Evaluation of Machine Learning Models for a Chipless RFID Sensor Tag

Nadeem Rather[1], Roy B. V. B. Simorangkir, John Buckley, Brendan O'Flynn, and Salvatore Tedesco
Wireless Sensor Network group, Tyndall National Institute, University College Cork, Cork, Ireland, [1]nadeem.rather@tyndall.ie

***Abstract*— Radar cross section (RCS) is a measure of the reflective strength of a radar target. Chipless RFID tags use this principle to create a tag that can be read at a distance without needing a power-hungry radio transceiver chip and/or battery. A chipless tag consists of a pattern of conductive and dielectric materials that backscatter electromagnetic (EM) waves in a distinctive pattern. A chipless tag can be read and identified by analysing the reflected waves and matching it with a predefined EM signature. In this paper, for the first time, several regression-based machine learning (ML) models are evaluated to detect identification and sensing information for an RCS-based chipless RFID tag. The simulated EM RCS signatures containing an 8-bit identification code and six capacitive sensing values are evaluated. The EM RCS signatures are evaluated within the UWB frequency band from 3.1 to 10.6 GHz. A dataset of 1,530 simulated signatures with relevant features are utilised for model training, validation, and testing. Root mean square error (RMSE) is used as the quantitative metric to evaluate their performance. It is found that Support Vector Regression (SVR) models provide the minimum RMSE for the identification code. At the same time, the Gradient Boosted Trees (GBT) regression model performed better in detecting the sensing information.**

***Index Terms*— Chipless RFID, Electromagnetic Signatures, Machine Learning, Radar Cross Section, Regression, Supervised Learning.**

I. Introduction

Passive RFID can be divided into traditional chip-based and chipless RFID (CRFID). The main difference is that with a traditional RFID tag, sensor integration, wireless data modulation, and data storage are all handled on-chip [1-5]. In contrast, a CRFID tag does not require a chip or battery to function. By altering the tag's structural features, which also regulate its resonant characteristics, data can be encoded, creating a means for passive data encryption in the backscattered Electromagnetic (EM) waves [6-8]. The reader's data processing methods are the only means through which the encoded information on the tag may be retrieved (See Fig. 1). Numerous applications that call for low-cost, reliable, and fully printable identifying and sensing tags have succeeded with CRFID technology [9, 10]. The absence of chip, battery, and other electronics provides CRFID tags with added versatility in harsh and dynamic environments as compared to chipped counterparts [11]. Based on the design and detection mechanism, CRFID tags are categorized into time-domain reflectometry-based, spectral-signature-based, and hybrid-based tags [6]. In this paper, the Radar Cross Section (RCS) spectral-signature based tags are considered.

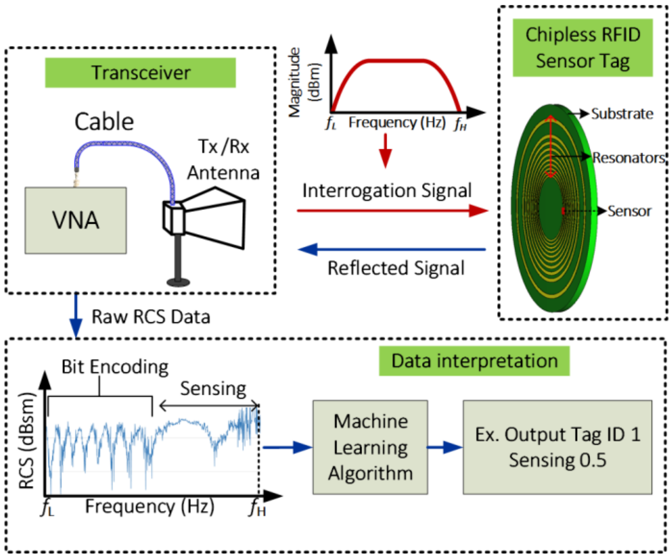


Fig. 1. Overview of the machine learning assisted RCS based CRFID system. (VNA: Vector Network Analyzer)

RCS is a measure of the reflective strength of a radar target and is measured in square meters (sm). In an RCS CRFID system, the tag acts as the radar target. When a plane wave interrogates such tag consisting of an array of resonant elements, a frequency selective reflection is achieved, creating peaks and troughs in the frequency response. The presence or absence of resonance peaks associated with the resonators of the tag is used to encode bit information '0' or '1' (see Fig. 1). A binary sequence can thus be generated and utilised to create different Identification codes (IDs). The operating frequency band, tag size, number of resonators, and encoding techniques determine the number of IDs that can be encoded using a CRFID tag. Sensing features can also be added by loading one or more resonators with certain materials which have electrical properties that vary in response to external stimuli or surrounding environmental conditions, resulting in a change in the tag's overall RCS response [9]. The backscattered signals of the tags are typically prone to background clutter, noise and interference affecting the RCS response (i.e., resonant frequencies, magnitudes, and phases) [12, 13]. This may lead to an inaccurate detection of the information carried by the tags.

In order to develop a reliable identification and sensing system, accurate detection of the tags backscattered signal is of paramount.

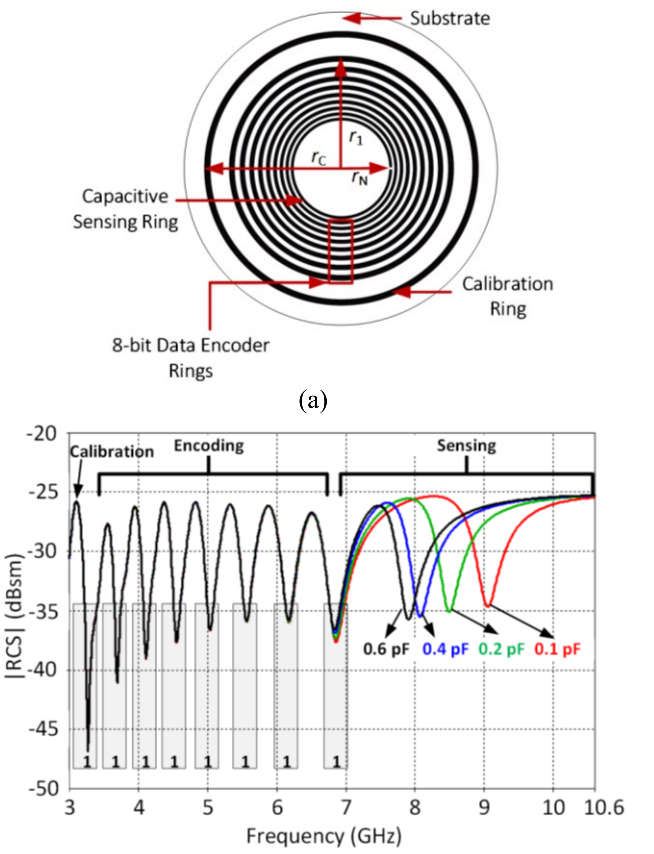
Several studies have applied supervised machine learning and deep learning algorithms to overcome the above challenge. The pattern classification capability of such techniques can mitigate the challenge mentioned above, paving a way to a robust chipless reading system [12-18]. For example, in [16], the authors use an RCS-based tag for gesture recognition. The tag is developed using three split ring resonators to create a specific 3-bit EM signature in the backscattered EM wave. A Feed Forward Artificial Neural Network (ANN) is utilised to train the models and to classify the eight possible combinations. Further, in [17], the authors utilize a deep learning-based security model to provide a high accuracy of above 93%, when classifying a cloned tag from the genuine CRFID tag even in the presence of additive RF interference in real-time.

In [12], the authors use RCS-based tags to develop a segregation system for a plastic recycling application. A Random Forest classifier is used and trained with a dataset consisting of 300 data points to detect two plastic types (2 IDs). The authors achieved an accuracy of 90% when classifying the two plastic types from homogenous bales. In another scenario, for non-homogeneous bales an accuracy of 65% was achieved. These papers, however, are limited to the number of IDs used for training. These papers also lack an efficient model evaluation for a given sensing technique. However, to develop high data storage capacity CRFID sensor tags for various IoT (Internet of Things) applications, it is essential to evaluate cases with relatively higher number of ID's and sensing capability.

In this paper, for the first time, machine learning models are trained to detect a higher bit capacity CRFID sensor tag compared to works already presented in the literature. This work considered an 8-bit capacity (ID 1 - 255) tag with six possible sensing states (0.1, 0.2, 0.3, 0.4, 0.5, and 0.6 pF) for each ID. In this work, unlike above papers, the dataset is built using 1,530 different EM signatures. Also, unlike [12] this work does not use ML methods to classify between two classes (i.e. material types), but relies on regression methods to directly predict continuous values (IDs and sensing value). The objective of the paper is to assess and compare various ML regression models to investigate which one best fits the provided EM signatures and patterns.

II. CRFID Tag Topology

For this study, the RCS-based CRFID tag design that we previously developed in [8] is used. As shown in Fig. 2(a), the tag consists of a calibration ring at the outermost edge, a sensing ring at the innermost edge, and eight rings in between for 8-bit data encoding. The tag is printed on a 1-mm thick FR4 substrate (ε_r = 4.6 and tan δ = 0.025) having a copper thickness of 0.035 mm. The information from the tag is retrieved by observing its RCS. As introduced in [8], the presence of the ring and its associated trough is utilised to represent bit '1' and their absence is used to represent bit '0'. CST Microwave Studio software was used to simulate the tag [19]. A vertically polarized plane wave was used to excite the tag remotely, and an RCS probe was used to obtain the RCS of the tag. One example of the tag RCS response is shown in Fig. 2(b) for a tag carrying an encoded information '11111111' with four sensing states.



(a)

(b)

Fig. 2. (a) Topology of CRFID tag used in this study (r_c = 10.85 mm). (b) Simulated EM RCS response representing '11111111' with capacitive sensing mechanism.

As mentioned, the innermost ring, which is associated with the last trough, is used for sensing. The frequency shift in the location of the last trough is used to indicate the change in the target sensing parameter, which is demonstrated in this case by varying the value of the capacitor on the innermost ring. It can therefore be understood that two forms of data encoding techniques are used in this tag: Amplitude Shifts for the identification data and Frequency Shifts for the sensing data. More details on the tag design and its working principles can be acquired from [8].

III. Research Methodology

In supervised learning, a model is trained to utilise direct feedback from provided labelled datasets and outputs. Applications for supervised learning include outcome prediction and data classification. Based on the task that the ML model needs to perform, one can either use classification or regression models.

In this paper, regression models are used for our dataset as the ML model is asked to predict a numerical value from some given inputs. To solve this task, the learning algorithm is asked to output a function f: $R_n \rightarrow R$ [20]. Four different popular ML models were evaluated for our dataset, namely, Support Vector Regression (SVR), Decision Trees (DT), Gradient Boosted Trees (GBT), and Random Forest (RF) [21].

A. Machine Learning

The input dataset for our ML models was attained from CST EM simulations. All 1,530 EM signatures were obtained from simulations consisting of ID 1 – 255 and six sensing states for each ID (corresponding to 0.1 – 0.6 pF capacitances). Tag ID-0 is not considered in the dataset, as it would be specific to only sensing points. EM signatures distributed along 3.1 to 10.6 GHz band were normalized into 1001 frequency sampling points, and the RCS magnitude was measured in decibels as dBsm. Further, the magnitude RCS response was used to acquire several relevant features to increase the data for training the models. Features such as min, max, mean, standard deviation, skewness, peaks, kurtosis, coefficient of variation (CV), root mean square (RMS), peak-to-peak, and signal magnitude area (SMA) were considered for feature extraction.

These features were first extracted for all samples for each EM signature, and then for samples relevant to the frequency samples needed for sensing and identifying each bit (Shown in Fig. 3). In accordance with the encoding principle, the samples relevant to identify the presence and absence of troughs, which corresponds to bit-1 and bit-0 respectively, were chosen. Therefore, a total of eleven features for each sampling band (e.g., eight bands for the IDs, and one band for the sensing value) plus for the whole signal were obtained. An example of the sampling bands are shown in Fig. 3 for Tag ID 255 with a sensing value of 0.6 pF.

As a result, the input dataset consists of 110 relevant features extracted from the RCS magnitude EM signatures. Python is chosen to develop the models. A total of 1,530 EM signatures were used in the final dataset. The dataset was split into 80% for the development set, and 20% for final testing, with the development set split 75/25 in training and validation sets, as shown in Fig. 4. The values of the dataset's numeric columns were then converted to a standard scale by normalization, which prevented information loss or distortion due to variances in the ranges of values. The standardization of value is achieved by equation (1), where 'N' and 'A' are the normalized and actual values, respectively.

$$N = \frac{(A - \text{Mean})}{\text{Standard Deviation}} \tag{1}$$

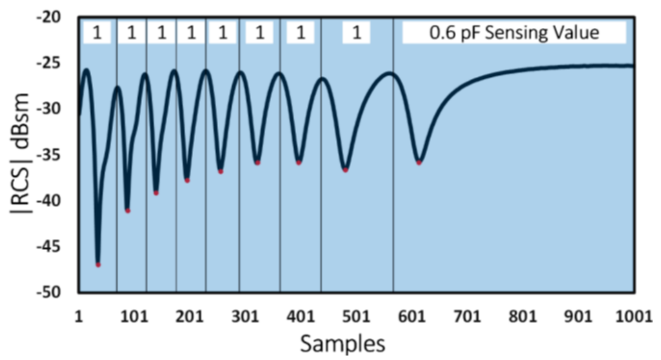


Fig. 3. Representation of the simulated RCS EM Signature sampling blocks that were considered during feature extraction.

The following process is followed to develop our ML regression models, where 'X' represents the input RCS magnitude values, 'y' represents the output labels consisting of the tag ID (1-255) and the sensing capacitance values (0.1 to 0.6 pF for each ID):

- The model is trained on the training dataset (X_train, y_train);
- A grid search is used during training to select the optimal hyper-parameters and features (Fig. 4) for each model;
- The model is checked for overfitting using the validation dataset (X_val);
- If overfitting does not occur, the model is trained on the development set using the optimal hyper-parameters defined in the previous step (X_dev, y_dev);
- The model makes predictions on the unknown test dataset (X_test) and the predictions are evaluated using RMSE scores.

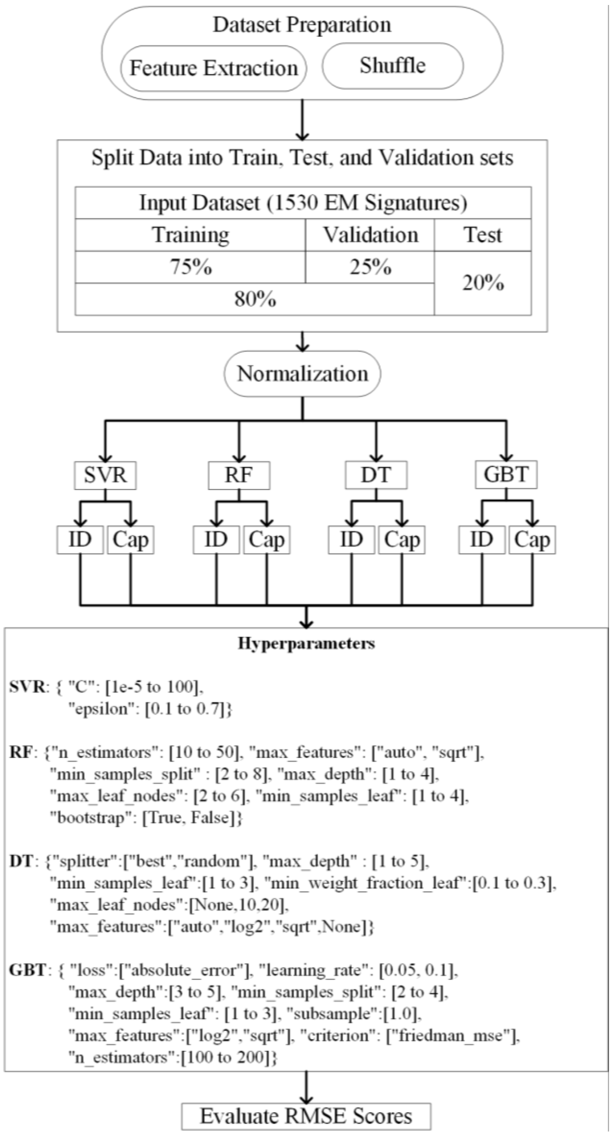


Fig. 4. Basic flow diagram of the overall machine learning process.

The validation set was used to make sure the models did not overfit, e.g. the model becomes too specific to the training dataset and does not generalise well to new unseen data. Due to the two different encoding techniques, two separate models are developed for the tag ID and capacitive sensing. Feature selection was included in each model, so that only the most relevant features from the input dataset were utilised for developing the model and avoid overfitting.

In order to fit the data and eliminate the irrelevant features, the Recursive Feature Elimination with Cross Validation (RFECV) approach is used. Cross-validation is used with RFE to score several feature subsets and choose the top scoring collection to determine the ideal number of features. Further, a grid search approach with a 10-fold cross-validation was performed to use the best hyper-parameters for each model (Fig. 4). As shown in Table I, the quadratic scoring metric known as RMSE given by Eq. (2) was utilised to evaluate model performance. In terms of outliers, RMSE penalizes their existence by generating huge errors. Hence, as in our case, it is helpful when huge inaccuracies are not desired.

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{n}}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (2)$$

Where n is the number of observations, y_i is the actual value, and $\hat{y}_i$ is the predicted value.

RMSE calculates the average magnitude of the errors. It is the average of the squared deviations between the predicted results and actual observations. Each model was evaluated for RMSE on all training, validation, and test sets [21].

IV. Results

Based on the training and validation dataset RMSE scores the best hyper-parameters were selected to evaluate the RMSE for test dataset (Table I). It is observed that the SVR model performed well on the ID and sensing data. The results show that the model can accurately predict the unknown test data set. A minimum error of 1.32 on the test set was observed for the tag ID task, which means the models are predicting quite reasonably. Indeed, the normalized RMSE (using the max and min output values, respectively 255 and 1) is calculate as 1.32/254 = 0.52%. Further, the DT and RF models performed poorly with the given data. It is observed that the error increased significantly for ID detection, while there was a slight increase in sensing error. Lastly, GBT regression was utilised, and as seen the results for the sensing improved significantly to 0.032 pF. The normalized RMSE in this case was approximately 6.4%. This is vital in developing models with high resolution sensing. The RMSE for ID, however, increased compared to SVR. It can be concluded that two different independent models are required to detect ID and sensing (capacitance) values on the adopted chipless RFID sensor tag. Moreover, it is seen that all the models were able to generalize to new data well, implying a good progress towards a machine learning assisted chipless RF identification and sensing detection system.

Table I. Summary of RMSE results for the evaluated ML models

Model	Tag ID			Capacitive Sensing (pF)		
	Training	***Val****	***Test***	***Training***	***Val****	***Test***
SVR	**1.56**	**1.84**	**1.32**	0.062	0.065	0.069
DT	9.27	9.06	9.14	0.098	0.111	0.119
RF	11.34	11.25	12.9	0.089	0.094	0.092
GBT	1.60	2.24	2.34	**0.019**	**0.038**	**0.032**

Best results are highlighted in bold (*Validation)

V. Conclusion and Future Work

In this paper, the performance of four different well-known ML regression models were evaluated. For the first time reported in the literature, these regression models are evaluated to detect 8-bit ID and sensing information for an RCS based CRFID sensor tag. Each ID has six sensing states contained in the same EM signature. The EM signatures in the given input dataset consisted of peaks and troughs in the simulated RCS signature. The trough's presence, absence, and frequency shifting represented the logic-'1', '0', and sensing points, respectively.

It is observed that low RMSE values are obtained by training two different ML models for ID and sensing values. The SVR model performed better for ID detection with an RMSE of 1.32, while GBT achieved the lowest RMSE for sensing detection at 0.032 pF. The next step is to evaluate these models with measured datasets. Further, a measured dataset of all the tag IDs and sensing states will be utilised for the training and testing of the models. The intention is to utilize this knowledge to develop and adopt several other ML models, including neural networks, and probabilistic models like Gaussian process to reduce detection errors.

The future models will also include the measured data such that the models can be developed with several changing parameters (i.e., read range, tag alignment, RF noise, etc.). The goal is to show that the tag can be detected with the correct ID and high-resolution sensing information on several IoT platforms (when the sensor tag is placed in close proximity with objects of varying dielectric properties such as conductivity, permittivity, and permeability) where the effects of clutter, frequency, phase, and amplitude shifts are also taken into consideration.

Acknowledgement

This paper has emanated from research funding provided by Science Foundation Ireland (SFI) as part of the SFI Centre VistaMilk (SFI 16/RC/3835). Aspects of this work have been supported under the following Grant Numbers: CONNECT Centre for Future Networks and Communications (13/RC/2077), Insight Centre for Data Analytics (SFI/12/RC/2289) and 16/RC/3918-CONFIRM, as well as the Enterprise Ireland funded Holistics DTIF (Disruptive Technologies Innovation Fund) (EI-DT20180291-A). All work is co-funded under the European Regional Development Fund.

REFERENCES

[1] K. Finkenzeller, *RFID handbook: fundamentals and applications in contactless smart cards, radio frequency identification and near-field communication*. John wiley & sons, 2010.

[2] D. R. Gawade *et al.*, "A battery-less NFC sensor transponder for museum artefact monitoring-a review of NFC sensor technology and a proposed solution," 2019.

[3] C. Peres *et al.*, "Theoretical models for underwater RFID and the impact of water salinity on the design of wireless systems," *International Journal on Advances in Networks and Services,* vol. 13, no. 34, pp. 45-59, 2020.

[4] L. Catarinucci, S. Tedesco, D. De Donno, and L. Tarricone, "Platform-robust passive UHF RFID tags: A case-study in robotics," *Progress in Electromagnetics Research C,* vol. 30, pp. 27-39, 2012.

[5] L. Catarinucci, S. Tedesco, and L. Tarricone, "On the use of UHF RFID antenna systems customized for robotic applications," in *Proceedings of the 2012 IEEE International Symposium on Antennas and Propagation*, 2012, pp. 1-2: IEEE.

[6] S. Preradovic and N. C. Karmakar, "Chipless RFID: Bar code of the future," *IEEE microwave magazine,* vol. 11, no. 7, pp. 87-97, 2010.

[7] N. Rather, R. B. Simorangkir, J. Buckley, and B. O'Flynn, "Flexible and Semi-transparent Chipless RFID Tag based on PDMS-Conductive Fabric Composite," in *2022 International Workshop on Antenna Technology (iWAT)*, 2022, pp. 33-36: IEEE.

[8] N. Rather, J. Buckley, B. O'Flynn, and M. Pigeon, "A novel RCS based CRFID tag design," in *2022 16th European Conference on Antennas and Propagation (EuCAP)*, 2022, pp. 1-5: IEEE.

[9] N. C. Karmakar, E. M. Amin, and J. K. Saha, *Chipless RFID sensors*. John Wiley & Sons, 2016.

[10] A. Subrahmannian and S. K. Behera, "Chipless RFID Sensors for IoT-Based Healthcare Applications: A Review of State of the Art," *IEEE Transactions on Instrumentation and Measurement,* 2022.

[11] C. Herrojo *et al.*, "Time-domain-signature chipless RFID tags: Near-field chipless-RFID systems with high data capacity," *IEEE Microwave Magazine,* vol. 20, no. 12, pp. 87-101, 2019.

[12] F. Villa-Gonzalez, R. Bhattacharyya, and S. Sarma, "Single and bulk identification of plastics in the recycling chain using Chipless RFID tags," in *2021 IEEE International Conference on RFID (RFID)*, 2021, pp. 1-8: IEEE.

[13] S. Jeong, M. M. Tentzeris, and S. Kim, "Machine learning approach for wirelessly powered RFID-based backscattering sensor system," *IEEE Journal of Radio Frequency Identification,* vol. 4, no. 3, pp. 186-194, 2020.

[14] L. M. Arjomandi and N. C. Karmakar, "An enhanced chipless RFID system in 60 GHz using pattern recognition techniques," in *2018 48th European Microwave Conference (EuMC)*, 2018, pp. 973-976: IEEE.

[15] L. M. Arjomandi, G. Khadka, Z. Xiong, and N. C. Karmakar, "Document verification: A cloud-based computing pattern recognition approach to chipless RFID," *IEEE Access,* vol. 6, pp. 78007-78015, 2018.

[16] M. Hajizadegan and P.-Y. Chen, "Gesture recognition using a portable and flexible meta-atom panel and machine learning," in *2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, 2019, pp. 637-638: IEEE.

[17] G. Khadka, B. Ray, N. C. Karmakar, and J. Choi, "Physical Layer Detection and Security of Printed Chipless RFID Tag for Internet of Things Applications," *IEEE Internet of Things Journal,* 2022.

[18] D. Nastasiu *et al.*, "A New Method of Secure Authentication Based on Electromagnetic Signatures of Chipless RFID Tags and Machine Learning Approaches," *Sensors,* vol. 20, no. 21, p. 6385, 2020.

[19] D. Systemes. (2021). *CST STUDIO SUITE ELECTROMAGNETIC FIELD SIMULATION SOFTWARE*. Available: https://www.3ds.com/products-services/simulia/products/cst-studio-suite/

[20] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[21] A. Geron, "Hands-On Machine Learning with Scikit-Learn & Tensorflow O'Reilly Media, Inc," *O'Reilly Media, Inc,* vol. 1005, p. 564, 2017.