People Flow Estimation with a Wi-Fi-Based Passive Radar

Laurent Storrer*‡, Dejvi Cakoni*, Hasan Can Yildirim*†, Martin Willame†,
Jérôme Louveaux†, Philippe De Doncker*, Sofie Pollin‡, François Horlin*

*Université Libre de Bruxelles - †Université Catholique de Louvain - ‡Katholieke Universiteit Leuven
{*laurent.storrer,dejvi.cakoni,hasan.can.yildirim*}*@ulb.be*, {*martin.willame,jerome.louveaux*}*@uclouvain.be*,
sofie.pollin@kuleuven.be, {*pdedonck,fhorlin*}*@ulb.ac.be*

***Abstract*—We investigate crowd monitoring with a Wi-Fi-based passive radar in the context of large events with multiple areas connected by alleys or streets, such as events in city centres. We derive an average people flow expression in people per second, away from a radar and towards it, and propose a processing scheme to estimate this flow with a Wi-Fi-based passive radar. It relies on splitting the range-Doppler map (RDM) in its negative and positive Doppler speeds parts, corresponding to the flow away from the radar and towards it respectively, and combining people counting and average people's speed estimation on each RDM part. A flow estimation error metric is introduced, and our proposed flow estimation framework is experimentally validated with a Wi-Fi-based passive radar setup using High-Efficiency Long Training Fields from the 802.11ax standard and built with Universal Software Radio Peripherals. A successful flow estimation is achieved, by obtaining a flow estimation error significantly lower than the true flow averaged on all measurements.**

***Index Terms*—People flow, Wi-Fi, passive radar, crowd monitoring, Convolutional Neural Network, CFAR.**

I. Introduction

Crowd monitoring is critical for public space and events supervision, to avoid overcrowded situations leading to danger or even deadly accidents, such as the recent overcrowding situation that led to the death of 10 people at the Astrowold festival in Houston, USA, in 2021, or the large-scale crowd crush in Seoul, South Korea, that caused the death of 156 people and left 173 injured [1]. A key point in crowd monitoring to avoid such disasters is the ability to monitor people's movement and count their number in real time.

Radars are well-suited for this task, thanks to their ability to measure the speed of people, and their anonymity preservation. Namely, Wi-Fi-based passive radar leverage the large bandwidth of Wi-Fi signals compared to other communications signals of opportunity [2]. Several recent works in the Wi-Fi-based passive radar literature for people monitoring focus primarily on people counting algorithms, usually following the same pattern: the signals of opportunity are processed in various ways to obtain range and Doppler information, or signal strength information, from which counting features can be extracted and fed to a counting algorithm [3]–[5]. Counting is often tackled as a classification problem, where each class corresponds to a given number of people or an interval of numbers of people.

In [3], the channel state information obtained from a Multiple-Input-Multiple-Output Wi-Fi setup is used to estimate the Doppler spectrum at multiple antennas. Statistical features are extracted from the Doppler spectrum and given as input to a Naive Bayes Classifier whose classes correspond to numbers of people. In [5], no counting task is performed, but two classes to estimate are defined, corresponding to the absence or presence of human targets in the scene, respectively. A Software Defined Radio is used to monitor the reference channel, *i.e.* to capture the transmitted Wi-Fi signal from an access point, while a second SDR monitors the surveillance channel, *i.e.* it measures the target echoes. The Doppler spectrogram is computed from Cross-Ambiguity Function processing, and derived features are given as input to a Support Vector Machine. In [4], range-Doppler maps (RDMs) and Doppler spectrograms are given as input to a Convolutional Neural Network (CNN) for counting, but with a limited number of 4 people. In [6], features are computed via a Principal Component Analysis followed by a Discrete Wavelet Transform on the Channel State Information matrix over multiple transmitter (TX) and receiver (RX) antenna pairs and fed to a Long Short-Term Memory network followed by a CNN.

In this work, we provide an exploratory study of another potential research track, being the estimation of people movement *flows*, in view of the scenario illustrated on the left side of Fig. 1. It is a schematic view of a location with multiple wide spaces (in blue) connected by alleys or streets (in grey). For example, it can be a city centre with squares connected by streets, or an event with multiple scenes connected by alleys. In this scenario, radars applying a counting scheme as discussed above would give information about the number of people gathered on the wide spaces/squares, while radars located in the streets/alleys (in orange in the figure) would give information about the flow of people transitioning from one square to another. A corresponding graph view could be built, with the counting numbers on the wide places as node values, and the flow numbers in the streets/alleys as edge values. Hence, having access to the counts of people and to the flows would give a complete overview of the system. To the best of our knowledge, no such work has been performed before.

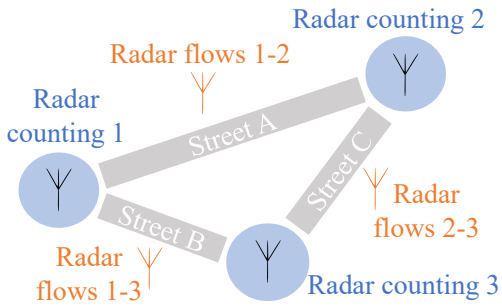

Fig. 1. Global scenario illustration with both counting and flow estimation.

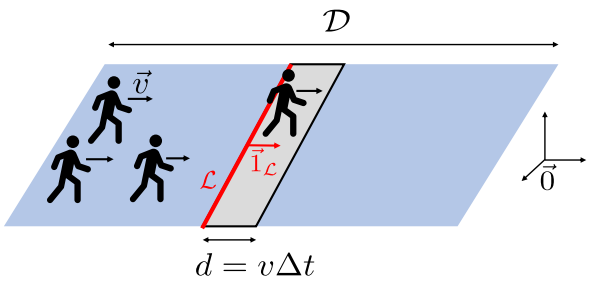

Fig. 2. Flow definition illustration.

Our contributions in this paper are the following:

- We propose a generic use case of Wi-Fi-based passive radars for crowd monitoring, and highlight the importance of measuring people flows.
- We define formally what is a flow of people, and propose a processing framework to estimate it with a Wi-Fi-based passive radar. It revolves around splitting the RDM in its negative and positive Doppler speeds parts, corresponding to the flow away from and towards the radar respectively, and combining people counting and people's average speed estimation on each part. We also introduce an error metric to measure the performance of the flow estimation.
- We experimentally validate our framework with a passive radar setup built with Universal Software Radio Peripherals (USRPs) and based on High-Efficiency Long Training Fields (HE-LTFs) from the latest Wi-Fi standard, 802.11ax. We show that our proposed error metric reaches a value being one order of magnitude lower than the true averaged flow.

In Section II, we define formally what a flow of people is. In Section III, the system model of a Wi-Fi-based passive radar based on channel estimation is presented, along with the radar processing to obtain a RDM. In Section IV, we explain our proposed flow estimation processing framework. In Section V, the proposed scheme is validated with experimental measurements. We conclude this paper in Section VI.

II. Flow definition

We wish to compute the flow of people, *i.e.* the number of people crossing a line $\mathcal{L}$, with a normal vector $\vec{1}_{\mathcal{L}}$, per unit of time [7]. Let us consider a number P of people moving at identical speed, with a speed vector $\vec{v} = v\vec{1}_v$ of norm $v = ||\vec{v}||$ and direction $\vec{1}_v$. Considering that the people are spread on a total ground distance $\mathcal{D}$ with a homogeneous people surface density, this density writes

$$\eta = \frac{P}{\mathcal{L}\ \mathcal{D}} \quad \left[\frac{\text{people}}{\text{m}^2}\right]. \tag{1}$$

Let us first assume that the speed direction and the direction normal to the line are identical, *i.e.* that $\vec{1}_v.\vec{1}_{\mathcal{L}} = 1$. We consider the situation at a given time instant and compute the number of people that have crossed the line at a time instant Δt later. They are contained in a rectangle of length $\mathcal{L}$ and width $d = v\Delta t$, hence of surface $\mathcal{L}d = \mathcal{L}v\Delta t$, as illustrated in Fig. 2. The number of people in the rectangle is [7]

$$\tilde{P} = \eta\mathcal{L}d = \eta\mathcal{L}v\Delta t. \tag{2}$$

The number of people in the parallelepiped per unit time, *i.e.* the people flow, is

$$\phi = \frac{\tilde{P}}{\Delta t} = \eta\mathcal{L}v. \tag{3}$$

Using the people density expression (1) yields

$$\phi = \frac{P}{\mathcal{L}\ \mathcal{D}}\mathcal{L}v = \frac{Pv}{\mathcal{D}} \tag{4}$$

Its units are

$$[\phi] = \frac{\text{people}\frac{\text{m}}{\text{s}}}{\text{m}} = \frac{\text{people}}{\text{s}}. \tag{5}$$

In the generic case where the people speed direction is not identical to the line normal direction, *i.e.* $\vec{1}_v.\vec{1}_{\mathcal{L}} \neq 1$, the flow writes $\phi = Pv/\mathcal{D}\ \vec{1}_v.\vec{1}_{\mathcal{L}}$ [7]. In our crowd monitoring scenario, we consider two cases: $\vec{1}_v = \vec{1}_{\mathcal{L}}$ and $\vec{1}_v = -\vec{1}_{\mathcal{L}}$, *i.e.* the case when people walk away from the origin or towards it respectively, referring to Fig. 2. Hence, we define a flow *towards* ϕ_t with P_t people, from left to right in Fig. 2, and a flow *away* ϕ_a with P_a people, from right to left in Fig. 2:

$$\begin{cases} \phi_t = & P_t\ v\ /\ \mathcal{D} \\ \phi_a = & -P_a\ v\ /\ \mathcal{D} \end{cases} \tag{6}$$

Let us now consider the case where each person has an individual speed. We simplify it by defining the *average towards speed* $\bar{v}_t$ as the average speed amplitude of people moving towards the origin, and the *average away speed* $\bar{v}_a$ as the average speed amplitude of people moving away from the origin. Note that $\bar{v}_a$ is a positive quantity and that the negative sign associated with its movement away from the origin is already included through the product $\vec{1}_v.\vec{1}_{\mathcal{L}}$. This way, we redefine ϕ_t and ϕ_a as the average flow towards and away respectively:

$$\begin{cases} \phi_t = & P_t\ \bar{v}_t\ /\ \mathcal{D} \\ \phi_a = & -P_a\ \bar{v}_a\ /\ \mathcal{D} \end{cases} \qquad \bar{v}_t,\ \bar{v}_a \geq 0. \tag{7}$$

$\mathcal{D}$ is the ground distance from (1) on which the flow is studied. Here we study the situation depicted in Fig. 1, hence $\mathcal{D}$ is the length L_{street} of the considered street, provided that the

radar maximal range $d_{\max}$ is higher than this length, otherwise it is the radar maximal range. Thus, it writes

$$\mathcal{D} = \min(L_{\text{street}}, d_{\max}). \tag{8}$$

A homogeneous people surface density is considered in the above flow development. In real conditions, the people density is likely to not be homogeneous if the considered distance is large. Therefore, we consider ϕ_t and ϕ_a as averaged flow from a second point of view, *i.e.* as flows computed for an averaged people density η. Hence, the flows ϕ_t and ϕ_a are averaged both on the speeds and on the people density. In the rest of this paper, for the sake of simplicity, we will refer to the average flow simply as flow.

III. System model & Processing

We consider a passive radar RX at the origin of the Cartesian coordinates system of Fig. 2 and a TX of opportunity next to the passive radar RX. From (7), we see that to estimate the flows, the radar must estimate the number of people coming towards it and their average speed, and the number of people moving away from it and their average speed. The radar is particularly well-suited for this task, since it can estimate the speed information and count people based on the range-Doppler map (RDM) that it produces, as detailed below.

A. Signal and channel model

We consider a continuous stream of N 802.11ax compliant HE-LTF symbols, hence modulated with Orthogonal Frequency-Division Multiplexing (OFDM). The Q OFDM subcarriers are indexed with $q = -Q/2, \ldots, Q/2-1$. In the time domain, each HE-LTF contains $Q + L_{cp}$ samples, where L_{cp} is the CP length. Considering a signal bandwidth B, each HE-LTF signal is sampled with a sampling time $T_s = 1/B$, also denoted as *fast time*, and indexed with $i = 0, \ldots, Q-1$. The time between the reception of two HE-LTF is $T = (Q + L_{cp})T_s$, also known as the *slow time*. It is indexed with $k = 0, \ldots, N-1$.

The transmitted signal from TX is reflected on moving people and static elements, and the corresponding echoes are measured at the passive radar receiver RX. These echoes are embedded in a frequency-domain Channel Transfer Function (CTF), at each slow time instant kT. We consider a geometry-based single bounce model with one antenna at TX and one antenna at RX. With the speed of light in vacuum c and the carrier frequency f_c, we consider a CTF matrix $\boldsymbol{H}$ whose elements are [8], [9]:

$$H[q,k] = \sum_{r=0}^{N_r} a_r \; e^{-j2\pi(f_c + q\frac{B}{Q})\tau_r} \; e^{j2\pi f_r k T}. \tag{9}$$

where the index r denotes each of the N_r objects or targets in the environment. For each target, the corresponding echo is characterised by (i) its complex amplitude a_r, (ii) its propagation delay τ_r to RX, corresponding to the bistatic distance $d_r = d_{\text{TX-target}} + d_{\text{target-RX}}$ with the relation $d_r = c\tau_r$, (iii) its Doppler frequency shift $f_r = \frac{2v_r}{\lambda}\varepsilon_r$ seen at RX, where v_r is the target's speed, $\lambda = c/f_c$ is the wavelength and $-1 \leq \varepsilon_r \leq 1$ is a projection factor depending on the bistatic geometry, whose absolute value is maximal when the target moves perpendicularly to the baseline TX-RX [10]. The index $r = 0$ refers to the direct path between TX and RX. TX and RX are static, hence $f_0 = 0$ Hz. The other indices $r > 0$ refer to the objects or targets in the environment. From this model and the received signals at RX, we obtain a RDM as follows.

B. Radar processing

The first processing step is time synchronisation, performed by correlating in the time domain the known transmitted HE-LTF symbols with the received symbols and finding the maximal correlation value. It corresponds to the index of the direct path between TX and RX, chosen as propagation delay 0s reference. Then, the CTFs at each slow time instant can be estimated, through an element-wise division of the received signal and the known transmitted signal (HE-LTF) in the frequency domain. This is a so-called frequency domain least-squares channel estimation [8] yielding an estimate $\hat{\boldsymbol{H}}$ of $\boldsymbol{H}$, with elements $\hat{H}[q,k]$. The estimated CTFs, *i.e.* the columns of the matrix, can be averaged by groups of M, yielding a matrix $\hat{\boldsymbol{H}}_a$ with elements $\hat{H}_a[q,k']$, where $k' = 0, \ldots, N/M-1$ is the number of remaining CTFs. We note $N' = N/M$. An Inverse Fast Fourier Transform (IFFT) can be applied along the first dimension of $\hat{\boldsymbol{H}}$, indexed by q, to obtain the so-called range-slow time map, noted $\hat{\boldsymbol{R}}$. A Fast Fourier Transform (FFT) can be applied along the second dimension of the range-slow time map $\hat{\boldsymbol{R}}$, indexed by k, to extract the Doppler frequency shift. This yields the so-called *range-Doppler map* (RDM) $\hat{\boldsymbol{D}}$, a 2D map with amplitude peaks at the propagation delay and Doppler speed of the targets or static objects. The elements of $\hat{\boldsymbol{D}}$ write

$$\hat{D}[i,n] = \frac{1}{\sqrt{QN}} \sum_{k=0}^{N-1} \sum_{q=-Q/2}^{Q/2-1} \hat{H}_a[q,k'] e^{j2\pi\frac{iq}{Q}} e^{-j2\pi\frac{k'n}{N'}}, \tag{10}$$

with $n = -N'/2, \ldots, N'/2-1$ the Doppler/speed bins index. The negative speed bins form the left part of the RDM and correspond to the people moving away from the radar, hence to the flow ϕ_a. The positive speed bins form the right part of the RDM and correspond to the people moving towards the radar, hence to the flow ϕ_t. An example of the two parts of the RDM is given at the end of this paper in Fig. 6. We consider to have at our disposal a number U of RDMs, originating from U different measurements, and indexed by $u = 0, \ldots, U-1$.

IV. Flow estimation

From the above discussion in Section III-B, we see that we can divide the RDM into its right and left parts, and perform counting and speed estimation on these parts separately. First, before counting or speed estimation, the RDM is cut at the range bin corresponding to the maximal distance $\mathcal{D}$ from (8).

A. People counting

Then, a counting algorithm is applied separately on the left and right parts of the RDM. We cannot simply use the number of peaks in the RDM as estimate of the number of people,

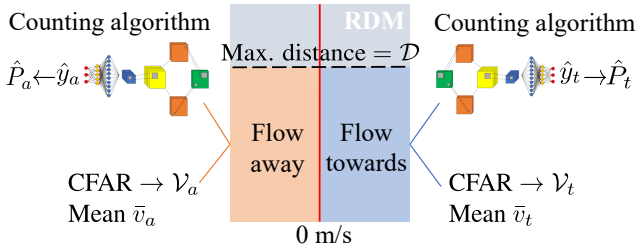

Fig. 3. Flow processing scheme on the left (orange) and right (blue) parts of the RDM, corresponding to the flows away and towards respectively. The scheme below the "Counting algorithm" text is an illustration of the CNN.

due to the limited range and speed resolutions that prevent the radar from resolving closely located targets, and due to the fact that people can mask each other. Therefore, as discussed in the Introduction, the counting is tackled as a classification problem. A counting class label $y = n_c$ corresponds to a number of people being in a given interval, for which we define $P_{n_c}^{(1)}$ as the lower bound of the people interval of class n_c and $P_{n_c}^{(2)}$ as the upper bound. We employ a Convolutional Neural Network (CNN) as classifier, as it avoids the problem of non-exhaustive handcrafted features and it exploits the spatial 2D shape of the RDM [11]. It consists of:

- 3 convolutional blocks: each block first consists of two sets of convolution kernels. The first set contains small-size kernels (squares of side equal to 2), and the second contains large-size kernels (squares of side equal to 7). The outputs of the convolution with the small and large kernels are concatenated along the depth dimension. This process is known as depth concatenation [12]. Then, a ReLU activation function is applied, followed by 2×2 average pooling. The output of the last convolutional block is vectorised before the subsequent network layers.
- One fully connected layer following the vectorisation, with 177 neurons.
- One softmax layer as output layer, giving the probability of the input RDM to belong to one of the classes. The class for which the estimated probability is the highest is assigned as the estimated class label for the RDM.

The values of the CNN hyperparameters were tuned with a Bayesian Optimisation procedure, which is out of the scope of this work. The interested reader is referred to [13]. Applying the CNN on the right and left parts of the RDM yields the counting class labels $\hat{y}_t$ and $\hat{y}_a$ corresponding to an estimated number of people $\hat{P}_t$ and $\hat{P}_a$ respectively. The CNN is trained and validated respectively on U_{train} and U_{val} RDMs of the U RDMs in the dataset, and it is tested on a part U_{test}. The repartition key is 70%, 20%, 10 %.

B. Average speed estimation

Separately from the counting, the average speed estimation is performed as follows. The RDM cells containing targets can be identified through a detector. The Ordered Statistic CFAR (OS-CFAR) is chosen here for its detection capabilities in the presence of multiple targets [14]. Each detected target r is identified by detection coordinates in terms of range and speed bin indexes, (i_r, n_r). This corresponds to a speed $v_r = n_r v_{res}$, where v_{res} is the speed resolution of the radar. Hence, considering the set of all detections $\{(i_r, n_r) \mid r = 1, \ldots, N_r\}$, there is a corresponding set of speeds

$$\mathcal{V} = \{v_r = n_r v_{res} \mid r = 1, \ldots, N_r\}. \tag{11}$$

The set of positive speeds (for the flow towards) and the set of negative speeds (for the flow away) are

$$\mathcal{V}_t = \{v_r \in \mathcal{V} \mid v_r \geq 0\} \text{ and } \mathcal{V}_a = \{v_r \in \mathcal{V} \mid v_r < 0\} \tag{12}$$

respectively. The average speed amplitude of the towards flow is the average of the set $\mathcal{V}_t$, and the average speed amplitude of the away flow is the absolute value of the average of the set $\mathcal{V}_a$, since the negative speed sign was already taken into account by the scalar product $\vec{1}_v.\vec{1}_{\mathcal{L}}$ when deriving the flow equation. They write

$$\bar{v}_t = \frac{1}{\text{card}(\mathcal{V}_t)} \sum_{v_r \in \mathcal{V}_t} v_r \text{ and } \bar{v}_a = \left| \frac{1}{\text{card}(\mathcal{V}_a)} \sum_{v_r \in \mathcal{V}_a} v_r \right|, \tag{13}$$

where $\text{card}(\mathcal{V}_t)$ is the cardinality of the set $\mathcal{V}_t$, *i.e.* the number of elements in the set, and idem for $\mathcal{V}_a$. The flow processing scheme is summarised in Fig. 3.

C. Performance metrics

From (7), the flow estimation heavily relies on the people counting. Hence, it is relevant to use a confusion matrix applied to our counting as a performance metric. The flow estimation also relies on the average speed estimation. This estimation depends on the OS-CFAR detection and the radar speed resolution. However, the speed resolution is known, the performance of the OS-CFAR has already been studied thoroughly [10], [14], [15], and a performance analysis of it would be out of the scope of this work. Moreover, having access to the ground truth of individual people's speeds for comparison in a measurement scenario with many people would be very challenging. Therefore, we consider the estimated speed as the ground truth speed, *i.e.* we neglect the speed estimation error, to simplify the analysis. This is a realistic assumption thanks to the fine speed resolution, detailed in Section V-A. Hence, we focus on the counting confusion matrix as an error metric, and we introduce an error metric for flows, computed on the testing set and derived from the counting accuracy, as follows. By considering the true class labels y and the lower and upper bounds of the number of people in each class interval, $P_{n_c}^{(1)}$ and $P_{n_c}^{(2)}$ respectively, we can write

$$y = n_c \Leftrightarrow P \in \left\{P_{n_c}^{(1)} \ldots, P_{n_c}^{(2)}\right\} \tag{14}$$

The interval can be represented by its mean $\bar{P}$, *i.e.*

$$y = n_c \Leftrightarrow \bar{P} = \left(P_{n_c}^{(1)} + P_{n_c}^{(2)}\right) / \, 2, \tag{15}$$

meaning that each class label can be associated with an average corresponding number of people. This average number of people can be associated with a corresponding mean flow

$$\bar{\phi}_t = \bar{P}_t \; \bar{v}_t \; / \; \mathcal{D} \quad \text{or} \quad \bar{\phi}_a = -\bar{P}_a \; \bar{v}_a \; / \; \mathcal{D}, \tag{16}$$

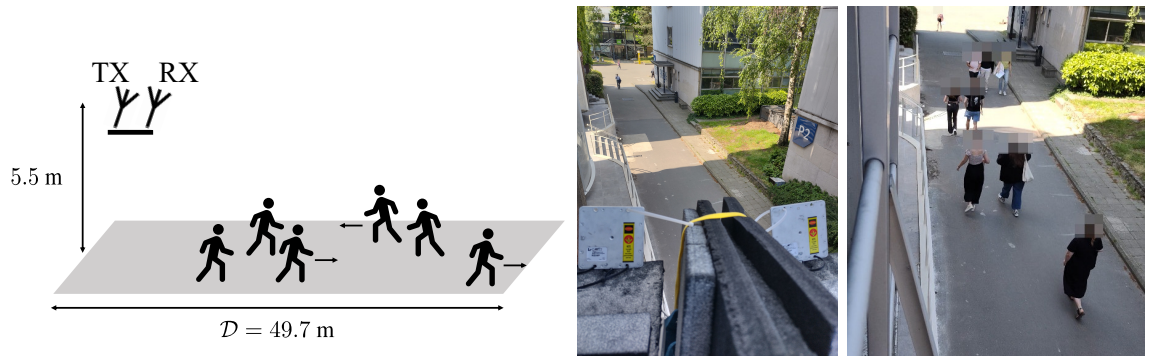

Fig. 4. Measurement setup scheme (left), picture from the balcony (middle), and flow situation example (right).

where $\bar{P}_t$ is equal to the $\bar{P}$ obtained when the right part of the RDM is considered, *i.e.* when $y = y_t$, and $\bar{P}_a$ is equal to the $\bar{P}$ obtained when the left part of the RDM is considered, *i.e.* when $y = y_a$. Averaged mean flows can be computed as

$$\bar{\phi}_t^{av} = \frac{1}{U_{\text{test}}} \sum_{u=0}^{U_{\text{test}}-1} \left(\bar{\phi}_t\right)_{|_u} \text{ and } \bar{\phi}_a^{av} = \frac{1}{U_{\text{test}}} \sum_{u=0}^{U_{\text{test}}-1} \left(\bar{\phi}_a\right)_{|_u} \quad (17)$$

on the U_{test} RDMs in the testing set. The subscript $|_u$ indicates that the mean flows are computed for the RDM u, with its positive speed part for the towards flow and its negative speed part for the away flow. In a similar way as for the true class labels, the corresponding estimated quantities are derived from the estimated labels $\hat{y}$:

$$\hat{y} = n_c \Leftrightarrow \hat{\bar{P}} = \left(P_{n_c}^{(1)} + P_{n_c}^{(2)}\right) / \, 2 \quad (18)$$

$$\hat{\bar{\phi}}_t = \hat{\bar{P}}_t \; \bar{v}_t \; / \; \mathcal{D} \quad \text{or} \quad \hat{\bar{\phi}}_a = -\hat{\bar{P}}_a \; \bar{v}_a \; / \; \mathcal{D} \quad (19)$$

$$\hat{\bar{\phi}}_t^{av} = \frac{1}{U_{\text{test}}} \sum_{u=0}^{U_{\text{test}}-1} (\hat{\bar{\phi}}_t)_{|_u} \text{and } \hat{\bar{\phi}}_a^{av} = \frac{1}{U_{\text{test}}} \sum_{u=0}^{U_{\text{test}}-1} (\hat{\bar{\phi}}_a)_{|_u} \quad (20)$$

A global flow error metric on all U_{test} testing data points can thus be formed as

$$\varepsilon_\phi = \frac{1}{2} \left(|\bar{\phi}_t^{av} - \hat{\bar{\phi}}_t^{av}| + |\bar{\phi}_a^{av} - \hat{\bar{\phi}}_a^{av}| \right). \quad (21)$$

Intrinsically, the only source of error considered here is the counting error, so this metric does not carry more information than the counting classification error. However, it gives us an error metric in terms of flows in people/s rather than in terms of classification error, which is more representative physically.

V. Experimental results

A. Scenario and experimental setup

The measurement scenario considered here is illustrated on the left and middle of Fig. 4. A SISO passive radar setup is used here. The setup consists of two Universal Software Radio Peripherals (USRPs) X310, connected to L-com RE09P 2.4-2.5 GHz panel antennas via Sucoflex 126E cables with SubMiniature version A (SMA) connectors. The first USRP is equipped with a single antenna and acts as TX Wi-Fi AP. The second USRP is equipped with one antenna and acts as RX. TX and RX are disposed in a quasi-monostatic configuration.

TABLE I
Measurement parameters & Counting class labels

f_c	B	N'	M	Q	L_{cp}	T
2.55 GHz	100 MHz	256	300	1024	64	10.88 μs

	$y=1$	$y=2$	$y=3$	$y=4$
$P_{n_c}^{(1)}$	0	1	3	5
$P_{n_c}^{(2)}$	0	2	4	9

The USRPs are connected via 10 Gigabit Ethernet cables to one single computer equipped with two 10Gtek X520-10G-2S-X8 10-Gigabit Ethernet cards and one 12-core AMD Ryzen 9 3900X CPU clocked at 3.8 GHz. The clocks of TX and RX are shared using an Octoclock CDA-2990, to avoid non-idealities such as Carrier Frequency Offset (CFO), since this is not the focus of this work. TX is transmitting HE-LTFs from the 802.11ax standard, and measurements at RX are processed with the steps described in Section III-B.

The setup is placed on a balcony on the side of a street of length $\mathcal{D} = 49.7$ m. The height of TX and RX w.r.t. the ground is 5.5 m. This positioning above the ground level prevents blocking phenomenons, at the cost of a vertical Doppler projection factor. 662 measurements are collected, hence $U = 662$ RDMs. The number of people in each direction, towards TX/RX or away from TX/RX, is between 0 and 9. An illustration of a measurement situation is shown on the right of Fig. 4.

The measurement parameters are given in the top part of Table I. A carrier frequency $f_c = 2.55$ GHz was chosen to avoid interferences from public Wi-Fi networks. High values of N' and M are chosen to ensure a fine speed resolution for the speed estimation. With these values, the speed resolution is $v_{res} = 0.07$ m/s. The counting labels are structured in 4 classes, corresponding to the intervals of numbers of people given in the bottom part of Table I.

B. Results

The averaged and best confusion matrices on 10 dataset partition realisations are also given in Fig. 5. They are computed on the testing set for all counting classifications regardless of the towards or away direction, hence on $2U_{\text{test}} = 134$ RDMs

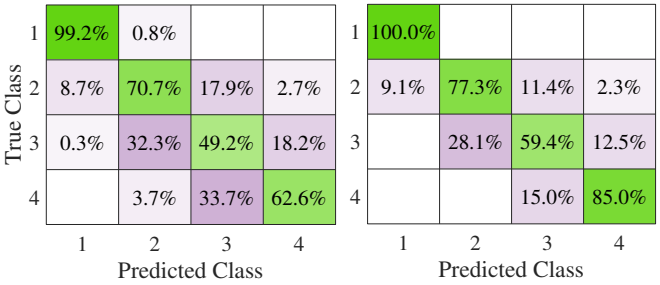

Fig. 5. Confusion matrices of the counting associated with flow estimation. Averaged (left) and best (right).

TABLE II
FLOW MEASUREMENT RESULTS, IN [PEOPLE/S].

	Towards	Away
True	$\bar{\phi}_t^{av} = 0.0376$	$\bar{\phi}_a^{av} = -0.0540$
Predicted	$\hat{\bar{\phi}}_t^{av} = 0.0348$	$\hat{\bar{\phi}}_a^{av} = -0.0531$

parts. We see that the classification of the first class, *i.e.* the class with 0 people in one direction, is almost perfect. This is because an RDM coming from a measurement with no targets does not contain any amplitude peaks. Hence, it is easy to differentiate it from other RDMs containing targets, *i.e.* from other classes. The second and last classes perform moderately well, but the third class yields poor accuracy compared to the others. It is mainly confused with the second class, which is expected since the interval of number of people is really close between them. We see with the averaged confusion matrix that the last class can be confused with the third one. Indeed, the last class contains more samples closer to 5 people than to 9, making it more vulnerable to confusion with the third class. We notice that no extreme misclassification occurs between class 1 and class 4. This is good since it prevents confusing potential high-flow situations with situations without any flow.

Then, the true and estimated classification labels are combined with the average speed estimation, yielding the true and predicted averaged mean flow towards and away, given in Table II. They are used to compute the flow error metric:

$$\varepsilon_\phi = 0.0039 \quad \text{people/s}. \tag{22}$$

We see that the error on the flow estimation is small compared to the true flow values $\bar{\phi}_t^{av}$ and $\bar{\phi}_a^{av}$, which is a promising result. These flow estimation results on a moderate number of people walking in the street still have room for improvement, namely on the counting classification part, but the global scheme seems promising. An example of obtained result with a correct counting classification is given in Fig. 6. The RDM is separated into its negative speed part (flow away) and positive speed part (flow towards). The estimated and true flows are given above each RDM part.

VI. CONCLUSIONS

In this paper, we proposed a processing scheme for people flow estimation in a street, to be used in the full scenario of a city centre or an event. We defined the people flow, along with a way to estimate it with a Wi-Fi-based passive radar. We defined a related error metric, and provided promising experimental measurements performed in a street.

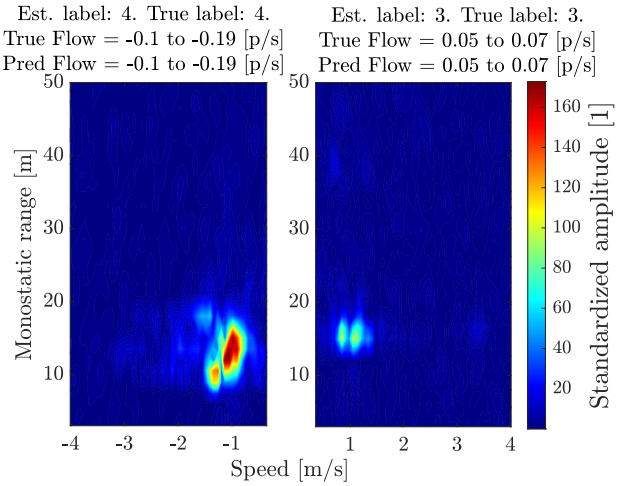

Fig. 6. Example of result of the flow estimation.

REFERENCES

[1] S.-J. Wang, "Survey of crowd crush disasters and countermeasures," *Prehospital and Disaster Medicine*, vol. 38, no. S1, p. s78–s78, 2023.

[2] P. Falcone, F. Colone, and P. Lombardo, "Potentialities and challenges of WiFi-based passive radar," *IEEE Aerospace and Electronic Systems Magazine*, vol. 27, no. 11, pp. 15–26, November 2012.

[3] S. Di Domenico, G. Pecoraro, E. Cianca, and M. De Sanctis, "Trained-once device-free crowd counting and occupancy estimation using WiFi: A Doppler spectrum based approach," in *2016 IEEE 12th Int. Conf. Wireless Mobile Comput., Netw. Commun.* IEEE, 2016, pp. 1–8.

[4] C. Tang, W. Li, S. Vishwakarma, K. Chetty, S. Julier, and K. Woodbridge, "Occupancy Detection and People Counting Using WiFi Passive Radar," in *2020 IEEE Radar Conference*. IEEE, 2020, pp. 1–6.

[5] W. Li, B. Tan, and R. J. Piechocki, "WiFi-based passive sensing system for human presence and activity event classification," *IET Wireless Sensor Systems*, vol. 8, no. 6, pp. 276–283, Dec. 2018.

[6] S. Liu, Y. Zhao, F. Xue, B. Chen, and X. Chen, "DeepCount: Crowd counting with WiFi via deep learning," *arXiv:1903.05316*, 2019.

[7] J. L. Meriam, L. G. Kraige, and J. N. Bolton, *Engineering mechanics: dynamics*. John Wiley & Sons, 2020.

[8] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM Wireless Communications with MATLAB*. Wiley Publishing, 2010.

[9] L. Storrer, H. C. Yildirim, M. Crauwels, E. I. P. Copa, S. Pollin, J. Louveaux, P. De Doncker, and F. Horlin, "Indoor tracking of multiple individuals with an 802.11 ax Wi-Fi-based multi-antenna passive radar," *IEEE Sensors Journal*, vol. 21, no. 18, pp. 20 462–20 474, Sept. 2021.

[10] M. Richards, W. Holm, and J. Scheer, *Principles of Modern Radar: Basic Principles*, ser. Electromagnetics and Radar. Institution of Engineering and Technology, 2010.

[11] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[12] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Comp. Soc. Conf. Comp. Vis. Patt. Recog.*, 2015, pp. 1–9.

[13] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," *Advances in neural information processing systems*, vol. 25, 2012.

[14] "Radar CFAR thresholding in clutter and multiple target situations, author=Rohling, Hermann," *IEEE transactions on aerospace and electronic systems*, no. 4, pp. 608–621, 1983.

[15] B. Magaz, A. Belouchrani, and M. Hamadouche, "Automatic threshold selection in OS-CFAR radar detection using information theoretic criteria," *Prog. Electromagn. Res. B*, vol. 30, pp. 157–175, 2011.