAV, since the signals are notably more narrowbanded, we used however $\delta_2 = 0.25$, such that we obtain a wider target area. The resulting binary matrix encodes signal regions with *1* and noise regions, as well as all other (interference) signals of non-interest, with *0*. In order to obtain a well defined signal region in both time and frequency dimension we rectify the signal areas. The rectification function essentially averages the frequency spread of each signal region, a detailed description of the rectification function can be found in [2]. The found approach was found to be well suited to quickly generate a dataset and was in the scope of this work, not further improved. A Phantom and Skywalker example of detected signal pixels and the rectified version of the spectrogram is shown in Fig. 3.

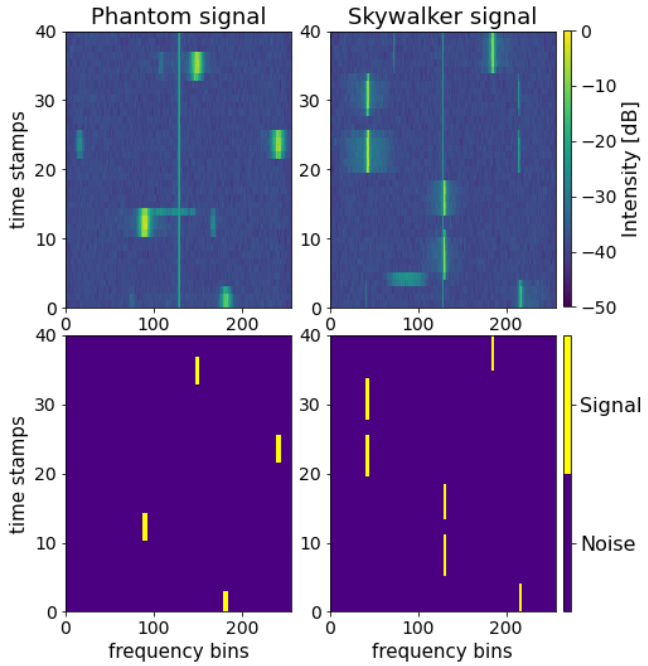

Fig. 3. Annotation of the signal regions. First row shows the original recording, second row the obtained annotated outcome y. Left column an example of an Phantom signal, right column the Skywalker signal.

TABLE I
ANNOTATION PARAMETER

	Phantom signal.	Skywalker signal
δ_1	90	125
δ_2	0.5	0.25

Since the target y takes only two instances at each entry, we aim to solve a multivariate binary-classification problem to perform the signal prediction.

The estimated signal-to-noise ratio (SNR) of the data was found to be 33.8 dB and 34.5 dB for the Phantom and Skywalker respectively. The SNR of a X sample was estimated

with

$$\mathrm{SNR}_{\mathrm{dB}}(X) = 10\log_{10}\left(\frac{\max(X)}{\mathrm{median}(X)}\right), \tag{2}$$

which can be justified by the fact that $< 1\%$ of signal pixels make up a spectrogram. The given SNR estimate for a UAV is the mean SNR over all respective samples.

Since the background measurements possess no signal of interest, they were not annotated, but otherwise processed in the same way as the drone signal measurements.

IV. Training

Convolutional neural networks (CNNs) are a common choice for model architectures in image recognition problems, which is also a classification task. To perform our classification, we use a 12 layer CNN architecture which is sketched in Fig. 4. A more detailed explanation of the architecture is given in [2].

The rough architecture of this model was inspired by the model presented [6], which was constructed for drone identification and classification based on the RF signature. Out of the 12 layers of our CNN, eight are trainable, i.e. model weights are optimized during the training process. The model was implemented with Python's PyTorch library. Note that, by construction, the model yields a 1D output $\hat{y}'$, with $\hat{y}' \in \mathbb{R}^k$, $k = 10240$. But it can always be seen as a flattened version of the 2D target sample $y \in \mathbb{R}^{40\times256}$ (number of time stamps $\times$ number frequency bins), and as such be reshaped. Thus, we regard the models output representation, 1D or 2D, as equivalent.

The small kernel size of two in the convolutional layer was found to be good in the edge detection, and ultimately in the future signal prediction. Besides the dropout layer (10), we used early-stopping as regularization method with a patience of 30 epochs, [2].

The training and testing sets were obtained by a random 70-30 split of the respective datasets. Additionally, we defined a validation set from the training data by segregating randomly 10% from it. Per epoch, the loss on this validation set was computed to determine the early-stopping point.

Binary cross entropy (BCE) with integrated sigmoid was used as a loss function for the backpropagation process. Binary cross entropy is a standard choice for a classification problems. The integration of the sigmoid function, $\frac{1}{1+e^{-z}}$, into BCE is a feature of the Pytorch library to ensure numerical stability. With this in mind, the output of the CNN depicted by Fig. 4 yields only the class probability per pixel, if an activation function, with property to map continuously $\mathbb{R} \to [0,1]$, is applied on the individual entries. The model parameter optimization was performed with mini-batch gradient descent with a batch size of 30.

V. Evaluation

After 755 epochs the training was stopped according to the early-stopping criterion. The criterion compared the average loss of the past 30 iterations, previous to the average loss over the patience. In order to quantify the model performance

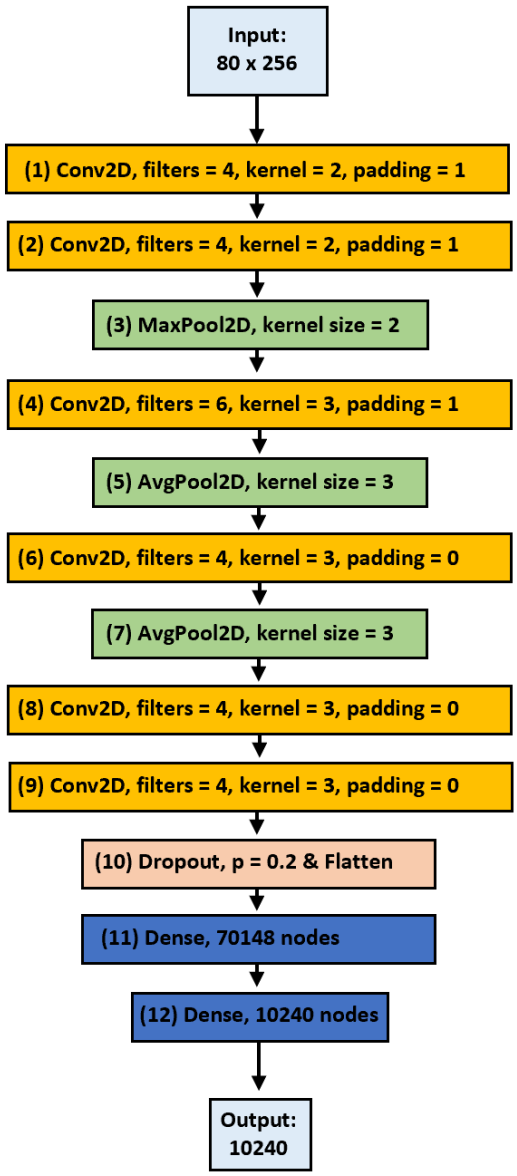

Fig. 4. CNN architecture of the constructed model

we propose a specially designed score measure, the *LP-score*, named after the authors initials. This score has as image map $\mathbb{R}^k \times \mathbb{R}^k \to [0,1]$, with $k = 10240$. It takes as inputs $y, \hat{y} \in \mathbb{R}^k$, where y is the ground truth and $\hat{y}$ the predicted future signal appearance. The entries of $\hat{y}$ are computed element wise, denoted by the index i, from the model output $\hat{y}'$ with

$$\hat{y}_i = \begin{cases} 1 & \text{if } \frac{1}{1+e^{-\hat{y}'_i}} \geq p \\ 0 & \text{else} \end{cases}. \tag{3}$$

We chose $p = 0.5$ as probability threshold. This threshold parameter was found to be nearly ideal, and was not further

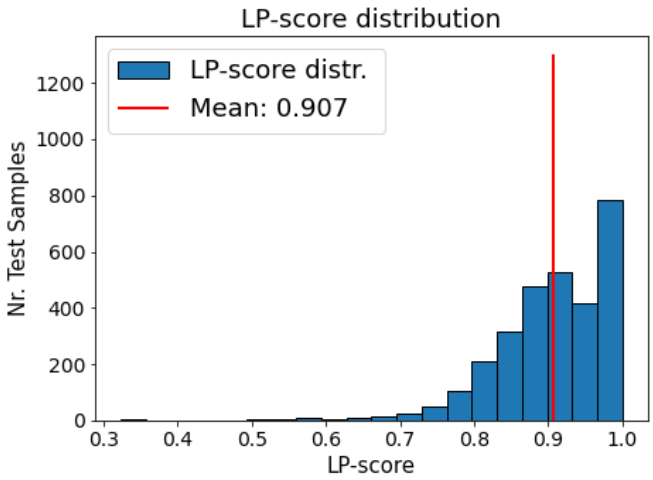

Fig. 5. LP-score distribution on X_{test}, with highlighted mean score, the σ is 0.081.

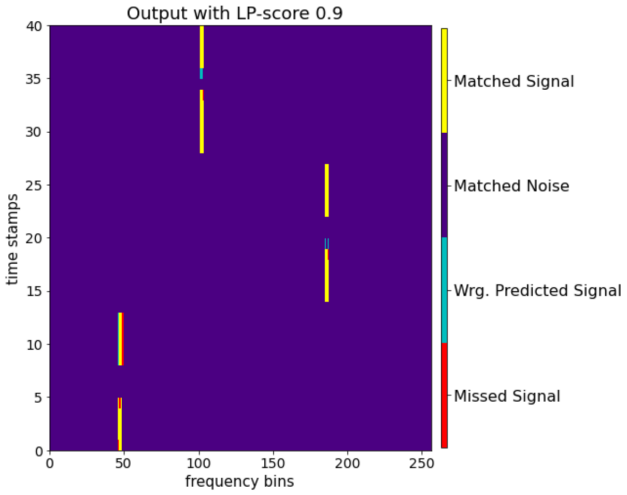

Fig. 6. Model output with score 0.9 compared to ground truth. Skywalker FHSS signal sample from test set.

optimized. The LP-score is defined as

$$\text{LP-score}(y,\hat{y}) = \begin{cases} 1 & \text{if } \forall i \;\; y_i = \hat{y}_i = 0 \\ \frac{\langle y,\hat{y}\rangle}{\|y\|\|\hat{y}\|} & \text{if } \exists i,j \text{ s.t. } y_i \neq 0, \hat{y}_j \neq 0 \\ 0 & \text{otherwise} \end{cases}, \tag{4}$$

where $\langle *,* \rangle$ denotes the standard scalar product and $\|*\|$ is the Euclidean vector norm. This score is a modification of the correlation matrix distance between two matrices, or their vectorized representation. The correlation matrix distance was used in [13], [14] as a measure to quantify the overlap between two peak regions of two matrices in an otherwise mostly flat (zero) landscape. The LP-score compares only signal regions, ignoring true negative instances (regions with no signal of interest), in the case a target signal is present. In the case if no signal is present, and the model predicts only its absence, a score of 1 gets returned. With this score, any model which randomly guesses a certain number of (possibly grouped) signal pixels, would yield almost certainly a score of 0. A slightly better naive approach would be a model which predicts the absence of signal in any case. This model would then yield an LP-score equivalent to the number of samples without any signals, over the amount of total samples, i.e. approximately 0.2 in our case.

Figure 5 shows the LP-score distribution obtained from the test set. With a mean LP-score of 0.906 and standard deviation σ of 0.081, one can consider the model to be well trained. To give a qualitative understanding what this score corresponds to, see Fig. 6 which displays the output of a sample yielding a score of 0.9 and compares it to the ground truth y. We found that all background samples were correctly identified to contain no signals of interest, this contributes to the relatively high number of samples yielding a score between 0.95 and 1. To differentiate the result better, we asses individually for both FHSS sources the mean score. It was found to be 0.872 $\pm$ 0.079 for the Phantom signal and 0.905 $\pm$ 0.057 for the Skywalker signal, where we use σ as error estimate. This was a first step to check the feasibility of the proposed approach for different FHSS signals, more data with various FHSS patterns would be needed to train a more complex model, which is ultimately able to predict all the different patterns. Care needs to be taken to construct a balanced dataset, especially if more complex and longer patterns are under consideration.

With the trained model, the prediction time of the future FHSS signal evolution is essentially limited by the acquisition time of 50 ms. Even though our approach provides only a short time signal estimation, it is by two orders of magnitude faster than the signal prediction described in [9] which extracts the entire hopping sequence. The observation window might be larger in the described approach, than in [15], depending on the exact hop duration, but posses the advantage of an easier implementation due to the fixed observation length.

VI. Conclusion

In this paper, we described the data acquisition of two commercially available UAVs, DJI Phantom and Skywalker, both utilizing FHSS for the command and control link. We described in detail how data can be processed and labeled to become suitable for a supervised learning process. A model with CNN architecture was trained to predict the following 25 ms of the signal in a time-frequency representation based only on a small observation window of 50 ms. To assess the model performance, we proposed a special score measure, based on the correlation distance measure, the LP-score. With this score, we show a robust prediction of the signal regions, with mean value of 0.907 and small standard deviation of 0.081. The results are recovered for both distinct FHSS signals independently.

Acknowledgment

This work received funding by the SCALA and MMCUAS project supported by the Austrian Austrian Research Promo-

tion Agency (FFG) under the programme KIRAS and of the Federal Ministry of Finance (BMF).

REFERENCES

[1] B. Canis, "Unmanned aircraft systems (uas): Commercial outlook for a new industry," 2015.

[2] P. L. Thiele, "Supervised machine learning for pattern recognition of frequency hopping spread spectrum communication," Universität Wien at AIT GmbH, master thesis, https://ubdata.univie.ac.at/AC16743185.

[3] E. Fernández de Gorostiza, J. Berzosa, J. Mabe, and R. Cortiñas, "A method for dynamically selecting the best frequency hopping technique in industrial wireless sensor network applications," *Sensors*, vol. 18, no. 2, 2018. [Online]. Available: https://www.mdpi.com/1424-8220/18/2/657

[4] A. Gök, S. Joshi, J. Villasenor, and D. Cabric, "Estimating the number of frequency hopping interferers using spectral sensing with time and frequency offset measurements," in *MILCOM 2009 - 2009 IEEE Military Communications Conference*, 2009, pp. 1–7.

[5] O. O. Medaiyese, A. Syed, and A. P. Lauf, "Machine learning framework for rf-based drone detection and identification system," *arXiv*, 2020.

[6] S. Al-Emadi and F. Al-Senaid, "Drone detection approach based on radio-frequency using convolutional neural network," in *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, 2020, pp. 29–34.

[7] I. Nemer, T. Sheltami, I. Ahmad, A. U.-H. Yasar, and M. A. R. Abdeen, "Rf-based uav detection and identification using hierarchical learning approach," *Sensors*, vol. 21, no. 6, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/6/1947

[8] H. Shin, K. Choi, Y.-S. Park, J. Choi, and Y. Kim, "Security analysis of fhss-type drone controller," in *WISA*, 2015.

[9] M. Song and T. Allison, "Frequency hopping pattern recognition algorithms for wireless sensor networks," in *PDCS*, 2005.

[10] F. Yin and J. W. C. Guo, Eds., *Advances in Neural Networks - ISNN 2004*. Springer, 2004.

[11] D. Löschenbrand, M. Hofer, L. Bernadó, G. Humer, B. Schrenk, S. Zelenbaba, and T. Zemen, "Distributed massive mimo channel measurements in urban vehicular scenario," pp. 1–5, 2019.

[12] M. S. Allahham, M. F. Al-Sa'd, A. Al-Ali, A. Mohamed, T. Khattab, and A. Erbad, "Dronerf dataset: A dataset of drones for rf-based detection, classification and identification," *Data in Brief*, vol. 26, 2019.

[13] M. Herdin, N. Czink, H. Ozcelik, and E. Bonek, "Correlation matrix distance, a meaningful measure for evaluation of non-stationary mimo channels," in *2005 IEEE 61st Vehicular Technology Conference*, vol. 1, 2005, pp. 136–140 Vol. 1.

[14] L. Bernadó, T. Zemen, F. Tufvesson, A. F. Molisch, and C. F. Mecklenbräuker, "The (in-) validity of the wssus assumption in vehicular radio channels," in *2012 IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications - (PIMRC)*, 2012, pp. 1757–1762.

[15] T. Elhabian, B. Zhang, and D. Shao, "Fast de-hopping and frequency hopping pattern (fhp) estimation for ds/fhss using neural networks," in *Advances in Neural Networks - ISNN 2004*, F.-L. Yin, J. Wang, and C. Guo, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 248–253.