People counting using IR-UWB radar sensors and machine learning techniques

Ange Joel Nounga Njanda, Jocelyn Edinio Zacko Gbadoubissa, Emanuel Radoi, Ado Adamou Abba Ari, Roua Youssef, Aminou Halidou

Systems and Soft Computing 6 (2024) 200095

Contents lists available at ScienceDirect

Systems and Soft Computing

journal homepage: www.elsevier.com/locate/soft-computing-letters

People counting using IR-UWB radar sensors and machine learning techniques

Ange Joel Nounga Njanda [a], Jocelyn Edinio Zacko Gbadoubissa [a], Emanuel Radoi [b,*], Ado Adamou Abba Ari [c,d,e], Roua Youssef [b], Aminou Halidou [f,g]

[a] African Institute for Mathematical Sciences, Crystal Garden, 608, Limbé, Cameroon
[b] Univ Brest, Lab-STICC, UMR CNRS 6285, CS 93837, 6 Avenue Le Gorgeu, 29238, Brest Cedex 3, France
[c] LaRI Lab, University of Maroua, 814, Maroua, Cameroon
[d] DAVID Lab, Universite Paris Saclay, University of Versailles Saint-Quentin-en-Yvelines, 45 Avenue États-Unis, 78035, Versailles cedex, France
[e] CREATIVE, Institute of Fine Arts and Innovation, University of Garoua, 346, Garoua, Cameroon
[f] Department of Computer Science, Faculty of Science, University of Yaoundé I, 337, Yaoundé, Cameroon
[g] Department of Mechanical and Industrial Engineering Technology, University of Johannesburg, 524, Johannesburg, South Africa

ARTICLE INFO

Keywords:
Multi-human detection
Machine learning
Ultra-wideband radar
Feature engineering
Wavelet transform

ABSTRACT

This study aims to detect and count people using impulse radio ultra-wideband radar and machine learning algorithms. However, the data quality, difficulty distinguishing human signals from noise and clutter, and instances where human presence is not detected make it challenging to count multiple humans. To overcome these challenges, we apply wavelet transformation to reduce signal size and use simple moving averages to eliminate noise. Next, we create features based on statistical and entropic properties of the signal and apply several classification algorithms, including ANN, Random Forest, KNN, XGBOOST, and multiple linear regression, to predict the number of people present. Our findings reveal that using the ANN classifier with the Daubechies 4 (db4) wavelet provides better results than other classifiers, with an accuracy rate of 99%. Additionally, filtering the data improves accuracy, and labeling the data after extracting essential characteristics significantly improves the model's accuracy.

1. Introduction

1.1. Background

Ultra-wideband radar systems attracted much attention in the scientific and commercial fields due to their remarkable advantages resulting from their notable wide bandwidth, good penetration, multipath immunity, and resolution [1,2]. Indeed, the impulse radio ultrawide band (IR-UWB) [3] has become a popular technology mainly due to the widespread deployment of Wireless Local/Personal/Body Area Networks (WLAN/WPAN, WBAN) and to the emergence of the Internet of Things (IoT) [4] and smart factories (Industry 4.0) [5]. IR-UWB has no harmful effect on the human body and employs extremely short pulses with a low-power level [6], which makes it possible to operate in a large part of the radio spectrum without disturbing the narrowband systems working in different frequency bands. Other benefits of this technology are its robustness in harsh environments, high precision ranging, low power consumption and high penetration capabilities. No particular pulse shape or modulation format is required; the only constraints related to the power spectral density masks are defined by the regulatory bodies.

IR-UWB sensors, their associated signals and information processing algorithms are successfully used in many applications, like indoor and machine-to-machine communications [7], detection and localization [8], human gait analysis [9], through-the-wall vital sign detection [10]. We are interested in human detection using machine-learning techniques. The detection refers to determining whether a specific object class instance is present in a given image (or signal) [11]. The human detection task involves locating all instances of persons present in an image (or signal) while producing as few false detections as possible [12]. The applications of human detection include activity recognition, people counting, real-time occupancy, etc. [1,2,13].

People counting refers to enumerating people in a defined location, and its various models can be grouped as device-based and device-free. The device-based scenarios require individuals to wear specific sensor devices, which will help to count the number of people accurately. Therefore, their performance depends strongly on the ability

* Corresponding author.
E-mail address: emanuel.radoi@univ-brest.fr (E. Radoi).

https://doi.org/10.1016/j.sasc.2024.200095
Received 20 October 2023; Received in revised form 28 February 2024; Accepted 1 April 2024
Available online 4 April 2024

and willingness of individuals to wear the sensor devices. In contrast, device-free systems consist of indoor sensors, whose inputs are analyzed to count people present in a specific location.

Our study relies on IR-UWB radar data to focus on device-free multi-human detection for people counting. The analysis of the backscattered IR-UWB signals enables us to discriminate those reflected by people from those reflected by artificial objects (clutter) and noise. This analysis remains a significant challenge since, depending on the coverage zone and its population density, the following issues may arise [14]: multipath signals are detected like backscattered signals from humans in densely populated areas, some signals reflected by humans are not detected in large areas where people are widely dispersed, and it can be difficult in some cases to separate human signals from noise and clutter.

The IR-UWB-based multi-human detection is often flawed by multiple false alarms due to the multipath components and the background clutter. Some proposed Constant False Alarm Rate (CFAR) algorithms achieved a few improvements [15]. Nevertheless, these models cannot be considered entirely reliable for real-life scenarios, as they suffer from the minimum required separation distance between humans (30 cm) and the number of false alarms generated [15].

1.2. Related work

Several research studies focusing on the people detection and counting have already been carried out. Most of them are interested in applying signal processing and using machine learning techniques to detect and count people. S. Chang et al. [16] introduced a version of CLEAN algorithm to detect the presence of a walking human in an urban environment populated by other moving objects such as cars and trucks. Their work uses parameters such as signal amplitude, Root Mean Square (RMS) range and speed characteristics to distinguish humans from other objects. The authors in [17] proposed a combination of the human kinetic Thalman model and a detector to better characterize the target and improve the signal-to-noise ratio as well as the detection probability.

Other related works leverage machine learning techniques to perform multi-human detection. In [18], people detection and tracking in a video stream obtained from a fixed or moving camera was proposed. V. Nogueira et al. [19] performed people counting using a neural network model with input data from RGB cameras. In [20], the authors used closed-circuit television (CCTV) footage to enumerate people in a region of interest (ROI) and eliminate those in the non-interest region. They propose a gap regularizer combined with a deep convolutional neural model to predict the number of people in the region of interest. The performance metrics used for prediction are the mean absolute error (MAE) and the root-mean-square error (RMSE). L Wang et al. proposed in [21] a multiscale attention network based on spatial position to capture the useful information needed for people counting. To count people in dense environments, the research work [22] suggested a multi-segment analysis. The first segment fetches information related to the density of the local (or global) space. A feature enhancement segment helps to refine local and global contextual features and, finally, the fusion segment is used to merge density-specific features. The authors in [23,24] proposed a learning algorithm based on Riemannian manifolds with Support Vector Machine (SVM), RandomForest and Adaboost classifiers for pedestrian detection. C.Wang et al. [25] performed people counting in a dense environment using a deep convolutional neural network trained with images.

Machine learning techniques are extensively used in signal processing for various applications ranging from signal detection to people counting. One can use wavelet entropy to extract meaningful features from those signals and then apply a machine-learning model to learn these features. Wavelet entropy is widely used for signal processing and analysis that involves decomposing a signal into wavelet coefficients and calculating entropy measures based on these coefficients. This approach has been applied in various fields, including biomedical signal analysis, financial time series analysis, image processing, and speech recognition. Using the continuous wavelet transform and spectral entropy measurement, the research work presented in [26] provides fast detection of whale vocalizations in the presence of noise. D. Dwivedi et al. [27] used wavelet space for identifying groundwater source boundaries to interpret potential field anomalies in the geophysical framework. In [28] presents a quantitative analysis of the brain's electrical multifrequency signals. The authors applied an orthogonal discrete wavelet to extract wavelet entropy features. I. Wijayanto et al. [29] proposed a random forest classification of epileptic electroencephalography signals using a feature extraction method based on wavelet energy calculation. The wavelet packet decomposition can be combined with a hybrid convolutional long-term memory neural network model to examine the central effects of transcutaneous electroacupuncture stimulation (TEAS) at different frequencies on the brain, as presented in [30]. [31] introduces a timely monitoring of chatters in industrial production by analyzing the characteristics of chatters' sound signals via spectral analysis and the wavelet packet transform. Finally, in [32], wavelet analysis was applied to examine the dominant periodicity in annual rainfalls of a given region.

1.3. Research contributions

Our proposal aims to enhance the accuracy of people counting by utilizing a combination of ultra-wideband pulse radars, signal processing tools, statistics, and machine learning techniques. We use the wavelet transform method to extract valuable information from reflected signals and reduce signal size, enabling faster and more efficient processing. Our approach also involves the application of advanced entropy and statistical measures to create salient signal features. Moreover, we have also implemented various data labeling strategies to further improve the accuracy of people counting.

1.4. Organization of the work

The rest of the paper is divided into three main sections. First, in Section 2, we give a comprehensive description of our contributions. This includes presenting the model architecture and describing each layer of the proposed model. Second, Section 3 covers the model evaluation. Finally, we conclude the work in Section 4.

2. Proposed model for improved detection

This section discusses the proposed model for improved multi-human detection. This work proposes a three-layer model for people counting using data measured by an impulse radio ultra-wideband radar. As shown in Fig. 1, the first layer performs a dimensionality reduction via a discrete wavelet transform. At the second layer, the resulting approximate coefficients undergo another transformation in the feature engineering layer, which reduces the input IR-UWB signal to a few but relevant features. The second layer of our model uses statistics and wavelet entropy to extract new features. The features generated (refer to Table 1) are used as input for machine-learning models in the third layer. In Section 3, we evaluate the efficiency of these features using well-known machine-learning algorithms.

2.1. Dimensionality reduction with wavelet transform

The wavelet transform is based on introducing new basis functions that can be expanded or compressed to capture a signal's low and high-frequency components. These components are determined by the wavelet scaling parameter [33,34]. Any signal $x(t) \in L^2(\mathbb{R})$ can be represented as a wavelet expansion, as shown in Eq. (1), using the mother wavelet $\psi : \mathbb{R} \to \mathbb{C}$. Factors such as the number of vanishing moments and wavelet compactness are considered when choosing the

Table 1
Features extracted from a ith wavelet transformed signal.

Signal	max	min	μ	σ^2	Median	$E_{s'_i}$	Quartile	Third-quartile	ϵ_s	ϵ_r	ϵ_t
s''_i	x''_1	x''_2	x''_3	x''_4	x''_5	x''_6	x''_7	x''_8	x''_9	x''_{10}	x''_{11}

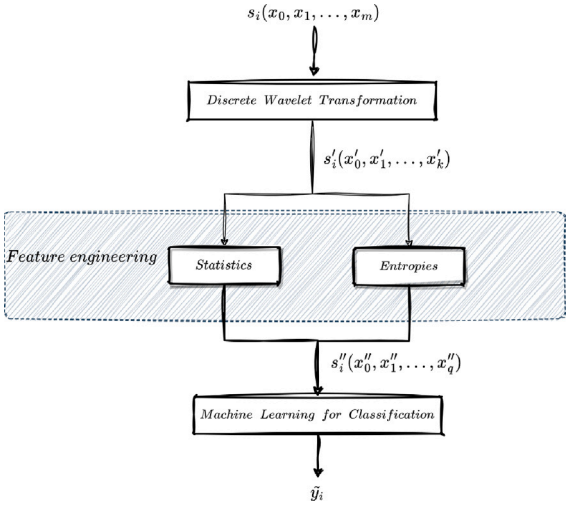


Fig. 1. Flow diagram of the proposed model.

mother wavelet. For our application, we found that the Daubechies mother wavelet with four vanishing moments (db4) is the most suitable. This wavelet is optimized to provide the most compact support for this degree of regularity. It is worth mentioning that this mother wavelet has been used in similar applications in the past [35].

$$x(t) = \sum_m \sum_n \left(\int_{-\infty}^{+\infty} x(t)\psi_{m,n}(t)dt \right) \psi_{m,n}(t), \tag{1}$$

$$\psi_{m,n}(t) = 2^{-\frac{m}{2}} \psi(2^{-m}t - n), \forall m, n \in \mathbb{Z} \tag{2}$$

2.2. Statistics-based features

Considering only the signal amplitude as a feature for the classification may lead to poor performance on unseen datasets. Therefore, we exploit its statistical properties in addition to the signal amplitude. Signal statistics provide us with a substantial amount of useful information for the classification. The statistical measures like maximum (max), minimum (min), mean (μ) and variance (σ^2) of the signal help to quantify signal behavior. The statistical mean μ gives the central tendency of the signal data. It is often used for noise removal. The metrics max, min and μ operate on the signal amplitude, while σ^2 capture its power characteristic.

2.3. Wavelet entropy-based features

The entropy of the approximated coefficient s'_i is computed to extract the quantity of information within the signal. It describes the uncertain distribution and complexity characteristics, which may quantitatively refer to the internal information characteristics contained in the signal. The wavelet entropy metric that we make use of in this work combines the wavelet transform and entropy to estimate the degree of uncertainly of a signal with a high time-frequency resolution [36–38]. For the three wavelet-based entropy measures considered in this project (Shannon, Renyi and Tsallis), we define the probability P_i as in (3).

$$P[s'_i] = \frac{E_{s'_i}}{E_{s_i}}, \tag{3}$$

where E_{s_i} and $E_{s'_i}$ are the energies of the ith raw s_i and wavelet transformed s'_i, respectively. The variable x''_i denotes the ith generated features from s'_i, as shown in Fig. 1 and Table 1.

The Shannon entropy is the commonly used entropy measure [36–39]. This entropy measure provides the variability associated with different frequency bands [38]. The expression of Shannon wavelet-based entropy is given by (4).

$$\epsilon_s(s'_i) = -P[s'_i]\log(P[s'_i]) \tag{4}$$

Rényi entropy can be seen as a generalization of the Shannon entropy. Renyi wavelet-based entropy of order $0 < \alpha < \infty$ is defined in (5).

$$\epsilon_r(s'_i) = \frac{1}{1-\alpha}\log\left(P^{\alpha}[s'_i]\right) \tag{5}$$

Tsallis wavelet-based entropy, given in (6), is a generalized form of Boltzmann–Gibbs entropy.Tsallis wavelet entropy can reduce the harmful effect of wavelet folding on the accuracy of feature extraction and extract the feature of power system transient signal accurately [38,39]. It could explain some anomalous phenomena, including the complexity of non-additive systems, which cannot be explained by the theory of expensive entropy [38].

$$\epsilon_t(s'_i) = \frac{1}{\gamma-1}\log\left(1 - P^{\gamma}[s'_i]\right) \tag{6}$$

3. Model evaluation and discussion

3.1. Data description

The dataset used in this work and the corresponding acquisition methodology are provided in [40]. It comprises at least 376000 reflected signals from 0 to 20 individuals. The data cover three scenarios, including 0-20 people randomly walking in a constrained area with densities of 3 (scenario 1) and 4 (scenario 2) persons per square meter and up to 15 people standing in a queue with a mean distance of 10 centimetres (scenario 3).

8000 radar signals were collected during each scenario, with 200 received signals recorded for each measurement. Each signal sample contains 1280 sampling points representing the 5 meters' detection range (spatial resolution of 0.0039 m). A total of 248000 radar signals were generated in the first and second scenarios and 128000 radar signals in scenario 3.

3.2. Machine learning models

In this section, we will provide a brief overview of the machine learning algorithms used in our analysis. The models we used include decision trees, random forests, eXtreme Gradient Boosting (XGBoost), k-nearest neighbors (KNN), multiple linear regression, and artificial neural networks (ANN).

The decision tree learning is a recursive partitioning of the feature space based on criteria to create a tree structure. It involves selecting the best splits, creating child nodes, and assigning predictions to the leaf nodes. Predictions are made by traversing the tree based on input features, and pruning can be applied to avoid overfitting.

Decision trees are beneficial because they are straightforward, concise, sensitive to context, and flexible [41]. They allow different attributes to be conditional on the outcomes of earlier tests. Decision tree-based models can easily handle continuous and discrete attributes, making them highly adaptable [42].

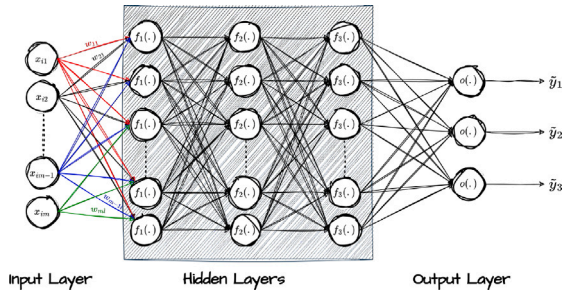


Fig. 2. Architecture of an Artificial Neural Network with three hidden layers.

Random Forest is an ensemble learning method that combines the predictions of multiple decision trees to make more accurate predictions [43]. Randomly sampling data and features reduces overfitting and improves generalization performance. Each decision tree is trained on a subset of the training data and features. At each split node, a random subset of features is considered.

The eXtreme Gradient Boosting (XGBoost) is a machine learning framework that uses Boosting Tree models [44,45]. It is known for its high performance, scalability, and ability to handle various data types and sizes. The mathematical formulation of XGBoost involves defining an objective function, which consists of a loss function and a regularization term, to be optimized. The algorithm then updates the model at each iteration [45].

The K-nearest neighbors (KNN) algorithm is a popular and straightforward supervised learning method for predicting qualitative and quantitative variables. It is a non-parametric technique that relies on the similarity between input and known labeled data to make predictions. This method is commonly used in data mining and pattern recognition for classification purposes [46–48]. Before applying the algorithm, it is necessary to determine the number of nearest neighbors, k.

A multiple linear regression model is a statistical method used for data with multiple predictor variables $(y_1, \ldots, y_i, \ldots, y_n)$ and one outcome $(x_{i1}, \ldots, x_{ik}, y_i)$, for $i = 1, 2, \ldots, n$ units of observation [49]. The model aims to find the relationship between multiple independent variables x_i and a single dependent variable y_i. This relationship is expressed in (7), where each β_j represents the impact of an independent variable on the dependent variable, the constant term β_0 is called the "y-intercept", and e_i denotes the difference between the predicted value $\tilde{y}_i$ and the actual value y_i. The objective of the model is to estimate the β_j parameters so that the sum of the squared errors e_i is minimized.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_m x_{im} + e_i. \tag{7}$$

An Artificial Neural Network (ANN) is a mathematical model miming biological neural networks [50]. It comprises interconnected neurons grouped into three categories of layers: input layer, hidden layers, and output layer (as shown in Fig. 2). Its activation functions convert input features into outputs, introducing non-linearity into the network [51,52]. GELU and Softmax are commonly used activation functions. The ANN optimizers adjust the model weights by minimizing the error or loss function. Gradient Descent, Adam, and RMSprop are widely used optimizers.

3.3. Evaluation metrics

This section will present the different metrics used to evaluate machine learning models in our work. These include Mean Square Error (MSE), Coefficient of Determination (R^2), Accuracy, Confusion Matrix, Precision, Recall, and Area Under the ROC Curve (AUC-ROC) [53,54].

MSE measures how well the regression line fits the data and accurately makes predictions. R^2 measures the accuracy of our model in predicting a dependent variable. Accuracy provides a general measure of how well the model is performing. A good accuracy score should be above 80%. Precision is the ratio of correctly predicted positive examples divided by the total number of predicted positive examples. The Confusion Matrix compares the observed and predicted values in a tabular visualization, providing a more complete perspective than just observing the accuracy. Recall indicates the number of correct positive predictions made out of all positive predictions that could have been made. The AU-ROC metric evaluates how well the model can distinguish between the two classes. It assumes values between 0 and 1, where 0 is the worst result and 1 is the best result.

Table 2
Random forest model.

Number of trees	Accuracy	Wavelet
100	72.3%	Daubechies 4

Table 3
Xgboost model.

Number of trees	Accuracy	Wavelet
150	72.3%	Daubechies 4

3.4. Feature engineering results

Based on the model architecture shown in Fig. 1 and thanks to Python libraries such as PyWavelets, Sklearn and TensorFlow, we performed feature extraction on the IR-UWB radar data obtained from Scenario 3. This involved extracting statistical and entropy features, resulting in a total of 11 features. These features were then used to apply machine learning algorithms for predicting the number of individuals present in each reflected signal.

The selected graph displays the various box sizes for each scenario, indicating the differences between the subgroups within each scenario. In addition, it is obvious that the upper quartile amplitude of the entropies features (see Fig. 4(a)) decreases as the number of people in the scenario increases. The statistical features are shown in Fig. 3(a). The upper-quartile amplitude of scenarios with 3 and 9 people is observed to be greater than that of other scenarios. This suggests that statistical features can effectively differentiate between scenarios with 3 and 9 people from other scenarios.

However, it is challenging to distinguish between scenarios with 3 and 9 people using only statistical features. To overcome this, we combine entropy-based features with statistical ones to improve prediction and better distinguish between different scenarios.

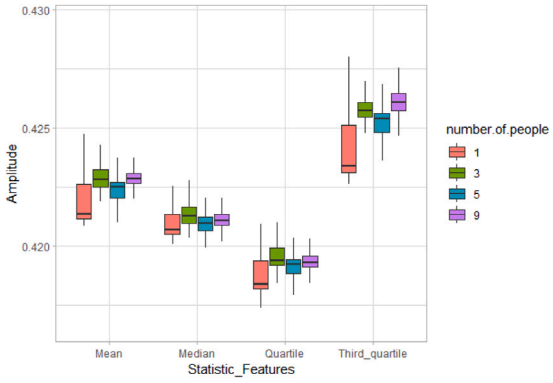
3.5. Models evaluation results

In this section, we will present the results obtained by different models in the case of the third scenario. We will plot the signals obtained during the recordings containing one person, five persons and 12 persons as presented in Fig. 5 below.

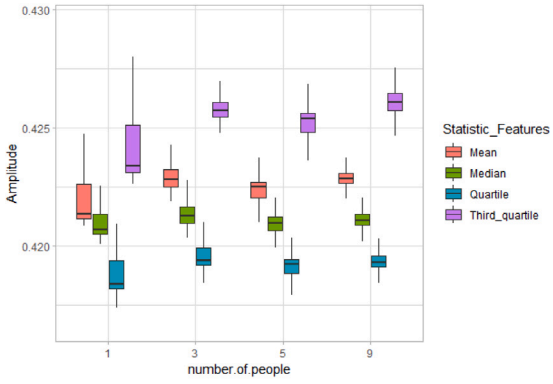
Fig. 5 shows the reflected signals from the different recordings. We have the reflected signal in each image and the processed reflected signal. We have performed a simple moving average to reduce the noise and smooth the amplitude of the reflected signals obtained. In addition, we split the data into 80% training and 20% testing sets and present the results below.

Fig. 6 indicates that the model's accuracy improves slightly with filtered data (72.3%) compared to raw data (70.85%). The details of the model are defined by the Tables 2 and 3.

The K-nearest neighbors with 7 Neighbors give an accuracy of 51.57%, and the coefficient of determination of Multiple linear regression is 45%. Despite data processing and feature engineering techniques, the models have difficulty approaching an accuracy of 84%.

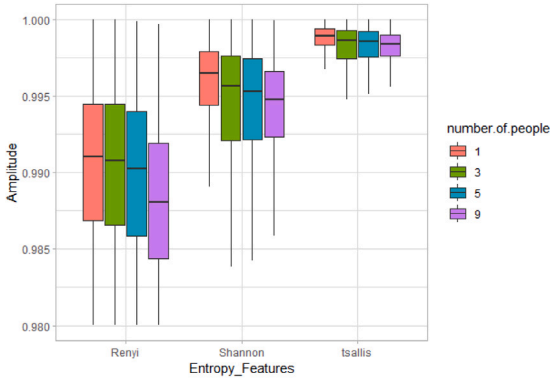


(a) Statistics Features in x-axis

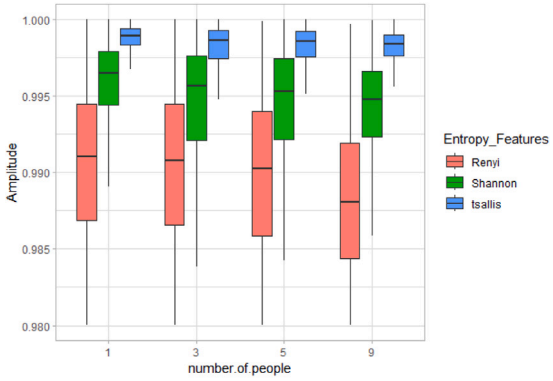


(b) Number of People in x-axis

Fig. 3. Features statistics.

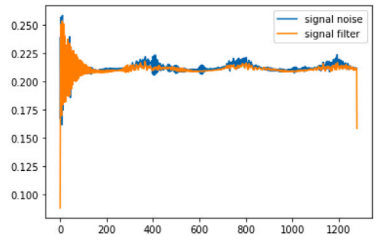


(a) Entropy Feature in x-axis

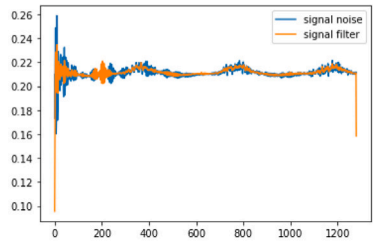


(b) Number of People in x-axis

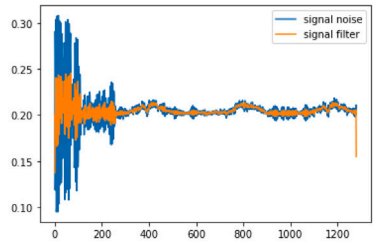
Fig. 4. Features entropies.



(a) 1 person

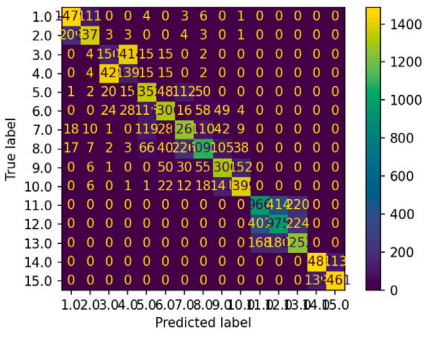


(b) 5 persons

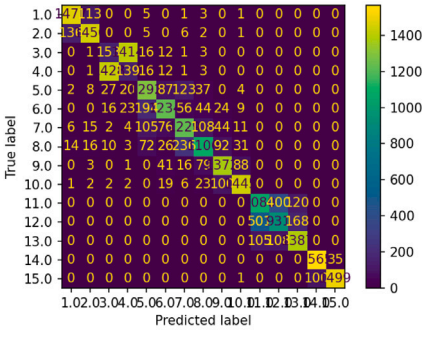


(c) 12 persons

Fig. 5. Reflected signals from one to 12 persons.

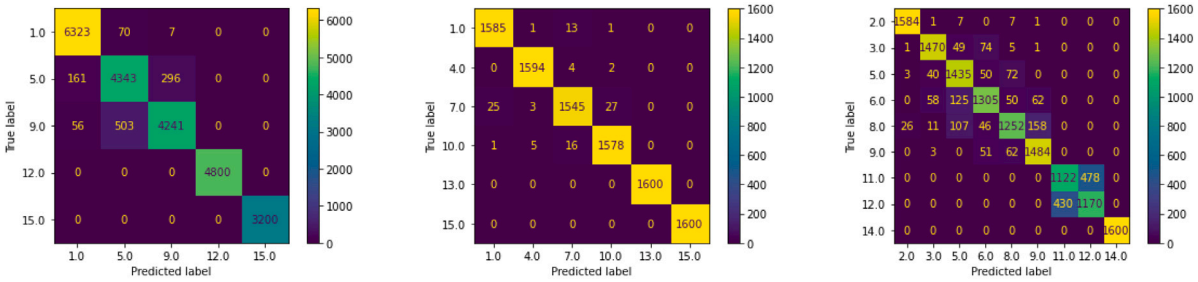


(a) data without filter



(b) data filter

Fig. 6. Confusion Matrix.



(a) Labeling #1 (b) Labeling #2 (c) Labeing #3

Fig. 7. Confusion Matrix after labeling technique.

Table 4
Sets of labels created with different labeling strategies.

Labeling	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
#1	1	1	1	1	5	5	5	9	9	9	12	12	12	15	15
#2	1	1	1	4	4	4	7	7	7	10	10	10	13	13	15
#3	2	2	2	3	5	6	7	8	9	10	11	12	12	14	14

This may be due to the quality of the data or a defect in the sensor during measurement. We also notice from Fig. 6 that it is difficult for the models to make a difference between the reflected signal containing 1 person and that containing 2 people. Therefore, we will apply a labeling technique to improve the performance of the models.

3.6. Models evaluation results with the labeling technique

Due to the low data quality, it is challenging to differentiate between signals corresponding to one and two people respectively. This is the reason why we implement a labeling technique during the feature engineering process. Labeling enables us to group signals within ranges and is essential for accurately identifying reflected signals reflected by individuals. For example, we can gather reflected signals from 1 to 3 people within the group label 1. Then, we train and test the models with the modified data to identify the range in which a signal falls. The process ensures that if a signal contains multiple people, we can accurately identify and label each individual. In other words, if a signal is classified as one person, it may contain 1, 2 or 3 people, giving us a more reliable model.

We apply different labeling strategies, which are described as follows. The reflected signals containing 2, 3, and 4 people are labeled as signals containing one person, then signals 6, 7, and 8 are labeled as signals containing 5 people, 10 to 9, then 11, 13 to 12 and 14 to 15 people. This labeling leads to five new labels: 1, 5, 9, 12,15. Table 4 contains the various labels. The confusion matrix obtained after these labeling techniques is presented in Fig. 7 below.

The Table 5 shows how accurate the model is with the new labels. That also says much about the data and the importance of good preprocessing techniques. Now, we offer the possibility of providing a detection margin that improves the model's accuracy. Compared with other works tackling the same problem, we noticed in [40] that the maximum precision obtained is 97%. The overall precision of the models is not specified in this work. However, the work in [6,55,56] is based more on a statistical study, and the detection of individuals needs to present a clear classification case compared to that which we propose. We summarize the results obtained by the XGBOOST and KNN model in the Tables 6, 7.

The results of the Multiple Linear Regression model are as follows: the coefficient of determination is 71.47%, the mean square error is 5.33, the mean absolute error is 5.33, and the root mean square error is 2.31. In the Tables 5,6, 7,8, Time represents the execution time of a single sample and the RAM values inserted take into account the model's training process.

In comparison to similar studies such as [6,40,55,56], it was found that the highest precision achieved was 97%. It is important to note that KNN, XGBOOST, and Multiple Linear Regression were not utilized for classification in these studies. However, our XGBOOST model yielded a higher accuracy rate.

3.7. Performance of the artificial neural network model

The architecture of our ANN model used consists of 3 hidden layers of 512 units and 3 dropout layers to eliminate overfitting with a probability greater than 0.5. The activation function used on each layer is the GELU function, and the output layer uses the softmax function. Nadam algorithm is used as an optimizer.

The predicted result is a vector of size 16, which contains probability values. Before the performance analysis, we transform these probabilities to binary. We define a threshold, then set to 1 all the probabilities bigger than the threshold and 0 otherwise. The results of different metrics are presented in Table 8.

When simulated on Google Colab, our model requires an estimated RAM (Random Access Memory) capacity of 8 GB and an estimated model runtime of 60 h, taking into account the training process of the model, the execution time of a sample is evaluated at 3800 ms (ms). Compared to other proposed models, our ANN classifier provides better accuracy and higher prediction rates of the different classes. Thus, the closest reported research using the same dataset [40] relies on the curvelet transform to extract hybrid features, resulting in an accuracy of 97.8% and 98.7% obtained with Neural Networks and random forest classifiers, respectively.

4. Conclusion

This paper presents an enhanced performance detection system which combines signal processing tools with machine learning techniques for people counting using IR-UWB radar. The proposed system performs the discrete wavelet transform of IR-UWB reflected signals to decrease their dimensions, while keeping the salient information. Then, we consider entropy and statistical measures from the transformed signals to extract relevant features that enable more accurate machine learning models in counting the number of people present in the region of interest. We finally improve the classification results by utilizing data labeling methods tailored to our use case.

Our sample dataset consists of 15 individuals who were waiting in line (queuing). After applying the moving average to eliminate noise and clutter from the signal, we extract 11 features from the discrete wavelet transform of the IR-UWB received signal. These features include 8 statistical features and 3 entropy features. We then use 5 classifiers to predict the number of present people. The counting accuracy is more significant than 99% with the ANN model. In addition, the impact of the moving average on the signal is proven by showing a slight improvement in counting accuracy by the classifier compared to the case where it is not applied. The standard machine learning

Table 5
Random forest model after labeling technique.

Labeling	Number of trees	Accuracy	Wavelet	RAM	Time
#1		95.44%			
#2	100	98.98%	Daubechies 4	3.3 GB	3900 ms
#3		86.26%			

Table 6
XGBOOST model result.

Labeling	Number of trees	Accuracy	Wavelet	RAM	Time
Without		72.3%			
#1	150	95.39%	Daubechies 4	3.3 GB	3950 ms
#2		98.61%			
#3		86.27%			

Table 7
KNN model result.

Labeling	Number of neighbors	Accuracy	Wavelet	RAM	Time
Without		51.57%			
#1	7	69.3%	Daubechies 4	3.3 GB	3800 ms
#2		90.22%			
#3		86.28%			

Table 8
Performance results of the ANN model.

Epochs	Accuracy	Loss	Auc	Precision	Recall	TP	FN	Time	RAM
5000 × 5	99.35%	0.1084	99.98%	91.04%	98.59%	32776	3224	2500 ms	8 GB

models perform poorly despite feature engineering, partly due to data quality or IR-UWB radar data acquisition. We encountered difficulty distinguishing between scenarios with a difference of one person, so we improved the model's accuracy with some labeling techniques. As future work, we plan to test the same strategy in other scenarios to validate the performance of the proposed method. We also plan to observe how the model behaves when new environmental factors are added, in terms of accuracy and computation time.

CRediT authorship contribution statement

Ange Joel Nounga Njanda: Conceptualization, Formal analysis, Investigation, Software, Visualization, Writing – original draft, Writing – review & editing, Data curation. **Jocelyn Edinio Zacko Gbadoubissa:** Conceptualization, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Writing – original draft, Writing – review & editing, Data curation. **Emanuel Radoi:** Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing. **Ado Adamou Abba Ari:** Formal analysis, Methodology, Supervision, Validation. **Roua Youssef:** Formal analysis, Methodology, Supervision, Validation. **Aminou Halidou:** Formal analysis, Methodology, Supervision, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

We used a publicly shared database.

References

[1] H. Elkhoukhi, Y. NaitMalek, A. Berouine, M. Bakhouya, D. Elouadghiri, M. Essaaidi, Towards a real-time occupancy detection approach for smart buildings, Procedia Comput. Sci. 134 (2018) 114–120, http://dx.doi.org/10.1016/j.procs.2018.07.151, URL: https://www.sciencedirect.com/science/article/pii/S1877050918311141, The 15th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2018) / The 13th International Conference on Future Networks and Communications (FNC-2018) / Affiliated Workshops.

[2] K.C.J. Simma, A. Mammoli, S.M. Bogus, Real-time occupancy estimation using WiFi network to optimize HVAC operation, Procedia Comput. Sci. 155 (2019) 495–502.

[3] V. Niemelä, J. Haapola, M. Hämäläinen, J. Iinatti, An ultra wideband survey: Global regulations and impulse radio research based on standards, IEEE Commun. Surv. Tutor. 19 (2) (2017) 874–890, http://dx.doi.org/10.1109/COMST.2016.2634593.

[4] G. Oguntala, R. Abd-Alhameed, S. Jones, J. Noras, M. Patwary, J. Rodriguez, Indoor location identification technologies for real-time IoT-based applications: An inclusive survey, Comp. Sci. Rev. 30 (2018) 55–79.

[5] G. Berardinelli, N.H. Mahmood, I. Rodriguez, P. Mogensen, Beyond 5G wireless IRT for industry 4.0: Design principles and spectrum aspects, in: 2018 IEEE Globecom Workshops, GC Wkshps, IEEE, 2018, pp. 1–6.

[6] E.C. Lee, S.H. Cho, Performande enhancement of multi-human detection using an IR-UWB radar by an adaptive thresholding algorithm, in: 2016 IEEE International Conference on Network Infrastructure and Digital Content, IC-NIDC, IEEE, 2016, pp. 476–480.

[7] P.K. Verma, R. Verma, A. Prakash, A. Agrawal, K. Naik, R. Tripathi, M. Alsabaan, T. Khalifa, T. Abdelkader, A. Abogharaf, Machine-to-machine (M2M) communications: A survey, J. Netw. Comput. Appl. 66 (2016) 83–105, http://dx.doi.org/10.1016/j.jnca.2016.02.016, URL: https://www.sciencedirect.com/science/article/pii/S1084804516000990.

[8] T. Ha, J. Kim, Detection and localization of multiple human targets based on respiration measured by IR-UWB radars, in: 2019 IEEE SENSORS, 2019, pp. 1–4.

[9] S.P. Rana, M. Dey, M. Ghavami, S.E.M. Dudley, Non-contact human gait identification through IR-UWB edge-based monitoring sensor, IEEE Sens. J. 19 (2019) 9282–9293.

[10] W. Wang, D. qiao Wang, B. Zhang, T. Li, S. Jiang, Through-wall multistatus target identification in smart and autonomous systems with UWB radar, IEEE Internet Things J. 5 (2018) 3278–3288.

[11] M.A. Ansari, D.K. Singh, Human detection techniques for real time surveillance: A comprehensive survey, Multimedia Tools Appl. 80 (2021) 8759–8808.

[12] J.W. Davis, V. Sharma, A. Tyagi, M. Keck, Human detection and tracking, in: S.Z. Li, A. Jain (Eds.), Encyclopedia of Biometrics, Springer US, Boston, MA, 2009, pp. 708–712, http://dx.doi.org/10.1007/978-0-387-73003-5_35.

[13] X. Yang, W. Yin, L. Zhang, People counting based on CNN using IR-UWB radar, in: 2017 IEEE/CIC International Conference on Communications in China, ICCC, IEEE, 2017, pp. 1–5.

[14] N. Patwari, A.O. Hero, M. Perkins, N.S. Correal, R.J. O'dea, Relative location estimation in wireless sensor networks, IEEE Trans. Signal Process. 51 (8) (2003) 2137–2148.

[15] J.W. Choi, S.S. Nam, S.H. Cho, Multi-human detection algorithm based on an impulse radio ultra-wideband radar system, IEEE Access 4 (2016) 10300–10309.

[16] S. Chang, N. Mitsumoto, J.W. Burdick, An algorithm for UWB radar-based human detection, in: 2009 IEEE Radar Conference, IEEE, 2009, pp. 1–6.

[17] S.Z. Gurbuz, W.L. Melvin, D.B. Williams, Comparison of radar-based human detection techniques, in: 2007 Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers, IEEE, 2007, pp. 2199–2203.

[18] B. Wu, R. Nevatia, Detection and tracking of multiple, partially occluded humans by bayesian combination of edgelet based part detectors, Int. J. Comput. Vis. 75 (2) (2007) 247.

[19] V. Nogueira, H. Oliveira, J.A. Silva, T. Vieira, K. Oliveira, RetailNet: A deep learning approach for people counting and hot spots detection in retail stores, in: 2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images, SIBGRAPI, IEEE, 2019, pp. 155–162.

[20] B. Pardamean, F. Abid, T.W. Cenggoro, G.N. Elwirehardja, H.H. Muljo, Counting people inside a region-of-interest in CCTV footage with deep learning, PeerJ Comput. Sci. 8 (2022) e1067.

[21] L. Wang, J. Li, S. Zhang, C. Qi, P. Wang, F. Wang, Multi-scale and spatial position-based channel attention network for crowd counting, J. Vis. Commun. Image Represent. 90 (2023) 103718.

[22] H. Kushwaha, S.K. Gupta, Counting of density crowd using multi-segment analysis, Available at SSRN 4327729.

[23] O. Tuzel, F. Porikli, P. Meer, Pedestrian detection via classification on riemannian manifolds, IEEE Trans. Pattern Anal. Mach. Intell. 30 (10) (2008) 1713–1727.

[24] P.F. Felzenszwalb, R.B. Girshick, D. McAllester, D. Ramanan, Object detection with discriminatively trained part-based models, IEEE Trans. Pattern Anal. Mach. Intell. 32 (9) (2009) 1627–1645.

[25] C. Wang, H. Zhang, L. Yang, S. Liu, X. Cao, Deep people counting in extremely dense crowds, in: Proceedings of the 23rd ACM International Conference on Multimedia, 2015, pp. 1299–1302.

[26] M.W. Rademan, D. Versfeld, J.A. du Preez, Soft-output signal detection for cetacean vocalizations using spectral entropy, k-means clustering and the continuous wavelet transform, Ecol. Inform. (2023) 101990.

[27] D. Dwivedi, A. Chamoli, S.K. Rana, Wavelet entropy: A new tool for edge detection of potential field data, Entropy 25 (2) (2023) 240.

[28] O.A. Rosso, S. Blanco, J. Yordanova, V. Kolev, A. Figliola, M. Schürmann, E. Başar, Wavelet entropy: a new tool for analysis of short duration brain electrical signals, J. Neurosci. Methods 105 (1) (2001) 65–75.

[29] I. Wijayanto, S. Rizal, S. Hadiyoso, Epileptic electroencephalogram signal classification using wavelet energy and random forest, in: AIP Conference Proceedings, vol. 2654, (no. 1) AIP Publishing LLC, 2023, 020026.

[30] Ç. Uyulan, D. Mayor, T. Steffert, T. Watson, D. Banks, Classification of the central effects of transcutaneous electroacupuncture stimulation (TEAS) at different frequencies: A deep learning approach using wavelet packet decomposition with an entropy estimator, Appl. Sci. 13 (4) (2023) 2703.

[31] K. Lu, X. Wang, X. Chen, X. Pang, F. Gu, Experimental study on entropy features in machining vibrations of a thin-walled tubular workpiece, J. Dyn. Monit. Diagn. 2 (1) (2023) 61–68.

[32] E. Thomas, I. Joseph, N.P. Abraham, Wavelet analysis of annual rainfall over Kerala and sunspot number, New Astron. 98 (2023) 101944.

[33] M. Vetterli, C. Herley, Wavelets and filter banks: Theory and design, IEEE Trans. Signal Process. 40 (ARTICLE) (1992) 2207–2232.

[34] S. Radhakrishnan, Introductory chapter: Understanding wavelets, in: S. Radhakrishnan (Ed.), Wavelet Theory and Its Applications, IntechOpen, Rijeka, 2018, http://dx.doi.org/10.5772/intechopen.78388.

[35] J. Chen, G. Li, Tsallis wavelet entropy and its application in power signal analysis, Entropy 16 (6) (2014) 3009–3025.

[36] J. Li, H. Zhang, J. Ou, W. Wang, A radar signal recognition approach via IIF-net deep learning models, Comput. Intell. Neurosci. 2020 (2020).

[37] J. Li, J. Guo, A new feature extraction algorithm based on entropy cloud characteristics of communication signals, Math. Probl. Eng. 2015 (2015) 1–8.

[38] R. Sharma, R.B. Pachori, U.R. Acharya, An integrated index for the identification of focal electroencephalogram signals using discrete wavelet transform and entropy measures, Entropy 17 (8) (2015) 5218–5240.

[39] Z. Zhang, Y. Li, S. Jin, Z. Zhang, H. Wang, L. Qi, R. Zhou, Modulation signal recognition based on information entropy and ensemble learning, Entropy 20 (3) (2018) 198.

[40] X. Yang, W. Yin, L. Li, L. Zhang, Dense people counting using IR-UWB radar with a hybrid feature extraction method, IEEE Geosci. Remote Sens. Lett. 16 (2019) 30–34.

[41] J.R. Quinlan, Decision trees and decision-making, IEEE Trans. Syst. Man Cybern. 20 (2) (1990) 339–346.

[42] O.Z. Maimon, L. Rokach, Data Mining with Decision Trees: Theory and Applications, vol. 81, World scientific, 2014.

[43] G. Biau, E. Scornet, A random forest guided tour, Test 25 (2016) 197–227.

[44] W. Li, Y. Yin, X. Quan, H. Zhang, Gene expression value prediction based on xgboost algorithm, Front. Genet. 10 (2019) 1077.

[45] T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.

[46] A. Pandey, A. Jain, Comparative analysis of KNN algorithm using various normalization techniques, Int. J. Comput. Netw. Inf. Secur. 11 (11) (2017) 36.

[47] J.E.Z. Gbadoubissa, A.A.A. Ari, A.M. Gueroui, Efficient k-means based clustering scheme for mobile networks cell sites management, J. King Saud Univ.-Comput. Inf. Sci 32 (9) (2020) 1063–1070.

[48] J.E.Z. Gbadouissa, A.A.A. Ari, C. Titouna, A.M. Gueroui, O. Thiare, HGC: HyperGraph based Clustering scheme for power aware wireless sensor networks, Future Gener. Comput. Syst. 105 (2020) 175–183.

[49] L.E. Eberly, Multiple linear regression, Top. Biostat. (2007) 165–187.

[50] A. Krenker, J. Bešter, A. Kos, Introduction to the artificial neural networks, in: Artificial Neural Networks: Methodological Advances and Biomedical Applications, InTech, 2011, pp. 1–18.

[51] S. Sharma, S. Sharma, A. Athaiya, Activation functions in neural networks, Towards Data Sci. 6 (12) (2017) 310–316.

[52] D. Hendrycks, K. Gimpel, Gaussian error linear units (gelus), 2016, arXiv preprint arXiv:1606.08415.

[53] G.S. Handelman, H.K. Kok, R.V. Chandra, A.H. Razavi, S. Huang, M. Brooks, M.J. Lee, H. Asadi, Peering into the black box of artificial intelligence: Evaluation metrics of machine learning methods, AJR Am. J. Roentgenol. 212 (1) (2018) 38–43.

[54] Ž. Vujović, et al., Classification model evaluation metrics, Int. J. Adv. Comput. Sci. Appl. 12 (6) (2021) 599–606.

[55] J.W. Choi, S.H. Cho, A new multi-human detection algorithm using an IR-UWB radar system, in: Third International Conference on Innovative Computing Technology, INTECH 2013, IEEE, 2013, pp. 467–472.

[56] J.W. Choi, D.H. Yim, S.H. Cho, People counting based on an IR-UWB radar sensor, IEEE Sens. J. 17 (17) (2017) 5717–5727.