Applicability of different neural network architectures in UWB signal processing for different object classification

Kristaps Greitans
Institute of Electronics and Computer science
Riga, Latvia
kristaps.greitans@edi.lv

Modris Greitans
Institute of Electronics and Computer science
Riga, Latvia
modris_greitans@edi.lv

***Abstract*—In the paper, the performance of different Artificial neural network (ANN) architectures - CNN, RNN, Transformers, CNN/LSTM, ResNet, and MLP is discussed in UWB impulse radio radar signal classification. The signals are obtained by reflecting and passing UWB pulses through different material objects that give different information for the classification purpose. The ANN architectures are compared on their classification precision, training time, and memory requirements. The training data consists of 144 objects including regular and crashed PET bottles, glass bottles, and metal cans. The results show the accuracy's of classification if mono-static (reflected signals are analyzed), bi-static (propagated trough signals), and multi-static setups of UWB radar are used. For single-channel cases GRU (99.65%), Resnet (99.69%), Transformer (99.66%) architectures are preferable, while the highest multi-static evaluation accuracy reaches 99.90% for the Transformer. As expected the more dense ANN networks perform better classification.**

I. INTRODUCTION

A newer technology, ultra wide-band impulse radio (UWB-IR) radar, is being used for indoor location, medical applications, security [1-4], and non-contact, non-destructive evaluations of material's visible and non-visible properties [4-6]. Different antenna arrangements, including bi-static, mono-static (more in the text, single-static), and multi-static, can be used to generate UWB-IR radar signals. In a mono-static radar configuration, there is still just one receiving antenna, but it is placed in a different location from the transmitting antenna, such as directly behind the target object. As a result, a signal that has passed through the target is collected and processed. Because there are multiple receiving antennas dispersed across numerous locations in addition to one emitting antenna in this arrangement, a multi-static system is more challenging. The multi-static arrangement enables simultaneous recording of signals that are propagated through and reflected in various combinations. In [6,7], where the classification approach is based on a hybrid neural network that includes Convolution layers for feature recognition, the benefits of a multi-static UWB-IR radar configuration are covered.

Different Artificial neural network (ANN) architectures have different advantages. For example, Convolutional neural networks (CNNs) achieved enormous success in computer vision, while recurrent neural networks (RNNs) gained popularity in speech recognition. Regarding the processing of UWB-IR radar signals, it is not known which type of ANN architecture is appropriate for certain applications. The impact of different architectures for UWB-IR radar performance to classify objects from different materials has not been analyzed so far. Therefore, this paper aims to investigate, evaluate (accuracy and training time) and compare the use of different neural networks in the feature recognition layer for both single-static and multi-static setups of UWB-IR radar. With a further possibility to understand how significantly each of the antennas contributes to the overall accuracy of the classification of objects from different materials. The ANN architectures used in the research in this article are: a multi-layer perceptron, CNN, a long-short time memory, a CNN-LSTM, Transformer, GRU, ResNet. The Artificial neural network methods for classification are used instead of some more classical classification methods because of the supremacy when working with big data sets and increased complexity with multiple signals or different materials in the future. For now this paper only looks at the basic case of four receiving antennas and three different materials and two differant shapes, keeping in mind this possible expansions of measurements to understand which network could be used. Given the above, the following paper is structured as follows: The next section provides information on the feasibility study as well as information about the different ANNs. Chapter 3 describes the experimental UWB-IR radar set up and the signal processing approach used to evaluate the efficiency of different ANN architectures. The results are presented in Section IV, followed by a final section with conclusions and ideas for further research.

II. PRELIMINARIES

A. *Use of multi-static UWB-IR radar*

In general, UWB impulse radio radar sends and receives very brief impulses on the order of sub-nanoseconds. In [8], the author analyzes how the relative dielectric constant and material losses affect the peaks in the reflected and transmitted through UWB pulse signal wave forms. Consequently, classification by visual means or utilizing only one time-domain method was not possible. The number of layers or shape abnormalities of the item increases the complexity of signal processing [9]. However, the signal characteristics stated above often only differ by a little amount.

On the other hand, when measuring a crushed or ruff surfaced object, there can be a possibility that a single antenna reflects from the object at an undesirable angle, thus making the measurement undetectable [10]. When UWB signals come into contact with the object under study, occurs not only reflection but also transition. UWB pulse, which is propagated through the object, carries a different kind of information about the object and can be used as an additional component for signal processing [11]. The analysis of the propagated through signals at different angles and reflected signal from the object in combination with an appropriate signal processing can give more information about the object (e.g. thickness, fill density, material, and cracks) and gain accuracy [12]. However, the propagated trough signals not only have different peaks, signal phases, and shapes but a slight delay in respect to the reflected signal as well.

The aspects described above about the nature of UWB-IR signals, especially in the case of multi-static architecture, illustrate the complexity of the development of appropriate signal processing algorithms. The use of ANN is one of the techniques typically used in problem solution in similar cases [13]. ANN can detect differences between signal profiles that are too complex for convenient signal processing methods.

B. Architectures of ANN

1) MLP architecture: The multi-layer perceptron (MLP) is the simplest and most used building block for any time series categorization application . It is a straightforward fully connected feed-forward network made up of nodes that calculate the output value from the input layer through a number of hidden layers. It served as both an output layer from the other feature recognition layers and a distinct feature recognition layer in this study [14].

2) CNN architecture: The Convolution Neural Network architecture is a widely used layer for image recognition and other 3D data classification. CNN is typically divided into two parts: the first is the convolution and pooling of data, and the second is the MLP layers for classification based on the extracted features in part 1. The convolutions layer uses convolution filters to perform convolution operations on the preceding layer's time series. The filter parameters are chosen based on prior domain knowledge or by using hyper-parameter search. In the classification task, the most common method is to use the maximum output neuron as the class label of the input time series [15].

3) LSTM architecture: RNN networks, or more precisely Long Short-Term Memory (LSTM) networks, are the most popular sequence classifiers. Their key advantage is their internal memory, which recalls the past inputs to create accurate output values. Recurrent networks operate by iteratively processing each element of a vector, storing the result, and categorizing the value. Due to their lengthy training periods and heavy computational demands, they are at a disadvantage. The vanishing gradient problem, which plagues general RNNs, is avoided by LSTM networks by adding gating functions into their state dynamics. An LSTM keeps track of a memory vector m and a hidden vector h that are in charge of managing state changes and outputs at every time step [16].

4) GRU architecture: Recurrent neural networks use gated recurrent units (GRUs) as their gating mechanism. The GRU is similar to an LSTM with a forget gate, but because it lacks an output gate, it has fewer parameters than an LSTM. It was discovered that GRU and LSTM performed similarly on some polyphonic music modeling, speech signal modeling, and natural language processing tasks. Given that our circumstance involves a small data set of only 100+ unique items, GRUs have been demonstrated to perform better on some smaller and less frequent data sets [17].

5) CNN/LSTM architecture: The CNN/LSTM architecture combines LSTMs to facilitate sequence prediction with CNN layers for feature extraction on input data. The advantages of both network types are used in this style of architecture. CNN/LSTMs were created for challenges involving the prediction of visual time series and for the purpose of producing textual descriptions from sequences of images (e.g. videos). However, this can be modified for vector classification in 1D and ought to be better than the two different kinds of designs [18].

6) Transformer architecture: The RNN/LSTM design is improved via transformer networks. Tokens are processed by LSTMs. This architecture keeps track of the complete sequence it has seen in a concealed state that is updated with each new input token. Transformers, in contrast, preserves direct connections to all earlier timestamps, enabling information to spread over a lot longer sequences [19].

7) ResNet architecture: ResNets serve as an example of a discriminative deep learning method. They discover the relationship between the raw input and the classes that might be assigned to those input records in the event of classification. ResNets are CNN versions that improve model performance by introducing linear shortcut connections between blocks of convolutional layers [20].

III. Data acquisition and processing

This section describes the approach for data capturing simultaneously by one or more UWB-IR receivers and the methodology for appropriate signal processing.

Figure 1 shows the block diagram of the signal processing approach proposed in [7] and is used in described research as well. This structure consists of four UWB receivers for signal acquisition each followed by their own ANN feature block, the feature blocks then are combined into a single output with a concatenate layer, to then the single output to be classified using two dense layers. The basic structure of the ANN block involves an input layer followed by one of the described ANN architecture [II-B] feature layers, dropout, and max-pooling layer. Dropout layers, dense layers have been further applied

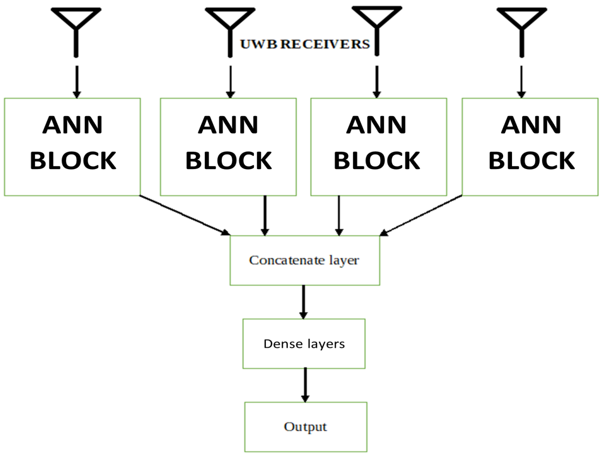

Fig. 1. Block diagram of the signal processing.

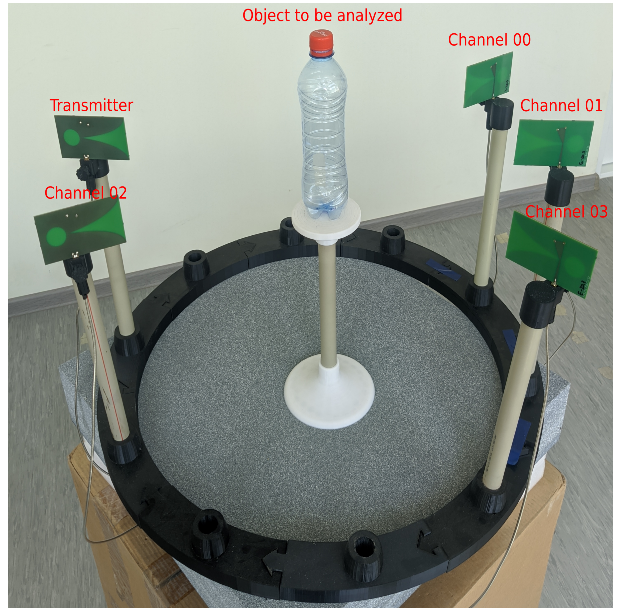

Fig. 2. Experimental set-up for the data acquisition.

or tweaked to regularize each of the different networks in their own way avoiding over-fitting. In particular, it is known that deep ANN architectures can achieve high accuracy in forecasting but, due to the high number of parameters involved, there is a high risk of over-fitting, with a consequent decrease of out-of-sample accuracy. Dropout is a technique that efficiently solves the above problem.

In general, additional pre-processing blocks can be added to signal processing before data are passed to the input layer, which would filter out various artifacts that may occur in the case of UWB receivers due to nearby working WI-FI or cellular network devices.

The proposed solution is designed in a way that the input channel count can be easily increased or decreased, and the results from different antennas or its combination can be compared. If an antenna input channel data is zero, that channel is ignored and calculates the output from the available input channels.

A. Data Acquisition

To compare the applicability of ANNs, which have been described within Subsection II-B, for the classification of objects from different materials, a multi-static experimental UWB-IR radar setup has been used for data acquisition [7]. The setup is shown in Figure 2. and it consists of UWB-IR radar [21], RF antenna switch, a single 2.0GHz Vivaldi transmitting antenna, and 4 exact receiving antennas. As it can be seen in Figure 2. the reflected antenna (Channel 02) is placed directly next to the transmitting antenna but the antennas (Channel 00, Channel 01, and Channel 03), for propagated through signals, are placed directly across from the transmitting antenna. The two side propagation antennas are placed at a distance of 36 degrees from the directly propagated trough antenna (Channel 1). This was done to cover quite a large area of the object while having no interference between the antennas.

The UWB radar is configured to sample 1023 signal points with a sampling interval equal to 20 ps. The UWB radar transmitter and receiver bandwidth rates are 0.1 – 4.5 GHz and the antenna parameters are: Frequency Range = 1.4 - 4.7 GHz and Gain @ 3.30GHz = 6.6 dB. 100 frames are recorded to ensure the possibility to eliminate background noise by averaging the data. The recorded data frames (1023 samples) are processed using the approach which is illustrated in Figure 1. No other signal processing methods for the signals were used. Parameters of the ANNs training phase are the same as in [7]. The parameters have been optimized using the Grid search method [22]. An early stopping function has been used for maximum accuracy. The total data set used to obtain the results of the presented research has been recorded from 114 objects, including 27 regular and 43 crashed PET bottles, 20 metal cans, and 24 glass bottles. All of the objects were taken with similar shape in mind to possibly eliminate that the ANN classifies the materials by shape and not the material.

B. Data processing

First, the data recorded from all objects were divided into two groups: 1) data for ANN training; 2) data for evaluation, which is not used in the training and has never been seen by any neural network.

Regarding the training, data of each object measurement has been averaged over 25 frames from the 100 recorded frames. Thus from the raw data frames, 288 averaged data sequences have been formed and used in the training process. Each data sequence remains 1023 data long. The averaging is done to eliminate artifacts in signals and stop the possibility of over-fitting. The training data for each channel consists of two parts:

1) 85% of recorded data as training data-set (244 data sequences);

2) 15% of recorded data as validation data split (44 data sequences).

The evaluation data is recorded from 16 regular and 26 crushed PET bottles, which were unseen for training. This data set consists of 100 frames for each object that are not averaged and forms 4200 data sequences). This allows testing the true accuracy of the used ANNs with unseen and from artifacts that influence unimproved data.

IV. Results and discussion

The results for the analysis of the applicability of ANNs are obtained in two sub-cases. Firstly, only one of four channels, which receives propagated trough (Channel 1) or reflected (Channel 2) UWB-IR pulses has been used for data acquiring. Secondly, multi-static case, where all four channels together have been exploited as data sources.

This was chosen to assess the accuracy improvement of using multiple channels in multi-static setup of the radar for different ANN architectures and different types of objects.

It is worth noting that all network models have been optimized with a grid search over hyper-parameters, such as network topology and depth, e.g. number of layers, neurons and activation functions, learning rate batch size, and epochs. Hyper-parameters optimization is by far the most demanding procedure in obtaining the best possible accuracy. An extensive grid search may take weeks.

Because of the nature of ANN training, the results presented in the paper are obtained by averaging 20 signal processing cycles. This is done to avoid the effects of specific training cases. Each cycle is done using different data for training and evaluation because in the training process Neural Networks every new cycle starts with random seeds.

The training and evaluation procedures have been performed using a PC with an A8-6500 AMD processor and Nvidia GTX780 graphics card.

1) Reflected and Propagated trough case results: Table 1. summarizes these classification results for various ANN architectures for single channel cases. Evaluation accuracy shows the percentage of correctly classified objects. The median of the values is used because it represents better the "typical" precision than mean value, which is which is significantly reduced in the event of disturbances and artifacts during the recording of an certain individual signal, because signals have not been averaged during evaluation phase. The amount of trainable parameters shows the number of elements in the architecture, neurons that are affected by the back propagation.

TABLE I
Evaluation results single channel

Architecture	Propagated trough (%)	Reflected (%)	Parameters
MLP	98.43	97.76	262 020
CNN	99.00	97.98	398 048
LSTM	98.21	98.26	1 571 580
GRU	99.60	99.65	133 524
CNN/LSTM	99.17	98.57	654 724
Transformer	99.64	99.66	10 793
ResNet	99.69	99.67	41 728

TABLE II
Wrongly classified objects in single channel case

Architecture	Glass bottle Mistakes	Crushed PET Mistakes	total mistakes
MLP	65	1	66
CNN	1	41	42
LSTM	53	22	75
GRU	11	9	20
CNN/LSTM	1	34	35
Transformer	6	9	15
ResNet	4	9	13

The results show that the worst ANN architecture was MLP with Propagated - 98.43% and reflected - 97.76% accuracy, this was expected as the architecture is the most basic one and does not extract or remember any features from previous layers.

The best performing ANN architecture is ResNet with channel 1 - 99.69% and channel 2 - 99.67% accuracy.

Comparing the two channels it can be seen that the networks with a memory like LSTM, GRU, and Transformer preform better on the reflected signals. As well as GRU network preforms much better than LSTM with a smaller data amount. Comparing the parameter count the better performing networks have a lower amount of parameters, for example, ResNet has only 41 728 but the second-worst network LSTM has 1 571 580 parameters.

To investigate the exact accuracy of the different ANN architectures confusion matrices are used. In Table 2 is shown how many times a regular PET bottle was mistaken with a glass bottle, with a crushed PET bottle and total mistakes made. No other classification errors were done by the architectures, the networks never classified one of the PET bottles as a metal can or crushed PET as a regular PET bottle. For this only Propagated trough data was looked at as it was an overall better-performing model.

It is observable that architectures with memory made more mistakes with crushed PET bottles than glass bottles. But for more traditional networks the glass bottle mistakes were higher. By changing the architectures we can see that the improvement from simple architecture like MLP (66 mistakes) to more dense networks like Transformer (15 mistakes) and ResNet (13 mistakes) decrease the total mistake count around 4.7 times. The mistakes in evaluation data could be explained by the less dense networks or networks without memory networks, did not have enough information or features to extract to correctly classify different object materials.

2) Multi-static case results: The multi-static case has been investigated a little bit deeper because of the better performance, by looking additionally at the mean and standard deviation of the results in addition to the ones done for single-static cases. Table 3. summarizes these classification results for various ANN architecture cases for UWB multi-static radar.

These results show that differently as in the single channel cases the highest median evaluation is 99.90%, for the Transformer network but the lowest accuracy is achieved by MLP with 99.57% accuracy. The Transformer network is the slowest

TABLE III
EVALUATION RESULTS - MULTI-STATIC CASE

Architecture	Evaluation accuracy Median(%)	Evaluation accuracy Mean(%)	Standard deviation (%)	Training time (s)	Parameters
MLP	99.57	94.39	4.96	60	3014282
CNN	99.62	99.5	1.27	70	657020
LSTM	99.79	93.63	5.99	1500	2090986
GRU	99.71	96.20	4.25	1200	1182348
CNN/LSTM	99.74	96.60	3.57	2000	654724
Transformer	99.90	98.45	1.02	21056	10793
ResNet	99.83	97.8	1.7	75	44886

TABLE IV
WRONGLY CLASSIFIED MATERIAL OBJECT TABLE FOR MULTI-STATIC CASE

Architecture	Glass bottle Mistakes	Crushed PET Mistakes	Total mistakes
MLP	10	8	18
CNN	9	7	16
LSTM	6	3	9
GRU	2	10	12
CNN/LSTM	3	8	11
Transformer	1	3	4
ResNet	3	4	7

network to train with an average of 21056 seconds but the MLP network was the fastest to train with 60 seconds per cycle. The high accuracy of the transformer network could be identified that the Attention mechanism inside the transformer enables the transformers to have extremely long-term memory. A transformer model can “attend” or “focus” on all previous tokens that have been generated and thus work better with the more inputs it can receive. As we can see from the parameter column the highest count is for the less dense network MLP with 3 014 282 trainable parameters in context to other architectures.

Table 4. shows numbers of the wrongly classified objects for multi-static case. The table having the same heading as single-static case.

Results from Table 4. illustrates, that that architectures with memory made more mistakes with crushed PET bottles than glass bottles. But for more traditional networks the glass bottle mistakes were higher. By changing the architectures we can see that the improvement from simple architecture like MLP (18 mistakes) to more dense networks like Transformer (4 mistakes) and ResNet (7 mistakes) decrease the total mistake count by around 4 times. The mistakes in evaluation data could be explained by the less dense networks or networks without memory networks, did not have enough information or features to extract to correctly classify different object materials.

V. CONCLUSION

The presented research shows that the different deep ANN architectures can correctly process UWB-IR radar signal processing to classify objects from different materials and can distinguish different shapes of the objects.

The accuracy of the classification of the objects from different materials is not only determined by the architecture of the network used, but also by the selection of the channels used to obtain the data. For some ANN architectures, the difference in accuracy between the multi-static and single-static setup is small, for the better performing networks Transformer and ResNet the difference is 0.26% and 0.16%, while for the other the difference can reach up to 1.58% for the LSTM architecture. In real life, any improvement in accuracy means a significant decrease of wrongly classified objects. For single-channel case it would be even 62 objects (difference in LSTM and ResNet classifications). In the multi-static setup, this difference is much smaller - the Transformer incorrectly classifies 14 objects less than MLP.

Looking at multi-static accuracy standard deviation (SD) we can note that all values are above 1% but with a more denser network and higher accuracy, the SD lowers. For example transformer network was the most consistent with 1.02% SD while the LSTM network had the lowest consistency with 5.99%. The high SD value was more associated with networks that have memory. A possible explanation could be that one iteration network remembers different information from other.

The different ANN architectures proved to improve the classification accuracy with different benefits, like training time range from 600 seconds to 2000 seconds or the parameter count which represents the usage of computer processing power. With more parameters, there are more calculations.

With taking into account all this information the most favourable architecture for UWB signal classification would be ResNet because of its high performance in multi-static and single-static cases while still having a small training time of 75 seconds.

With Artificial neuron networks we can distinguish response signals on various materials and correctly classify them and if they are crushed or normal. This research helps distinguish which ANN architecture would be more favourable for certain an application with certain constraints to complexity, training and processing time, and resources. Material classification has multiple further research purposes in industry, for example, to help at recycling plants, quality control, non-destructive control of defects in different objects. This phenomenon is due to different material electrical properties, which when in contact with UWB signal changes the properties of the signal, like amplitude, frequency, phase. From a signal processing point of view, the research can be extended to investigate how a multi-static UWB radar setup can help detect holes and cracks inside the material, multi-layer structures, their size, and other parameters.

ACKNOWLEDGMENT

This research is funded by the Latvian Council of Science, project AI-based analysis of multi-static UWB-IR radar signals for non-destructive estimation of materials and structure (MIRSA), project No. lzp-2020/2-0270

REFERENCES

[1] Alarifi, Abdulrahman, et al. ”Ultra wideband indoor positioning technologies: Analysis and recent advances.” Sensors 16.5 (2016): 707.

[2] Gu, Changzhan. ”Short-range noncontact sensors for healthcare and other emerging applications: A review.” Sensors 16.8 (2016): 1169.
[3] Bari, Bifta Sama, et al. ”Ultra wide band (uwb) based early breast cancer detection using artificial intelligence.” InECCE2019. Springer, Singapore, 2020. 505-515.
[4] D. R. Gibbins, T. N. T. Henriksson, Ian J. Craddock “Bulk Material Dielectric Property Measurement by an Adapted Free-Space UWB Inverse Scattering Method” 2017 11th European Conference on Antennas and Propagation (EUCAP).
[5] R. Salman, T. Schultze, I. Willms “UWB Material Characterisation and Object Recognition with Applications in Fire and Security”, Proceedings Of The 2008 IEEE International Conference On Ultra-Wideband (ICUWB2008), Vol. 20.
[6] Aristov, V., Greitans, M. Determination of the Electrophysical Parameters of Dielectric Objects via the Processing of Ultra-Wideband Pulse Radar Signals. Aut. Control Comp. Sci. 55, 577–587 (2021).
[7] K. Greitans and M. Greitans, ”Multi-static UWB radar for classification of objects from different materials,” 2021 IEEE Microwave Theory and Techniques in Wireless Communications (MTTW), 2021, pp. 7-11.
[8] V. S. Chernyak, Fundamentals of Multisite Radar Systems, U.K., London:Gordon and Breach Science Publisher, 1998.
[9] S. Sardar and A. K. Mishra, ”UWB based dielectric material characterization using hardware/software co-design based ANN,”2013 IEEE International Conference on Industrial Technology (ICIT), Cape Town, 2013, pp. 1196-1200.
[10] M. Ho, ”Simulation of electromagnetic pulse propagation through dielectric slabs with finite conductivity using characteristic-based method,” PIERS ONLINE, vol. 2, no. 6, pp. 609-613, 2006.
[11] MAKARATAT, K., STAVROU, S. AND BROWN, T. Simulation of UWB distortion combined with indoor spatio-temporal channels. IET Seminar Digest. 2008. 1 - 5.
[12] A. Gulck, T. Lehmann and R. Knochel, ”Characterisation of dielectric obstacles using ultra-wideband techniques,” European Radar Conference, 2005. EURAD 2005., 2005, pp. 259-262.
[13] Li, X.; He, Y.; Jing, X. A Survey of Deep Learning-Based Human Activity Recognition in Radar. RemoteSens. 2019, 11, 1068.
[14] Pal, Sankar K., and Sushmita Mitra. ”Multilayer perceptron, fuzzy sets, classifiaction.” (1992).
[15] Cheng, W., Wang, Y., Peng, Z., Ren, X., Shuai, Y., Zang, & Wu, J. (2021). High-efficiency chaotic time series prediction based on time convolution neural network. Chaos, Solitons & Fractals, 152, 111304.
[16] Dey, R., & Salem, F. M. (2017, August). Gate-variants of gated recurrent unit (GRU) neural networks. In 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS) (pp. 1597-1600).
[17] Sherstinsky, A. (2020). Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D: Nonlinear Phenomena, 404, 132306.
[18] Kim, T. Y., & Cho, S. B. (2019). Predicting residential energy consumption using CNN-LSTM neural networks. Energy, 182, 72-81.
[19] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., & Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems (pp. 5998-6008).
[20] Targ, S., Almeida, D., & Lyman, K. (2016). Resnet in resnet: Generalizing residual architectures. arXiv preprint arXiv:1603.08029.
[21] https://www.edi.lv/en/solutions/uwb-impulse-radar-kit
[22] Liashchynskyi, Petro, and Pavlo Liashchynskyi. ”Grid search, random search, genetic algorithm: A big comparison for NAS.” (2019).