ROMANIA

MINISTRY OF NATIONAL DEFENSE

MILITARY TECHNICAL ACADEMY „FERDINAND I”

FACULTY OF COMMUNICATIONS AND ELECTRONIC SYSTEMS FOR DEFENSE AND SECURITY

MASTER PROGRAM: COMMUNICATION SYSTEMS ENGINEERING AND ELECTRONIC SECURITY

”THE ANALYSIS OF SIGNALS FROM ULTRA-WIDEBAND SYSTEMS”

”ANALIZA SEMNALELOR PROVENITE DIN SISTEME UWB”

Coordinating professor:

Cpt. conf. univ. dr. ing. Angela DIGULESCU-POPESCU

Master student:

Slt.ing. Paraschiva-Cristina POPOVICI

Contains: ……..……..

Inventoried under number: ……..….

Position in the indicator: ……...….

Storage period: ……....…….

Bucharest 2022

Chapter 3

People detection and counting using and IR-UWB radar based on Artificial Intelligence Algorithm

This chapter presents an application of UWB signals. The application refers to people detection and counting using an IR-UWB radar based on Artificial Intelligence Algorithms such as K-nearest neighbours (KNN), Support Vector Machine (SVM), Multilayer Perceptron Neural Network (MLP) and Convolutional Neural Network (CNN).

Three methods are proposed for people classification. First approach implies the use of complex feature-extraction algorithm such as Curvelet transform for frequency-based features, the segmented-based feature extraction method for time-based features and Principal Component Analysis applied on the final set of features. For these sets of features, K-nearest neighbours, Support Vector Machine and Multilayer Perceptron Neural Network algorithms are used for classification, with the mention that both the scenario and the number of persons in the radar range are considered in the label set.

The second approach is based on a Convolutional Neural Network applied on the raw dataset of radar samples, without prior feature extraction.

The third methods proposes the use of Principal Component Analysis for data compression, both of the entire set of data and the set of extracted hybrid features, using KNN, SVM, MLP and CNN. For the two datasets, two architectures are proposed for CNN, with the mention that only the number of persons is considered in the label set.

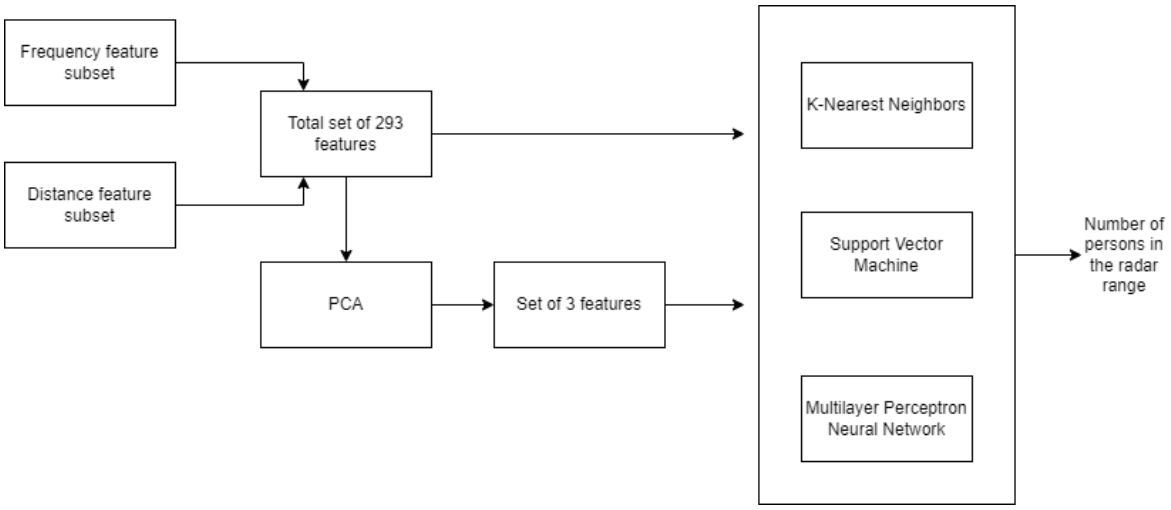


Figure 12. Block diagram of the first approach

3.1 State of the art of Human Detection and Classification Methods using radar technology

In literature, some methods are proposed for human detection, human activity classification, human gesture recognition, and moving target tracking. Usually, people detection and further classification is done using micro-Doppler signatures [31] [41] [14] [16] or by detecting the local maxima in the received signals [24] as extracted features.

Paper [31] uses micro-Doppler signatures of human walking gait, where a roof-top X-band radar is used for measurements. The body Doppler shift and micro-Doppler modulation of the arms` movement is detected by applying STFT to the input signals [31]. The human gait is highlighted using the superposition of the time-frequency signatures over all range cells, which gives a full time-frequency signature of a walking person.

In paper [24], a 2D IR-UWB Radar System is used for the detection, localization and tracking of a moving target. As discussed in this paper, some methods suggested for clutter and noise filtering are: Exponential Averaging (EA)

which uses the previously estimated signals for clutter estimation, and Singular Value Decomposition (SVD) which is an efficient method for through-wall applications and tracking systems, its disadvantage being an expensive computational cost, and the Kalman Filter (KF) which is finally implemented.

The targets are detected due to the comparison of signal strength with a threshold. Two decisions are made regarding the presence of the target, depending on the signal strength. The presence of the target is given by a greater signal strength, the absence being given by a weak signal strength, relative to a threshold [24].

The target distance, material and shape negatively influence the pulse waveform, which affects the matching ratio of the echo signal and a template pulse [24]. To overcome this problem, Constant False Alarm Rate (CFAR) is proposed in [42], its disadvantage being the difficulty of choosing the threshold, depending on the clutter distribution [24]; and the CLEAN algorithm which provided the optimum results, presuming the comparison of the cross-correlation between the echo and the template signal with a given threshold.

A Specular Multi-Path Model (SMPM) together with Multi-target Tracking Technique (MTT) for human body scattered UWB waveforms characterization, and human presence detection, respectively, is proposed in the paper [21]. The authors concluded that UWB radar-based human detector provides accurate results, with more than 80% detection probability and with 1.58% false alarm rate in a realistic outdoor environment. A way to improve the proposed method is the use of a hidden Markov model [21].

Paper [41] shows the effectiveness of using deep neural networks (DNN) for recognizing the micro-Doppler radar signals generated by human walking and background noises, proposing two types of DNN classifier: a binary classifier, with 97.5% classification accuracy and a DNN (Deep Neural Network) multiclass classifier, showing 95.6% accuracy.

A new approach is made in paper [14], where deep convolutional neural networks (DCNNs) are applied to raw micro-Doppler spectrograms for human

detection and activity classification. The measurements are done by a Doppler radar, in an outdoor environment and under line-of-sight conditions. It was concluded that, without any knowledge of the input data features, DCNNs provide high accuracy, with results of 97.6 % for human detection and 90.9% for human activity classification [14]. The vulnerabilities of this method could be a degrading performance due to irregularity in motions and the computational complexity in real-time applications [14].

An automotive radar of 79 GHz is used in paper [16] for pedestrian detection using the micro-Doppler signatures of the human body in the near field (0-15 m). A multi-objective optimization based on Genetic Algorithm (GA) and Random Search (RS) methods is considered for the radar parameters, improving the quality of the radar signal, by improving the velocity resolution of 0.12m/s. The Support Vector Machine (SVM) method is used for object classification: pedestrian or non-pedestrian, providing an accuracy of 99.5%

Conventional detection methods propose the comparison of the received signal power with a fixed or adaptive threshold [41]. The disadvantages of these methods are a high false alarm rate due to noise sources and a radar cross-section fluctuation of targets which degrade the detection performance [41]. New methods of human detection based on machine learning are recently developed, showing high detection accuracy.

3.2. People detection and counting based on Artificial Intelligence Algorithms

This subchapter presents the original dataset of radar samples, the data preprocessing, the feature extraction methods and the classification methods mentioned in the beginning of the chapter.

3.2.1 Dataset

The paper [18] is used as a reference, where dense human counting is performed using an NVA-R661 IR-UWB radar, with an emission pulse having a central frequency of 6.8 GHz and a bandwidth of 2.3GHz. The echo signals are received at a sampling frequency of 39GHz. Curvelet transform and Distance bin method are used as feature extractors, resulting a set of 300 features. The machine learning algorithms used for people counting are Decision Tree, AdaBoost, Random Forest and a Neural Network. The highest performance is given by the RandomForest algorithm, in the case of the second scenario, detailed below, where the accuracy is 98.7%.

A database of radar signals is constructed in [18], corresponding to four scenarios: first scenario when there are 0 to 10 people randomly walking in a radar range of 5 meters, the second scenario where there are 11 to 20 persons with a density of 3 persons per squared meter, the third scenario where there are 11 to 20 persons with a density of 4 persons per squared meter and the last scenario where 0-15 persons are standing in a queue, at a distance of 10 cm one to another.

Table 2 Dense people counting scenarios

Scenario	Number of persons	Action
1	0 to 10	Randomly walking in the radar range
2	0 to 15	Standing in a queue with an average of 10 cm between persons
3	11 to 20	Randomly walking in the radar range, with a density of people of 3 persons per squared meter
4	11 to 20	Randomly walking in the radar range, with a density of people of 4 persons per squared meter

The entire dataset is composed of 319000x1280 samples, meaning 6380 radar samples. A radar sample consists of 50 received signals, where a single received signal is a vector of 1x1280 samples. One radar sample of 50 received signals corresponds to a range of 5 meters on the x-axis and 1.25 seconds on the y-axis.

3.2.2 Data preprocessing

The data are preprocessed to remove the noise. First, the direct current component is removed from all the signals, by extracting the mean of the signal's values, and then a digital bandpass filter is applied, with the cutoff frequencies of 5.65GHz and 7.95GHz. On the resulted filtered signals, the Running Average method is applied to remove the clutter in the radar sample.

The Running Average method implies a series of averages of subsets of data with a fixed length from the total dataset. The subsets are shifted forward, excluding the first values of the series and including the next ones. Running average method acts as a low-pass filter for the data, where data become smoother.

Suppose p_j data points from a dataset, where $j = \overline{1,n}$. The mean of the last k data points SMA_k is computed as:

$$SMA_k = \frac{1}{k}\sum_{i=n-k+1}^{n} p_i \tag{26}$$

The next mean $SMA_{k,next}$ has the same number of data points in the mean calculation and its based on the previous computed mean:

$$SMA_{k,next} = \frac{1}{k}\sum_{i=n-k+2}^{n+1} p_i = SMA_{k,prev} + \frac{1}{k}\left(p_{n+1} - p_{n-k+1}\right) \tag{27}$$

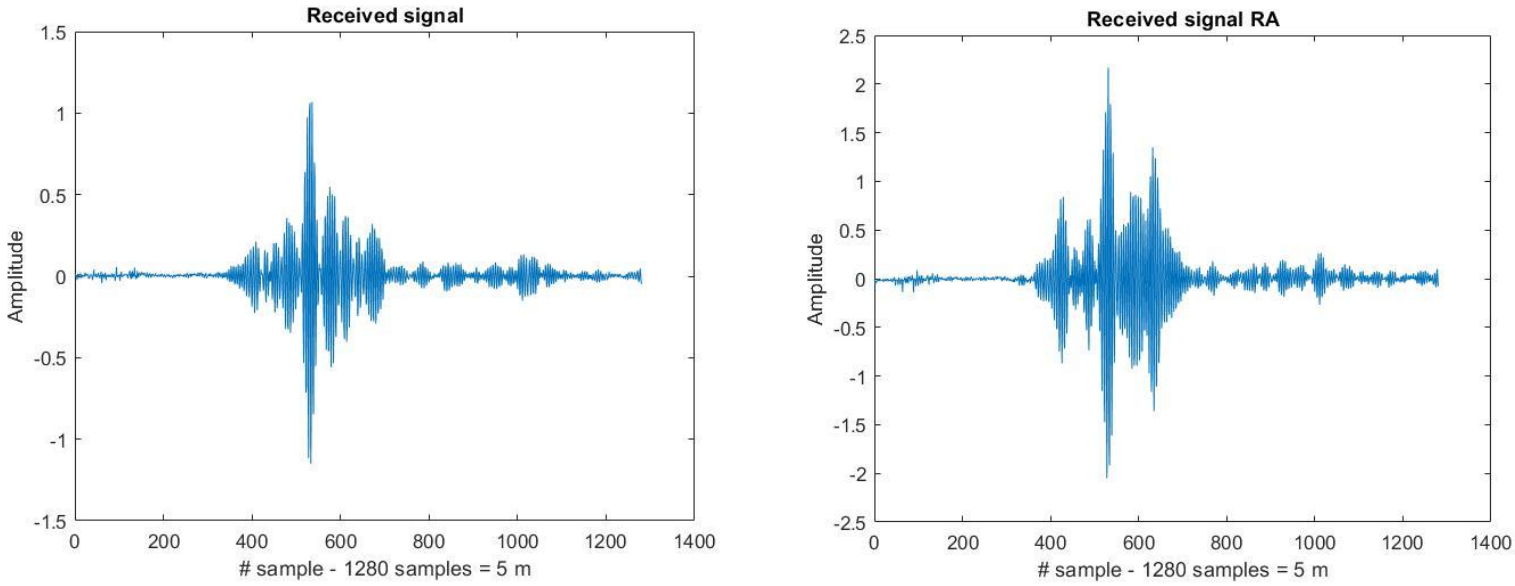
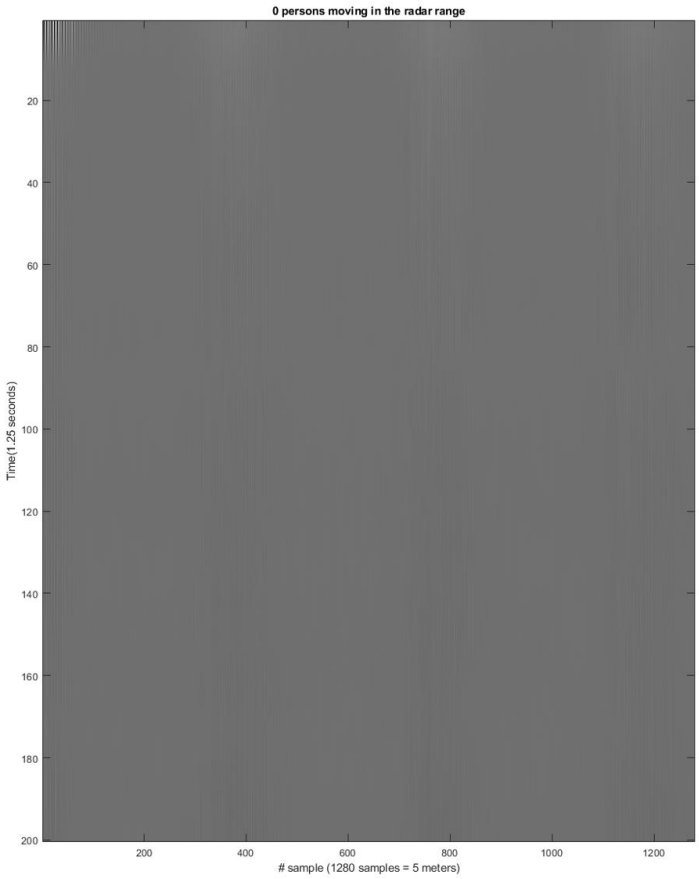


Figure 13 Received signal and clutter removed received signal

Taking as an example one signal, there can be seen that the data become smoother, the clutter is removed and the signal to noise ratio increases.

The figure below represents two cases for the first scenario: the image on the left represents a radar sample of 0 persons moving in the radar range, and the radar sample on the right represents the echo signal for 10 persons randomly walking in the radar range.



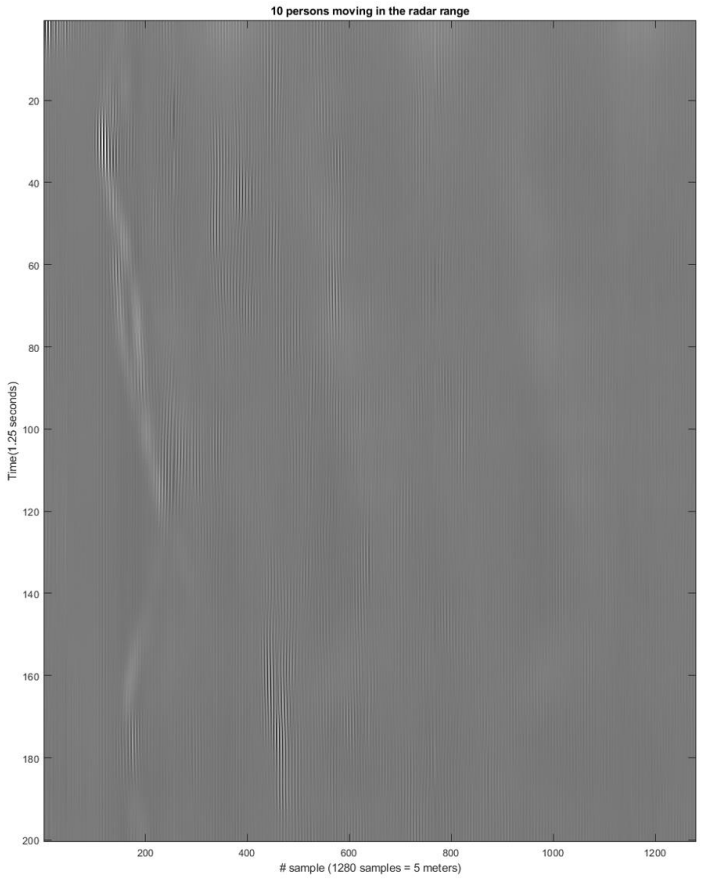


Figure 14 a) 0 persons moving in the radar range b) 10 persons moving in the radar range

The next figure represents two radar samples from the second scenario. The image on the left represents the echo signal for 5 persons standing in the queue, with 10 cm distance between them, and the right image represents the echo signal for 15 persons standing in the queue, with 10 cm distance between them.

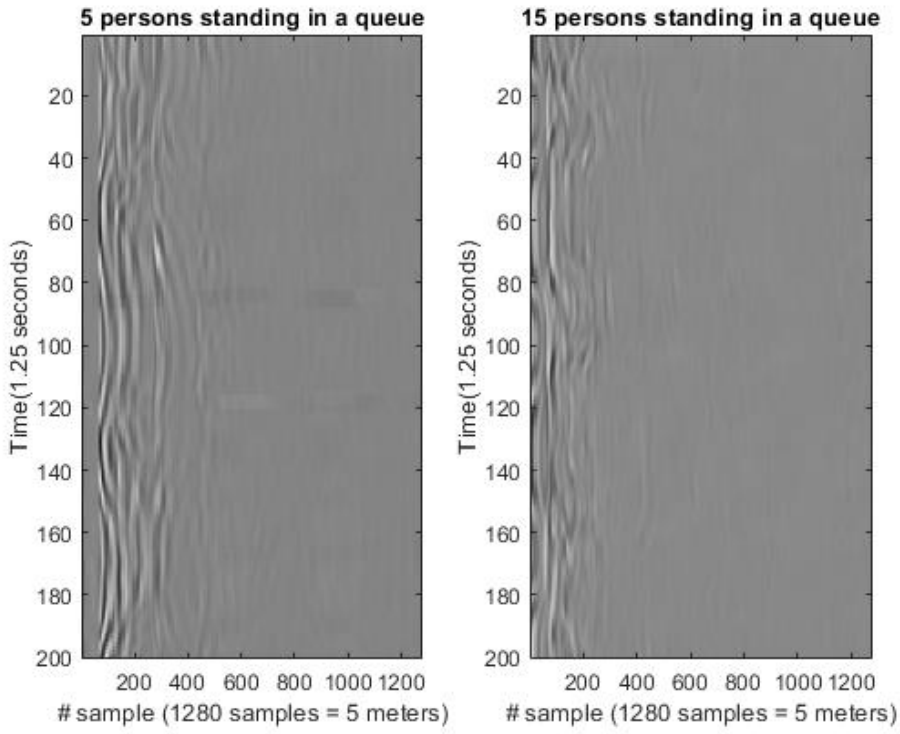


Figure 15 a) 5 persons standing in a queue b) 15 persons standing in a queue

The next figure represents two radar samples from the third scenario. The image on the left represents the echo signal for 11 persons walking in the radar range, with a density of three persons per square meter, the right image represents the echo signal for 20 persons walking in the radar range, with a density of three persons per squared meter.

The figures below represent two radar samples from the last scenario. The image in the left represents the echo signal for 11 persons walking in the radar range, with a density of four persons per square meter, the right image represents the echo signal for 20 persons walking in the radar range, with a density of four persons per squared meter.

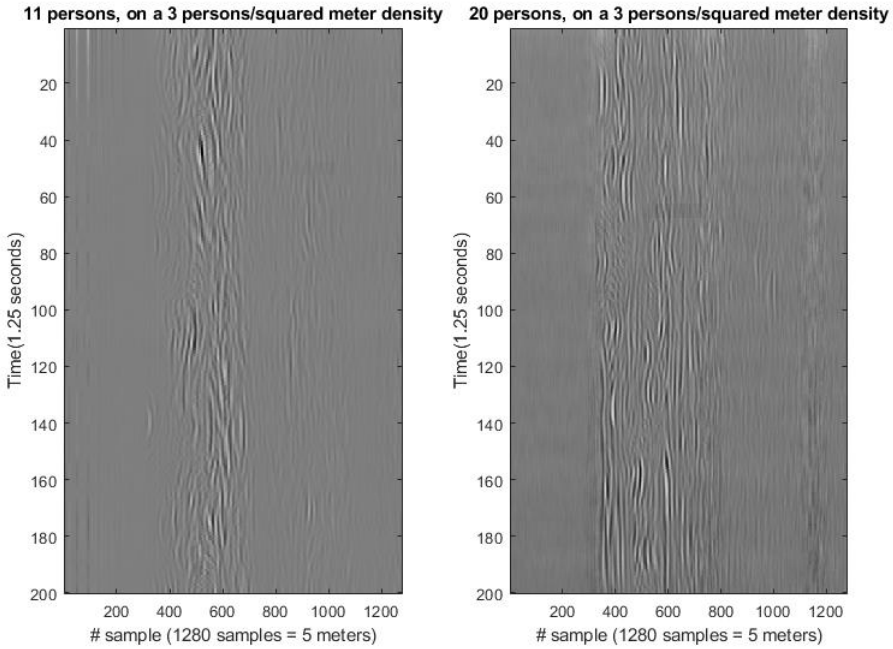


Figure 16 11 and 20 persons walking in the radar range, with a density of three persons per squared meter

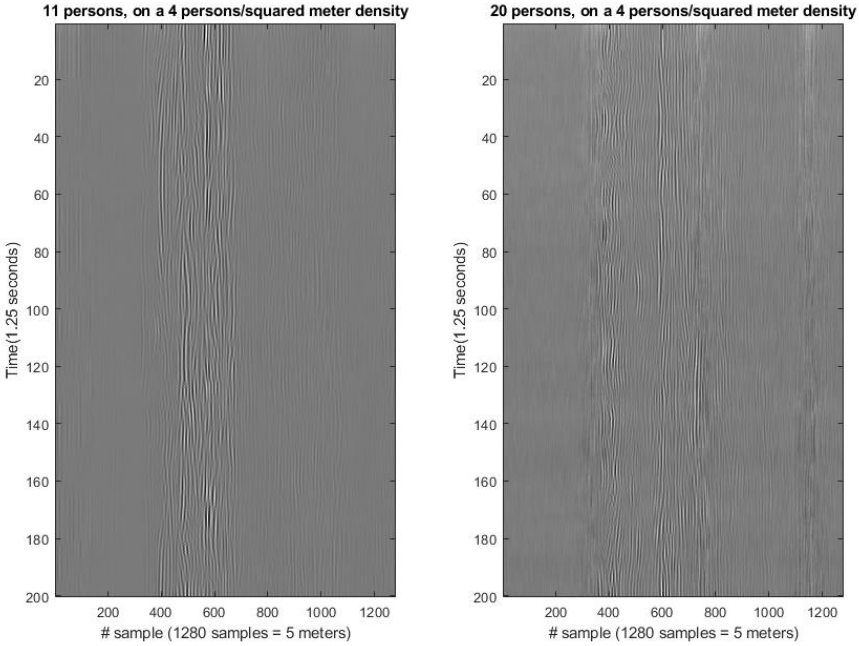


Figure 17 11 and 20 persons walking in the radar range, with a density of four persons per squared meter

3.2.3 Feature extraction

The Curvelet transform feature extraction implies Fast Discrete Curvelet Transforms computation, using the wrapping method [19], on the preprocessed radar samples. Energy and amplitude features are extracted from the coarse, detail and fine layers.

The segmented-based feature extraction method feature extraction method implies the segmentation of each received signal in fixed-length segments of 32, 64, 128 samples (corresponding to a 125, 250 and 500 mm distance). Energy and amplitude features are extracted from every segment, both from the preprocessed signals and the clutter removed signals.

3.2.3.1 The Curvelet transform

Curvelet transform represents a multi-scale orientation-selective analysis algorithm, first proposed by Candes and Donoho, and it is mostly used for image processing due to its capability of anisotropy, strong direction and efficiency of image edge representation [43].

The curvelet transform equation involves three parameters: the scale j, the direction l and the position k:

$$C(j,l,k) = \int \hat{f}(\omega)\widehat{U}_j(S_{\theta_l}^{-1}\omega)e^{i<S_{\theta_l}^{-T}b,\omega>}d\omega \quad (28)$$

Where the terms represent the following:

f the input Fourier samples of the data in the Cartesian coordinate system

$\widehat{U}_j$ the frequency rectangle window for each scale

S_{θ_l} the shear matrix with orientation θ_l, where $S_{\theta_l} = \begin{bmatrix} 1 & 0 \\ -tan\theta_l & 1 \end{bmatrix}$

$b \coloneqq (k_1 2^{-j}, k_2 2^{-j/2})$ where k_1, k_2 are the translation parameters taking values on a rectangular grid.

The curvelet coefficients are computed using the Inverse Fourier Transform of the multiplication of the Fast Fourier Transform of the curvelet function and the FFT of the original image. The convolution operation in the spatial domain corresponds to multiplication in the frequency domain.

$$C = IFFT\{FFT(Curvelet)\ FFT(Image)\} \tag{29}$$

Hard thresholding on energy is used in this case, and it implies energy thresholding, where the cumulative sum of the energy is computed to apply the established threshold at which the values below it would be dropped out. In our case, the curvelet coefficients whose energy is below 90% of the energy of the total coefficients, are dropped.

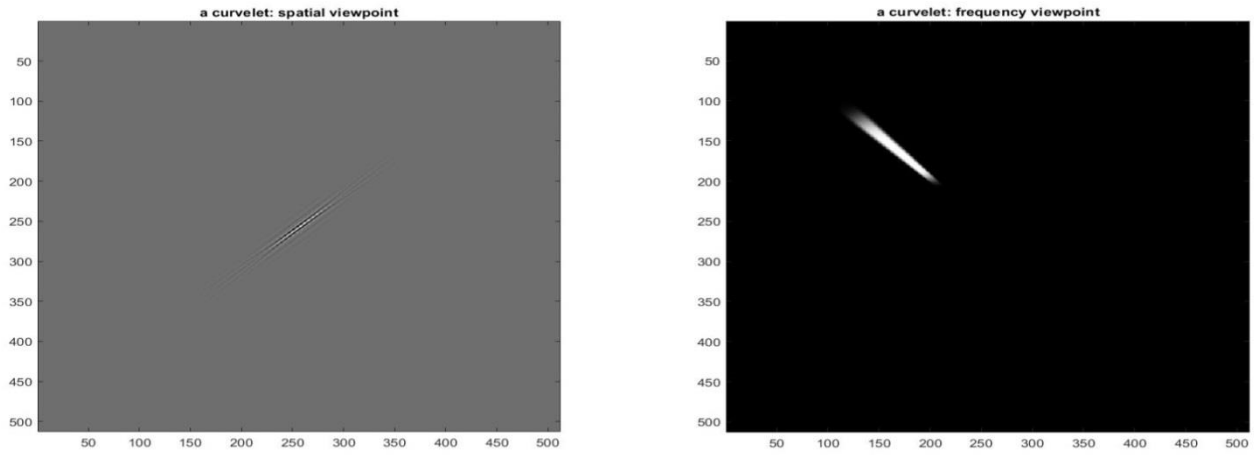


Figure 18 A curvelet in the spatial and frequency domain

The coefficients are divided into the finest layer, corresponding to the high frequencies, the detail layer and the coarse layer corresponding to low frequencies. The coarse layer provides general information about the input signal and the detail layer contains high-frequency information which provides the details [43].

Figure 19 Original image and Curvelet coefficients of the original image

The right image in Figure 27 shows the graphical representation of all the curvelet coefficients of the original image, at all scales, orientations and directions. The approximation coefficients are represented in the center of the image.

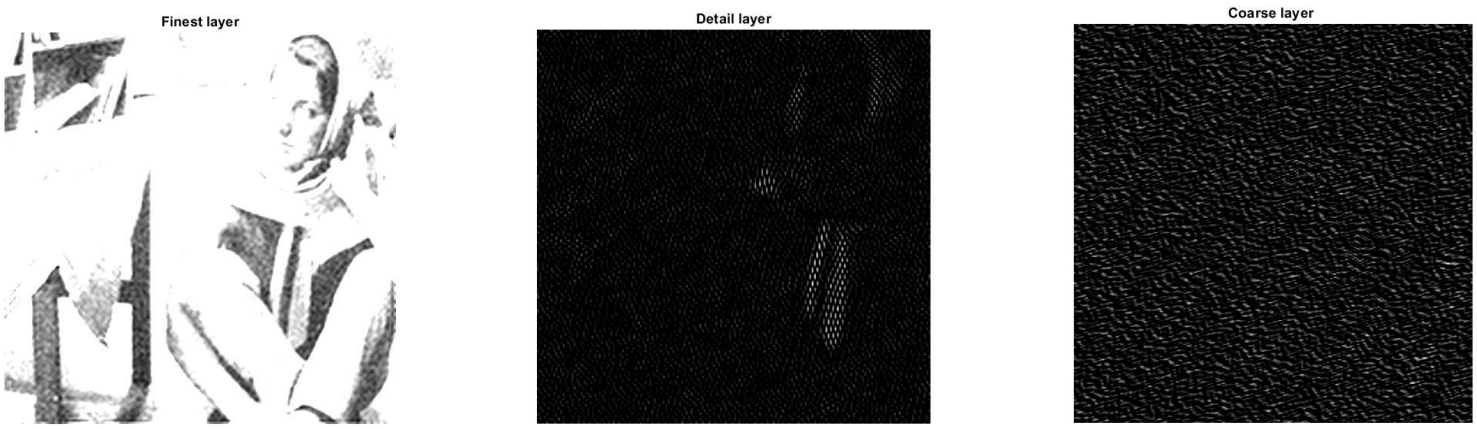


Figure 20. Coefficients on the finest, detail and coarse layer

Traditional wavelets have strong limitations concerning their effectiveness in a higher dimension than 1, because they rely on isotropic elements occurring at all scales and locations and do not describe the anisotropic elements [44]. The advantage of curvelet transform in front of wavelet transform is that it is more efficient when it comes to compression, de-noising, structure and edge extraction for features with line and surface singularities [45].

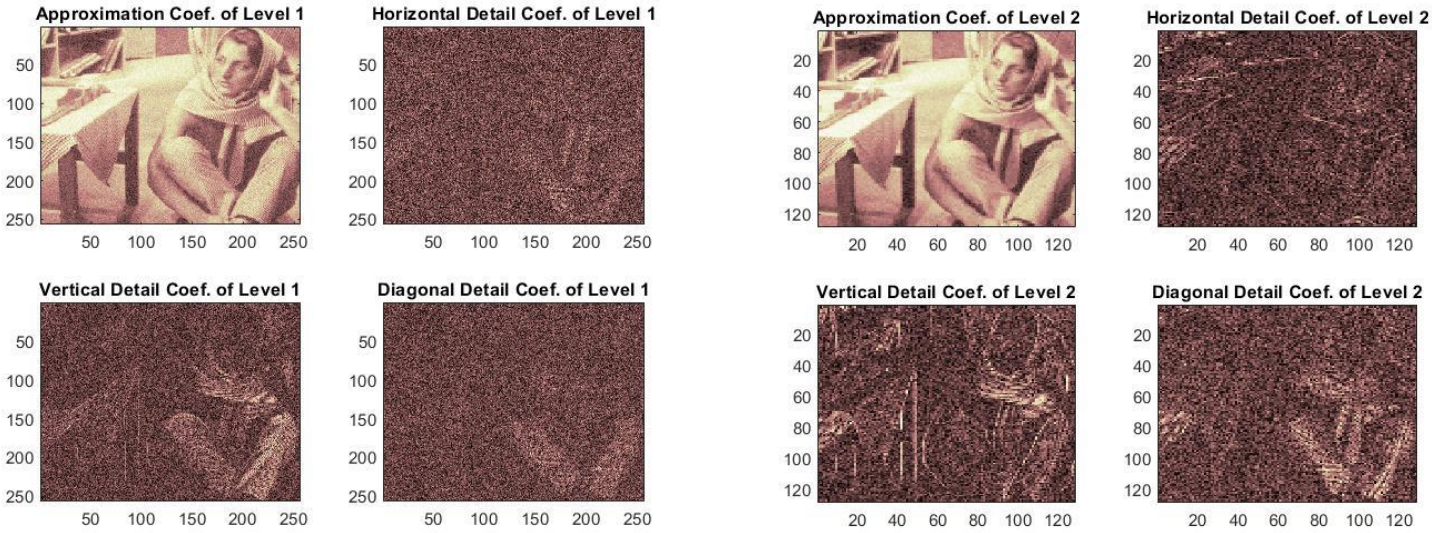


Figure 21 Coefficients of Level 1 and Level 2 using wavelet transform

Figure 22. Image reconstruction and the correspondent SNR

The original image is reconstructed and de-noised both using curvelet transform and wavelet decomposition. The signal to noise ratios in the reconstructed and de-noised images, 25.1316 for curvelet transform and 13.591 dB for wavelet transform show the effectiveness of the curvelet transform.

The FDCT via wrapping method implies the wrapping of the data, in the frequency domain, into a rectangle, initially inside a parallelogram $P_{j,l}$, which will be then centered at the origin [19].

First step is the 2D FFT computation of the data, obtaining the Fourier samples $\hat{f}[n_1, n_2]$, and for each scale j and angle l a parallelogram P_j is constructed, inside where the rectangle window $\overline{U_{j,l}}[n_1, n_2]$ with dimensions $L_{1,j} = 2^j, L_{2,j} = 2^{j/2}$ is supported [19].

$$P_j = \{(n_1, n_2): n_{1,0} \leq n_1 < n_{1,0} + L_{1,j}, n_{2,0} \leq n_2 < n_{2,0} + L_{2,j}\} \quad (30)$$

Where $(n_{1,0}, n_{2,0})$ is the index of the bottom-left pixel of the rectangle. The data $d[n_1, n_2]$ are windowed by the discrete localizing window $\overline{U_{j,l}}[n_1, n_2]$, where:

$$d[n_1, n_2] = \overline{U_{j,l}}[n_1, n_2]\, \hat{f}[n_1, n_2] \quad (31)$$

The product is then wrapped around the origin, where $0 \leq n_1 < L_{1,j}$ and $0 \leq n_2 < L_{2,j}$:

$$\bar{f}_{j,l}[n_1, n_2] = W\left(\bar{U}_{j,l}\hat{f}\right)[n_1, n_2] \quad (32)$$

The discrete coefficients will be collected by applying the inverse 2D Fourier transform, 2D IFFT, for each $\bar{f}_{j,l}$.

3.2.3.2 The segmented-based feature extraction method

The segmented-based feature extraction method implies the segmentation of each received signal into segments of length 32, 64 and 128 samples corresponding to a 125, 250 and 500 mm distance.

The segmented-based feature extraction method is applied both on the band passed input signals and on the reconstructed signal. The energy and the maximum amplitude are extracted as features from every segment.

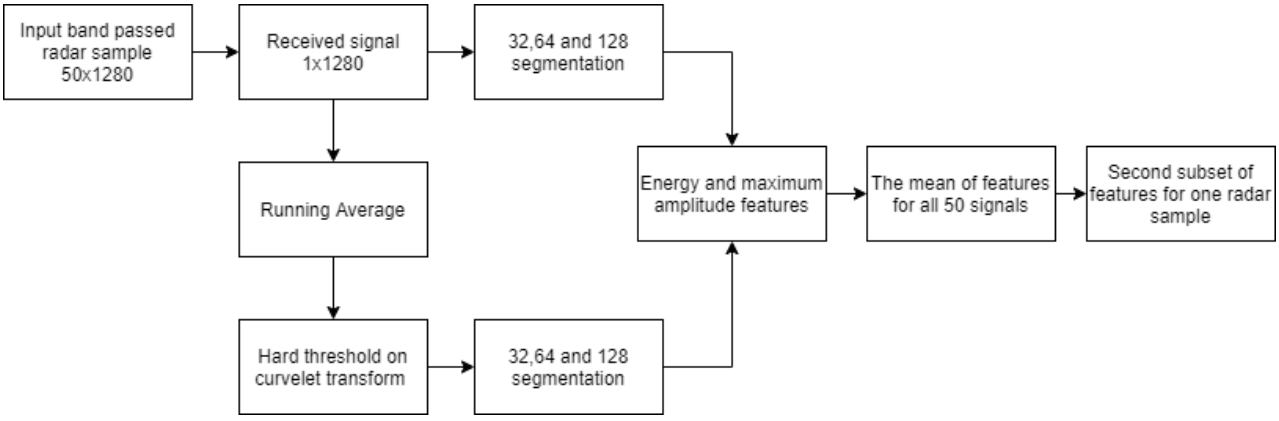


Figure 23 Block diagram the segmented-based feature extraction method

3.2.3.3 Principal Component Analysis (PCA)

The principal component analysis is a method used for dimension reduction of a high dimensional dataset by keeping the most important information from the data as a set of components called principal components, which are a linear combination of the original data [46]. PCA method aims to find directions along which the variation in the data is maximal and along which the data can be projected, with minimal loss of information [46].

PCA method could be used in data which are correlated (meaning high redundancy), thus the data can be reduced to new uncorrelated variables, named principal components. The purpose of this method is to reduce the data dimensionality by removing the noise and redundancy, finding hidden patterns and identifying correlations in the dataset [46].

The main advantages of this method are that it provides uncorrelated vectors as outputs (principal components) and that it reduces the dimensionality of the data, by keeping most of the energy, mainly compressed in the first principal component.

The Principal Component Analysis relies on the computation of the eigenvector and eigenvalues of the covariance matrix of the original data. The covariance matrix is a symmetric matrix that shows correlations in the data. The main diagonal of this matrix represents the variances of the variables in the data.

The first principal component is the line that maximizes the variance of the data (the mean of the squared distances from the projected points to the origin [47]). The second principal component will be orthogonal to the first one, and it has the next higher variance.

The eigenvectors of the covariance matrix give the direction of the axes which has the most variance and the eigenvalues give the amount of the information in each principal component [47]. Each eigenvector has an eigenvalue and the highest eigenvalue corresponds to the first principal component, and so on,

in descending order. The amount of information in each principal component is given by the division of the eigenvalue of the principal component and the sum of all eigenvalues.

Suppose the original data matrix X, the covariance matrix C_X is given by:

$$C_X = E[(X - E[X])(X - E[X])^T] \quad (33)$$

Where $E[\,]$ represents the expected value, or the mean value and T is the transpose operation of a matrix.

"If T is a linear transformation from a vector space V over a field F into itself and v is a nonzero vector in V, then v is an eigenvector of T if T(v) is a scalar multiple of v. This can be written as: $T(v) = \lambda v$, where λ is a scalar in F, known as the eigenvalue, characteristic value, or characteristic root associated with v." [48]

To construct the feature vector E (given by the highest eigenvalues of the covariance matrix) and to reduce the dimensionality of the data, a threshold is used for discarding the principal components that have the lowest amount of information (given by the corresponding lowest eigenvalues).

The final dataset Y is constructed based on the multiplication of the transpose of the original dataset X and the transpose of the feature vector.

$$Y = X^T E^T \quad (34)$$

In our case, Principal Component Analysis is applied to the feature matrix extracted from the original dataset of the radar data to reduce the dimensionality of the data, avoid over-fitting and improve the computational cost and speed of the training process of the classifiers.

Suppose X – the original matrix of the non-normalized features of size $(M\ x\ N)$. The extracted features are normalized using the amount of energy within it. For each subset of features X_i, $i = \overline{1,n}$, n the number of subsets of features, the autocorrelation matrix is computed:

$$R_{xx} = \frac{1}{N}(X_i{}^T X_i) \quad (35)$$

Then a diagonal matrix K_n is constructed, having as elements the squared root of the inverse of the sum of the main diagonal elements of the autocorrelation matrix (representing the energy in the signal). The final normalized dataset will be $Y_i = X_i K_i$.

The K_i matrixes are saved for future normalization of the new data's subsets of features. The final normalized dataset is constructed by putting together all the Y_i matrixes.

3.2.4 People counting based on Artificial Intelligence Algorithms

Machine Learning is a branch of Artificial Intelligence (AI). Machine Learning algorithms learn from the data and provide accurate outputs when it comes to new data, based on what it has learned from the training data. The data are split into training data, from which the algorithm learns, data and test data, which are used to test the algorithm performance and evaluation.

In AI, there could be supervised and unsupervised learning algorithms, which consist of models trained to find patterns and correlations in the data. Briefly, supervised learning algorithms learn from the training data having the given targets or labels, while unsupervised learning algorithms find patterns in the training data and provide output classes on their own.

In our case, we use supervised learning for people counting. The final features dataset and its corresponding matrix of labels from 0 to 20 (corresponding to the scenario and number of persons) are used as input for the classifiers, to provide the number of persons in the radar range. The dimension reduction for the initial dataset improves the computational cost and speed of the training process and avoids over-fitting.

In this chapter, four Artificial Intelligence algorithms (K-nearest neighbours, Support Vector Machine, Multilayer Perceptron Layer and Convolutional Neural Network) used for people counting are briefly presented.

3.2.4.1 K-nearest neighbours

K-nearest neighbours is a supervised machine learning algorithm used both for regression and classification. This algorithm is based on similarity (or distance) between surrounding points, where Euclidean distance is mostly used. The K number of "neighbours" is given by the user, the bigger the K, the smaller the effect of the noise, but the lesser distinctions between classes. The algorithm's output for new input data relies on the dominant class among its neighbours (samples in the training set).

Suppose two points $A(a,b), B(c,d)$, the Euclidean distance $d(A,B)$ from A to B is the following:

$$d(A,B) = \sqrt{(a-c)^2 - (b-d)^2} \tag{36}$$

3.2.4.2 Support vector machine (SVM)

Support Vector Machine is a supervised machine learning algorithm used for classification and regression, widely used in industry and science. The algorithm relies on splitting the data with hyperplanes and projects them into higher dimensions.

There are several approaches for SVM: Linear SVM, Nonlinear SVM and Kernel Methods. SVM aims to optimize the largest margin between the data [49].

The Linear SVM constructs a hyperplane: $w * x + b = 0$, where each hyperplane has a different value of w and the constant b, which optimally separates the labelled data [49]. Given a new data point x_j, it can be classified by computing the sign of $w * x_j + b$.

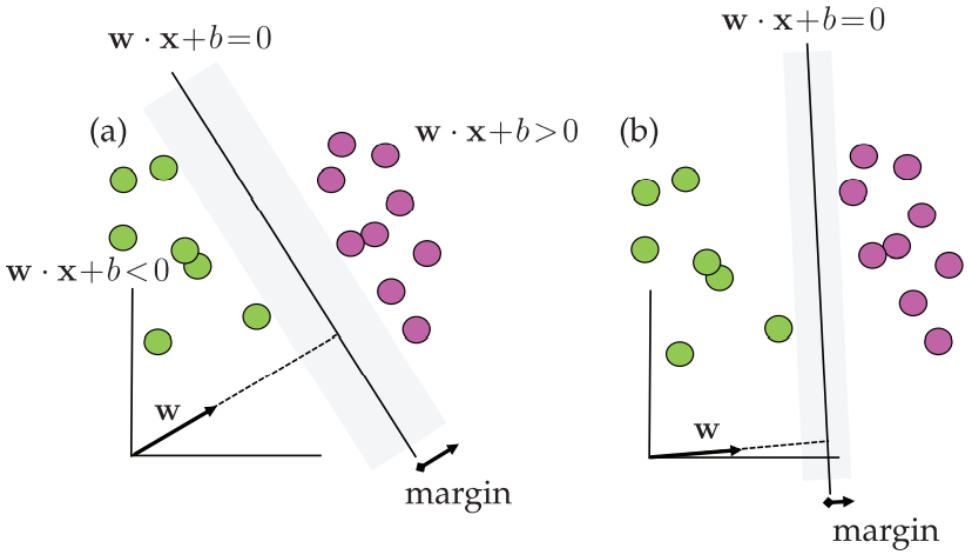


Figure 24 SVM Classification scheme [37]

The support vectors are given by the vectors which touch the edge of the grey regions.

The Nonlinear SVM is one of the most successful machine learning algorithms and it maps the data into a new, nonlinear, higher-dimensional space by including new enriched nonlinear features to build more sophisticated classification curves. The hyperplane function is given by :

$$f(x) = w\,\varphi(x) + b \tag{37}$$

Considering two data points $x = (x_1, x_2)$. The new space of data will be:

$$(x_1, x_2) = \ (z_1, z_2, z_3) \coloneqq (x_1, x_2, x_1^2 + x_2^2) \tag{38}$$

The new labelling function is given by $\overline{y_j} = sign(w\ \varphi(x) + b\ )$, where $\varphi(x)$ is the enriched space of the observed data [49].

To reduce the computation cost, kernel SVM methods are developed. The w vector is represented as $w = \sum_{j=1}^{m} \alpha_j \varphi(x_j)$, where α_j are the weights for the nonlinear observable functions $\varphi(x_j)$. The hyperplane function will be the following:

$$f(x) = \sum_{j=1}^{m} \alpha_j \varphi(x_j)\ \ \varphi(x) + b \tag{39}$$

The Kernel function is defined as:

$$K(x_j, x) = \varphi(x_j)\varphi(x) \tag{40}$$

The Kernel used in our case is the Radial basis function (RBF), defined as:

$$K(x_j, x) = \exp(-\gamma \|x_j - x\|^2) \tag{41}$$

Where γ is the width of the Gaussian kernel measuring the distance between data points and the classification line [34].

3.2.5. Multilayer Perceptron Neural Network

Multilayer Perceptron Neural Network is an algorithm which maps the input layer to the output layer through a function that is learned while the training process. Between the input and output layer, there could be more hidden layers consisting of neurons (weights) which are updated based on the mean squared error between the real output and the actual output. In Neural Networks, the mean squared error between the given targets and the actual outputs is optimized by backpropagation methods (example: Gradient Descent).

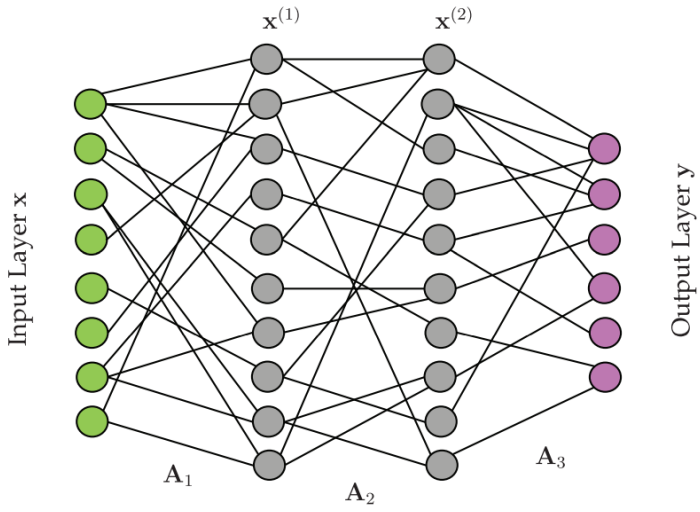


Figure 25. Two hidden layers of Neural Network Architecture [34]

In our case, the neural network has 4 layers of 50 neurons. The activation function is Rectified Linear Unit (ReLU). ReLU activation function is given as following:

$$f(x) = \max(0, x) \tag{42}$$

Where x is the neuron's input. The advantages of this activation function are that it has a sparse activation (reduces the computational cost because it makes some of the neurons inactive), better gradient propagation (for minimizing the

mean squared error between the output and the target output) and it is computationally efficient.

3.2.6. Convolutional Neural Network (CNN)

Convolutional Neural Networks are deep-learning algorithms that are inspired by the functioning of the human visual system and are commonly used for image classification problems. CNN architecture consists of convolutional layers, pooling layers, and fully connected layers.

The role of the convolutional layers is to extract features from the input data, which are tensors, meaning they have more than two dimensions. The results will be an activation map, or a feature map, which will be passed to the next layer. Commonly, the convolutional layers are followed by an activation function, which removes negative values from the feature map and set them to zero.

The pooling layer has the role of down-sampling the data, in order to reduce the spatial size of representation and the number of parameters. The most common pooling layer function is max pooling. The fully connected layers connect all the activations from the previous layer. The fully connected layer is usually followed by a logistic function such as SoftMax.

3.2.7 Metrics in model evaluation

To validate the model, some metrics are usually used in classification problems, such as accuracy, precision, recall, f1-score.

The accuracy is given by the ratio of correct predicted values and the total predicted values.

$$accuracy = \frac{nr.\ of\ correct\ predictions}{nr.\ of\ total\ predictions} \tag{43}$$

The precision is given by the ratio of true positive and the sum of true positive and false positive, where true positive refers to the case where the model correctly predicted the positive class and false positive refers to the case where the model incorrectly predicts the positive class.

$$precision = \frac{True\ positive}{True\ positive + false\ positive} \tag{44}$$

Recall function is given by the ratio of true positive and the sum of true positive and false negative, where false negative refers to the case where the model incorrectly predicts the negative class.

$$recall = \frac{True\ positive}{True\ positive + false\ negative} \tag{45}$$

F1- score function combines both the precision and recall function as follows:

$$f1score = \frac{2\ precison * recall}{precision + recall} \tag{46}$$

4.2 Results and discussion for people detection and counting

By applying the algorithm for the hybrid features, frequency-based and time-based, on the dataset of radar samples, a set of 293 features are resulted. 280 features are extracted using the segmented-based feature extraction method, which include the energy and the maximum amplitude from the segmented signals and 13 features using the frequency-based method, which consist of the curvelet coefficients. The total feature vector for each radar sample of 50 received signals consists of 293 features, a smaller number of features than used in the reference paper.

The FDCT via wrapping method is applied to the preprocessed dataset of radar samples without clutter removal, to fully extract the statistical features. The data matrix is decomposed into three layers: fine, detail and coarse layer. The coarse layer contains the low-frequency coefficients and the fine layer contains the high-frequency coefficients.

Energy and the mean value are extracted as features from the coarse layer, which generally describe the data.

The finer edge information is described by the fine layer, from which the first 5 peaks of high-frequency coefficient values and the fine layer energy are extracted as features.

The detail layer contains coefficients corresponding to different directions in the trajectories of people. The coefficients are divided into 16 directions. According to the paper [18], the people who move away from the radar are represented by the coefficients in panels 1,16,8 and 9, in a 45° direction. People who move towards the radar are represented by the 135° direction, meaning the panels 4,5,12 and 13. The persons standing still in the radar range are represented by the coefficients on the 90° directions, meaning the panels 6,7,14 and 15.

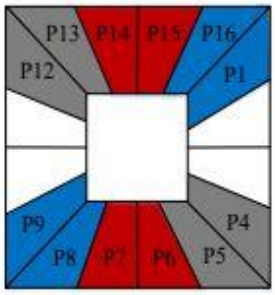


Figure 36. Index panel of directions from the detail layer [29]

The energy of the $90^{\circ}, 45^{\circ}, 135^{\circ}$ directions and high-frequency coefficients are extracted as features. The advantage of this feature extraction method is that more detailed features are provided, due to its anisotropic decomposition, providing detailed edge information.

The segmented-based features are extracted from each received signal of 1x1280 samples, thus for a radar sample of 50 x1280 values, the total vector of features will represent the mean of the extracted segmented-based feature extraction method features for each 1x1280 signal. For every received signal, 80 features concerning the energy and amplitude are extracted from the 32 length segment, 40 from the 64 length segment and 20 from the 128 length segment, both for band passed signals and for the signals without clutter. The advantage of this method is that it discriminates each person from the superposed signals, in a dense scenario, by choosing the length of each segment of 32, 64, or 128 samples, corresponding to physical parameters. Also, the segmented-based feature extraction method features are extracted from the band passed signals and the signals without noise and clutter, providing more detailed information.

The databased contains 6380 radar samples from which 293 features are extracted. The dataset was split in two parts: the training set and the test set. The training set was used for training the AI models, while the test set is used for model validation.

The results are given in the table below, in terms of accuracy, recall, f1-score, precision and cross-validation. We can observe that the greatest results are provided by Multilayer Perceptron Neural Network, followed by the K-nearest

neighbours algorithm. It is worth mentioning that KNN methods uses 5 neighbours, SVM uses the RBF kernel and MLP has 4 fully-connected layers of 50 neurons each.

Table 3. The classification performance for the 6380x293 set of features

Algorithm	Accuracy	Precision	Recall	F1 score
K-NN	94.08%	94.08%	94.08%	94.00%
SVM	85.12%	85.12%	85.12%	82.76%
MLP	99.85%	99.85%	99.85%	99.85%

The second approach is based on using a Convolutional Neural Network on the raw set of data. It is worth mentioning that the CNN was implemented using Pytorch, where a number of 100 000 parameters are constant updated, using 50 epochs (the iteration over the entire dataset). The purpose of the network is to minimize the error between the predicted data and the label, provided by a loss function, which is encoded using the one-hot encoding method. Based on the error, an optimizer is used to update the weights (the parameters). The optimizer used in the network is Adam optimizer, where the learning rate is adaptive. The loss function which is used is the Mean Squared Error which uses the L2 norm to compute the error between the predicted data and the target data.

The CNN is applied on the entire raw set of radar samples, and the results are provided in the Table 4.

Table 4. The performance of CNN for the set of raw data

Algorithm	Accuracy	Precision	Recall	F1 score
CNN	95.445 %	95.445 %	95.445 %	95.445 %

The architecture of the network is given in the figure below, whereas input it takes the radar sample of input 1x50x1280 samples:

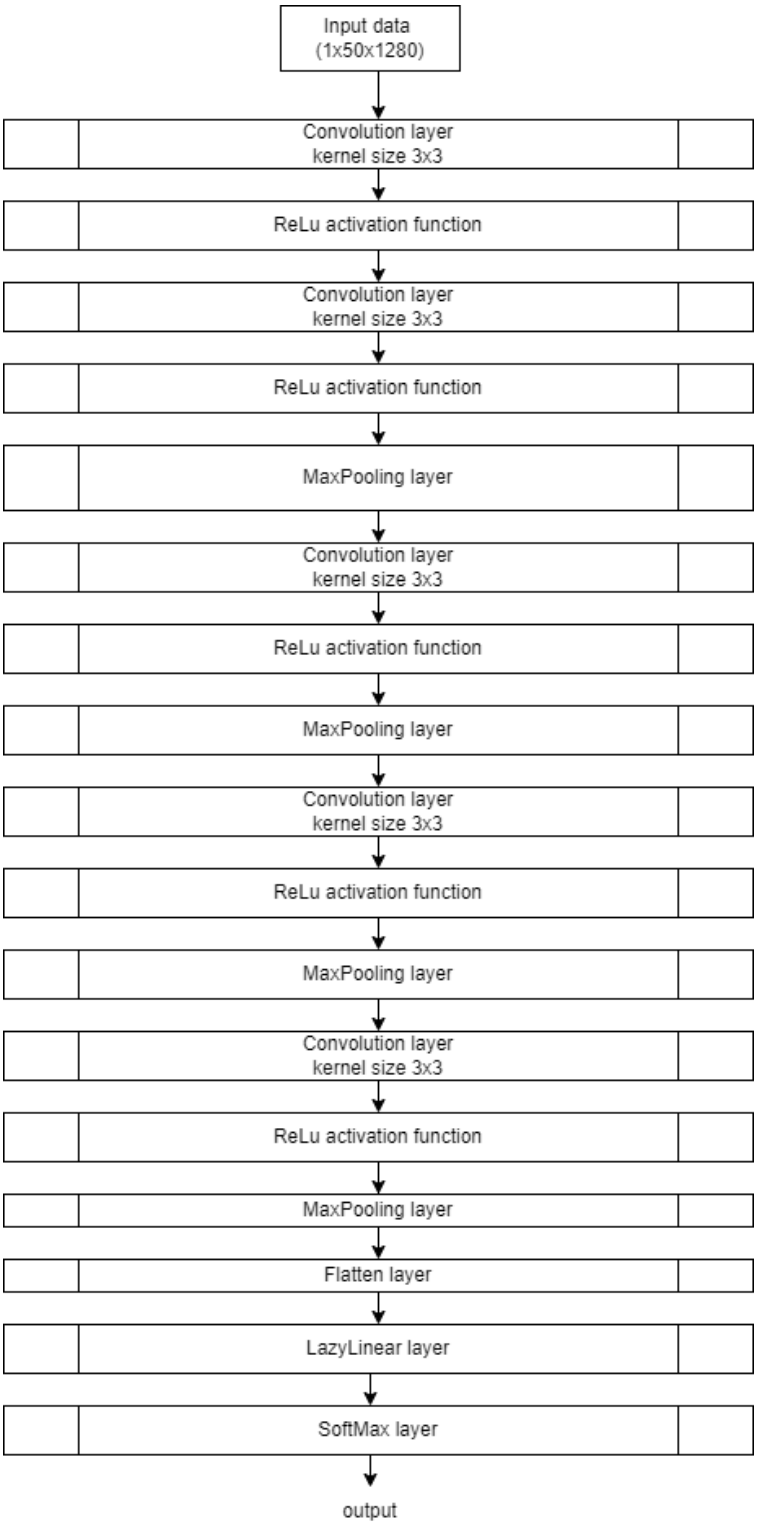


Figure 37. CNN architecture

80 din

The third approach proposes the Principal Component Analysis method for data compression, where the eigenvalues and eigenvectors of the covariance matrix of the dataset of 293 the already extracted hybrid features, are computed and sorted in descending order. A threshold of 0.95 energy percentage of the largest selected eigenvalues of the data covariance matrix C_Y is used to keep the most important principal components, to reduce the dimensions of the data.

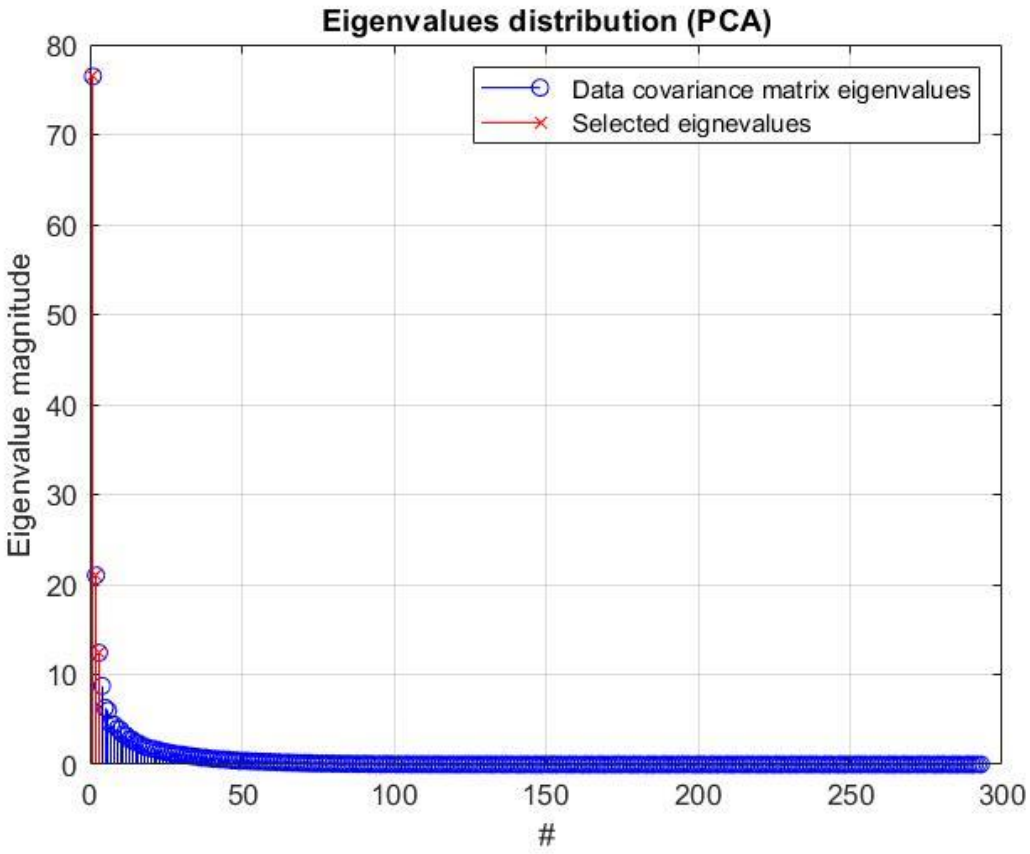


Figure 38. Eigenvalue distribution

By applying the threshold on the eigenvalues, only three of them are kept, corresponding to the three directions having the highest variance (amount of information), meaning that the new dataset has three dimensions in the projection space. The threshold influences the dimension of the new dataset and the loss of information.

The features extracted from the radar samples correspond to a fixed number of people in the radar range. PCA reduces the number of features, providing linear combinations of the initial ones, but the new variables correspond to the same specific number of persons in the radar range as the old features. For some random number of people, the data are visualized in three dimensions in the projection space, where N_{num} represents the number of persons:

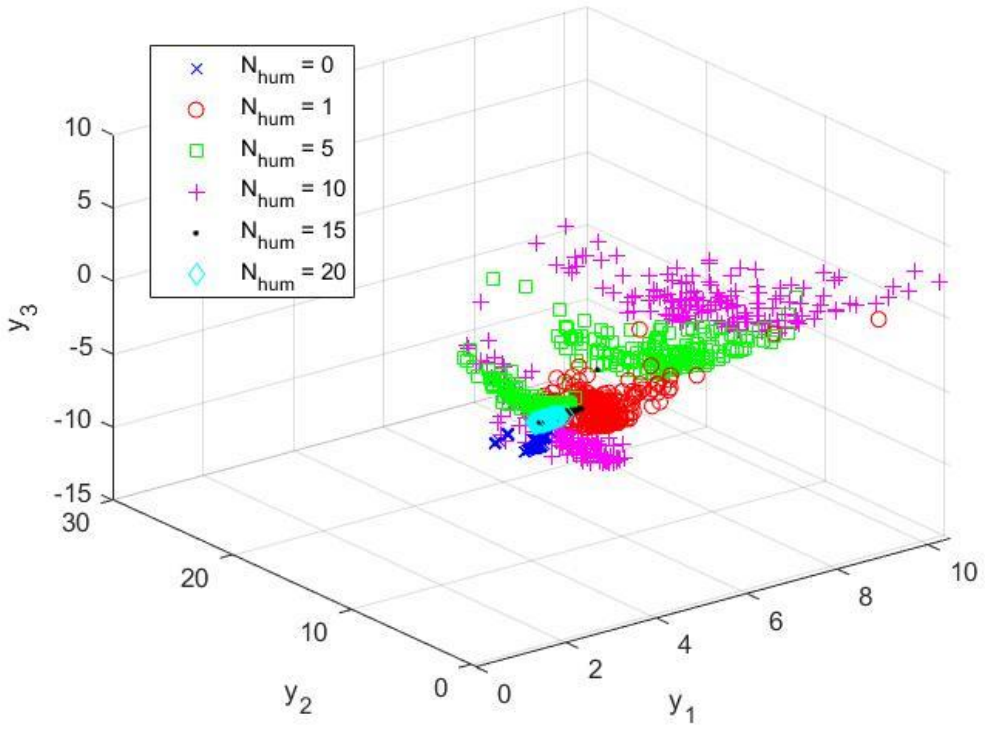


Figure 39 The projection space

The first principal component captures 36% of the information from the data matrix, while the second principal component captures 10% and the last 6% of the total energy (this percentage is given by the division of the corresponding eigenvalue to the sum of all eigenvalues).

The classification is done on the set of 6380x3 set of features, where the results are presented in table 5. We can observe that the highest performance is given by the K-nearest neighbours method, followed by MLP, which means that the architecture of MLP is not suitable for this dataset.

Table 5 The classification performance for the 6380x3 set of features

Algorithm	Accuracy	Precision	Recall	F1 score
K-NN	73.66%	73.66%	73.66%	73.56%
SVM	55.74%	55.74%	55.74%	53.43%
MLP	64.31%	64.31%	64.31%	62.90%

The PCA method is also applied on the entire set of data, where 6 principal components are extracted. CNN with only two layers of convolution is applied on the resulting dataset and the results are given in the table 6. It is worth mentioning that the network used almost 800 parameters to be updated, with the same hyper-parameters as the CNN applied on the entire raw set of data. The CNN architecture in this case is given as follows:

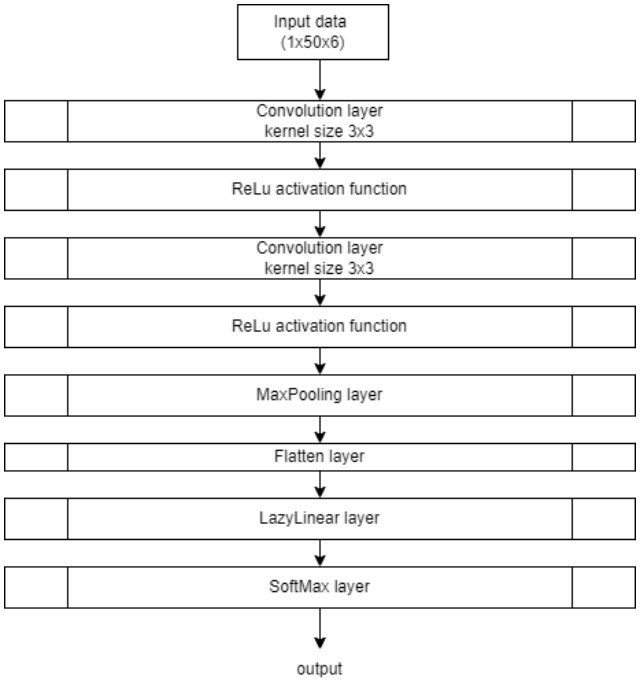


Figure 40. CNN architecture

Table 6 The performance of CNN for the set of PCA on the raw dataset

Algorithm	Accuracy	Precision	Recall	F1 score
CNN	95.41%	95.41%	95.41%	95.41%

Conclusions

In conclusion, this paper presents some signal processing methods for ultra-wideband signals such as Envelope Detection using Hilbert Transform, Power Spectral Density Estimation via Welch's method, Baseband conversion, Matched Filter and Search Subtract and Readjust algorithm and an application for UWB signals, which is People Detection and Counting using an IR-UWB radar, based on Artificial Intelligence algorithms.

On the experimental signals, the signal's period is determined by using Envelope Detection using Hilbert Transform, the signal's bandwidth is determined by using Power Spectral Density Estimation via Welch's method, following by signal filtration, baseband conversion and multipath components extraction by using SSR algorithm. The SSR method extracts the multipath components from the received signal and iteratively maximizes the signal to noise ratio and improves the time resolution in the signal, compared to Matched Filter, which maximizes the signal to noise ratio just once. It is observed that there are three multipath components on a signal period, from which two are superposed, having a small-scale fading, due to the multipath propagation. Also, the strongest path is not always the first ray that arrives at the receiver and the later the rays arrive, the lower power they have.

The application proposed for UWB signals is People Detection and Counting using an IR-UWB radar, based on Artificial Intelligence algorithms, where the aim is to compare three approaches for feature extraction.

The three approaches refers to the set of features extracted from the database, which first proposes a complex algorithm composed of the Curvelet Transform and the Segmented-based feature extraction methods for frequency-based and time-based feature extraction, respectively, resulting in a set of 293 features. The second method proposes a Convolutional Neural Network for feature extraction and the third proposes the Principal Component Analysis method for data compression.

In real-time systems, the model provided by the CNN applied on the set of 6 components extracted by PCA is recommended, due to its effectiveness in terms of speed and computational demand. CNN applied on the entire set of data provides good results and it's more computationally expensive than the first method, where it requires the training of 100 000 parameters instead of 800 parameters as with the use of PCA. The algorithm used for extraction of the hybrid features provides great results, but it's not computationally effective and the complexity is higher. Data compression provided by PCA provides fast training and testing, but the data insufficiency can degrade the performance, if not a great dataset is used – the case of the dataset consisted of 3 features.

UWB technology is suitable for people detection and counting applications due to its fine temporal resolution and artificial intelligence algorithms provide high performance in terms of classification.

Bibliography

[1] T. Kivinen, "datatracker.ietf.org," Internet Engineering Task Force (IETF), may 2017. [Online]. Available: https://datatracker.ietf.org/doc/html/rfc8137. [Accessed 15 06 2021].

[2] R. Yusnita, N. Razali, A. Tharek and H. P.S, "Ultra Wideband Technology and Its Applications," in *Wireless Communication Centre (WCC), Faculty of Electrical Engineering, Universiti Teknologi Malaysia 81310 Johor Bahru,Malaysia*, Department of Electronic, Electrical and Computer Engineering University of Birmingham Edgbaston Birmingham, B15 2TT United Kingdom.

[3] T. C. a. G. P. M. I. Marco Cavallaro, "Gaussian Pulse Generator for Millimeter-Wave," *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS, JUNE ,* Vols. NO. 6,VOL. 57, 2010.

[4] F. C. Commission, ""Revision of Part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission Systems"," Fist Report and Order, ET Docket 98-153, FCC 02-48,April, 2002.

[5] I. Opperman, M. Hamalainen and J. Iinatti, UWB Theory and Applications, 2004.

[6] M. Andreas, "Ultra-Wide-Band Propagation Channels," in *Proceedings of IEEE*, 2009.

[7] A. Molisch, D. Cassioli, C.-C. Chong, S. Emami, A. Fort, B. Kannan, J. Karedak, J. Kunisch, H. Schantz, K. Siwiak and M. Win, "A Comprehensive Standardized Model for Ultrawideband Propagation Channels," *IEEE Transactions on Antennas and Propagation,* Vols. 54, no.11, 2006.

[8] V. C. Chen, F. Li, S.-S. Ho and H. Wechsler, "Micro-Doppler Effect in Radar: Phenomenon, Model and Simulation Study," *IEEE Transanctions on aerospace and electronic systems,* Vols. 42 , no. 1, 2006.

[9] J. Kwon, S. Lee and N. Kwak, "Human Detection by Deep Neural Networks Recognizing Micro-Doppler Signals of Radar," in *Proceedings of the 15th European Radar Conference* , 2018.

[10] V.-H. Nguyen and J.-Y. Pyun, "Location Detection and Tracking of Moving Targets by a 2D IR-UWB Radar System," *Sensors,* 2015.

[11] J.-H. Choi, J.-E. Kim and K.-T. Kim, "People Counting Using IR-UWB Radar Sensor in a Wide Area," *IEEE,* pp. 2327 - 4662, 2020.

[12] F. Diaconescu, "Impulse Radio UWB blind detectiong using Cross Recurrence Plot," in *2020 13th International Conference on Communications (COMM)*, Bucharest, Romania, 2020.

[13] S. Lim, J. Jung and S.-C. Kim, "Deep Neural Network-Based In-Vehicle People Localization Using Ultra-Wideband Radar," *IEEE ,* vol. 8, 2020.

[14] K. Youngwook and M. Taesup, "Human Detection and Activity Classification Based on Micro-Doppler Signatures Using Deep Convolutional Neural Networks," *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS,* Vols. 13, NO. 1, January 2016.

[15] R. Qi, X. Li, Y. Zhang and Y. Li, "Multi-classification Algorithm for Human Motion Recognition Based on IR-UWB Radar," *IEEE Sensors Journal,* vol. 14, 2015.

[16] S. V. B. Joao, A. Zimmer and T. Brandmeier, "Pedestrian recognition using micro Doppler effects of radar signals based on machine learning and multi-objective optimization," *Expert Systems With Applications,* 2019.

[17] Y. Kim and H. Ling, "Human Activity Classification Based on Micro-Doppler Signatures Using a Support Vector Machine," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 47, 2009.

[18] X. Yang, W. Yin, L. Li and L. Zhang, "Dense People Counting Using IR-UWB Radar with a Hybrid Feature Extraction Method," *IEEE Geoscience and Remote Sensing Letters,* vol. 16, no. 1, pp. 30-34, 2018.

[19] E. Candes , L. Demanet, D. Donoho and L. Ying, "Fast Discrete Curvelet Transforms," March 2006.

[20] M. Hozhabri, Human Detection and Tracking with UWB Radar, Stockholm: E-Print AB, 2019.

[21] C. SangHyun, M. Naoki and B. Joel, "An algorithm for UWB radar-based human detection," in *IEEE International Conference on Ultra-Wideband (ICUWB)*, Sydney, NSW, Australia doi:10.1109/ICUWB.2013.6663820, 2013.

[22] I. Oppermann, M. Hamalainen and J. Iinatti, UWB Theory and Applications, 2004.

[23] K. Yakup, W. Henk, M. Arjan, B. J. Mark and S. G. William, "An Experimental Study of UWB Device-Free Person Detection and Ranging," in *IEEE International Conference on Ultra-Wideband (ICUWB)*, Sydney, NSW, Australia doi:10.1109/ICUWB.2013.6663820, 2013.

[24] N. Van-Han and P. Jae-Young, "Location Detection and Tracking of Moving Targets by a 2D IR-UWB Radar System," *Sensors,* 2015.

[25] B. Schleicher and H. Schumacher, Impulse Generator Targeting the European UWB Mask, 2010.

[26] "A Summary of Worldwide Telecommunications Regulations governing the use of Ultra-Wideband radio," Decawave APPLICATION NOTE: APR001, 2015.

[27] I. Oppermann, M. Hamalainen and J. Iinatti, UWB Theory and Applications, 2004.

[28] L. Jing and Z. Zhaofa, "Through wall detection of human being's movement by UWB radar," in *The Society of Exploration Geophysicists to Chinese Geophysical Society*, Beijing, China, 2011.

[29] Q. Rui, L. Xiuping, Z. Yi and L. Yubing, "Multi-classification Algorithm for Human Recognition Based on IR-UWB Radar," *IEEE Sensors Journal,* Vols. 14, no. 8 , 2015.

[30] E. Anwar , S. Timothy, W. Safwan and X. Tian, "Machine Learning for Respiratory Detection Via UWB Radar Sensor".

[31] C. C. Victor , L. Fayin, H. Shen-Shyang and W. Harry, "Micro-Doppler Effect in Radar: Phenomenon, Model, and Simulation Study," *IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS,* Vols. 42,NO.1, 2006.

[32] A. Molisch, "Ultra-Wide-Band Propagation Channels," *Proceedings of the IEEE,* vol. 97, no. 2 DOI:10.1109/JPROC.2008.2008836, 2009.

[33] A. Molisch, "Ultra-Wide-Band Propagation Channels," in *Proceedings of the IEEE*, Vol.97 no. 2 2009.

[34] A. Molisch, D. Cassioli, C.-C. Chong, S. Emami, A. Fort, B. Kannan, J. Karedal, J. Kunisch, H. Schantz, K. Siwiak and M. Win, "A Comprehensive Standardized Model for Ultrawideband Propagation Channels," *IEEE Transactions on Antennas and Propagation,* vol. 54 no. 11, 2006.

[35] Z. Cui, Y. Gao, J. Hu, S. Tian and J. Cheng, "LOS/NLOS Identification for Indoor UWB Positioning Based on Morlet Wavelet Transform and Convolutional Neural Network," *IEEE Communications Letters,* 2020.

[36] F. Grejtak and A. Prokes, "UWB - ULTRA WIDEBAND CHARACTERISTICS AND THE SALEH-VALENZUELA," *Acta Electrotechnica et Informatica,* vol. 13 no. 2, p. 32–38, 2013.

[37] T. Ulrich, "Envelope Calculation from teh Hilbert Transform," March 2006.

[38] E.-T. Lee and H.-C. Eun, "Structural Damage Detection by Power Spectral Density Estimation Using Output-Only Measurement," *Hinwawi , Shock and Vibration,* 2016.

[39] J. C. Bancroft, "Introduction to matched filters," CREWES Research Rerport Volume 14, 2002.

[40] C. Falsi, D. Dardari, L. Mucchi and M. Win, "Time of Arrival Estimation for UWB Localizers in Realistic Environments," *EURASIP Journal on Applied Signal Processing,* vol. 2006, pp. 1-13.

[41] K. Jihoon, L. Seungeui and K. Nojun, "Human Detection by Deep Neural Networks Recognizing Micro-Doppler Signals of Radar," in *Proceedings of the 15th European Radar Conference*.

[42] D. Kocur, J. Gamec, M. Svecova, M. Gamcova and J. Rovnakova, "Imaginig Method: An efficient Algortihm for Moving Target Tracking by UWB Radar," *Acta Polytech.,* 2010.

[43] Z.-y. Zhang, X.-d. Zhang, H.-y. Yu and X.-h. Pan, "Noise suppression based on a fast discrete curvelet transform," *Journal of Geophysics and Engineering,* vol. 7, pp. 105-112, 2010.

[44] J. Fadili and J.-L. Starck, "Curvelets and Ridgelets," *R.A. Meyers, ed. Encyclopedia of Complexity and Systems Science ,* vol. 14, pp. 1718-1738, 2009.

[45] J. Ma and G. Plonka, "The Curvelet Transform," *IEEE Signal Processing Magazine,* 2010.

[46] Kassambara, "www.sthda.com," 23 09 2017. [Online]. Available: http://www.sthda.com/english/articles/31-principal-component-methods-in-r-practical-guide/112-pca-principal-component-analysis-essentials/. [Accessed 09 04 2021].

[47] Z. Jaadi, "https://builtin.com," 01 04 2021. [Online]. Available: https://builtin.com/data-science/step-step-explanation-principal-component-analysis. [Accessed 12 04 2021].

[48] "www.wikipedia.com," Wikipedia, 09 04 2021. [Online]. Available: https://en.wikipedia.org/wiki/Eigenvalues_and_eigenvectors. [Accessed 12 04 2021].

[49] S. Brunton and N. Kutz, Data-Driven Science and Engineering, Cambridge University Press,DOI: 10.1017/9781108380690, 2019.

[50] B. Schleicher and H. Schumacher , Impulse Generator Targeting the European UWB Mask, 2010.

[51] X. Guanlei, W. Xiaotong and X. Xiaogang, "Generalized Hilber transform and its properties in 2D LCT domain," *Signal Processing 89,* pp. 1395-1402, 2009.

[52] "www.wikipedia.com," 20 03 2021. [Online]. Available: https://en.wikipedia.org/wiki/Decision_tree. [Accessed 23 04 2021].

[53] B. Schleicher and H. Schumacher, Impulse Generator Targeting the European UWB Mask, 2010.

[54] F. Grejtak and A. Prokes, "UWB - ULTRA WIDEDBAND CHARACTERISTICS AND THE SALEH-VALENZUELA," *Acta Electrotehnica et Informatica,* vol. 13 . no. 2, pp. 32-38, 2013.