IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 9, 2018, accepted May 17, 2018, date of publication June 7, 2018, date of current version June 20, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2843391

Dictionary Learning With Low Computational Complexity for Classification of Human Micro-Dopplers Across Multiple Carrier Frequencies

SHELLY VISHWAKARMA, (Student Member, IEEE), AND SHOBHA SUNDAR RAM, (Member, IEEE)
Indraprastha Institute of Information Technology, Delhi, New Delhi 110020, India

Corresponding author: Shelly Vishwakarma (shellyv@iiitd.ac.in)

This work was supported in part by the Department of Science and Technology, Government of India, in part by the DST Inspire Fellowship, and in part by the Air Force Office of Scientific Research (AFOSR), AOARD, under Grant 5IOA036 FA23861610004.

ABSTRACT Recently, several machine learning algorithms have been applied for classifying micro-Doppler signatures from different human motions. However, these algorithms must demonstrate versatility in handling diversity in test and training data to be used for real-life scenarios. For example, situations may arise where the propagation channel or the presence of interference sources in the test site will permit only specific frequency bands of radar operation. These bands may differ from those used previously while training. In this paper, we examine the performances of three sparsity driven dictionary learning algorithms—synthesis, deep, and analysis—for learning unique features extracted from training data gathered across multiple carrier frequencies. These features are subsequently used for classifying test data from another distinct carrier frequency. Our experimental results, from measurement data, show that the dictionary learning algorithms are capable of extracting meaningful representations of the micro-Dopplers despite the rich frequency diversity in the data. In particular, the deep dictionary learning algorithm yields a high classification accuracy of 91% with a very low computational time for testing.

INDEX TERMS Radar, micro-Dopplers, sparse coding, synthesis dictionary learning, deep learning, analysis dictionary learning, classification.

I. INTRODUCTION

Motions of non-rigid bodies give rise to frequency modulations, known as micro-Dopplers, to continuous wave radar signals [1]–[3]. These micro-Dopplers have been observed for different types of ground moving vehicles, humans and animals [4]–[7], wind turbines [8], helicopters and other airborne targets [9]–[11] as well as ships and underwater vehicles [12], [13]. Over the last decade, classification of human motions on the basis of their micro-Dopplers has attracted significant research focus for applications as diverse as law enforcement, indoor tracking, fall detection and assisted living [12]–[27]. However, in all of these cases, the radar system conditions and the propagation channel under training and test scenarios have been nearly identical. To demonstrate the actual usefulness of radar micro-Dopplers in real-world scenarios, we need to consider two factors - One, situations where the test conditions vary significantly from the training conditions. For instance, in indoor tracking, the presence of Wi-Fi or other wireless devices at specific bands (say 2.4 or 5.8GHz) may considerably interfere with the radar [28]. Alternately, the walls being dispersive mediums, may support certain frequencies over others. Therefore, a degree of reconfigurability or versatility is desired in the radar hardware parameters - such as the carrier frequency - which can be achieved using software-defined radar platforms. Simultaneously, the classification algorithms need to use non-heuristic methods to derive features to handle the diversity in the training and test data. Second, the classification must occur real-time. In this paper, we present a low computationally complex dictionary learning framework for classifying different human motions using diverse multi-frequency training and test data.

Micro-Doppler data from non-rigid targets have been successfully represented using data independent transforms such as DCT, Fourier, and wavelets for classification [9], [14], [19], [23]. However, these fixed dictionaries are not suitable for our objective of classifying radar data gathered from different training and test conditions. For instance, [14] used hand-crafted features such as the maximum and minimum spreads for classification. As a result, the training and test features can significantly differ when they are gathered at different carrier frequencies, even within the same target motion class. The same limitations apply to non-heuristic feature extraction techniques such as principal component analysis (PCA) and independent component analysis (ICA) [29], [30]. Moreover, most of the existing works utilize the spectrograms of the micro-Doppler returns to classify the data. Here, the short time window (also known as dwell time) is a critical parameter, which is usually heuristically selected to represent data from a specific type of dynamic target. Representation learning [31] provides an alternative to canonical feature extraction techniques. Here, unique dictionaries are learned directly from the raw radar data from each target class. The algorithms attempt to extract fundamental features of particular target motion and are therefore less likely to be influenced by the changes in radar system conditions between training and test scenarios.

Dictionary learning (DL) is a technique that falls under the broader category of representation learning where a set of basis are customized to different classes of signals in a sparse fashion. As a result, the dictionaries can extract meaningful features from the raw radar returns, usually time-domain, and capture their uniqueness. DL has shown promising results for diverse applications such as denoising [32], audio processing [33], [34], energy disaggregation [35] and classification [36]–[39]. In the radar community, [40] used a robust dictionary learning method, Label Consistent K-SVD (LC-KSVD), for classifying the micro-Doppler signatures of helicopters, and [41] used a label discriminative sparse representation classifier for joint fall and aspect angle recognition. In a previous paper, we studied the performance of dictionary learning for single-channel source separation of radar micro-Dopplers [42]. In this paper, we examine three recent dictionary learning algorithms for classifying human micro-Doppler data gathered across multiple carrier frequencies. These are the synthesis dictionary learning (SDL) [43]–[45], the deep dictionary learning (DDL) [46] and the analysis dictionary learning (ADL) algorithms [47]–[50]. This paper is an extension to [51] where we presented preliminary classification results for SDL. In SDL, we learn to express the training signals from each class using a linear combination of a few dictionary atoms. These dictionaries are fine-tuned to the underlying signals and are useful signatures for discerning the right target during classification. The SDL uses a single layer sparse representation for each target class. In emerging learning methodologies, the depth of representation is perceived as a key aspect towards successful classification. Therefore, we extend the shallow dictionary learning, in SDL, into multiple levels also known as deep dictionary learning. Here, the representation from each layer acts as an input to the subsequent layer. Each succeeding layer, thus, requires fewer features. The classification is carried out using the representation from the last layer and thus requires a much lower computation time than SDL. The ADL is an alternative paradigm to the SDL. Here, a dictionary directly operates on the data to obtain its sparse representation. The key difference is that during the test phase, the sparse features required for classification are directly obtaining without requiring any inverse operation. As a result, the ADL involves the least computational cost and time. All three algorithms, trained with data from a set of carrier frequencies, are used to classify test data from another distinct carrier. We benchmark the performance of the proposed data features with others such as LC-KSVD [40], PCA [29], physical features (PF) [14], DCT coefficients [52], and cepstral coefficients [53] in terms of their classification accuracy and computational time.

The paper is organized in the following manner. In the succeeding section, we describe the three different dictionary learning algorithms. In Section.III, We present the experimental set up that was used for measurement data collection using a reconfigurable micro-Doppler radar. The classification results across multiple carrier frequencies are presented in Section.IV along with detailed analysis. We also discuss the performances of the proposed algorithms in Section.IV and conclude in Section.V.

II. THEORY

In this section, we present the synthesis, deep and analysis dictionary learning algorithms for real-time classification of micro-Doppler data across multiple carriers.

A. SYNTHESIS DICTIONARY LEARNING (SDL)

The objective of the algorithm is to first learn to represent the time-domain micro-Doppler signal, y_c, of c^{th} dynamic target with only few basis vectors, D_c, and a sparse coefficient vector, z_c

$$y_c = D_c z_c \tag{1}$$

Subsequently, D_c will be used for classification purposes - both at the training and at the test stages. Fig. 1 provides a pictorial representation of the equation. The red columns of dictionary matrix D_c, are linearly combined with coefficient values, colored red in z_c, to represent a signal vector y_c.

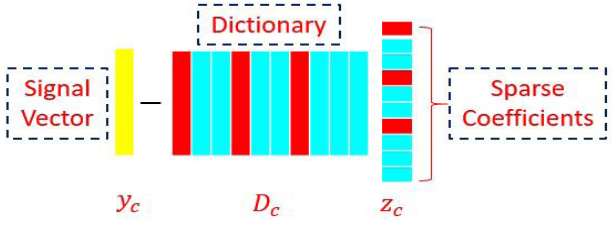

FIGURE 1. Synthesis dictionary learning framework.

As indicated in the figure, z_c is sparse. The learning problem in (1) is non-convex and solved using an alternating minimization technique. Consider a training data matrix $Y_c \in \Re^{N\times M}$. These are micro-Doppler measurements captured at multiple carrier frequencies for c^{th} target class. Each of the M columns represents an independent N-dimensional time-domain radar signal measurement. Learning the corresponding dictionary, $D_c \in \Re^{N\times K}$, and sparse code, $Z_c \in \Re^{K\times M}$, fundamentally involves minimizing the objective function, $J(D_c, Z_c)$, shown in (2).

$$J(D_c, Z_c) = \min_{D_c, Z_c} \|Y_c - D_c Z_c\|_F^2 \quad s.t.\ \|Z_c\|_0 \le \tau \tag{2}$$

Here, K defines the number of atoms in the dictionary D_c and τ, is the sparsity level in Z_c which is controlled by the l_0-norm to ensure sparse representation of Y_c. l_0-minimization is NP-hard [54]. Z_c can be updated using greedy matching pursuit algorithms such as orthogonal matching pursuit (OMP) [55], [56]. Alternately, (2) can be relaxed to a higher order convex l_1-minimization problem shown in (3).

$$J(D_c, Z_c) = \min_{D_c, Z_c} \|Y_c - D_c Z_c\|_F^2 + \lambda \|Z_c\|_1 \tag{3}$$

$\lambda \in \Re$ is the regularization parameter that balances the trade-off between the data representation accuracy and sparsity. We solve (3) using a two-stage iterative procedure. First, we initialize the dictionary matrix using randomly selected signal vectors from the training data, Y_c. Given D_c, the coefficient matrix Z_c is updated using (4) which is known as Least Angle Shrinkage and Selection operator (LASSO) [57].

$$Z_c = \min_{Z_c} \|Y_c - D_c Z_c\|_F^2 + \lambda \|Z_c\|_1 \tag{4}$$

The literature has a plethora of techniques for solving a l_1-minimization problem. In this paper, we solve (4) using the iterative soft-thresholding algorithm (ISTA) discussed in [58]. Once Z_c is obtained, estimating D_c reduces to a least squares problem [59] which is solved using (5).

$$D_c = \min_{D_c} \|Y_c - D_c Z_c\|_F^2 \quad s.t.\ \left\|d_{c,m}\right\|_2^2 \le 1 \quad \forall m = 1, 2, ...M \tag{5}$$

At each iteration, the columns of the dictionary are normalized to have a unit norm to prevent scale ambiguities arising due to differences in strengths of the received signals from targets of varying radar cross-sections. We alternate between (4) and (5) till the algorithm converges.

Then, we concatenate the dictionaries from all of the C classes to form a single, over-complete $D \in \Re^{N\times KC}$ shown in (6).

$$D = [D_1\, D_2\, D_3\, ..D_C] \tag{6}$$

This aggregate dictionary is used to generate the sparse coefficient matrix, $\tilde{Z}_c \in \Re^{KC\times M}$, for each class, c

$$\tilde{Z}_c = \min_{\tilde{Z}_c} \left\|Y_c - D\tilde{Z}_c\right\|_F^2 + \lambda \left\|\tilde{Z}_c\right\|_1 \tag{7}$$

Note that $\tilde{Z}_c$ is distinct from Z_c obtained from (4). We hypothesize that if the target belongs to target class c, the representation must be a linear combination of few atoms of the corresponding dictionary D_c. This is shown in Fig. 2 Where the red columns show the active atoms of the aggregate dictionary and corresponding red rows in $\tilde{Z}_c$ indicate the non-zero values of the coefficient matrix. Therefore, $\tilde{Z}_c$ will mostly have significant values at similar positions for target class c while the remaining values will be either zero or negligibly small. In other words, $\tilde{Z}_c$ will exhibit distinct row sparsity patterns for different target classes c amid the cluster of coefficients belonging to all classes. The columns of sparse coefficient matrices, $\tilde{Z}_c$, from all the target classes are used as input training features for a support vector machine (SVM) classifier with a linear kernel function (consistent across all the algorithms). The SVM is a popular classifier that has been extensively used for numerous machine learning applications.

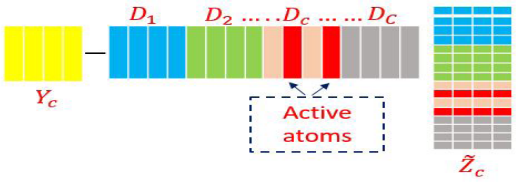

FIGURE 2. Synthesis dictionary learning framework for multiclass classification.

Therefore, z_{test}, is classified by the SVM which is trained using $\tilde{Z}_c$.

Each radar test signal, y_{test}, is a single time-domain micro-Doppler measurement at a distinct carrier frequency from those used while training. We find the sparse features vector, $\tilde{z}_{test} \in \Re^{KC\times 1}$, from y_{test} using D as shown in (8).

$$\tilde{z}_{test} = \min_{\tilde{z}_{test}} \|y_{test} - D\tilde{z}_{test}\|_F^2 + \lambda \|\tilde{z}_{test}\|_1 \tag{8}$$

Similar to the training phase, the intuition here is that if the test sample y_{test} belongs to c class, the signal will be sparsely represented using few atoms of D_c. As a result, the entries in the sparse coefficient vector, $\tilde{z}_{test}$, corresponding to the other classes, will be either zero or negligibly small. Meaning, the sparsity pattern in $\tilde{z}_{test}$ is most likely to resemble columns of $\tilde{Z}_c$ among all the classes.

B. DEEP DICTIONARY LEARNING (DDL)

The SDL represents the micro-Doppler data through a single layer representation. Recent research on deep learning, mostly in neural networks, suggest that we can extract more fundamental or meaningful features through more profound representations of data. On similar lines, we extend the single layer dictionary framework to multiple layers of dictionaries [46]. The first step in the DDL algorithm is identical to the SDL algorithm. Training samples, Y_c, are used to learn the dictionary and coefficient matrices, D_c^1 and Z_c^1, for the first level of DDL using the alternating minimization technique

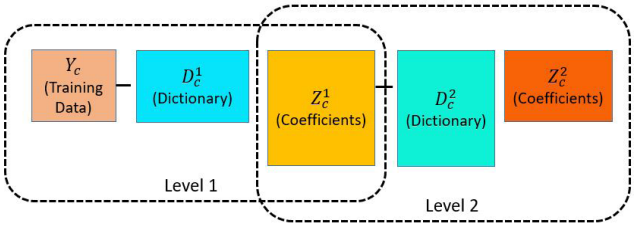

FIGURE 3. Deep dictionary learning framework.

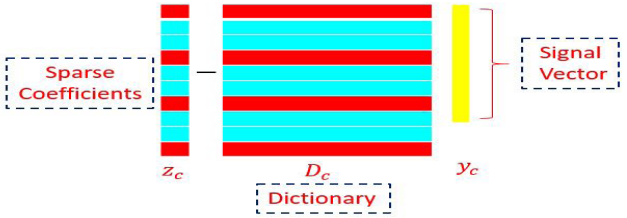

FIGURE 4. Analysis dictionary learning framework.

described in (4) - (5).

$$J(D_c^1, Z_c^1) = \min_{D_c^1, Z_c^1} \left\| Y_c - D_c^1 Z_c^1 \right\|_F^2 + \lambda_1 \left\| Z_c^1 \right\|_1 \quad (9)$$

For each successive n^{th} layer ($n = 2$ on wards to N total layers), we learn the corresponding dictionary and coefficient matrices from the coefficients of the previous layer such that $Z_c^{n-1} = D_c^n Z_c^n$ as shown in Fig. 3.

$$J(D_c^n, Z_c^n) = \min_{D_c^n, Z_c^n} \left\| Z_c^{(n-1)} - D_c^n Z_c^n \right\|_F^2 + \lambda_n \left\| Z_c^n \right\|_1 \quad (10)$$

The class dictionary, D_c, is formed as a product of the multi-level dictionaries shown in (11).

$$D_c = D_c^1 \times D_c^2 \times D_c^3 ..... \times D_c^N \quad (11)$$

The size of the class dictionary, D_c, learned using an N-layer deep architecture is substantially reduced from the single layer class dictionary learned in SDL. In other words, as we opt for deeper networks, we need fewer features for representing the radar signals. The computation time in the training stage of DDL is greater than SDL due to the incorporation of the additional layers of learning. The class dictionaries from multiple classes are concatenated to form D which is subsequently used for classification in a similar manner to (6)-(8). The sparse coefficient vector $\tilde{Z}_c$ obtained from the training micro-Doppler data Y_c, is used to train the SVM. Once the model is trained, $\tilde{z}_{test}$, is classified similarly to SDL. Due to the reduced dimensionality of $\tilde{z}_{test}$ compared to SDL, the deep learning framework is faster during the test stage.

C. ANALYSIS DICTIONARY LEARNING (ADL)

The algorithms discussed in Sections.II-A and II-B belong to the synthesis framework where a signal y_c is synthesized with the linear combination of only a few atoms of the dictionary, leading to its sparse representation. An alternate generative framework - known as the analysis framework - is where a dictionary operates on y_c to generate its sparsest form $z_c = D_c y_c$ as shown in Fig.4. Unlike SDL which focuses on classification based on sparsity (number and position of non-zero) patterns in the representation of the signals, here, the co-sparsity patterns (number of zeros) in z_c are utilized for distinguishing between multiple classes. The blue rows indicate these in the figure. The hypothesis here is that the signals belonging to different target classes will exhibit unique co-sparsity patterns and thus belong to distinct subspaces from which they can be classified. The objective of ADL is to first learn a unique dictionary $D_c \in \Re^{P \times N}$ for each target class from the corresponding training data $\hat{Y}_c \in \Re^{N \times M}$. When D_c operates on $\hat{Y}_c$, it produces sparse features $Z_c = D_c \hat{Y}_c$. As mentioned before, M indicates the number of independent time-domain measurements consisting of N samples. To obtain high co-sparsity in Z_c, the rows of D_c should exhibit high linear dependencies. Learning D_c involves the minimization of the following objective function -

$$J(D_c, \hat{Y}_c) = \min_{D_c, \hat{Y}_c} \left\| Y_c - \hat{Y}_c \right\|_F^2 \quad s.t. \left\| D_c \hat{Y}_c \right\|_0 \leq \tau \quad (12)$$

We relax the l_0 constraint to its nearest convex form by taking l_1 norm.

$$J(D_c, \hat{Y}_c) = \min_{D_c, \hat{Y}_c} \left\| Y_c - \hat{Y}_c \right\|_F^2 + \lambda \left\| D_c \hat{Y}_c \right\|_1 \quad (13)$$

We propose to solve (13) using a variable splitting technique. To solve the problem more efficiently, we introduce a proxy variable Z_c such that the new objective function becomes

$$J(D_c, \hat{Y}_c) = \min_{D_c, \hat{Y}_c} \left\| Y_c - \hat{Y}_c \right\|_F^2 + \lambda \left\| Z_c \right\|_1 \quad s.t. \; Z_c = D_c \hat{Y}_c \quad (14)$$

Here, solving exact Lagrangian is not desired.Therefore we formulate the final objective function as augmented Lagrangian [50] shown in (15)

$$J(D_c, \hat{Y}_c, Z_c) = \min_{D_c, \hat{Y}_c, Z_c} \left\| Y_c - \hat{Y}_c \right\|_F^2 + \lambda \left\| Z_c \right\|_1 + \mu \left\| Z_c - D_c \hat{Y}_c \right\|_F^2 \quad s.t. \; \left\| d_{c,m} \right\|_2^2 \leq 1 \quad \forall m = 1, 2, \ldots M \quad (15)$$

Here, λ acts as a regularizer trading off between representation error and sparsity. μ is a hyperparameter that controls the equality between Z_c and its representation $D_c \hat{Y}_c$. For a high value of μ, equality is enforced. Otherwise, the constraint is relaxed. We further ensure that each row of D_c is constrained to have a unit norm to prevent scale ambiguities due to variations in the signal strengths. Similar to synthesis framework we solve (13) using a two-stage iterative framework. We initialize D_c and $\hat{Y}_c$, using randomly selected samples from the training data gathered at multiple carriers.

The co-sparse coefficient matrix is obtained by $Z_c = D_c\hat{Y}_c$. During the first stage of the iterative framework, Z_c and $\hat{Y}_c$ are held constant and D_c is updated using least squares. In the second stage, we solve for $\hat{Y}_c$ using the least squares as shown in (16).

$$\min_{\hat{Y}_c} \left\| Y_c - \hat{Y}_c \right\|_F^2 + \mu \left\| Z_c - D_c\hat{Y}_c \right\|_F^2 \tag{16}$$

This is equivalent to solving the least squares problem defined in (17)

$$\min_{\hat{Y}_c} \left\| \begin{pmatrix} Y_c \\ \sqrt{\mu}Z_c \end{pmatrix} - \begin{pmatrix} I \\ \sqrt{\mu}D_c \end{pmatrix} \hat{Y}_c \right\|_F^2 \tag{17}$$

Finally, we update Z_c using the soft-thresholding method shown in (18).

$$Z_c = soft(D_c\hat{Y}_c, \gamma) \tag{18}$$

Here, $\gamma = \lambda/2\mu$, is the threshold that we select for the co-sparse coefficient vector. The function above is defined as

$$soft(D_c\hat{Y}_c, \gamma) = sign(D_c\hat{Y}_c) \times \max(0, \left| D_c\hat{Y}_c \right| - \gamma) \tag{19}$$

The iterative procedure is continued until the objective function $J(D_c, \hat{Y}_c, Z_c)$ converges to some local minimum. The class dictionaries from different target classes are then concatenated to form an aggregate dictionary D. We train SVM using the co-sparse features $\tilde{Z}_c = DY_c$, corresponding to each class. During the test phase, the SVM classifies the co-sparse feature vector $\tilde{z}_{test} = Dy_{test}$ of the test micro-Doppler data y_{test}. Note that the ADL is significantly faster in generating features at test time as compared to its synthesis counterpart as the feature generation in ADL involves only a simple product operation instead of the inverse operation in (8). This makes the algorithm more suited for real-time applications.

III. MEASUREMENT DATA COLLECTION

In this section, we describe the experimental set up used to generate measurement micro-Doppler data at multiple carrier frequencies. All of the measurements are conducted in indoor, line-of-sight conditions.

A. MEASUREMENT SET UP

Fig. 5 shows the narrowband monostatic radar set up used for measurement data collection. The radar consists of an N9926A Field-fox vector network analyzer (VNA) and two double ridged linearly polarized broadband horn antennas (HF907). The two horns are connected to the two ports of the VNA. We measure the time-domain S_{21}, the inversion loss scattering parameter, which consists of scattered returns off the targets along with noise and background clutter. The transmitted power from the radar is +3dBm. The VNA is configured to the narrow band mode with a bandwidth of 10Hz and a center frequency of any one of five carrier frequencies − 2.4, 3, 4, 4.5 and 5.8 GHz. We considered a set of lower carrier frequencies to avoid aliasing issues since the sampling frequency is limited to 370Hz by the VNA hardware.

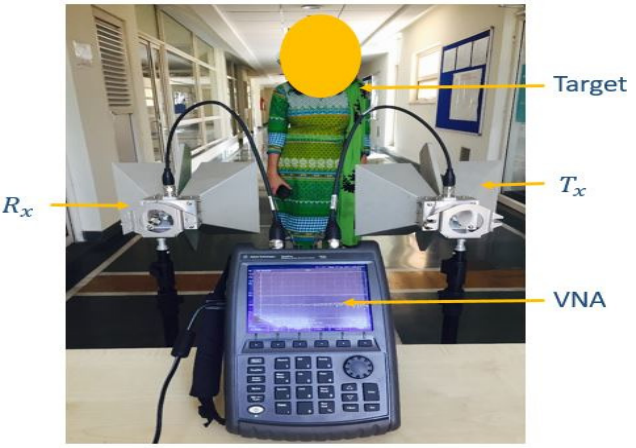

FIGURE 5. Monostatic continuous wave radar configured using a vector network analyzer and two linearly polarized horn antennas for five distinct carrier frequencies − 2.4, 3, 4, 4.5 and 5.8GHz.

TABLE 1. Description of test and training scenarios across multiple carrier frequencies.

Folds	Test Data Carrier Frequencies	Training Data Carrier Frequencies
Fold 1	2.4GHz	3, 4, 4.5, 5.8GHz
Fold 2	3GHz	2.4, 4, 4.5, 5.8GHz
Fold 3	4GHz	2.4, 3, 4.5, 5.8GHz
Fold 4	4.5GHz	2.4, 3, 4, 5.8GHz
Fold 5	5.8GHz	2.4, 3, 4, 4.5GHz

The duration of each measurement is 27 seconds. The measured data are further segmented to 10 shorter signals, each of 2.7 seconds duration.

We consider four distinct target classes. Three target classes consist of human motions. The fourth target class is a table fan (TF) with rotating blades. We have specifically chosen this distinct fourth target class since a table fan gives rise to micro-Doppler signatures that can contribute to significant clutter in radar tracking of humans in indoor environments. The three human motion categories that we consider are - human walking in front of the radar (HW), two humans walking before the radar (TH) and a person standing and boxing his arms (HB). For each of these categories, we conducted measurements with 20 subjects of different heights, gait patterns, velocities, gender, and ages. These measurements were repeated for each of the five different carrier frequencies mentioned above. The human motions were completely unrestricted. They, therefore, consist of motion transitions such as starting from rest, acceleration, turns, slowing down to halt, etc. Next, measurements were carried out with the table fan at different angular speeds, distances, and orientations concerning the radar at the five carrier frequencies. There are therefore a total of 100 measurements for each motion category of which 80 measurements corresponding to 4 carrier frequencies are used for training the dictionary learning algorithms. The remaining data (20 measurements) corresponding to the fifth carrier frequency are used as test data.

TABLE 2. Measurement Data Description

Target Scenario and Description	Target Parameters	Training Data	Test Data
Two humans walking (TH), Human boxing (HB), Human walking (HW),	Number of humans: 20 Target Heights: 5ft to 6ft Target Velocities: 0.6m/s to 1.5m/s Number of measurements = 100 (5 measurements from each of the 20 subjects)	80 (5 measurements from each of 4 carrier frequencies)	20 (5 measurements from the fifth carrier frequency)
Table fan (TF)	Number of fans: 1 Angular velocities: 1400 rpm, 2000 rpm Locations from the radar : 10 Number of measurements = 100 (5 measurements from each of the 20 cases)	80 (5 measurements from each of 4 carrier frequencies)	20 (5 measurements from the fifth carrier frequency)

Table.2 summarizes the entire data collection. The data are available, for interested readers, on the URL: https://goo.gl/dMz1Ko. The noise floor of the VNA is approximate -110dBm. The targets move before the radar between 1m and 10m distance. If we assume the human to have an average radar cross-section of $1m^2$, then the approximate dynamic range in the measurement data is 30dB (between -68 and -98 dBm). The radar cross-section of the table fan is much lower and can, therefore, result in weaker returns, sometimes close to the noise floor, especially when the fan is at an inclination away from the radar at 10m.

B. MICRO-DOPPLER SPECTROGRAMS ACROSS MULTIPLE CARRIERS

The classification algorithms are posed with some unique challenges when the training and test measurement data are gathered under different conditions. We illustrate these challenges by presenting the micro-Doppler spectrograms for these motions for two carrier frequencies in Fig. 6. The spectrograms are generated using the short-time Fourier transform with a dwell time of 0.05 seconds. The figures on the left column correspond to data generated at 2.4GHz while the figures from the right column correspond to 5.8GHz data. These are the lowest and highest carrier frequencies that we had selected for our experimental purposes. The first row - Fig. 6(a) and Fig. 6(b) - show the spectrograms corresponding to two humans walking before the radar. One human walks towards the radar, with mostly positive micro-Dopplers, while the second walks away from the radar, with mostly negative micro-Dopplers. The legs give rise to higher Dopplers than the arms or torso. These leg micro-Dopplers can occasionally be aliased to lower frequencies especially at 5.8GHz as seen in Fig. 6(b). The next two figures, Fig. 6(c) and Fig. 6(d), correspond to the human walking before the radar. Again, in these figures, we observe that the strongest returns arise from the torso while the limbs contribute to much weaker returns. This is especially evident in Fig. 6(d). This is because of the lower cross-section of the limbs and due to shadowing of one limb by another. Fig. 6(e) and Fig. 6(f) show the micro-Dopplers from a human standing still and boxing his arms. Here the torso Doppler is mostly around 0Hz since there is no translational motion of the human. Similarly, the legs do not have a distinct Doppler. Instead, we observe both the positive and negative Dopplers arising from the swinging motion of the boxer's arms. The last two figures, Fig. 6(g) and Fig. 6(h), are derived from the rotation of the three blades of the table fan. The rotation motion of the blades of the fan gives rise to micro-Dopplers that are a function of the number of blades, the angular speed of rotation, the orientation of the blades with respect to the radar and the length of the blades. Due to the low sampling frequency, the spectrograms of the table fans show significant aliasing across all the carrier frequencies.

These figures highlight the key challenges before the classification problem. The algorithm, trained with data gathered at particular carrier frequencies (say 2.4GHz to 4.5GHz data) must successfully classify data belonging to 5.8GHz. Some of the important factors to be taken into consideration are listed below.

1) The figures demonstrate that the Doppler spectrograms corresponding to high carrier frequency data (in the right column) show finer frequency resolution than those from low carrier frequency data (in the left column). This is particularly evident in the case of the three human motion classes.
2) The sampling frequency is identical for all the measurements. Therefore, there is a much higher probability of aliasing to occur at 5.8GHz (especially if the human moves at a high velocity) when compared to the lower carriers. This problem of aliasing is particularly evident in the case of the table fan spectrograms as pointed out before. This limitation may give rise to some errors in classification.
3) There are differences in the micro-Doppler patterns within the same motion categories arising due to the variation in gait patterns across individuals (due to their height, weight, age, gender, fitness, and mood).
4) Measurement data have issues arising from shadowing of a target or parts of a target by the environment or the presence of other targets.

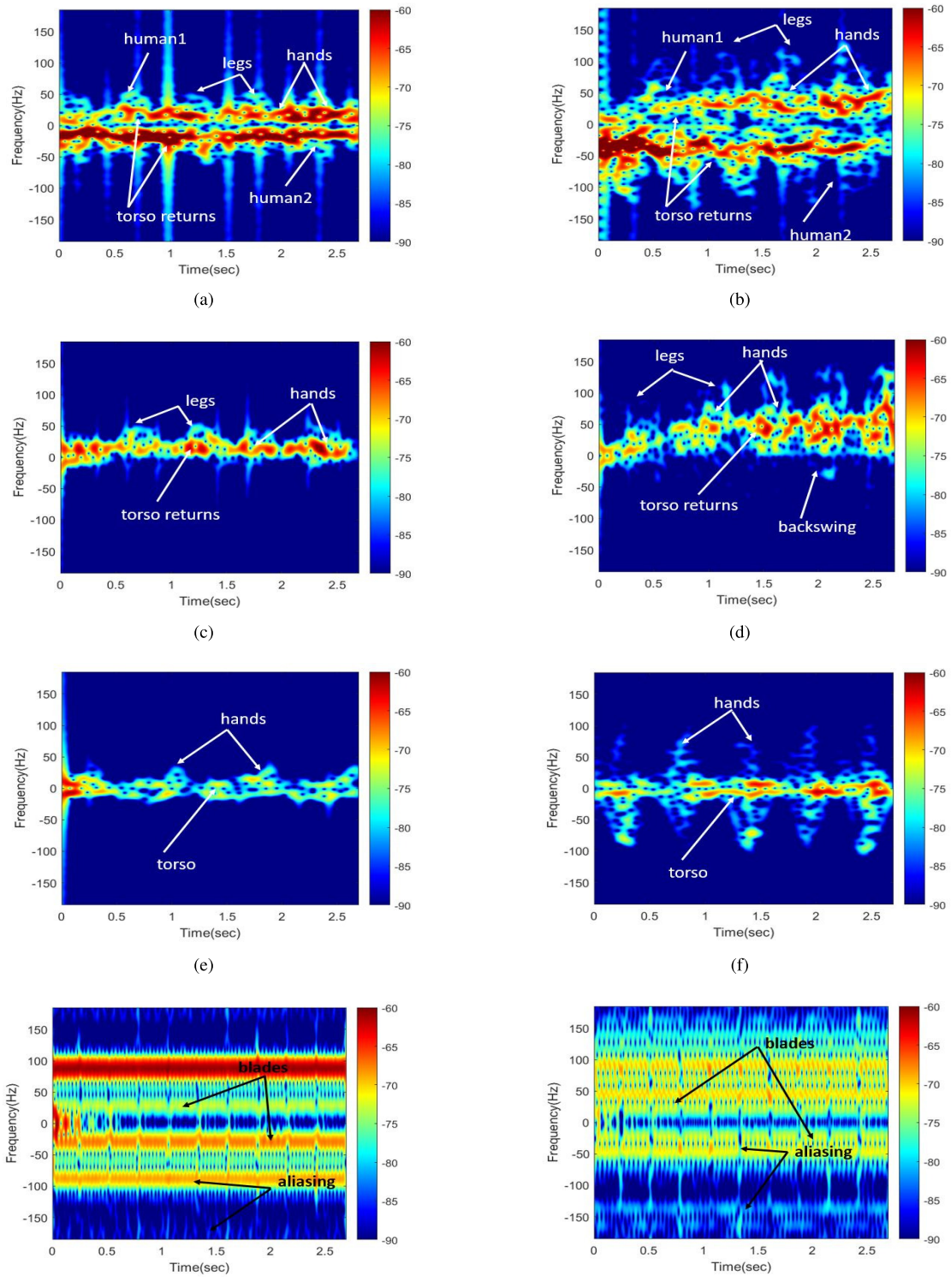

FIGURE 6. STFT spectrograms of (a, b) two walking humans - one walking towards the radar and the other walking away from the radar, (c, d) a human walking towards the radar, (e, f) a human boxing and (g, h) a rotating table fan. The figures in the left columns are generated from data collected at 2.4GHz and figures in the right columns are generated from data collected at 5.8GHz.

5) Even the measurement data gathered from a single individual, show micro-Doppler variations due to motion transitions.

IV. MEASUREMENT RESULTS AND ANALYSES

The measurement data consists of four dynamic target classes- two humans walking (TH), human boxing while

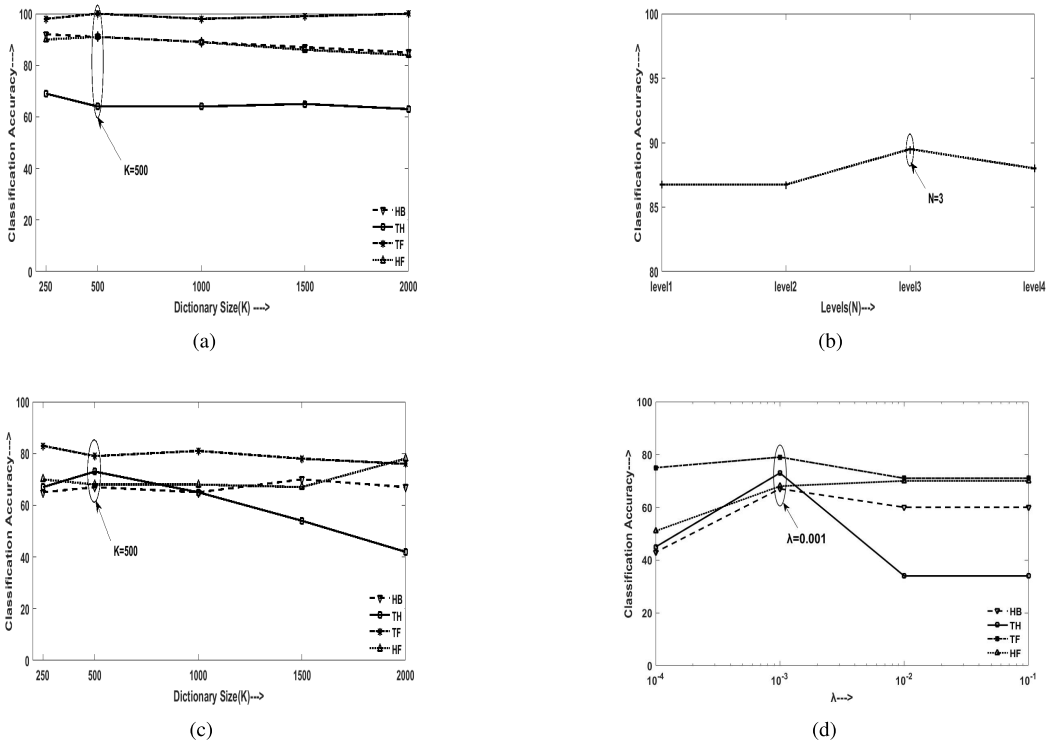

FIGURE 7. Variation of classification accuracy with (a) dictionary size for SDL, (b) depth in dictionary layers in DDL, (c) dictionary size for ADL and (d) Sparsity parameter (λ) for ADL.

standing still (HB), a single human walking (HW) and a rotating table fan (TF). First, we consider the single carrier case. Here, the training and test micro-Dopplers are gathered at the same carrier frequency. Then we discuss the more challenging multi-carrier case. Here, the training features are obtained from micro-Doppler data from 4 out of 5 carrier frequencies while the test features are derived from the fifth carrier frequency.

We apply the three dictionary learning frameworks discussed in Section.II-A, II-B, and II-C to the measurement data to study their effectiveness at classifying micro-Dopplers. We compare their performances with four non-DL based feature extraction methods that have been used for micro-Doppler signature classification in recent literature - physical features (PF), DCT coefficients, cepstral features (CF), and principal components analysis. We also consider another DL based algorithm that was tested recently - the label consistent KSVD (LC-KSVD). The LC-KSVD is also based on the synthesis learning framework. Unlike the SDL, where class dictionaries are learned individually, the LC-KSVD learns multi-class dictionaries jointly with regularizers for class-wise sparsity and inter-class discrimination in its objective function. Secondly, the algorithm uses l_0-norm rather than l_1 minimization techniques. For a detailed description of LC-KSVD, we refer readers to [40]. The algorithms are run in MATLAB 2015b on an Intel(R) Core(TM) i7-5500U CPU running at 2.40 GHz; 16-GB RAM, Windows 10 (64 b).

A. PARAMETER SELECTION FOR DICTIONARY LEARNING

We consider a training matrix of size [1000 × 200] for each target class by randomly repopulating the original measurements. Each signal vector has 1000 time domain samples over a duration of 2.7 seconds. We use the SDL technique described in Section.II-A, to learn under-complete class dictionaries each of size $[1000 \times K]$. These dictionaries from the four classes are then concatenated to form a single aggregate over-complete dictionary of size $[1000 \times 4K]$. The concatenated dictionary is used to generate a class-specific sparse features matrix of size $[4K \times 200]$ which is used to train a support vector machine (SVM) classifier. We hypothesize that radar signals from different classes will exhibit distinctive patterns in this sparse features matrix. In the test phase, we use a single $[1000 \times 1]$ micro-Doppler signal to generate a sparse feature vector of size $[4K \times 1]$. This vector is classified as one of the four classes by the SVM. We examine different sizes of class dictionaries, K, and the results are presented in Figure.7(a). The results show that the algorithm is not very sensitive to the size of K provided the concatenated dictionaries from the four classes (of size $4K$) are over-complete. We choose $K = 500$.

In the DDL framework, we learn an N-layer deep network with each successive layer having a reduced dictionary size. We start with a dictionary size of [1000 × 500] to match that of SDL. We considered different depths as shown in Figure.7(b). Our results show that there is an improvement

TABLE 3. Comparison of average classification accuracy across multiple algorithms when training and test micro-Doppler data are gathered at same carrier frequencies.

Cases	SDL	DDL	ADL	LC-KSVD	PF	CF	DCT	PCA
TH	100	100	100	100	95	100	100	100
HB	100	100	100	100	90	100	100	100
HW	100	100	94	100	80	100	100	100
TF	100	100	100	100	100	100	100	100
Average	100	100	98.5	100	91.25	100	100	100

TABLE 4. Comparison of average classification accuracy across multiple algorithms when training and test micro-Doppler data are gathered at different carrier frequencies.

Cases	SDL	DDL	ADL	LC-KSVD	PF	CF	DCT	PCA
TH	67	68	73	13	89	38	69	58
HB	88	98	67	81	47	63	93	72
HF	93	97	68	25	52	40	56	59
TF	95	99	79	98	100	46	97	37
Average	85.75	90.5	71.5	54.25	72	46.75	78.75	56.5

when $N = 3$ after which we lose the advantage of abstraction. Subsequent dictionary sizes are $[500 \times 250]$ and $[250 \times 100]$. The sparse feature matrix used for training the SVM is of size $[100 \times 200]$ and the sparse feature vector used for testing is of size $[100 \times 1]$. This size is considerably smaller than the SDL. Since the deep learning architecture requires fewer features for classification as we go down the hierarchy the computational complexity during the test phase is reduced.

Finally, in ADL, we learn a class dictionary of size $[K \times 1000]$ that will operate on the training matrix. Again, we examine the effect of different values of K on the performance of the algorithm. The results are shown in Figure.7(c). The results show that the algorithm's performance is not very sensitive to K. We choose $K = 500$ to make it consistent with the other two dictionary algorithms. The concatenated dictionary from all four classes, of size $[2000 \times 1000]$, is used both for training and test. The size of the dictionaries is thus consistent across all the algorithms. The co-sparse feature matrix used for training the SVM is of size $[2000 \times 200]$ and the co-sparse feature vector, used while testing, is of size $[2000 \times 1]$. The choice of λ dictates the trade-off between representation and sparsity error. We show how the performance varies as a function of λ in Figure.7(d). We choose $\lambda = 0.001$ for the best results.

We fix three additional parameters in the LC-KSVD algorithm - the inter-class discrimination that is the weights for label constraint term (α), weights for classification error term (β) and sparsity prior (S) regularizers. The variation of these parameters results in minor changes in the classification accuracy (around 3 to 4 %). We select values $\sqrt{\alpha} = 0.001$, $\sqrt{\beta} = 0.001$ and $S = 10$ that yield the best results.

B. CLASSIFICATION BASED ON SINGLE-CARRIER DATA

The classification results for the five-fold single-carrier frequency case are presented in Table.3. The training and test for each fold are carried out on data from the same carrier. The classification accuracy for each entry in this table is obtained by the average across five folds. The results show 100% average classification success across five folds in all three synthesis algorithms (SDL, DDL, and LC-KSVD) and slightly lower performance for the ADL (average classification accuracy of 98.5%). The results compare favorably with other classification works on micro-Doppler data presented in literature (PF, CF, DCT, and PCA). The results show that the micro-Dopplers for these motions are *distinctive* and any conventional feature extraction technique is suitable for classification when we consider single carrier data.

C. CLASSIFICATION BASED ON MULTIPLE CARRIER DATA

Next, we consider the more challenging scenario - when the training and test data are gathered at different carrier frequencies. We again consider five fold classification. Each fold now represents results for training data gathered from four out of five carrier frequencies and test data from the fifth (remaining) carrier. This is fully detailed in Table.1. The resulting classification accuracies are presented in Table.4.

The results show a significant deterioration in the classification performance of all the algorithms in comparison to the single carrier case. This means that the micro-Dopplers from the multiple motions are no-longer distinctive. Among the non-DL algorithms, the DCT coefficients are most successful at classification. However, the performance is inferior to SDL and DDL algorithms.

To understand this further, we visualize the data reduced to a two-dimensional space using the t-distributed Stochastic Neighbor Embedding technique [60]. We consider the scatter plots of the raw data, the features extracted by DDL and the handpicked features extracted by a non-DL method (PF) as shown in Figure.8. Ideally, the four target class data must belong to four distinct clusters. However, this is not the case for the raw data in Figure.8(a). Figure.8(b), for PF case, shows a very distinct cluster for TF but the features from the three remaining motions overlap considerably. In the case

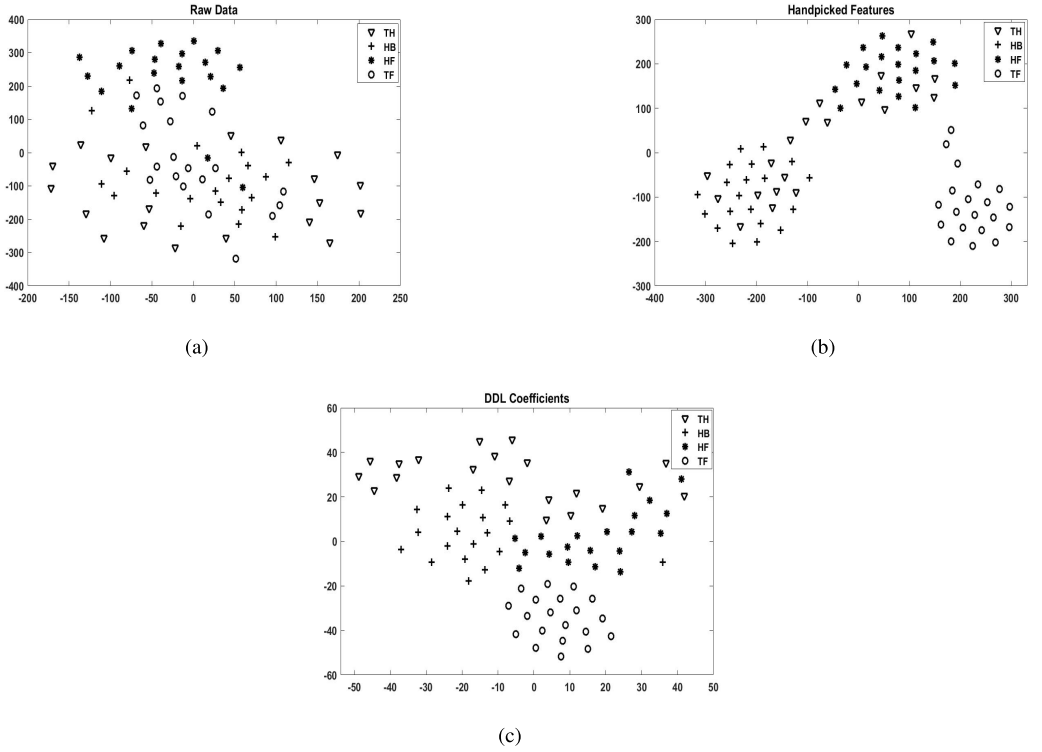

FIGURE 8. Scatter plot to visualize (a) raw data, (b) handpicked features (PF) and (c) features extracted using DDL .

of DDL, in Figure.8(c), The features from the four classes are mostly well separated - except for few cases corresponding to TH and HF. These result in the incorrect classifications seen in Table.6.

We examine the performance of SDL, DDL, and ADL in greater detail in the following sections.

1) SDL

Table.5 shows the confusion matrix of the classification results for the SDL algorithm. Here, the row entries under the header "Test Cases" are the true class labels for the test micro-Doppler data and the column headers are the class labels to which the data are classified. The diagonal entries, therefore, indicate the correct classification results.

The overall accuracy for HB, HW, and TF are superior to TH. TH is mostly confused with HW in all of the five folds. It may be because the radar returns from the second human may be much weaker than the first, in some instances, due to its greater distance from the radar or because of the first subject shadows the second subject. There are also some instances when both the humans move in a synchronized manner with respect to the radar giving rise to overlapped micro-Doppler returns. TH is confused with HB to a lesser extent. This is likely since both of these signatures show both positive and negative micro-Dopplers spreads. For the same carrier frequency, the micro-Doppler spread of the TH is higher than that of the HB due to the absence of micro-Dopplers from legs in the latter case. However, these spreads could be similar when we consider data from different carrier frequencies. This is why the confusion between the TH and HB did not occur in the single carrier frequency case while it occurs in the multi-carrier frequency case. HW and HB are at times confused. Here the confusion arises due to the backswing motion of the arms and legs while walking

TABLE 5. SDL results when training and test micro-Doppler data are gathered at different carrier frequencies.

Folds	Test cases	TH	HB	HW	TF
Fold 1	TH	75	10	15	0
	HB	0	70	25	5
	HW	0	20	80	0
	TF	0	0	5	95
Fold 2	TH	65	5	30	0
	HB	0	95	5	0
	HW	0	15	85	0
	TF	0	0	0	100
Fold 3	TH	45	10	45	0
	HB	0	90	10	0
	HW	0	0	100	0
	TF	0	0	0	100
Fold 4	TH	75	15	10	0
	HB	0	90	5	5
	HW	0	0	100	0
	TF	0	0	0	100
Fold 5	TH	75	0	25	0
	HB	5	95	0	0
	HW	0	0	100	0
	TF	0	0	20	80
	Recall Accuracy	67	88	93	95

TABLE 6. DDL results when training and test micro-Doppler data are gathered at different carrier frequencies.

Folds	Test cases	TH	HB	HW	TF
Fold 1	TH	75	15	10	0
	HB	0	100	0	0
	HW	0	0	100	0
	TF	0	0	0	100
Fold 2	TH	70	5	25	0
	HB	0	100	0	0
	HW	0	5	95	0
	TF	0	0	0	100
Fold 3	TH	55	10	30	5
	HB	0	95	5	0
	HW	0	0	100	0
	TF	0	0	0	100
Fold 4	TH	70	5	25	0
	HB	0	100	0	0
	HW	0	5	90	5
	TF	0	0	0	100
Fold 5	TH	70	0	30	0
	HB	5	95	0	0
	HW	0	0	100	0
	TF	0	0	5	95
	Recall Accuracy	68	98	97	99

TABLE 7. ADL results when training and test micro-Doppler data are gathered at different carrier frequencies.

Folds	Test cases	TH	HB	HW	TF
Fold 1	TH	70	10	20	0
	HB	0	55	25	20
	HW	10	25	60	5
	TF	0	0	5	95
Fold 2	TH	90	0	10	0
	HB	0	70	25	5
	HW	5	25	50	20
	TF	0	30	5	65
Fold 3	TH	75	10	15	0
	HB	0	75	20	5
	HW	0	0	60	40
	TF	0	0	0	100
Fold 4	TH	50	15	35	0
	HB	0	60	30	10
	HW	5	10	80	5
	TF	0	0	5	95
Fold 5	TH	80	5	15	0
	HB	20	75	5	0
	HW	10	0	90	0
	TF	0	0	60	40
	Recall Accuracy	73	66	68	79

that results in some negative Dopplers (when the human is walking towards the radar). TF shows a poorer performance in the fifth fold. This can be attributed due to the similarity of micro-Doppler spreads from both cases due to aliasing of HW at high carrier frequencies.

2) DDL

We present the results of the DDL algorithm in Table.6.

The DDL shows an overall superiority to SDL across the classes - especially HB, HW, and TF. The TH shows the poorest performance and is confused mostly with HW and to a lesser extent with HB. Again, this poor performance can be attributed to the underlying challenge in distinguishing two targets with a radar system of limited dynamic range and frequency resolution. The superiority of the performance of DDL to SDL shows that the deeper representations lead to extraction of more fundamental features from measurement data.

3) ADL

The performance of the ADL is very poor compared to SDL and DDL across all classes except TH as seen in Table.7.

It is an important observation since the TF class has hitherto been classified successfully (above 90%) by the other dictionary learning algorithms. The ADL results in this table mark a significant departure from the previous ADL results reported in Table.3 (measurement data with identical training and test scenarios). From these observations, we infer that the ADL algorithm is heavily impacted by the aliasing in the measurement data, both during training and testing. In particular, aliasing occurs for the TF data across all carriers. The human motions data, on the other hand, are impacted more severely at the higher carrier frequencies. The performance

TABLE 8. Comparison of overall classification accuracy and computational complexity.

Algorithm	Training time (seconds)	Test time (seconds)	Classification accuracy (%)
SDL	414.9	3.4	85.75
DDL	631.5	0.2	90.50
ADL	291.3	0.07	71.50
LC-KSVD	1815.3	0.2	50.25
PF	232.9	0.9	72
CF	2.96	0.4	46.75
DCT	4.15	0.76	78.75
PCA	3.1	0.47	56.5

of the ADL depends on the ability of the algorithm to generate a unique co-sparse representation of the data from each class. To ensure rich co-sparsity in the signal representation, the class dictionaries must exhibit a high degree of row-wise linear dependency. The algorithm fails to achieve this when there are overlaps in the micro-Doppler signal spreads due to aliasing. This problem does not occur in the single carrier case.

Table.8 also shows the computational time required during the test and training phases for all the algorithms. The training phase includes the time taken to learn the training features matrices for the SVM (for SDL, DDL, ADL, PF, CF, and DCT) as well as training the SVM. The training time for the DL algorithms is considerably higher than the non-DL methods since they involve learning class specific dictionaries. In particular, the LC-KSVD takes the longest training time since it requires a l_0-norm computation using a greedy approach. The SDL, DDL and ADL algorithms use the faster l_1-minimization operation instead. The table shows that the DDL algorithm takes longer than the SDL during the training phase since generating class specific dictionaries

includes learning at multiple layers. The ADL and PF take equivalent training time while CF, DCT, and PCA have the lowest training times as these use fixed dictionaries and do not involve any inverse operation. In case of PF, the features are extracted from micro-Doppler spectrograms whose generation consumes most of the training time. The training time, though an important consideration for practical deployment, does not factor into actual radar operation which depends only on the test time.

The test time includes the time taken to generate the test features and the time used by the SVM to classify these features in all the algorithms. The second step is mostly identical across all the algorithms and is approximately 0.01 seconds. We observe that the ADL requires the lowest computational time. It is because the test feature extraction in ADL uses a single matrix multiplication operation which is computationally much more straightforward than the matrix inversion operation in SDL and DDL (8). The DDL takes less time than SDL because the inversion operation in the case of DDL involves a matrix of reduced dimensionality due to the multi-layer dictionary synthesis. The baseline algorithms such as LC-KSVD, PF, CF, DCT, and PCA are also computationally inexpensive. The DDL is therefore comparable to the non-DL algorithms. In conclusion, the ADL algorithm offers some exceptional advantages regarding computational time and complexity. However, its performance, in our study, is limited by radar system issues such as the low sampling frequency. The CF, PCA also lend themselves to real-time operation - but they are not very successful when there are considerable variations between test and training data. The SDL and DDL algorithms, on the other hand, successfully learn unique dictionaries from multi-carrier data despite the system challenges. The DDL, in particular, is suited for real-time micro-Doppler classification due to its short testing time and high classification accuracy.

V. CONCLUSION

We demonstrate three sparse coding-based dictionary learning techniques - SDL, DDL, and ADL - to classify micro-Doppler data from dynamic indoor targets. These algorithms facilitate the representation of radar signals using unique basis vectors even when the training data are gathered from multiple carrier frequencies. As a result, these class dictionaries successfully classify test data from a different carrier frequency from those used while training. This capability makes these algorithms suitable for re-configurable radar platforms for human tracking under diverse operating situations. The SDL and DDL use unique sparsity patterns while the ADL uses the unique co-sparsity patterns of the representations of the radar signals for classification. The computational complexity of the ADL, in the test phase, is much lower than the SDL and DDL. However, in our study, the ADL's performance was limited by radar system issues. As a result of using multi-layered dictionary learning in DDL, the algorithm has better classification accuracy and a much shorter computation time during testing in comparison to the single layer SDL thus making it an ideal candidate for real-world scenarios where low computational complexity and high accuracy are desired.

REFERENCES

[1] V. C. Chen and H. Ling, *Time-Frequency Transforms for Radar Imaging and Signal Analysis*. Norwood, MA, USA: Artech House, 2002.
[2] V. C. Chen, F. Li, S.-S. Ho, and H. Wechsler, "Analysis of micro-Doppler signatures," *IEE Proc.-Radar, Sonar Navigat.*, vol. 150, no. 4, pp. 271–276, Aug. 2003.
[3] V. C. Chen, F. Li, S.-S. Ho, and H. Wechsler, "Micro-Doppler effect in radar: Phenomenon, model, and simulation study," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 1, pp. 2–21, Jan. 2006.
[4] A. Ghaleb, L. Vignaud, and J. M. Nicolas, "Micro-Doppler analysis of wheels and pedestrians in ISAR imaging," *IET Signal Process.*, vol. 2, no. 3, pp. 301–311, Sep. 2008.
[5] J. A. Nanzer and R. L. Rogers, "Bayesian classification of humans and vehicles using micro-Doppler signals from a scanning-beam radar," *IEEE Microw. Wireless Compon. Lett.*, vol. 19, no. 5, pp. 338–340, May 2009.
[6] Y. Li, L. Du, and H. Liu, "Hierarchical classification of moving vehicles based on empirical mode decomposition of micro-Doppler signatures," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 3001–3013, May 2013.
[7] L. Du, L. Li, B. Wang, and J. Xiao, "Micro-Doppler feature extraction based on time-frequency spectrogram for ground moving targets classification with low-resolution radar," *IEEE Sensors J.*, vol. 16, no. 10, pp. 3756–3763, May 2016.
[8] R. Nepal, J. Cai, and Z. Yan, "Micro-Doppler radar signature identification within wind turbine clutter based on short-CPI airborne radar observations," *IET Radar, Sonar Navigat.*, vol. 9, no. 9, pp. 1268–1275, Dec. 2015.
[9] T. Thayaparan, S. Abrol, E. Riseborough, L. Stankovic, D. Lamothe, and G. Duff, "Analysis of radar micro-Doppler signatures from experimental helicopter and human data," *IET Radar, Sonar Navigat.*, vol. 1, no. 4, pp. 289–299, Aug. 2007.
[10] M. K. Bączyk, P. Samczyński, K. Kulpa, and J. Misiurewicz, "Micro-Doppler signatures of helicopters in multistatic passive radars," *IET Radar, Sonar Navigat.*, vol. 9, no. 9, pp. 1276–1283, Dec. 2015.
[11] P. Molchanov, K. Egiazarian, J. Astola, A. Totsky, S. Leshchenko, and M. P. Jarabo-Amores, "Classification of aircraft using micro-Doppler bicoherence-based features," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 1455–1467, Apr. 2014.
[12] Z. Zhang, P. O. Pouliquen, A. Waxman, and A. G. Andreou, "Acoustic micro-Doppler radar for human gait imaging," *J. Acoust. Soc. Amer.*, vol. 121, no. 3, pp. EL110–EL113, 2007.
[13] A. Balleri, K. Chetty, and K. Woodbridge, "Classification of personnel targets by acoustic micro-Doppler signatures," *IET Radar, Sonar Navigat.*, vol. 5, no. 9, pp. 943–951, 2011.
[14] Y. Kim and H. Ling, "Human activity classification based on micro-Doppler signatures using a support vector machine," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1328–1337, May 2009.
[15] B. G. Mobasseri and M. G. Amin, "A time-frequency classifier for human gait recognition," *Proc. SPIE*, vol. 7306, p. 730628, May 2009.
[16] B. Lyonnet, C. Ioana, and M. G. Amin, "Human gait classification using microdoppler time-frequency signal representations," in *Proc. IEEE Radar Conf.*, May 2010, pp. 915–919.
[17] I. Orović, S. Stanković, and M. Amin, "A new approach for classification of human gait based on time-frequency feature representations," *Signal Process.*, vol. 91, no. 6, pp. 1448–1456, 2011.
[18] J. D. Bryan, J. Kwon, N. Lee, and Y. Kim, "Application of ultra-wide band radar for classification of human activities," *IET Radar, Sonar Navigat.*, vol. 6, no. 3, pp. 172–179, Mar. 2012.
[19] D. P. Fairchild and R. M. Narayanan, "Classification of human motions using empirical mode decomposition of human micro-Doppler signatures," *IET Radar, Sonar Navigat.*, vol. 8, no. 5, pp. 425–434, 2014.
[20] F. Fioranelli, M. Ritchie, and H. Griffiths, "Classification of unarmed/armed personnel using the NetRAD multistatic radar for micro-Doppler and singular value decomposition features," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 9, pp. 1933–1937, Sep. 2015.
[21] M. G. Amin, F. Ahmad, Y. D. Zhang, and B. Boashash, "Human gait recognition with cane assistive device using quadratic time-frequency distributions," *IET Radar, Sonar Navigat.*, vol. 9, no. 9, pp. 1224–1230, Dec. 2015.

[22] F. Fioranelli, M. Ritchie, and H. Griffiths, "Aspect angle dependence and multistatic data fusion for micro-Doppler classification of armed/unarmed personnel," *IET Radar, Sonar Navigat.*, vol. 9, no. 9, pp. 1231–1239, 2015.
[23] F. H. C. Tivive, S. L. Phung, and A. Bouzerdoum, "Classification of micro-Doppler signatures of human motions using log-Gabor filters," *IET Radar, Sonar Navigat.*, vol. 9, no. 9, pp. 1188–1195, Dec. 2015.
[24] B. Tekeli, S. Z. Gürbüz, and M. Yuksel, "Information-theoretic feature selection for human micro-Doppler signature classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2749–2762, May 2016.
[25] Y. Kim and T. Moon, "Human detection and activity classification based on micro-Doppler signatures using deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 1, pp. 8–12, Jan. 2016.
[26] Y. Shao, Y. Dai, L. Yuan, and W. Chen, "Deep learning methods for personnel recognition based on micro-Doppler features," in *Proc. 9th Int. Conf. Signal Process. Syst.*, 2017, pp. 94–98.
[27] S. Abdulatif, Q. Wei, F. Aziz, B. Kleiner, and U. Schneider. (2017). "Micro-Doppler based human-robot classification using ensemble and deep learning approaches." [Online]. Available: https://arxiv.org/abs/1711.09177
[28] A. D. Singh, S. Vishwakarma, and S. S. Ram, "Co-channel interference between wifi and through-wall micro-Doppler radar," in *Proc. IEEE Radar Conf. (RadarConf)*, May 2017, pp. 1297–1302.
[29] J. Zabalza, C. Clemente, G. Di Caterina, J. Ren, J. J. Soraghan, and S. Marshall, "Robust PCA micro-Doppler classification using SVM on embedded systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 3, pp. 2304–2310, Jul. 2014.
[30] V. C. Chen, "Spatial and temporal independent component analysis of micro-Doppler features," in *Proc. IEEE Int. Radar Conf.*, May 2005, pp. 348–353.
[31] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.
[32] M. Elad and M. Aharon, "Image denoising via learned dictionaries and sparse representation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, Jun. 2006, pp. 895–900.
[33] M. Zibulevsky and B. Pearlmutter, "Blind source separation by sparse decomposition in a signal dictionary," *Neural Comput.*, vol. 13, no. 4, pp. 863–882, Apr. 2001.
[34] D. Barry, E. Coyle, D. Fitzgerald, and R. Lawlor, "Single channel source separation using short-time independent component analysis," in *Audio Engineering Society Convention 119*. New York, NY, USA: Audio Engineering Society, 2005.
[35] J. Z. Kolter, S. Batra, and A. Y. Ng, "Energy disaggregation via discriminative sparse coding," in *Proc. Adv. Neural Inf. Process. Syst.*, 2010, pp. 1153–1161.
[36] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, 2009, pp. 689–696.
[37] J. Yang, K. Yu, Y. Gong, and T. Huang, "Linear spatial pyramid matching using sparse coding for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2009, pp. 1794–1801.
[38] I. Ramirez, P. Sprechmann, and G. Sapiro, "Classification and clustering via dictionary learning with structured incoherence and shared features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2010, pp. 3501–3508.
[39] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
[40] F. K. Coutts, D. Gaglione, C. Clemente, G. Li, I. K. Proudler, and J. J. Soraghan, "Label consistent K-SVD for sparse micro-Doppler classification," in *Proc. IEEE Int. Conf. Digit. Signal Process. (DSP)*, Jul. 2015, pp. 90–94.
[41] Q. Chen, M. Ritchie, Y. Liu, K. Chetty, and K. Woodbridge, "Joint fall and aspect angle recognition using fine-grained micro-Doppler classification," in *Proc. IEEE Radar Conf. (RadarConf)*, May 2017, pp. 0912–0916.
[42] S. Vishwakarma and S. S. Ram, "Detection of multiple movers based on single channel source separation of their micro-Dopplers," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 1, pp. 159–169, Feb. 2018.
[43] K. Engan, S. O. Aase, and J. Hakon Husoy, "Method of optimal directions for frame design," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5. Mar. 1999, pp. 2443–2446.
[44] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
[45] Q. Zhang and B. Li, "Discriminative K-SVD for dictionary learning in face recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2010, pp. 2691–2698.
[46] S. Tariyal, A. Majumdar, R. Singh, and M. Vatsa, "Deep dictionary learning," *IEEE Access*, vol. 4, pp. 10096–10109, 2016.
[47] R. Rubinstein, T. Peleg, and M. Elad, "Analysis K-SVD: A dictionary-learning algorithm for the analysis sparse model," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 661–677, Feb. 2013.
[48] S. Nam, M. E. Davies, M. Elad, and R. Gribonval, "Cosparse analysis modeling—Uniqueness and algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 5804–5807.
[49] S. Shekhar, V. M. Patel, and R. Chellappa, "Analysis sparse coding models for image-based classification," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 5207–5211.
[50] S. Singh and A. Majumdar, "Analysis co-sparse coding for energy disaggregation," *IEEE Trans. Smart Grid*, to be published, doi: 10.1109/TSG.2017.2743763.
[51] S. Vishwakarma and S. S. Ram, "Dictionary learning for classification of indoor micro-Doppler signatures across multiple carriers," in *Proc. IEEE Radar Conf. (RadarConf)*, May 2017, pp. 0992–0997.
[52] P. Molchanov, J. Astola, K. Egiazarian, and A. Totsky, "Ground moving target classification by using dct coefficients extracted from micro-Doppler radar signatures and artificial neuron network," in *Proc. Microw., Radar Remote Sens. Symp. (MRRS)*, Aug. 2011, pp. 173–176.
[53] E. Hughes, M. Lewis, and E. Reid, "The application of speech recognition techniques to radar target Doppler recognition: A case study," in *Proc. IET Seminar High Resolution Imag. Target Classification*, 2006, pp. 145–152.
[54] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, vol. 24, no. 2, pp. 227–234, 1995.
[55] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
[56] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4680–4688, Jul. 2011.
[57] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Stat. Soc., B (Methodol.)*, vol. 58, no. 1, pp. 267–288, 1996.
[58] I. W. Selesnick, "Sparse signal restoration," in *Proc. Connexions*, 2006, pp. 1–16.
[59] C. M. Bishop, "Pattern recognition," *Mach. Learn.*, vol. 128, pp. 1–58, 2006.
[60] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.

SHELLY VISHWAKARMA received the B.Tech. and M.Tech. degrees in electronics and communication engineering from Guru Gobind Singh Indraprastha University, Delhi, in 2011 and 2013, respectively. She is currently pursuing the Ph.D. degree in electrical engineering with the Indraprastha Institute of Information Technology Delhi. Her current interests include radar signal processing and machine learning.

SHOBHA SUNDAR RAM received the master's and Ph.D. degrees in electrical and computer engineering from The University of Texas at Austin in 2006 and2009, respectively. She is currently an Assistant Professor with the Indraprastha Institute of Information Technology Delhi. Her research interests are in electromagnetic sensor conceptualization, design, and modeling. She has received two student paper awards at the IEEE Radar Conference in 2008 and 2009, respectively, for her work in through-wall radar tracking of humans. She is a Department of Science and Technology India Inspire Fellow.