Comparison of ML algorithms to distinguish between human or human-like targets using the HOG features of range-time and range-Doppler images in through-the-wall applications

Yunus Emre ACAR[1,*] ![ORCID], **İsmail SARITAŞ**[1] ![ORCID], **Ercan YALDIZ**[2] ![ORCID]

[1]Department of Electrical and Electronics Engineering, Faculty of Technology,
Selçuk University, Konya, Turkey
[2]Department of Electrical and Electronics Engineering, Faculty of Engineering and Natural Sciences,
Konya Technical University, Konya, Turkey

Abstract: When detecting the human targets behind walls, false detections occur for many systematic and environmental reasons. Identifying and eliminating these false detections is of great importance for many applications. This study investigates the potential of machine learning (ML) algorithms to distinguish between the human and human-like targets behind walls. For this purpose, a stepped-frequency continuous-wave (SFCW) radar has been set up. Experiments have been carried out with real human targets and moving plates imitating a regular breath of a healthy human. Unlike conventional methods, human and human-like returns are classified using range-Doppler images containing range and Doppler information. Then, the histogram of oriented gradients (HOG) features of the range-Doppler images are extracted, and the number of these features is reduced by principal component analysis (PCA). Finally, popular ML algorithms are executed to distinguish the human and human-like returns. The performances of the ML algorithms are compared for both range-time and range-Doppler images with or without HOG features. Experiments have indicated that the HOG features of the range-Doppler profiles provide the best results with the support vector machine (SVM) classifier with an accuracy of 93.57%.

Key words: HOG feature, human detection, machine learning, through-the-wall, radar

1. Introduction

Radar systems have emerged at the beginning of the 20th century for maritime applications and gained great importance during World War I and World War II by evolving into supersecret technologies. Although radar systems have frequently used for military purposes until the 1970s, they have reached a wide range of civilian applications in line with the developments in electronics technology. In the last decade, remote sensing of humans has been one of the most popular applications standing out the radar technology. Many radar-based systems have been introduced for various human-targeted applications such as range detection, positioning, vital signal monitoring, and classification [1–5]. Continuous-wave (CW) radars have been used to monitor vital signs at first, whereas ultrawideband (UWB) radar has been extensively studied for distance detection [6–10]. On the other hand, frequency-modulated continuous-wave (FMCW) radars and stepped-frequency continuous-wave (SFCW) radars have come to the fore with the capability of distance and velocity detection recently [11–15]. In

---

*Correspondence: yacar@selcuk.edu.tr

SFCW radars, wider bandwidth provides higher range resolution, while FMCW radars can operate with shorter observation times due to shorter frequency sweeping periods [16]. Although these two structures have different advantages and disadvantages, they are very similar in hardware configuration. Both FMCW and SFCW radars can be built in homodyne or heterodyne structures. Heterodyne structures provide easier solutions to DC noise and clutter rejection by reducing the high-frequency return signal to an intermediate frequency, while homodyne structures are advantageous in hardware complexity, cost, and physical size since they reduce the return signal directly to the baseband [17].

In this study, considering the information given above, a radar structure that can detect human targets has been created for through-the-wall applications to distinguish the human and human-like returns. In our previous work [18], the radial distances and respiration rates of the targets have been determined for the line of sight scenarios. The structure has been rearranged for detecting targets behind obstacles by adding baseband amplifiers and filters. The main goal of this work is to distinguish between human and human-like target returns to eliminate the possible false detections. For this purpose, a moving plate platform imitating human chest wall motion has been used for the human-like returns. Human and moving-plate targets have been evaluated as two separate classes. Range-Doppler profiles of the targets have been recorded as images to be used in the classification process. Histogram of oriented gradients (HOG) has been extracted as features. PCA has been applied to these features to reduce the size keeping only the components that explain 95% of the variance. The performance of the popular ML algorithms has been compared among themselves.

The organization of the rest of the article is as follows. Section 2 presents the hardware and system parameters. The experiments and results are discussed in Section 3 with a comparison of the popular ML algorithms. Finally, Section 4 summarizes and concludes the paper.

2. Materials and methods

In this part of the paper, the hardware of the created radar structure and the signal processing steps are presented.

2.1. Hardware of the created radar structures

Distinguishing the returns of the human and human-like targets behind walls has been focused on as a two-class classification problem. Because most of the applications require the range and Doppler frequency of the targets, a method has been proposed for the SFCW radar structure providing these. The radar hardware is a modified version of the radar system created in [18]. Since through-the-wall scenarios are studied, it is required to enhance the returning signals. Two identical instrumental amplifiers have been added to amplify and filter the baseband signals to compensate for the weakening caused by the wall. The complex-baseband structure has been kept to utilize both magnitude and phase of the returns to prevent the null points and improve the noise figure [19, 20]. The hardware has been set up in homodyne configuration to lessen cost and complexity. The employed radar structure is illustrated in Figure 1.

A vector network analyzer (VNA) generates the required stepped-frequency RF signal. The frequency of the RF signal sweeps between 2 and 4 GHz as an increasing ramp form with a period of 0.7032 s for 101 sweep points. These setups bring a high FFT resolution for range-FFT while still adequate for Doppler FFT [18]. RF power divider splits the sweeping signal into two branches. One of the branches is used for the transmitting side, whereas the other one provides a reference for the receiving end for IQ demodulation. The frequencies of the receiving signals are directly downconverted to the baseband after the mixing stage. The resulting baseband
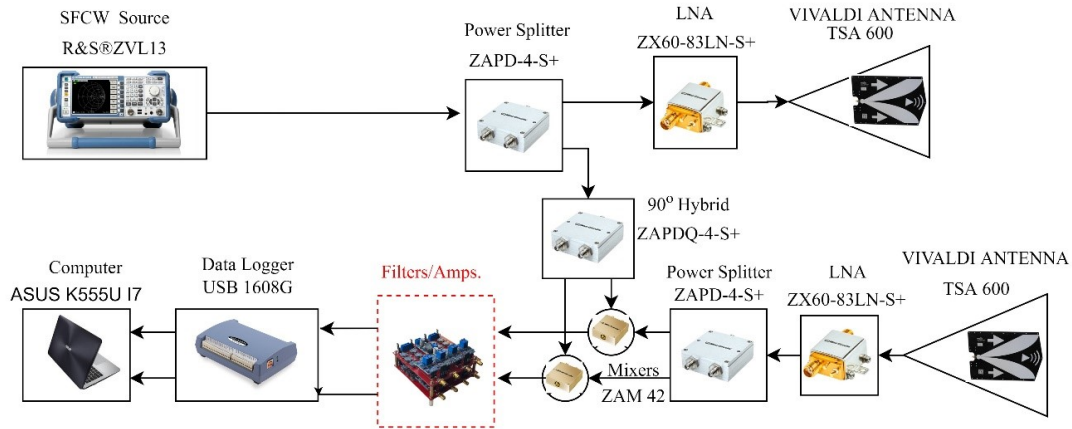
Figure 1. The experimental radar set-up.

signals are amplified around 20 dB by the instrumental amplifiers.

A moving plate platform is created to imitate a regular chest wall to obtain human-like returns. In literature, it is common to assume that the respiration rate is around 0.2 Hz–0.35 Hz for adults [21–23]. Thus, the platform has been set to move 1 cm back and forth with a period of 4s. The image of the platform is given in Figure 2.
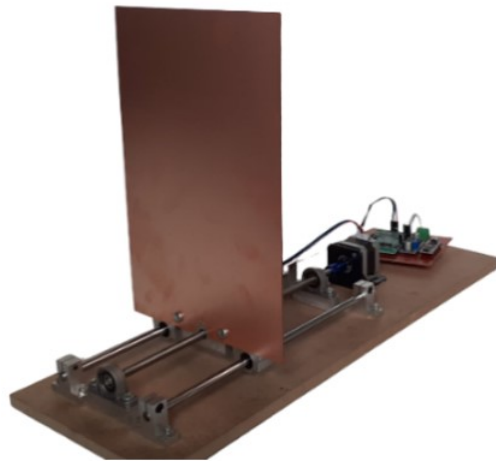

Figure 2. Moving-plate platform imitating the regular chest wall motion [18].

2.2. Proposed method

The signal processing steps of the proposed method begin after digitizing the baseband signals. For an observation time of T_o and sampling frequency of F_s, the length of the recorded baseband signals is $T_o \times F_s$. For the ith frequency return, the normalized baseband signals can be expressed as follows:

$$I_i[n] = cos\left(-\frac{4\pi}{\lambda_i}\left(R + x[n]\right) - \Delta\phi[n]\right), n = 1, 2, \cdots, T_o \times F_s, \tag{1}$$

$$Q_i[n] = sin\left(-\frac{4\pi}{\lambda_i}(R + x[n]) - \Delta\phi[n]\right), n = 1, 2, \cdots, T_o \times F_s, \tag{2}$$

where λ_i is the instantaneous operating wavelength, R is the nominal distance between the radar and the target, $x[n]$ is the micromotion of the target, and $\Delta\phi[n]$ is the residual phase noise. Since the same signal source is used in both transmitting and receiving sides, $\Delta\phi[n]$ can be omitted due to the range correlation effect [16]. Thus, the phase of the baseband signals becomes proportional to the range and micromotion. It is possible to create a complex-valued matrix by combining and reshaping the baseband signals. The phase of this signal, containing the range and micromotion information, can be extracted by applying IFFT (inverse fast Fourier transform) to the rows of this matrix. The created complex signal $C_i[n]$ can be expressed as in (3)

$$C_i[n] = e^{j\left(-\frac{4\pi}{\lambda_i}(R + x[n])\right)}, n = 1, 2, \cdots, T_o \times F_s. \tag{3}$$

It should be noted that the range resolution is directly proportional to the instantaneous bandwidth, which is too narrow in continuous-wave radars compared to the UWB radars. Fortunately, joint processing of discrete samples improves the range resolution by N in SFCW structures. The enhanced range resolution ΔR_e is expressed as [16]

$$\Delta R_e = \frac{c}{2N\Delta f}, \tag{4}$$

where c is the propagation speed of the wave, Δf is the frequency step for N sweeping frequency points. IFFT is one of the simplest and favorite methods for joint processing and phase extraction. First, the complex signal $C_i[n]$ is placed in the rows of the matrix in case each row corresponds to the consecutive sweep returns. The resulting matrix becomes a complex-valued matrix. The rows and columns of the matrix are sampled with ADC's sampling period and sweeping period, respectively. The range and micromotion information are extracted by applying IFFT to the rows of the matrix, and a range-slow time profile is obtained. By a further phase extraction, performing fast Fourier transform (FFT) through the slow-time axis of this profile, the range-Doppler profile is created. The steps from complex-valued matrix generation to obtaining the range-Doppler matrix are illustrated in Figure 3. As examples, range-Doppler and range-slow time images for a target located at 2.5 m are given in Figure 4.
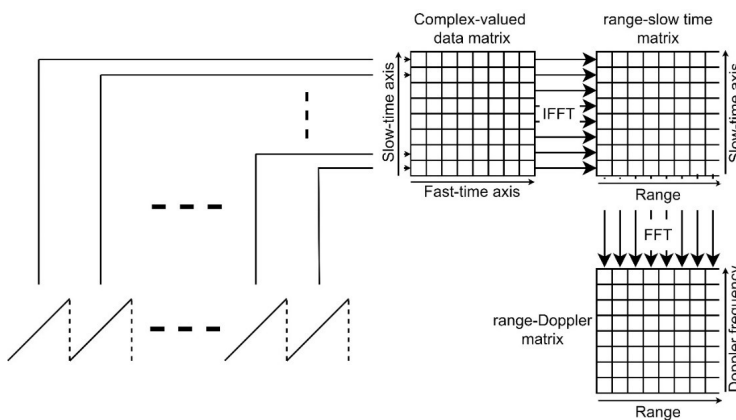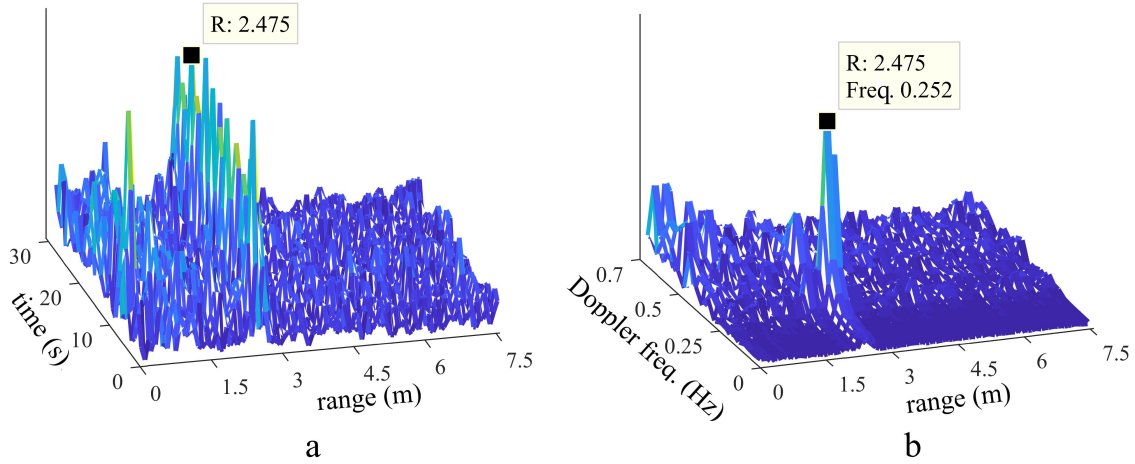

Figure 3. The steps to create the range-Doppler matrix.

Figure 4. (a) range-slow time image, (b) range-Doppler image for a target at 2.5 m

In the created range-Doppler matrix, the highest peak corresponds to the reflection from the wall. Together with this strong echo, the presence of noise and clutters also necessitate additional filtering operations. Two-line delay cancellation is applied to the range axis as a moving target indicator (MTI) to suppress the wall and emphasize the targets behind it. Sequentially, the slow time axis of the matrix is filtered with a 32nd order low pass filter. After filtering operations, the gray-colored radar images are obtained by normalizing the matrix values to the range of 0 to 1. The last step of the proposed method is to extract the features to input to the classifiers.

HOG, one of the popular tools by the nature of using the gradient distribution in complex form, is used for feature extraction. First, the radar images are binarized. Then, the binarized images are divided into subimages, and complex-valued gradients are calculated for each subimage as given in (5) and (6).

$$\nabla f(x,y) = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} f(x+1,y) - f(x-1,y) \\ f(x,y+1) - f(x,y-1) \end{bmatrix}. \tag{5}$$

$$g = g_x + jg_y, \tag{6}$$

where $f(x,y)$ represents any subimage, g is the complex-valued gradient, and g_x and g_y are the gradients in x and y directions, respectively. The histograms are formed as the distribution of the amplitudes of the complex-valued gradient g for predefined angle values. The final HOG feature is a vector that is the concatenation of each histogram. The whole signal processing steps of the proposed method are summarized in Figure 5.
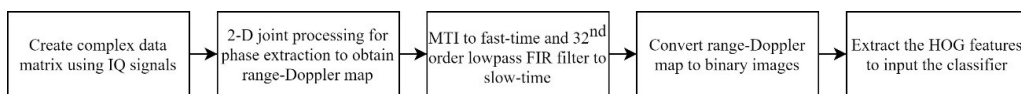
Figure 5. The signal processing steps of the proposed method.

3. Experiments and results

While detecting the targets behind walls, ghost targets may appear at locations where there should be none [18]. This paper investigates the discrimination performance of the ML algorithms for this problem. Synthetic ghost targets have been created with moving plates imitating human breathing. A total of 140 observations have been recorded at 70 different locations for human and human-like targets. A 20-cm wall has been created with gas concentrate blocks as the obstacle between the radar and the targets. The attenuation of the wall has been recorded as around 10 dB for the operating bandwidth of 2–4 GHz. Wall attenuation has been measured as given in Figure 6.

The radar was placed next to the wall and worked for target detection for an observation time of 30 s. The recorded baseband signals have been converted into radar images to be classified. First, popular ML algorithms have run with range-slow time images. Then, Doppler-FFT has been applied to range-slow time images as an additional phase extraction. The resulting range-Doppler images have been evaluated as inputs for the classifiers. Both image types have been used directly in their vectorized form or with their HOG features. The performances of the classifiers have been compared for each input type. A total of 80% of the input images have been used for training while the rest have been kept for the testing phase. Five-fold cross-validation has been preferred to validate that the classifiers learn all images. The experiment environment and measurement setup are shown in Figure 7.
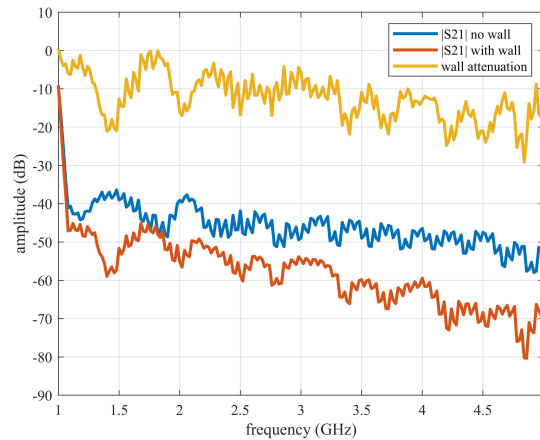

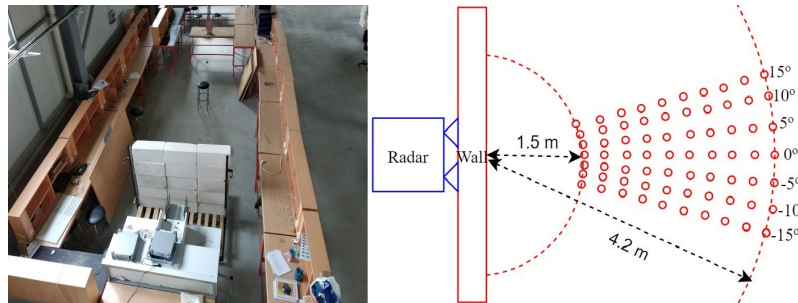**Figure 6**. The atteunation of the wall (1–5 GHz).

Figure 7. Experiment environment and the measurement setup (o: target locations)
.

It has been investigated how additional phase extraction and HOG+PCA affect classification performance and fit which ML method best. Classifiers have run for both range-slow time and range-Doppler images to see the effect of the additional phase extraction. The m-by-n image is converted into a vector with a length of m × n. Well accepted ML classifiers such as decision trees (DTs) [24], support vector machines (SVMs) [25], k-nearest neighborhoods (KNNs) [26], and ensemble approaches [27, 28] are used for distinguishing the human and human-like returns by using these vectors. The algorithms have been executed with different parameters. The settings providing the best results for each method are tabulated in Table 1.

Table 1. Parameter settings providing the best results for each ML algorithm.

	Decision tree	**SVM**	**KNN**	**Subspace discriminant**
Settings	maximum # of splits: ”4”	kernel function: ”quadratic”	# of neighbors: ”10”	learner type: ”discriminant”
	split criterion: ”Gini's diversity index”	box constraints level: ”1”	distance metric: ”cosine”	# of learners: ”30”
	surrogate decision splits: ”off”	multiclass method: ”one-vs-one”	distance weight ”equal”	subspace dimentsion: ”6912”

The performances of the classifiers are compared with common metrics like accuracy, specificity, sensitivity, and F1 score. These metrics are represented as given below.

•$Accuracy = (TP + TN)/(TP + FP + FN + TN)$

•$Specificity = TN/(TN + FP)$

•$Sensitivity = TP/(TP + FN)$

•$Precision = TP/(TP + FP)$

•$F1score = 2 \times (Sensitivity \times Precision)/(Sensitivity + Precision)$

Here, true positive (TP) and true negative (TN) values give the number of correctly classified human and human-like returns, respectively. On the other hand, false positive (FP) and false negative (FN) values correspond to the misclassified returns. The performances of the best classifiers in each ML group are given in Table 2.

Table 2 indicates the effect of additional phase extraction. The performances of the classifiers are shown for both input types. As seen from Table 2, the performances of the classifiers are improved with additional phase extraction, Doppler-FFT. Providing the range-Doppler images rather than range-slow time images has improved the accuracy and F1 score values with an average of %5.36 and %5.26, respectively. Although the accuracies and F1 scores seem satisfactory, differences between sensitivity and specificity values point to an unbalanced training for all methods and each input type. According to these results, additional phase extraction improves the accuracy of distinguishing the human and human-like targets but cannot solve the tendency to unbalanced learning.

In the second scenario, HOG features of the images are input to the classifiers rather than the direct use of the vectorized images. The cell size, which determines the size of the subimages, is determined as 4-by-4 by trial and error. The length of the feature vector is reduced by PCA by keeping only the components that explain 95% of the variance. The results are given in Table 3 for both input types.

Table 3 shows us that HOG feature extraction fits better with range-Doppler images rather than range-slow time images. Compared to the direct use of range-slow time images, the proposed approach increases

Table 2. Performance comparison of the MLs for two input types.

Input	ML	Accuracy	Sen.	Spec.	F1 Score	TP	TN	FP	FN
Range-ST	Coarse Tree	73.57	70.00	77.14	72.59	49	54	16	21
	SVM (Quadratic)	86.43	87.14	85.71	86.52	61	60	10	9
	KNN (Cosine KNN)	87.86	82.86	92.86	87.22	58	65	5	12
	Ensemble (Subspace Discriminant)	85.71	84.29	87.14	85.51	59	61	9	11
Range-Doppler	Coarse Tree	83.57	77.14	90.00	82.44	54	63	7	16
	SVM (Quadratic)	90.71	**88.57**	92.86	90.51	**62**	65	5	8
	KNN (Cosine KNN)	88.57	84.29	92.86	88.06	59	65	5	11
	Ensemble (Subspace Discriminant)	**92.14**	**88.57**	**95.71**	**91.85**	**62**	**67**	**3**	**8**

Table 3. Performance comparison of MLs when HOG+PCA applied to input images.

Input	ML	Accuracy	Sens.	Spec.	F1 Score	TP	TN	FP	FN
Range-ST with HOG+PCA	Coarse Tree	80.71	78.57	82.86	80.29	55	58	12	15
	SVM (Linear)	84.29	85.71	82.86	84.51	60	58	12	10
	KNN (Medium KNN)	81.43	82.86	80.00	81.69	58	56	14	12
	Ensemble (Subspace KNN)	85.00	85.71	84.29	85.11	60	59	11	10
Range-Doppler with HOG+PCA	Coarse Tree	80.71	84.29	78.67	77.14	59	54	16	11
	SVM (Linear)	**93.57**	**94.29**	**92.96**	**92.86**	**66**	**65**	**5**	**4**
	KNN (Medium KNN)	87.14	87.14	87.14	87.14	61	61	9	9
	Ensemble (Subspace KNN)	92.86	**94.29**	91.67	91.43	**66**	64	6	**4**

the obtained maximum accuracy from 87.86% to 93.57 %. Moreover, Table 3 also indicates that the proposed approach provides balanced learning for both classes. When evaluating Tables 2 and 3 together, it is seen that the feature extraction with HOG+PCA cannot improve the accuracy and F1 score for all methods. The approach fits best with SVM with range-Doppler images. The performances of the classifiers are shown in Figures 8 and 9 to visualize Tables 2 and 3.
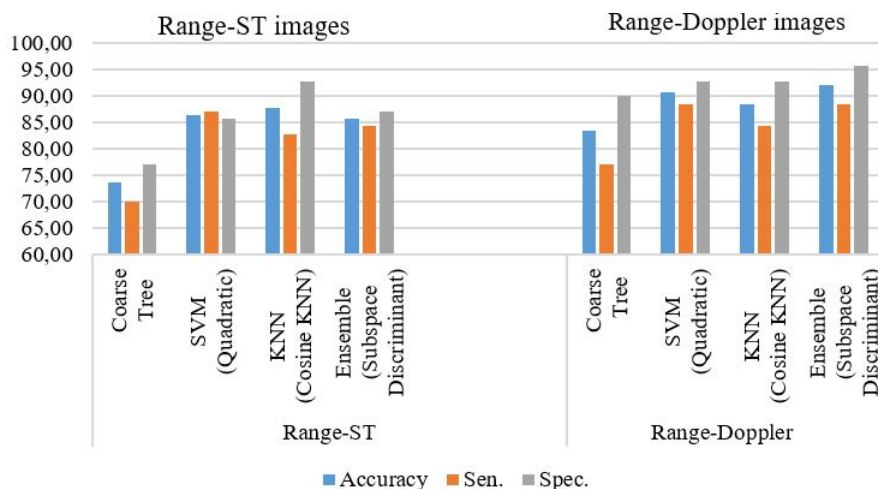
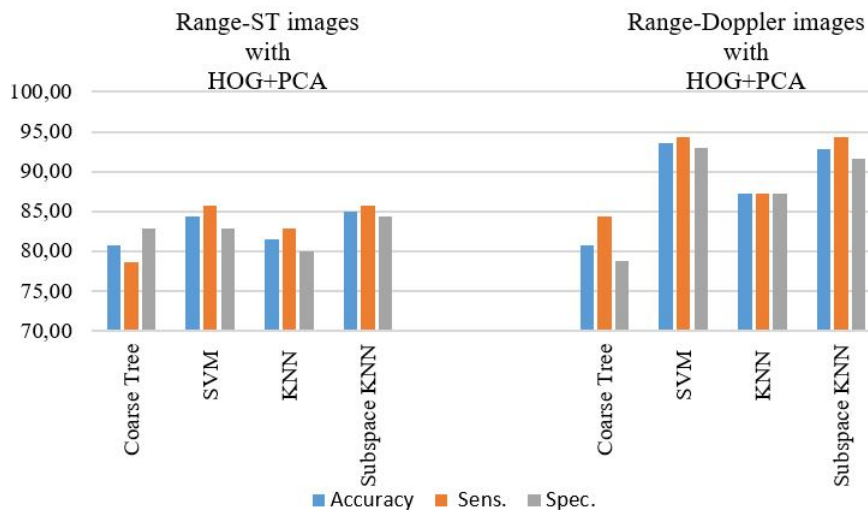**Figure 8**. Performance comparison of the MLs for two input types.

Figure 9. Performance comparison of MLs when HOG+PCA applied to input images.

4. Conclusion

This study presents an approach for distinguishing the human and human-like returns for through-the-wall detection. Unlike most classification studies, SFCW radar structure is preferred rather than CW radars that use micro-Doppler signatures. In CW radars, the time-frequency map is the radar output for the classifiers. Although the time-frequency map is very distinctive and suitable for classification problems, it does not contain the range information of the targets. This is an issue that limits the technology for the future work of multiple-target classification.

In this study, the potential of using the range-Doppler images has been investigated for classification problems. First, an SFCW radar structure has been created with IQ demodulation that helps to prevent the possible null points in range. Then, the processed baseband signals have been converted into a complex-valued matrix. After applying IFFT through the rows of this matrix, the range-slow time images are formed. KNN has provided the best accuracy of 87.86% for human and human-like targets classification. When input images have

been changed to range-Doppler images, the accuracy has increased to 92.14% with the ensemble of subspace discriminant classifiers. However, the unbalanced learning problem has continued. Additional feature extraction has been applied with HOG+PCA to solve this problem. According to the results, HOG+PCA fits best with SVM for the range-Doppler images. The accuracy has increased to 93.57%. While the accuracy has increased, the sensitivity and specificity have reached 94.29% and 92.96%, respectively. These results show that the classifier has learned both human and human-like returns. The study demonstrates using the range-Doppler images with HOG+PCA has a magnificent potential for future works of multiple target detection with the ability to distinguish the human and human-like targets by the range and Doppler information.

Acknowledgment

Contribution of authors

The contributions of the authors are as follows.

Yunus Emre Acar: Methodology, Software, Investigation, Data curation, Writing – original draft, Visualization, Funding acquisition.

İsmail Sarıtaş: Project administration, Funding acquisition, Supervision.

Ercan Yaldız: Conceptualization, Validation, Formal analysis, Writing – review and editing.

References

[1] Vaishnav P, Santra A. Continuous human activity classification with unscented kalman filter tracking using fmcw radar. IEEE Sensors Letters.2020; 4 (5):1-4.

[2] Su W-C, Juan P-H, Chian D-M, Horng T-SJ, Wen C-K et al. 2-d self-injection-locked doppler radar for locating multiple people and monitoring their vital signs. IEEE Transactions on Microwave Theory and Techniques 2020.

[3] Seflek I, Acar YE, Yaldiz E. Small motion detection and non-contact vital signs monitoring with continuous wave doppler radars. Elektronika ir Elektrotechnika.2020; 26 (3):54-60.

[4] Sadreazami H, Bolic M, Rajan S. Contactless fall detection using time-frequency analysis and convolutional neural networks. IEEE Transactions on Industrial Informatics.2021.

[5] Hashim HA, Mohammed SL, Gharghan SK. Accurate fall detection for patients with parkinson's disease based on a data event algorithm and wireless sensor nodes. Measurement.2020;107573.

[6] Kuutti J, Paukkunen M, Aalto M, Eskelinen P, Sepponen RE. Evaluation of a doppler radar sensor system for vital signs detection and activity monitoring in a radio-frequency shielded room. Measurement.2015; 68:135-142.

[7] Dremina MK, Anishchenko LN.Contactless fall detection by means of cw bioradar. In: Progress in Electromagnetic Research Symposium (PIERS) 2016.

[8] Şeflek I, Yaldiz E. Frekans modüleli sürekli dalga radarıyla simüle edilen hayati sinyallerin temassız tespiti.Avrupa Bilim ve Teknoloji Dergisi.2020; pp.72-77.

[9] Liang X, Wang Y, Wu S, Gulliver T. Experimental study of wireless monitoring of human respiratory movements using uwb impulse radar systems. Sensors. 2018; 18 (9):3065.

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

[11] Zhang Z, Nian Y, Chen J,He M. An Experimental Study to Optimize the Stepped-Frequency Continuous-Wave Radar Parameters for Noncontact Multi-target Vital Sign Monitoring. In: 2019 IEEE International Conference on Computational Electromagnetics (ICCEM). 2019; pp. 1-4.

[12] Shrestha A, Li H, Le Kernec J, Fioranelli F. Continuous human activity classification from FMCW radar with Bi-LSTM networks. IEEE Sensors Journal. 2020; 20 (22): 13607-13619.

[13] Acar YE, Yaldiz E, Saritas I. An algorithm to detect the vital signs of multiple humans in the presence of high static clutters. In: 8th International Conference on Advanced Technologies (ICAT'19).2019; Bosnia and Herzegovina.

[14] Han K, Hong S. Phase-extraction method with multiple frequencies of fmcw radar for human body motion tracking. IEEE Microwave and Wireless Components Letters.2020; 30 (9):927-930.

[15] Acar YE, Saritas I, Yaldiz E. An s-band zero-if sfcw through-the-wall radar for range, respiration rate, and doa estimation. Measurement. 2021; 186:110221.

[16] Amin M. Radar for indoor monitoring: Detection, classification, and assessment. CRC Press. 2017.

[17] Nguyen C, Park J. Stepped-frequency radar sensors: Theory, analysis and design. Springer.2016

[18] Acar YE, Saritas I, Yaldiz E. An experimental study: Detecting the respiration rates of multiple stationary human targets by stepped frequency continuous wave radar.Measurement. 2021; 167:108268.

[19] Ramasubramanian K. Using a complex-baseband architecture in fmcw radar systems. Texas Instruments.2017.

[20] Kearney F, Frizelle D. Complex rf mixers, zero-if architecture, and advanced algorithms: The black magic in next-generation sdr transceivers.2017. Visit analogdialogue com.15.

[21] Fouladi RF,Oncu A. Vital signs modeling for Doppler radar cardiorespiratory monitoring. In: 36th International Conference on Telecommunications and Signal Processing (TSP).2013; pp. 363-366.

[22] Acar YE, Seflek I, Yaldız E. Wavelet based denoising of the simulated chest wall motion detected by sfcw radar. Advanced Electromagnetics.2019; 8 (2):85-91.

[23] Seflek I, Yaldiz E. Düşük Maliyetli Sürekli Dalga Doppler Radarı İle Temassız Yaşamsal Belirti Ölçümü. Konya Mühendislik Bilimleri Dergisi.2020; 8 (9):9-14.

[24] Quinlan JR. Decision trees and decision-making. IEEE Transactions on Systems, Man, and Cybernetics.1990; 20 (2):339-346.

[25] Pisner DA, Schnyer DM. Support vector machine. Machine learning. Elsevier.2020; 101-121.

[26] Fix E, Hodges JL. Discriminatory analysis. Nonparametric discrimination: Consistency properties. International Statistical Review/Revue Internationale de Statistique. 1989; 57 (3):238-247.

[27] Ashour AS, Guo Y, Hawas AR, Xu G. Ensemble of subspace discriminant classifiers for schistosomal liver fibrosis staging in mice microscopic images. Health information science and systems.2018; 6 (1):1-10.

[28] Zhang Y, Cao G, Wang B, Li X. A novel ensemble method for k-nearest neighbor. Pattern Recognition.2019; 85:13-25.