MILCOM 2024 - 2024 IEEE Military Communications Conference (MILCOM) | 979-8-3503-7423-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/MILCOM61039.2024.10773904

Tracking and Identification of Targets via mmWave MIMO Radar

Alessandro Vaccari*, Moe Z. Win†, and Andrea Conti*

*Department of Engineering and CNIT, University of Ferrara, Ferrara, Italy

†Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, USA

***Abstract*—Sensing is essential to enable civil, industrial, and military applications that require situational awareness. Simultaneous tracking and identification of heterogeneous device-free targets (e.g., humans, robots, and vehicles) can provide information superiority for different types of operations and surveillance tasks. This paper presents a framework for tracking and identification of multiple device-free targets based on reflected radiofrequency signals. The proposed framework consists of (i) clutter mitigation and target detection relying on the estimated clutter intensity distribution in the environment; (ii) multitarget tracking relying on probabilistic data association; and (iii) neural network-based classification for target identification relying on time-domain representations of micro–Doppler signatures generated by target movements. We performed an experimentation, employing an frequency modulated continuous wave multiple-input–multiple-output radar at mmWaves, which validates the proposed framework. The experimental results, in terms of tracking and identification accuracies, show the benefits of using the proposed framework.**

***Index Terms*—Tracking, classification, data association, neural network, MIMO radar.**

I. Introduction

Sensing via reflected radiofrequency (RF) signals is a key enabler for emerging civil, industrial, and military applications, including human monitoring, factory of the future, and smart surveillance [1]–[5]. While several applications require positional information of connected targets [6]–[9], the 3rd Generation Partnership Project (3GPP) has recently proposed use cases regarding localization and identification of unconnected (i.e., device-free) targets in the Release 19 [10]–[12]. However, to achieve accurate sensing from samples of reflected RF signals is challenging, especially in complex wireless environments characterized by multipath propagation and clutter conditions [13]–[15].

Tracking and classification (e.g., target identification and activity recognition) are independent tasks that can benefit each other. Localization over time of device-free targets is typically performed in a recursive manner, referred to as multitarget tracking (MTT) filtering [16]–[20]. At any time instant, MTT filtering consists of a positional prediction phase, based on a motion model, and an update positional phase based on the collected measurements. Therefore, the knowledge of the target characteristics and activity is crucial for accomplishing a reliable prediction using a specific motion model. On the other hand, the knowledge of the target position reduces the classification space since some types of targets may not be allowed and some actions have a lower probability to be performed in a specific spatial area. In addition, tracking is essential to collect specific signal features from each target to use for target identification.

Multiple-input–multiple-output (MIMO) radars operating at millimeter waves (mmWaves) are effective for sensing based on reflected RF signals [21]–[23]. In particular, frequency modulated continuous wave (FMCW) MIMO radars enable the collection of accurate measurements to provide positional information and classification analytics. The use of probabilistic frameworks to perform MTT enables operation in complex wireless environments where measurements may be affected by clutter and multipath propagation, as well as managing measurements generated by close targets with overlapping trajectories. Neural networks (NNs) are promising to perform target identification based on features of reflected RF signals [24]–[26]. By iteratively updating the model parameters during the learning phase, NNs are able to recognize the most significant identification features from an input data. Then, NN-based approaches allow to achieve high accuracy target identification without requiring a mathematical characterization of the wireless propagation, which is challenging, especially in scenarios characterized by clutter and several scatterers.

The goal of this paper is to develop a tractable framework for tracking and identification of heterogeneous targets via a single FMCW MIMO radar operating at mmWaves. The key idea is to exploit positional information for extracting the time-domain representation of a Doppler signature from reflected chirp signals to exploit for recognizing target characteristics and activities. This paper presents an amenable processing framework for simultaneous tracking and identification involving (i) clutter mitigation and target detection relying on the offline learning of the environmental clutter intensity distribution; (ii) MTT via probabilistic processing; and (iii) NN-based classification of the time-domain representation for Doppler signatures. The key contributions of the paper can be summarized as follows:

- design of a clutter mitigation and target detection approach based on offline learning of the clutter distribution;
- development of a computationally efficient framework to track and identify heterogeneous device-free targets; and

The fundamental research described in this paper was supported, in part, by the Office of Naval Research under Grant N62909-22-1-2009, by the European Union under the NextGenerationEU Italian NRRP PE00000001 program RESTART, by the National Science Foundation under Grant CNS-2148251, and by federal agency and industry partners in the RINGS program.

- quantification of the proposed framework performance based on experimentation in an industrial environment.

We validate the proposed framework using measurements gathered in an industrial environment affected by heavy clutter and severe multipath, employing a single FMCW MIMO radar operating at mmWaves. The MTT accuracy is quantified using the generalized optimal subpattern assignment (GOSPA) metric, while classification in terms of target identification accuracy.

Notations: Random variables are displayed in sans serif, upright fonts; their realizations in serif, italic fonts. Vectors and matrices are denoted by bold lowercase and uppercase letters, respectively. For example, a random variable and its realization are denoted by x and x; a random vector and its realization are denoted by $\mathbf{x}$ and $\boldsymbol{x}$; a random matrix and its realization are denoted by $\mathbf{X}$ and $\boldsymbol{X}$, respectively. Sets and random sets are denoted by upright sans serif and calligraphic font, respectively. For example, a random set and its realization are denoted by X and $\mathcal{X}$, respectively. The m-by-m identity matrix is denoted by $\boldsymbol{I}_m$. The function $\boldsymbol{f}_{\mathbf{x}|\mathbf{y}}(\boldsymbol{x}|\boldsymbol{y})$ and, for brevity when possible, $\boldsymbol{f}(\boldsymbol{x}|\boldsymbol{y})$ denote the probability distribution function (PDF) of $\mathbf{x}$ conditioned on $\mathbf{y} = \boldsymbol{y}$. The function $\varphi(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the PDF of a Gaussian random vector $\mathbf{x}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Integrals are over the entire space of the integration variable.

II. System Model

At each discrete time instant k, consider a random number N_k of targets, which can appear, disappear, and move freely. The k-th target is described by its positional state $\mathbf{x}_k^{(i)} \in \mathbb{R}^{n_x}$, with $i \in \{1, 2, \ldots, N_k\}$ for a realization N_k of N_k. The target positional state $\mathbf{x}_k^{(i)}$ consists of parameters such as position, velocity, and acceleration. The target class is denoted by $\mathsf{c}_k^{(i)} \in \mathcal{C}_\mathrm{t}$, where $\mathcal{C}_\mathrm{t}$ is the set of all possible identity classes. Positional states and classes of tracked targets are denoted by the multitarget positional state $\mathbf{x}_k = \left[\mathbf{x}_k^{(1)\mathrm{T}}, \mathbf{x}_k^{(2)\mathrm{T}}, \ldots, \mathbf{x}_k^{(\mathsf{N}_k)\mathrm{T}}\right]^\mathrm{T}$ and multitarget class $\mathbf{c}_k = \left[\mathsf{c}_k^{(1)}, \mathsf{c}_k^{(2)}, \ldots, \mathsf{c}_k^{(\mathsf{N}_k)}\right]^\mathrm{T}$, respectively.

Consider M_k measurements obtained from sensors at the current time k. Each measurement $\mathbf{z}_k^{(j)} \in \mathbb{R}^{m_z}$, with $j \in \{1, 2, \ldots, M_k\}$ for a realization M_k of M_k, contains positional information to be employed as input for the MTT filtering. The measurement vector $\mathbf{z}_k = \left[\mathbf{z}_k^{(1)\mathrm{T}}, \mathbf{z}_k^{(2)\mathrm{T}}, \ldots, \mathbf{z}_k^{(\mathsf{M}_k)\mathrm{T}}\right]^\mathrm{T}$ contains all the detected measurements, including false alarms. In this work, we consider that: (i) all the measurements are independent of each other; (ii) a measurement is generated by at most one target; and (iii) a target can generate at most one measurement. In the remaining of this section, we describe the signal processing techniques, including clutter mitigation and target detection methods, for collecting a set of measurements when employing a FMCW MIMO radar.

A. FMCW Radar Signal Processing

A single measurement collected via an FMCW radar consists of multiple frequency modulated signals, referred to as chirp signals. The reflected chirp signals contain context information, including the time-domain representation of micro-Doppler shifts generated by the target movements. A time-domain sequence encapsulating the micro-Doppler signal features generated by the target activity is promising to perform classification. In fact, movements and activities generate specific sets of micro-Doppler shifts, which are referred to as Doppler signatures [27], [28]. The time-domain representation of a Doppler signature contains information regarding the target dynamics preserving time correlation. Compared to conventional classification features, such as radar cross section (RCS) and target velocity, the time-domain analysis of Doppler signatures allows to characterize even the micro-movements performed by a specific target and their correlations over time.

Consider an FMCW MIMO radar transmitting linear chirps, i.e., sinusoidal signals with frequencies increasing linearly from f_m to $f_\mathrm{M} = f_\mathrm{m} + B$ over time, where B is the signal bandwidth. The frequency of linear chirp signals increases in a time T_c, referred to as chirp duration, thus with a slope $S = B/T_\mathrm{c}$. A measurement via FMCW MIMO radar consists of $N_\mathrm{c} > 1$ chirp signals, each transmitted every T_c seconds. The chirp signal echoes after multipath propagation and target reflections can be collected at the receiver and then processed to obtain dynamics of the reflecting targets. The n-th received chirp at each antenna can be written as

$$s_\mathrm{R}^{(n)}(t) = A_\mathrm{R}^{(n)} \cos\left(2\pi f_\mathrm{m}(t_n - t_\mathrm{d}) + \pi S(t_n - t_\mathrm{d})^2\right) \quad (1)$$

where $A_\mathrm{R}^{(n)}$ is the n-th signal amplitude, $t_n = t - nT_\mathrm{c}$, and t_d is a time delay representing the time-of-flight (TOF). Specifically, the time delay is $t_\mathrm{d} = 2(r + vt)/c$, where r and v are the range and velocity of the scatterer that reflected the signal and c is the propagation speed of electromagnetic waves.

At the receiver of the FMCW radar, the transmitted and received waveforms are multiplied using a mixer and their product is then filtered using a low-pass filter (LPF). The resulting signal is referred to as intermediate frequency (IF) signal and represents the difference of instantaneous frequency between the transmitted and received waveforms. The IF signal obtained by the transmission of the n-th chirp is given by

$$s_\mathrm{IF}^{(n)}(t) = \frac{A_\mathrm{R}^{(n)}}{2} \cos\left(2\pi(f_\mathrm{p} t_n + f_\mathrm{D} T_\mathrm{c})\right) \quad (2)$$

where $f_\mathrm{p} = 2B\,r/cT_\mathrm{c}$ and $f_\mathrm{D} = 2f_\mathrm{m}\,v/c$. Each IF signal can be sampled with period T_s in N_s samples, while the entire IF signal sequence with period T_c in N_c samples, i.e., one for each transmitted chirp. The sampling processes with periods T_s and T_c are referred to as fast sampling (FS) and slow sampling (SS), respectively.

In MIMO radars, multiple antennas receive the same reflected signal with a time delay that depends on the angle-of-arrival (AOA) of the waveform. Therefore, samples of the IF signal sequence obtained via FS and SS can be arranged in a 3-D tensor, referred to as *radar data cube*, in which FS and SS represent the FS and SS dimensions, respectively, while the third dimension is due to multiple antennas.

By processing the radar data cube via discrete Fourier transform (DFT) we obtain the range-Doppler-angle (RDA)

map, which consists of energy bins for each frequency shift considered over FS, SS, and angular dimension. Let $\tilde{f}_{\mathrm{f}}$, $\tilde{f}_{\mathrm{s}}$, and $\tilde{f}_{\mathrm{a}}$ denote the frequency shifts over FS, SS, and angular dimension, respectively. The range is determined as $\hat{r} = cT_{\mathrm{c}}\,\tilde{f}_{\mathrm{f}}/(2B)$, velocity as $\hat{v}=c\,\tilde{f}_{\mathrm{s}}/(2f_{\mathrm{m}})$, and AOA $\hat{\theta}$ such that $\sin(\hat{\theta})=c\,\tilde{f}_{\mathrm{a}}/(2\pi d\,f_{\mathrm{m}})$, where d is the shift between antennas.

The marginalization of the RDA map with respect to the angular dimension provides a range-velocity representation of the environment, referred to as range-Doppler (RD) map. The RD map is promising to distinguish between energy bins (i.e., frequency shifts) generated by the target presence from those produced by clutter and multipath propagation.

B. Clutter Mitigation and Target Detection

In cluttered environments, target detection is challenging due to RD maps corrupted by signal distortion and noise. The proposed technique for clutter mitigation consists of an offline and an online phase. In the offline phase, the environmental clutter intensity distribution is estimated by the collection of measurements with no target in the area of interest. For each point with coordinate $\boldsymbol{p}$ in the RD or RDA map, a specific threshold is determined based on the clutter distribution previously estimated.[1] In the online phase, RD map elements are compared with the designed thresholds. If the magnitude of the RD map element under test is below the corresponding threshold, its value is forced to zero, otherwise is kept.

The threshold for a point at the coordinate $\boldsymbol{p}$ is defined by considering as Gaussian distributed its noise power $\mathsf{n}_{\boldsymbol{p}}$, i.e., $\mathsf{n}_{\boldsymbol{p}} \sim \varphi\big(n_{\boldsymbol{p}}; \bar{n}_{\boldsymbol{p}}, \sigma^2_{\boldsymbol{p}}\big)$. The distribution parameters can be estimated during the offline learning by the collection of multiple noisy measurements without any target in the monitored area. The threshold $\xi_{\boldsymbol{p}}$ for a point at the coordinate $\boldsymbol{p}$ is given by

$$\xi_{\boldsymbol{p}} = \bar{n}_{\boldsymbol{p}} + \alpha_{\boldsymbol{p}}\,\sigma_{\boldsymbol{p}} \tag{3}$$

where $\alpha_{\boldsymbol{p}} \in \mathbb{R}^+$ tunes the false alarm and misdetection probabilities. A higher $\alpha_{\boldsymbol{p}}$ reduces the false alarm probability, while increases the number of misdetections, and vice-versa.

The high range and velocity resolutions of mmWave FMCW radars and the noise introduced by multipath propagation generate a representation of each detected target as an extended object, i.e., characterized by a set of range-velocity coordinates corresponding to an over threshold magnitudes. The DBSCAN algorithm allows to group such coordinates in clusters by leveraging their density in the RD map [29]. A weighted mean of the cluster elements provides a single-value estimate, referred to as centroid, which consists of a range-velocity coordinate. In particular, by denoting as Ω_j the j-th cluster of coordinates, with $j \in \{1, 2, \ldots, M\}$ and M representing the number of clusters obtained by computing the DBSCAN algorithm, the centroid $\boldsymbol{\mu}_j$ is given by

$$\boldsymbol{\mu}_j = \frac{\sum_{\boldsymbol{p}\in\Omega_j} P_{\boldsymbol{p}}\,\boldsymbol{p}}{\sum_{\boldsymbol{p}\in\Omega_j} P_{\boldsymbol{p}}} \tag{4}$$

[1]For notational convenience, we denote $\boldsymbol{p} \in \mathbb{R}^2$ a RD map coordinate and a $\boldsymbol{p} \in \mathbb{R}^3$ RDA map coordinate.

where $\boldsymbol{p}$ denotes a range-velocity coordinate and $P_{\boldsymbol{p}}$ is the magnitude of the corresponding energy bin in the RD map. The centroid $\boldsymbol{\mu}_j$ is the estimated range-velocity coordinate for representing the cluster j. At a range-velocity coordinate $\boldsymbol{\mu}_j$, the reflection AOA is obtained by considering the angle with highest magnitude in the RDA map.

The knowledge of the target positional states at the previous time instant is exploited to further reduce the number of misdetections and false alarms. From the RDA map, a sub-tensor for each previously estimated positional state can be extracted and, then, processed via the DBSCAN algorithm to enlarge the set of detections. The M_k measurements are composed by the union of centroids collected evaluating both the RD map and the RDA map sub-tensors. Therefore, the set of measurements consist of range-velocity-AOA coordinates, which can be used as input of the MTT filtering.

III. Tracking and Classification Framework

This section presents the proposed framework to perform tracking and classification of device-free targets.

A. Tracking of Multiple Targets

Tracking of multiple device-free targets requires data association, which is the combinatorial problem of determining (i) which target generates a given measurement, (ii) if the measurement is due to false alarm, and (iii) if a target is misdetected. Let $\mathsf{\theta}_k^{(i)}$ be the data association variable denoting the assignment of the measurement j, with $j \in \{1, 2, \ldots, M_k\}$, to the target i, with $i \in \{1, 2, \ldots, N_k\}$, at the time instant k, as follows

$$\theta_k^{(i)} = \begin{cases} j & \text{if target } i \text{ is associated to a measurement } j \\ 0 & \text{if target } i \text{ is associated to no measurement.} \end{cases} \tag{5}$$

In the presence of multiple targets, we define the multitarget data association vector $\boldsymbol{\theta}_k = \left[\mathsf{\theta}_k^{(1)}, \mathsf{\theta}_k^{(2)}, \ldots, \mathsf{\theta}_k^{(N_k)}\right]^{\mathrm{T}}$ providing the measurement-to-target associations for all the tracked targets. In particular, the multitarget data association vector is such that $\boldsymbol{\theta}_k \in \Theta_k$, where Θ_k is the set containing all possible vectors of valid multitarget data associations. A valid multitarget data association requires that (i) each target is either associated to a measurement or misdetected and (ii) any pair of detected targets is not associated to the same measurement. To compute data association and MTT filtering analytically, we exploit the Kalman filter (KF) solution considering linear and Gaussian system models.[2]

Data association can be seen as an optimization problem, where a data association probability $P_{\mathrm{a}}\big(\theta_k^{(i)}\big)$, for linear and Gaussian modeling, is given by

$$P_{\mathrm{a}}\big(\theta_k^{(i)}\big) = \begin{cases} P_{\mathrm{d}}\Big(\boldsymbol{x}_k^{(i)}\Big)\,\varphi\Big(\boldsymbol{z}_k^{(\theta_k^{(i)})}; \boldsymbol{H}_k\tilde{\boldsymbol{x}}_k^{(i)}, \boldsymbol{\Sigma}_k\Big) & \theta_k^{(i)} \neq 0 \\ 1 - P_{\mathrm{d}}\Big(\boldsymbol{x}_k^{(i)}\Big) & \theta_k^{(i)} = 0 \end{cases} \tag{6}$$

[2]For nonlinear non-Gaussian system models, computationally feasible approximate algorithms include variants of the KF such as the extended KF and the unscented KF [30]. Both the extended KF and the unscented KF are based on the Kalman equations but perform different approximations of the system models.

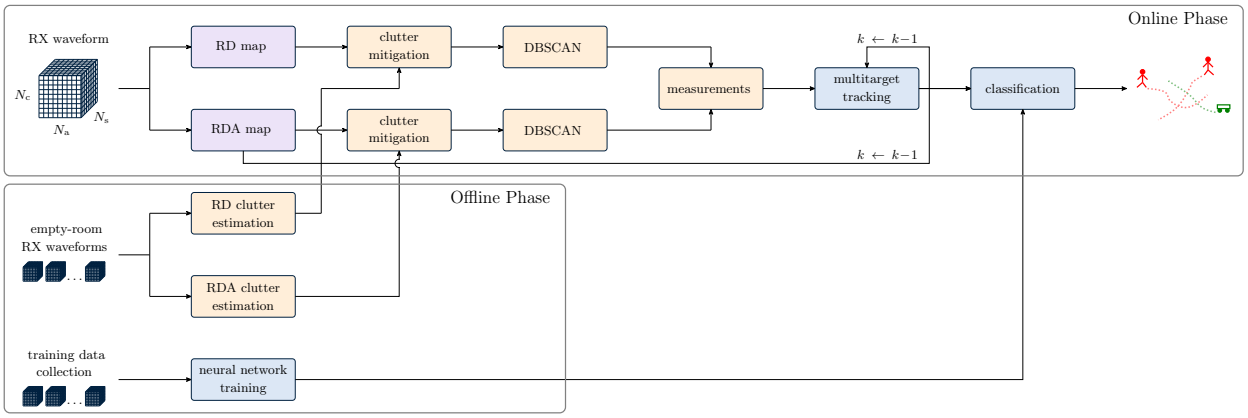

Fig. 1. Block diagram of the presented framework for tracking and classification via single FMCW MIMO radar operating at mmWaves. In particular, lilac blocks represent the signal processing phase, orange blocks the clutter mitigation and detection phase, and blue blocks the tracking and classification phase.

where $P_{\mathrm{d}}(\cdot)$ is the detection probability and $\boldsymbol{H}_k$ is the linear measurement model, while $\hat{\boldsymbol{x}}_k^{(i)}$ and $\boldsymbol{\Sigma}_k$ are the positional prediction and the innovation covariance of the KF, respectively. The data association costs

$$l_{i,j} = -P_{\mathrm{a}}\big(\theta_k^{(i)} = j\big) \tag{7}$$

can be arranged in a matrix $\boldsymbol{L}$, where rows represent targets and columns represent both measurements and misdetections. By processing $\boldsymbol{L}$ using the Murty's algorithm [31], we define $\hat{\Theta}_k \subseteq \Theta_k$ representing the set of the Q-best data association hypotheses obtained evaluating the data association costs.

The joint probabilistic data association (JPDA) algorithm leverages a soft-decision approach for data association by accounting for the uncertainty in matching measurements with tracked targets [32]. In particular, JPDA assigns an association score $\beta_k^{(i,j)}$ to each possible measurement-to-target association, based on how the measurement aligns with the predicted positional state of the target. The association score can be calculated as

$$\beta_k^{(i,j)} = \sum_{\boldsymbol{\theta}_k \in \left\{\boldsymbol{\theta}_k \in \hat{\Theta}_k,\, \theta_k^{(i)} = j\right\}} P_{\mathrm{a}}\left(\theta_k^{(i)}\right) \tag{8}$$

where $j \in \{0, 1, \ldots, M_k\}$. Based on the association score, the distribution of the multitarget positional state for independent targets is given by

$$f(\boldsymbol{x}_k|\boldsymbol{z}_{1:k}) = \prod_{i=1}^{N_k} \sum_{j=0}^{M_k} \beta_k^{(i,j)} f\left(\boldsymbol{x}_k^{(i)}|\boldsymbol{z}_k^{(j)}\right). \tag{9}$$

where $f\big(\boldsymbol{x}_k^{(i)}|\boldsymbol{z}_k^{(j)}\big)$ can be calculated employing the KF for linear and Gaussian models.

Appearing and disappearing targets require the initialization and the deletion of new and existing tracks, respectively. Track initialization is performed from a sequence of measurements collected over time as follows. At the first time instant, all the measurements are considered as potential tracks, i.e., each generates a track hypothesis. At the next time instants consider the greediest data association hypothesis. The measurements that are not assigned to any existing track are either associated to a track hypothesis or considered as new potential tracks. Finally, if a track hypothesis achieves a length of K measurements is validated as real track and processed via the JPDA algorithm. Given that the positional states are represented using a Gaussian density, the covariance matrix serves as a crucial indicator of the plausibility of the track hypothesis. In particular, until the trace of the covariance matrix is below a deleter threshold, the hypothesis is valid. By adjusting the deleter threshold, the same approach can be employed to delete existing tracks processed via the JPDA algorithm.

B. Identification of Multiple Targets

Identification requires the extraction of specific signal features for each target tracked. In particular, MTT filtering provides range-angle coordinates, which can be exploited to determine specific classification analytics for each target from reflected chirp signals. Consider a collected radar data cube, by applying the DFT to the FS and angular dimensions, we obtain range-angle coordinates, each characterized by the N_{c} samples of the SS dimension. Such samples provide a time–domain representation of the micro-Doppler shifts generated by movements at a specific range-angle coordinate.

We consider a classification problem for target identification taking as input a sequence of WN_{c} time-domain samples collected by exploiting target positions estimated via MTT. In particular, the parameter $W \in \mathbb{N}$ defines the duration of the observation window in which the classification analytics are collected. The entire sequence of WN_{c} samples is organized as a linear array to preserve the correlations over time. Increasing W enables performing a more accurate characterization of the target motion.

IV. Experimentation in Industrial Environment

The proposed framework to perform tracking and identification of multiple targets is validated via experimentation using a mmWave MIMO radar in the pilot line of the BI-REX, an Italian competence center for Industry 4.0.

A. System setting

The considered experimentation area is $8.3 \times 6.5\,[\mathrm{m}^2]$ and presents many scatterers, such as industrial machines, tables,

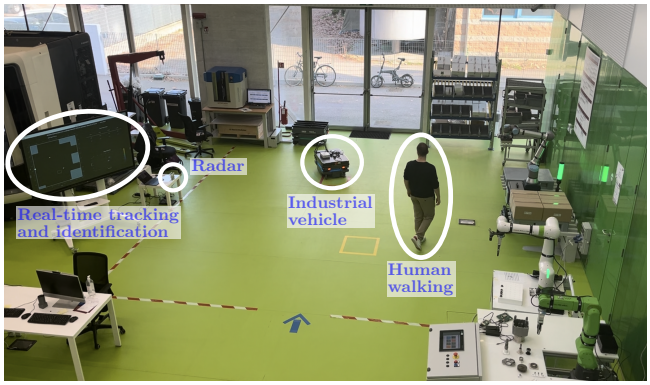

Fig. 2. Experimentation area in the pilot line of the BI-REX with a human walking and an industrial vehicle moving in the monitored area.

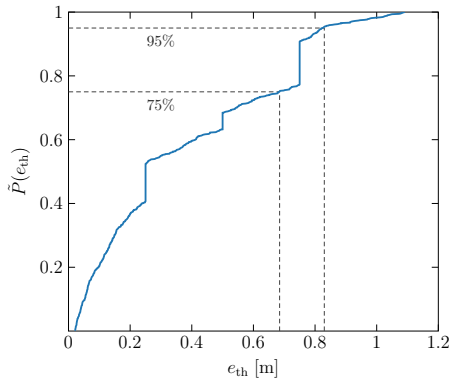

Fig. 3. ECDF of the tracking error evaluated using the GOSPA metric.

and other assets. Fig. 2 shows the scenario during the experimentation. The sensor employed to perform both tracking and identification is an off-the-shelf FMCW MIMO radar with 2 transmitting and 16 receiving antennas forming a linear array. The transmitted signals are linear chirps operating in the frequency range $76 - 77\,\text{GHz}$ with a duration $T_c = 120\,\mu\text{s}$. Each radar measurement consists in transmitting 128 chirp signals and collecting echoes generated by reflections at all the receiving antennas. Each reflected signal is composed by 512 samples. The entire chirp sequence is collected every $0.1\,\text{s}$ by multiple antennas and compose the radar data cube.

The collected radar data cubes are processed via DFT to obtain RD and RDA maps. All the elements in the RD and RDA maps are compared with a specific threshold to detect the target presence. The thresholds are defined during the offline phase, in which 1000 measurements are collected with no target in the environment. The processing of such measurements produces RD and RDA maps, that are exploited to estimate the Gaussian distribution parameters that best approximate the noise distribution at each range-Doppler and range-Doppler-angle coordinate. Then, for a point at the coordinate $\boldsymbol{p}$ in the RD and RDA maps, the threshold $\xi_{\boldsymbol{p}}$ is designed according to equation (3) by using the estimated mean $\bar{n}_{\boldsymbol{p}}$ and standard deviation $\sigma_{\boldsymbol{p}}$, together with a scaling parameter $\alpha_{\boldsymbol{p}} = 10$. The over threshold elements of the RD map are clustered by employing the DBSCAN algorithm, with a neighborhood searching radius $\epsilon = 0.05$ and considering clusters with a minimum of 2 elements. Centroids are obtained by equation (4) and compose the set of positional measurements.

B. Multitarget tracking

Target positional states and positional measurements are 4-dimensional vectors (i.e., $n_x = m_z = 4$), each one characterized by a 2D position and a 2D velocity. Track initialization requires that a track hypothesis achieves a length of $K = 3$ measurements, while both track hypotheses and confirmed tracks are deleted for values of the corresponding covariance matrix trace above $\xi_{\text{d}} = 2.5$. Data association is performed considering a gating region $\xi_{\text{g}} = 0.5\,\text{m}$ around the predicted positional state of each target. In particular, the target positional state prediction is performed considering the constant velocity model, which assumes no velocity variations in the interval between two measurements. The linear measurement model employed is the identity matrix, i.e., $\boldsymbol{H}_k = \boldsymbol{I}_4$.

The tracking accuracy is quantified using the GOSPA metric of order $p = 1$, cut off $c = \xi_{\text{g}}$, and normalization term $\alpha = 2$ [33]. In particular, localization error is evaluated computing the Euclidean distance and each misdetection introduces a penalty of $0.25\,\text{m}$ in the GOSPA metric. When a misdetection occurs the range-angle coordinate employed for collecting classification analytics is given by the predicted positional state. The target position is estimated every $0.1\,\text{s}$, while the target class every $2\,\text{s}$ aiming to determine if the target is a human or an industrial vehicle.

Fig 3 shows the empirical cumulative distribution function (ECDF) of the tracking error based on the GOSPA metric, which accounts both localization accuracy and misdetections. The proposed framework has been validated monitoring an area with moving humans and industrial vehicles by collecting 1200 measurements to cover a time frame of $120\,\text{s}$. The 99% of the measurements considers the presence of at least of 1 target, while the 85% of exactly 3 targets. At the 95th percentile, the presented approach for detection and tracking provides a localization error of $0.83\,\text{m}$, while, at the 75th percentile, the error is reduced to $0.68\,\text{m}$. The average number of false-alarms per measurement is 1.23. In particular, for 98% of the measurements at most 3 false-alarms are collected.

C. Multitarget identification

Classification is performed employing a convolutional neural network (CNN), which inputs $W = 20$ sequences of $N_{\text{c}} = 128$ samples collected relying on the estimated target positions and organized as a linear array for preserving the time correlations. The CNN architecture employed for processing the 1D time series consist of 6 layers: (i) 1D convolutional input layer with 128 filters and kernel size of 2; (ii) max pooling layer with pool size of 2; (iii) 1D convolutional layer with 32 filters and kernel size of 2; (iv) max pooling layer with pool size of 2; (v) a fully connected layer with 16 units; and (vi) the output layer. The activation function in the CNN inner layers is the rectified linear unit function, while the output layer employs the sigmoid function. Training is performed employing the categorical cross–entropy as empirical loss function and the Adam optimizer with a learning rate of 10^{-3}.

The training set consists of 1800 time series collected in the same experimentation environment, but with a different deployment of the industrial machines. The 1800 instances equally represent humans and industrial vehicles and are divided in 70% of training and 30% of validation. We apply data augmentation on the training set by performing a cyclic permutation of each training instance.

Identification has been performed for 137 time-domain sequences. The achieved identification accuracy during the testing phase is 92%. Wrong classification occurs according to misdetections and significative localization errors. In fact, the averaged localization error of sequences leading to a wrong classification is 30% greater than that of sequences producing a correct classification.

V. Conclusion

This paper presents a framework for simultaneous tracking and identification of device-free targets based on samples of reflected signals at mmWaves. The proposed framework employs tracking information to gather time-domain representation of micro-Doppler signatures specific for each target that are used as input of a CNN to perform target identification. The experimental results employing an FMCW MIMO radar validate the developed framework in an industrial environment showing a localization error below $0.83\,\mathrm{m}$ for 95% of the occasions and a target identification accuracy of 92%. The effectiveness of the proposed approach relies on the statistical characterization of the environmental clutter distribution to obtain reliable measurements for both tracking and identification even in cluttered environments. The proposed framework paves the way for the amenable implementation and deployment of sensing-based services in next-generation wireless networks.

References

[1] M. Chiani, A. Giorgetti, and E. Paolini, "Sensor radar for object tracking," *Proc. IEEE*, vol. 106, no. 6, pp. 1022–1041, Jun. 2018.

[2] G. Kwon, Z. Liu, A. Conti, H. Park, and M. Z. Win, "Integrated localization and communication for efficient millimeter wave networks," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 12, pp. 3925–3941, Dec. 2023.

[3] L. Pucci, E. Paolini, and A. Giorgetti, "System-level analysis of joint sensing and communication based on 5G new radio," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2043–2055, Jul. 2022.

[4] X. Li, Y. Cui, J. A. Zhang, F. Liu, D. Zhang, and L. Hanzo, "Integrated human activity sensing and communications," *IEEE Commun. Mag.*, vol. 61, no. 5, pp. 90–96, May 2023.

[5] X. Zhou, W. Liang, K. I.-K. Wang, H. Wang, L. T. Yang, and Q. Jin, "Deep-learning-enhanced human activity recognition for internet of healthcare things," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6429–6438, Jul. 2020.

[6] A. Conti, S. Mazuelas, S. Bartoletti, W. C. Lindsey, and M. Z. Win, "Soft information for localization-of-things," *Proc. IEEE*, vol. 107, no. 11, pp. 2240–2264, Nov. 2019.

[7] Y. Li *et al.*, "Toward location-enabled IoT (LE-IoT): IoT positioning techniques, error sources, and error mitigation," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4035–4062, Mar. 2021.

[8] A. Conti, G. Torsoli, C. A. Gómez-Vega, A. Vaccari, G. Mazzini, and M. Z. Win, "3GPP-compliant datasets for xG location-aware networks," *IEEE Open J. Veh. Technol.*, vol. 5, pp. 473–484, Apr. 2024.

[9] A. Conti, G. Torsoli, C. A. Gómez-Vega, A. Vaccari, and M. Z. Win, "xG-Loc: 3GPP-compliant datasets for xG location-aware networks," *IEEE Dataport*, Dec. 2023. [Online]. Available: https://dx.doi.org/10.21227/rper-vc03

[10] *Technical Specification Group Services and System Aspects; Feasibility Study on Integrated Sensing and Communication*, 3rd Generation Partnership Project 3GPP™ TR 22.837 V19.0.0, Jun. 2023, Release 19.

[11] C. A. Gómez-Vega, O. Y. Kolawole, M. Hunukumbure, T. Mach, M. Z. Win, and A. Conti, "Device-free localization: Outdoor 5G experimentation at mm-Waves," *IEEE Commun. Lett.*, vol. 27, no. 9, pp. 2353–2357, Sep. 2023.

[12] M. Shafiqul Islam, M. Humayun Kabir, M. Ali Hasan, and W. Shin, "Wi-MIR: A CSI dataset for Wi-Fi based multi-person interaction recognition," *IEEE Access*, vol. 12, pp. 67 256–67 272, 2024.

[13] S. Bartoletti, Z. Liu, M. Z. Win, and A. Conti, "Device-free localization of multiple targets in cluttered environments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 5, pp. 3906–3923, Oct. 2022.

[14] K. Witrisal *et al.*, "High-accuracy localization for assisted living," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 59–70, Mar. 2016.

[15] S. Bartoletti, A. Conti, A. Giorgetti, and M. Z. Win, "Sensor radar networks for indoor tracking," *IEEE Wireless Commun. Lett.*, vol. 3, no. 2, pp. 157–160, Apr. 2014.

[16] R. Mahler, *Advances in Statistical Multisource-Multitarget Information Fusion*. USA: Artech House, Inc., 2014.

[17] F. Meyer *et al.*, "Message passing algorithms for scalable multitarget tracking," *Proc. IEEE*, vol. 106, no. 2, pp. 221–259, Feb. 2018.

[18] D. Gaglione, P. Braca, G. Soldi, F. Meyer, F. Hlawatsch, and M. Z. Win, "Fusion of sensor measurements and target-provided information in multitarget tracking," *IEEE Trans. Signal Process.*, vol. 70, pp. 322–336, Jan. 2022.

[19] G. Vivone and P. Braca, "Joint probabilistic data association tracker for extended target tracking applied to x-band marine radar data," *IEEE J. Ocean. Eng.*, vol. 41, no. 4, pp. 1007–1019, Oct. 2016.

[20] L. Svensson, D. Svensson, M. Guerriero, and P. Willett, "Set JPDA filter for multitarget tracking," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4677–4691, 2011.

[21] J. Pegoraro, F. Meneghello, and M. Rossi, "Multiperson continuous tracking and identification from mm-Wave micro-Doppler signatures," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 2994–3009, Apr. 2021.

[22] X. Li, X. Wang, Q. Yang, and S. Fu, "Signal processing for TDM MIMO FMCW millimeter-wave radar sensors," *IEEE Access*, vol. 9, pp. 167 959–167 971, Dec. 2021.

[23] A. Shastri *et al.*, "A review of millimeter Wave device-based localization and device-free sensing technologies and applications," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 3, pp. 1708–1749, 3rd Quart. 2022.

[24] S. Abdulatif, Q. Wei, F. Aziz, B. Kleiner, and U. Schneider, "Micro-Doppler based human-robot classification using ensemble and deep learning approaches," in *IEEE Radar Conf.*, Oklahoma City, OK, USA, Apr. 2018, pp. 1043–1048.

[25] Z. Chen, G. Li, F. Fioranelli, and H. Griffiths, "Personnel recognition and gait classification based on multistatic micro-Doppler signatures using deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 669–673, Mar. 2018.

[26] H. Zou, Y. Zhou, J. Yang, H. Jiang, L. Xie, and C. J. Spanos, "Deepsense: Device-free human activity recognition via autoencoder long-term recurrent convolutional network," in *Proc. IEEE Int. Conf. Commun.*, Kansas City, MO, USA, May 2018, pp. 1–6.

[27] B. Vandersmissen *et al.*, "Indoor person identification using a low-power FMCW radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 7, pp. 3941–3952, Apr. 2018.

[28] A. G. Argüello and D. Berges, "Radar classification for traffic intersection surveillance based on micro-Doppler signatures," in *Eur. Radar Conf.*, Madrid, Spain, Sep. 2018, pp. 186–189.

[29] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise." AAAI Press, Aug. 1996, p. 226–231.

[30] J. LaViola, "A comparison of unscented and extended kalman filtering for estimating quaternion motion," in *Proc. American Control Conf.*, Denver, CO, USA, Jun. 2003, pp. 2435–2440.

[31] K. G. Murty, "An algorithm for ranking all the assignments in order of increasing cost," *Oper. Res.*, vol. 16, no. 3, pp. 682–687, Jun. 1968.

[32] S. H. Rezatofighi, A. Milan, Z. Zhang, Q. Shi, A. Dick, and I. Reid, "Joint probabilistic data association revisited," in *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 3047–3055.

[33] A. S. Rahmathullah, A. F. García-Fernández, and L. Svensson, "Generalized optimal sub-pattern assignment metric," in *Int. Conf. Inf. Fusion*, Xi'an, China, Jul. 2017, pp. 1–8.