This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Kianoush, Sanaz; Savazzi, Stefano; Beschi, Manuel; Sigg, Stephan; Rampa, Vittorio
A Multisensory Edge-Cloud Platform for Opportunistic Radio Sensing in Cobot Environments

Published in:
IEEE Internet of Things Journal

DOI:
[10.1109/JIOT.2020.3011809](https://doi.org/10.1109/JIOT.2020.3011809)

Published: 15/01/2021

Document Version
Peer reviewed version

Please cite the original version:
Kianoush, S., Savazzi, S., Beschi, M., Sigg, S., & Rampa, V. (2021). A Multisensory Edge-Cloud Platform for Opportunistic Radio Sensing in Cobot Environments. *IEEE Internet of Things Journal*, 8(2), 1154-1168. [9146837]. <https://doi.org/10.1109/JIOT.2020.3011809>

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

© 2021 IEEE. This is the author's version of an article that has been published by IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

A Multisensory Edge-Cloud Platform for Opportunistic Radio Sensing in Cobot Environments

Sanaz Kianoush, *member, IEEE*, Stefano Savazzi, *member, IEEE*, Manuel Beschi, *member, IEEE*, Vittorio Rampa, *member, IEEE*, Stephan Sigg, *member, IEEE*, and Sameera Palipana *member, IEEE*,

Abstract—Worker monitoring and protection in industrial production lines require advanced sensing capabilities and flexible solutions that are able to monitor the movements of the operator in proximity of moving robots. Considering that, in the context of human–robot collaborative (HRC) workplaces, no single technology can currently solve the problem of continuous worker monitoring, we propose the combination and transformation of multiple sensing technologies into a multisensory cloud-edge platform for passive sensing and perception. Multidimensional data acquisition, from different sources, pre-processing, manipulation, feature extraction, data distribution and fusion, along with distributed learning and computing methods are described. The platform proposes a practical solution for data fusion and analytics obtained from distributed heterogeneous sensor devices, that perform the opportunistic perception of the worker through radio signals, including WiFi, radars working in the sub-THz band and infrared sensors. The proposed platform performance is validated through real HRC use case scenarios inside a pilot industrial plant in which operator safety is guaranteed.

Index Terms—Passive radio sensing, cloud-assisted Internet of Things, real-time data analysis, semantic data models.

I. INTRODUCTION

NEXT generation manufacturing, Industry 4.0 (I4.0), must ensure workspace safety and production efficiency during human-robot cooperation (HRC), while continuous and accurate tracking and perception of workers activities is the key to achieve a high level of efficiency and flexibility as required by future applications [1], [?]. In particular, I4.0 is fascinated by the unlimited possibilities of re-organizing workflows and interactions offered by interconnected components. Dynamic environments entail reconfigurable layouts, high degrees of sharing of resources, and humans moving in loosely structured environments [2]. Interactions with multiple devices and more complex workflows is especially true for robots, in collaborative robot (cobot) applications specifically, where shared workspaces with humans is the regular condition. In this context, the reuse, or transformation [4], of sensors [6], wireless

S. Kianoush, S. Savazzi and V.Rampa are with the Institute of Electronics, Computer and Telecommunication Engineering (IEIIT) of Consiglio Nazionale delle Ricerche (CNR), p.zza Leonardo da Vinci, 20133 Milano, Italy, e-mail: {sanaz.kianoush, stefano.savazzi, vittorio.rampa}@ieiit.cnr.it.

M. Beschi is with the Institute of Sistemi e Tecnologie Industriali Intelligenti per il Manifatturiero Avanzato (STIIMA), of Consiglio Nazionale delle Ricerche (CNR), Milano, Italy, e-mail: {manuel.beschi@stiima.cnr.it}.

Stephan Sigg and Samira Palipana are with Aalto University, Department of Communications and Networking, Finland, e-mail: {sameera.palipana, stephan.sigg}@aalto.fi.

This work has been partially carried out in the framework agreement CNR-Regione Lombardia CYBER SORT project[†].

Fig. 1: Multisensory platform and collaborative HR workspace consisting of a robotic manipulator and a worker in a fenceless space.

technologies (5G and machine-type [7]) and their orchestration to support flexible worker monitoring represents the backbone of pervasive control methodologies. Wireless networks are thus expected to represent the principal component for more integrated and safer environments.

Recently, techniques to capture and process the wireless stray electromagnetic (EM) radiations originated from different radio sources are gaining increasing attention [55], [29], [18], [68]. In particular, these techniques can be exploited for advanced human-scale sensing, human behavior recognition, detection/localization, and crowd density estimation/mapping. Passive radio sensing, or vision [29] is an emerging paradigm that leverages different radio frequency (RF) technologies for sensing tasks. These solutions aim to track, recognize and predict human body motions in dynamic smart environments

through real-time collection and processing of heterogeneous RF stray radiations from dense multisensory networks. RF signals are perturbed by moving objects/bodies, and by changing scenario configurations, due to the propagation of reflected, scattered and diffracted EM waves. Hence, in addition to transporting modulated information, they can serve as virtual stimuli to infer a two- or three-dimensional (2D/3D) view of all objects traversed by the EM wavefield.

It is expected that the combination and transformation of multiple sensing technologies will be the key to meet the expected accuracy in cobot environments [3], [5], [6]. For example, previous works on sensing for HRC were usually focused on the design of individual technologies, such as wearables [3], vision [6], radars [19], or machine type connections for device free localization [33]. WiFi and multiple antenna technologies have been also assessed for multi-subject recognition [15] and localization [16].

The paper demonstrates the integration of heterogeneous RF sensing technologies into an industry-compliant edge-cloud computing architecture to provide augmented information about worker safety in the context of HRC manufacturing. With respect to previous literature in the field, it is specifically dedicated to motion sensing capabilities derived from the transformation of heterogeneous industrial wireless sensors that are integrated into a multisensory platform. Sensors are opportunistically transformed to monitor worker safety, namely to support human-robot distance monitoring, classification of potential unsafe conditions, localization of the worker in the shared workspace and detection of the number of workers moving in the workspace.

The paper is organized as follows....

II. LITERATURE REVIEW AND CONTRIBUTIONS

Opportunistic sensing targets the cross-fertilization of computing and communication technologies, leveraging opportunistically different ambient signals (radio, acoustic, light etc..) to uncover novel sensing modalities. The compound effect of heterogeneous sensors is expected to enable new capabilities to obtain the semantic awareness of surrounding environment. Technologies for opportunistic and transformative radio sensing, such as device free radio localization [13], [14], [54], activity recognition [40], people counting [11] typically focus on the augmentation and transformation of existing radio devices, such as WiFi, machine-type connectivity (MTC) or cellular- wide wireless area networks (WWAN) [17] into human-scale sensors. They generally exploit electromagnetic (EM) fields maintained by different radio sources (i.e., micro-wave, THz bands, infrared). Measurements of such EM fields are used to extract an image of the environment, or its alterations, for various motion perception tasks [19]. In particular, subject presence in the environment changes the EM propagation characteristics due to reflection, scattering and diffraction effects. In [11], [18] WiFi signal are used for target counting and activity recognition. In [8] a cloud-IoT platform is proposed to sustain device free human sensing by real time processing of channel quality information (CQI) time series. In order to manage high dimensional data processing and real

time analysis, a cloud platform along with machine learning tools are proposed. Multidimensional data analytics is therein implemented on WiFi signals, while the problem of data fusion and service orchestration are not discussed. In [33] a device free localization and fall detection system are proposed for HRC applications. It adopts unmodified machine type radio communications based on the IEEE 802.15.4 standard.

All above mentioned works focus on the evaluation of individual technologies for human sensing, besides it has been also verified that no single technology could provide robust and accurate information.

A. Human-Robot Cooperation and Sensors

(TO BE REVISED/MODIFIED, A DESCRIPTION OF A TYPICAL, BUT SIMPLIFIED, ROBOTIC CELL WOULD BE ALSO USEFUL).

Human-Robot Cooperation (HRC) is a key enabler in advanced manufacturing, allowing the unchallenged adaptability and dexterity of human workers to be assisted by the flexibility of robots (cobots) sharing the execution of complex repetitive workflows. Effectiveness in HRC tasks largely depends on the possibility to allow human operator great autonomy in decision making, task execution order and timing. Examples include, but are not limited to holding/resuming some tasks [??], temporarily leaving workstations for contingencies, buffering and re-sequencing production steps [??], etc. HRC shared spaces are highly dynamic due to mixed movements, and workers safety necessarily depends on (or greatly benefits from) *localization* services. In the following sections, we refer to localization in terms of anonymous (passive, or device-free) detection of motion, as opposed to *tracking* where identity tagging occur [3]. Localization services should be able to map increasingly complex figures related to human/robot motion, such as:

- *pose*: application-space current location coordinates of humans (and cobots);
- *trajectory*: 1st/2nd order kinematics used for tracing the projection of current pose and predicting space occupancy;
- *gesture*: trajectories of body parts associated with labeling and/or semantics;
- *behaviors*: full trajectory patterns associated with some specific worker activity and semantics.

To some extent, all are relevant for safety, with someone (pose, trajectory) being necessary when used for reactive actions like maintaining a minimum separation distance, while some other ones (gestures, behaviors) being used for anticipating potentially hazardous situations.

Unlike robot motion, detecting human position and movements requires the deployment of additional/dedicated sensors. Mainstream safety technology makes use of dedicated sensors for detecting presence, tracing displacements or distances. Optoelectronic devices [6] typically exploit the properties of reflectivity, while other technology uses emissivity of thermal images or acoustic fields. The major downsides of these systems are the full-accuracy ranges, the presence of occlusions, the environmental conditions (dust, fumes, etc) and, for those

based on vision tracking, privacy concerns. Orchestration of such sensors through collaborative approaches and advanced data fusion systems based on machine learning is a critical problem and still considered as open [1]. Besides cooperation of sensors and fusion, transformative, or opportunistic, sensing and computing [4] are also of interest to adapt the individual sensors to support advanced tasks, far beyond their original designs.

B. Contributions

The paper addresses the design of an ecosystem of heterogeneous RF sensors that combines and transforms multiple radio sources into different sensing modalities. These modalities are all instrumental to HRC workspace monitoring and enable: i) the classification of potential unsafe worker conditions, ii) the localization of the worker and iii) the HR distance monitoring in the shared workspace. WiFi signals are also used for monitoring the perimeter of the robotic cell and for the detection of the number of workers moving in the workspace. As depicted in Fig. 1, different sensors cooperate to perform the perception of the worker by analyzing radio signals opportunistically. These signals originate from unmodified multi-antenna WiFi radios, i.e. pre-existing in the workplace, a network of radars monitoring the sub-THz (100-122GHz) band and a THz imaging camera (100GHz). Using combinations of radio technologies working in different bands for passive human sensing is addressed here for the first time. Considering that multiple radio sources are analyzed, we also propose a practical solution for data fusion and analytics obtained from distributed heterogeneous sensor devices. Data analytics is carried out on a cloud unit and it is based on the processing of features, namely uniform representations of raw data from different sensors. Features are obtained by edge units that are thus responsible of data pre-processing and fusion of different sensor inputs for the considered robotic cell. Edge units are in turn composed by several data processing pipelines each managed by independent micro-servers in charge of feature computation. Depending on the chosen sensing modality, namely localization, distance monitoring, activity recognition etc., human sensing in the context of HRC can be transferred into a classification problem, while features from fused raw data and different sensors are used as input for inference. The proposed cloud-edge platform is designed so that any sensing task can be implemented both individually and through collaborative work among different sensors with similar characteristics. It is also designed to effectively distribute and process massive amount of (multi-dimensional) raw data [8] and implement dedicated distributed learning tools.

Components, edge-side data pre-processing, feature extraction and manipulation and data fusion are discussed, particularly in Sect. III. Data fusion approach is based on feature extraction and processing on raw data and it is implemented on a dedicated edge device, monitoring an individual robotic cell (Fig. 1). Edge node consists of multiple micro servers/data centers (micro-edges) for extracting, processing and fusing features from different sensors. Learning over time series

of informative features can be used for specific applications running on the cloud and eliminate high dimensional data analytics.

Experimental activity inside a robotic cell has been conducted to evaluate the proposed approach and performance by combining several sensors. Different deep learning strategies running inside the cloud unit are compared to find the best performance trade off in terms of accuracy, latency and computational complexity for real-time worker detection. A specific case study, Sect. IV focusing on the problem of worker localization and activity recognition (safe vs unsafe activity discrimination) is addressed in detail and verified in a representative industrial pilot plant.

III. OPPORTUNISTIC SENSORS AND PRE-PROCESSING

In this section we describe the multisensory platform hardware (sensors) and related technology-dependent data pre-processing. As highlighted previously, the proposed platform consists of an ecosystem of sensor technologies that are transformed for the purpose of human-robot workspace sharing. The platform architecture depicted in Fig. 2 assigns one edge unit for each individual robotic cells. DEFINE THE ROBOTIC CELL IN SECT 2. The edge unit organizes the processing of data obtained from sensors into several *data processing pipelines*. It is thus in charge of extracting high level features from the fusion/aggregation of the pipelines. Each pipeline is managed by a virtual micro-edge that is responsible of aggregating the corresponding pipeline and contribute to the output features.

Data pipelines consist in general of a group of M sensors whose raw data require the same pre-processing. Each pipeline thus implements several sensor technology-dependent stages including data abstraction from raw radio signals and denoising (i.e. background subtraction). We adopt a wide range of passive radio sensing devices operating in the sub-THz (100GHz -122GHz), microwave (2.4-5GHz) and infrared (IR), from which it is possible to infer the EM radiation as perturbed by the presence of the human subject, or the IR/thermal radiation as emitted by the human subject being measured, respectively. Considering in general M sensors and N pipelines, the raw data $X_{k,i}(t)$ at time t for sensor $k = 1, \dots, M$ and pipeline $i = 1, \dots, N$ follows ad-hoc pre-processing stages as detailed in the following sections. The corresponding outputs $\tilde{X}_{k,i}$ samples are transferred to the edge computing units, i.e. the micro-edges, to obtain the corresponding fused features. The related machine learning tools for abstracting features are described in Sect. IV. Feature computation depends on the HRC functions to be implemented and is in turn controlled by a cloud computing unit, via a data controller feedback link.

A. THz radars and detectors

Terahertz (THz) propagation is extremely sensitive to environmental changes in the surroundings [27] as such changes typically affect both EM spreading and absorption losses. Transforming deployed THz infrastructures into sensing modalities might therefore become a promising opportunity. THz band communications [24] lie in the frequency gap

Fig. 2: Multisensory system model and data analytics

between 0.1 THz and 10 THz and to be widely applied in automotive and industrial applications [25]. Compared to the microwave radiation in the 1 – 50 GHz band, THz frequency range achieves a fairly good spatial resolution for precise imaging. In particular, the sub-THz band, from 0.1 THz up to 1 THz is expected to provide reduced multipath effects, and in turn increase the accuracy of human sensing and radio imaging at the expense of a reduced coverage area. In what follows, we focus on two different technologies in sub-THz band, namely 100 – 122 GHz. The first technology is based on a network of Frequency Modulated Continuous Wave radars (FMCW) and the latter is based on IMPATT diode technology [19] and array of THz detectors monitoring the intensity (signal strength) of the THz radiation.

FMCW radars transmitting antennas radiate a swept modulated waveform with the bandwidth 6 GHz (carrier frequency 119 GHz) and ramp (pulse) duration of $T = 1\text{ms}$. The

radar echoes, reflected by moving objects, and the transmitted signal are mixed at the receiver to obtain the beat signal. Beat signal is converted in frequency domain (i.e., beat signal spectrum) by 512-point Fast Fourier Transform (FFT) and averaged over 8 consecutive frames (i.e., frequency sweeps or ramps). Fig. 4 on top shows FFT images of 6 radars versus time collected in empty and occupied environment. FFT images are clearly sensitive to the presence of the subject. For device k , pipeline $i = 1$ and time t , FFT samples are the raw data inputs, therefore

$$X_{k,1}(t) = [x_{k,t}(1), \dots, x_{k,t}(512)], \quad (1)$$

while such inputs follow a pre-processing steps, namely denoising and background $\bar{X}_k(\emptyset)$ subtraction []

$$\tilde{X}_{k,1}(t) = \mathbf{C}_k^{-\frac{1}{2}} [X_{k,1}(t) - \bar{X}_k(\emptyset)] \quad (2)$$

with $\bar{X}_k(\emptyset) = E[X_{k,t}(\emptyset)]$ being the (time) average beat signal spectrum $X_{k,t}(\emptyset)$ observed in the empty robotic cell ($\emptyset$), or

background, and covariance $\mathbf{C}_k = \text{diag}[\sigma_k^2(1), \dots, \sigma_k^2(512)]$ and $\sigma_k^2(h) = \text{E}_t[X_{k,t} - \bar{X}_k(\emptyset)]^2$ used for denoising stage. Covariance takes into account any environmental change (i.e., due to concurrently robot movements) that might alter the the back scattered wavefield, yet without constituting any alert for worker safety. Post-processed data $\tilde{X}_{k,1}(t)$ are fed back to the corresponding micro-edge handler inside the edge node for increasing abstractions. In particular, each individual radar can be designed to discriminate body movements from robots and, in turn, to detect the distance of the worker from the robot with the purpose of identifying potential unsafe conditions. Combining radar outputs with other RF sensors enable more advanced worker perception modalities as show in Sect. VI.

THz detectors for Line-of-Sight monitoring are adopted as an additional sub-THz radio source (about 100mW at 100GHz) that uses an IMPATT diode (impact ionization avalanche transit-time diode) technology [28] and a sub-THz camera consisting of $K = 1024$ detectors arranged over a 2D array (32×32). The camera, developed by [26], has a noise-equivalent power of $1nW/\sqrt{\text{Hz}}$ in the frequency range $0.01 - 1$ THz and an acquisition rate up to 50fps (frames per second). The sub-THz camera captures the presence of the subject moving in the surrounding of the line-of-sight (LOS) path connecting the sub-THz source and the camera. In addition it can discriminate specific body movements as changing the received signal intensity [19]. Fig. 4 shows an example of signal intensity images of THz detector where the subject is moving inside the monitoring area. Notice that the measured intensity is proportional to the electric field power of the received electromagnetic waves normalized to the $[0 - 1]$ value range. Considering pipeline $i = 2$ Let $X_{k,2}(t)$ be a measure of the $[0 - 1]$ normalized intensity, or strength, of the radiated THz field observed at the the THz detector $k \in [1, \dots, K]$ at time t and for pipeline $i = 2$. The sampling interval Δt is chosen according to some application-dependent criteria detailed in Sect. VI. The K detectors of the camera monitor independently the alterations of the received THz radiation, while body activities leave a characteristic footprint on such observed radiation. In general, the multi-ray propagation between the fixed sub-THz source and receiver consists of line-of-sight (LOS), reflected, scattered and diffracted rays [19]. When no obstacles are present near the LOS path and the THz setup is directional (e.g., using dielectric lenses as in the experimental setup), the propagation is mostly due to the main LOS ray that depends only on the free-space loss and the absorption loss. Therefore, the radiation intensity is simply modeled as

$$X_{k,2}(t) = S_k(t) + W_k(\emptyset), \quad (3)$$

where $S_k(t)$ corresponds to the body-induced signature and $W_k(\emptyset) \sim \mathcal{N}(b_k, \sigma_k^2)$ models the background radiation (i.e., observed in the *empty* ($\emptyset$) environment, namely with the worker outside the detection area). The background component W_k is modeled here as Gaussian with average intensity b_k and deviation σ_k^2 . Average component b_k is a function of LOS terms: absorption and free-space loss. The absorption loss accounts for the attenuation that a propagating wave

suffers because of molecular absorption and depends on the concentration of molecules encountered along the path. This is considered a constant at the given center frequency (here $f = 100\text{GHz}$) and static temperature [25]. The free-space loss accounts for the attenuation due to the expansion of the wave as it propagates through the medium. Reflection, scattering and diffraction effects due to a worker or an obstacle (e.g., robot) near the LOS path introduce time-varying changes in (3). Deviation σ_k^2 models such environmental changes. Likewise FMCW radars, post-processed data $\tilde{X}_{k,2}(t)$ consists of signatures that are estimated from the input radiation measurements $X_{k,2}(t)$ through subtraction

$$\tilde{X}_{k,2}(t) = \sigma_k^{-1}(X_{k,2}(t) - b_k). \quad (4)$$

These are fed back, following the corresponding pipeline, to the corresponding micro-edge. In Sect. VI we exploit the THz radio technology as an additional sensor device to implement a *virtual safety fence* with the goal of isolating a human worker from robots that are cooperating in implementing the same HRC task. In particular, we explore the problem of data abstractions, fusion with other sensors inside the edge unit and HRC application-specific classification inside the cloud.

B. IR array sensors and body-induced thermal signatures

The use of thermal sensors for human body sensing [20], [21] is becoming attractive in many IoT-relevant scenarios, such as smart spaces, assisted living and industrial automation [22]. Thermal vision and related computing tools enable the possibility of analyzing body induced thermal signatures for detection of body motions, as well as discriminating those signatures from the environment. The experimental validation scenario of Sect. V exploits $K = 3$ thermopile sensors. Each sensor consists of an array of 64 thermopile elements organized in 2D 8×8 images (Panasonic Grid-EYE model [23]). Each element independently measures the captured IR radiation and translates to temperature readings that are inputs of data pipeline $i = 3$ (in Fig. 2)

$$X_{k,3}(t) = [x_{k,t}(1), \dots, x_{k,t}(64)]. \quad (5)$$

Herein, body-induced thermal signatures are used as inputs for feature computation (Sect. IV). These signatures $\tilde{X}_{k,3}(t)$ measure the temperature increase as induced by body movements and are related with corresponding raw temperature readings $X_{k,3}(t)$ by linear model, therefore

$$\tilde{X}_{k,3}(t) = X_{k,3}(t) - W_k(\emptyset) \quad (6)$$

where now $W_k(\emptyset) = [x_{k,t}(1), \dots, x_{k,t}(64)]$ conveys information about stationary heat-sources (i.e., robots, other machinery) that are not caused by body movements but characterize the background temperature of the empty space $\emptyset$. Computer vision [21] or statistical approaches [22] can be adopted for tracking of body motions. In Sect. VI we verify that fusion of such information with features obtained from THz radars is effective in improving worker motion discrimination in proximity with the robotic manipulator.

C. WiFi radios for sensing (Stephan, Sameera)

Pipeline $i = 4...$

IV. MULTISENSORY EDGE-CLOUD PLATFORM AND DATA ANALYTICS

The multisensory platform is introduced in this section by focusing on the edge/cloud computing units that are in charge of processing each pipeline, as shown in Fig. 2. More specifically, the edge nodes apply feature extraction and fusion of data pipelines based on the specific HRC functions (data controller) and using ML tools. In cobot environments, the features are used for example to implement dedicated speed-separation monitoring (SSM) tasks, to estimate the worker position/activity and, if necessary, to alarm when the worker performs dangerous activities in the surrounding of the manipulator.

A. Edge node: data fusion and distribution

Focus here is on data distribution at the edge of the multisensory platform as well as from the edge to the cloud unit. In particular, a fog computing structure [64] is adopted where both edge and cloud devices can provide seamless integration of sensors and might implement partial/full data fusion approaches. As shown in Fig. 3, the edge node contains N gateways. Gateways are serving as edge micro data centers and are referred to as micro-edges as they act as dedicated server/broker for the corresponding data pipeline. Considering telemetry collection on each pipeline, the data and the corresponding time-stamp information $\tilde{X}(t)$ are represented by JavaScript Object Notation (JSON), being a compact serialized form frequently used for communication of raw sensor data on serial ports. For transport level, both RESTful web services and Message Queuing Telemetry Transport (MQTT) can be adopted inside each micro-edge. Other approaches can be found in [??]. Micro-edges serve as bridges for individual sensors. Therefore, when using HTTP transport they act as a dedicated servers exposing resources, instead when using MQTT transport they act as brokers accepting subscriptions from the specific pipeline for telemetry publishing. Finally, the micro-edges also provide compute, storage and caching functions [65] for the corresponding pipeline.

As introduced before, each i -th micro-edge is in charge of mid-level of data analytics. Beside bridging functions and data maintenance, the micro-edge main task is to process the corresponding data-sets $\{\tilde{X}_{1,i}(t), \dots, \tilde{X}_{M,i}(t)\}$ to obtain high level features $\mathbf{V}_i = \{\mathbf{V}_{1,i}, \dots, \mathbf{V}_{M,i}\}$ for the considered pipeline. In general, features take the following form

$$\mathbf{V}_i = f_i(\tilde{X}_{1,i}(t), \dots, \tilde{X}_{M,i}(t) | \mathbf{V}_h, \forall h \neq i), \quad (7)$$

being a non-linear, pipeline-dependent, transformation $f_i(\cdot)$ of inputs $\{\tilde{X}_{1,i}(t), \dots, \tilde{X}_{M,i}(t)\}$ and functions of the features $\mathbf{V}_h, \forall h \neq i$ obtained from the other pipelines (if available). In Sect. 4 we show that these serve as “low-dimensional” representations of sensor data. Features are obtained individually at the edge node, however the cloud unit might supervise such stage through a data controller feedback (Fig. 3). Data

controller acts as feature updater to control data selection and collection based on the HRC functions/tasks.

Edge devices have limited processing and computing capacity. Therefore, they need to interact with a cloud unit that applies real-time inference, via ML tools, using all fused high level features $\{\mathbf{V}_1, \dots, \mathbf{V}_N\}$ as inputs. As shown in Fig. 3, output features are again serialized by the edge node using JSON representation and sent to the cloud using MQTT as transport layer. In what follows we describe data processing inside the cloud unit.

B. Cloud unit and HRC functions

The cloud unit processes the features received from the edges to infer a hidden process (i.e., subject motion, position or activity closeby the manipulator). Data fusion can be done in the edge, as in eq. (7), to obtain the features, but also inside the cloud by applying ML tools on fused features. In general, the cloud unit could decide whether complete or partial data fusion should be carried out by the individual edge, based on the specific HRC function to be implemented. Case studies in Sect. 4 highlight different fusion approaches.

ML and deep learning using features as inputs implement Long-Short Term Memory (LSTM) and Convolutional Neural Networks (CNN) models (see Fig. 2). ML models are optimized to allow for real-time manipulation of heterogeneous feature streams. With respect to end-user interfacing, the cloud unit provides an intermediate, open layer between third party applications (end user app.) and the underlying data manipulation resources. This is based on MQTT transport layer. In other words, third party applications might subscribe to the cloud for a specific HRC function of interest, and wait for real-time results pushed back by the ML unit, through the cloud service broker. Different sensing tasks could be supported based on the specific cobot applications. In particular, in Sect. 4 localization, worker counting TO BE VERIFIED and activity recognition are analyzed in detail. Such information are typically useful for controlling the manipulator and replanning its activity when needed.

V. LATENCY IN HUMAN-ROBOT COOPERATION

VI. EXPERIMENTAL VALIDATION AND CASE STUDIES

The proposed platform validated through experimental activity inside the controlled site with hybrid sensors technologies including FMCW radar, WiFi, and sub-terahertz sensors. In the following section we describe the cobot environment for the test, and experimental layout. Also localization and posture recognition result considering both non cooperative sensing obtained from individual data pipeline and cooperative sensing with multiple data pipelines are discussed and results are compared.

Three use cases: 1) THz radars and detectors monitoring the space in the surrounding of the robotic manipulator, 2) THz radars, detectors and IR array monitoring the HR workflow (low HR distance): we propose to combine IR sensors and THz radars for tracking the worker occupancy (position, motion intentions) inside the cobot environment, 3) worker counting via WiFi devices deployed in the surrounding of the robotic cell

Fig. 3: Edge, cloud unit. Feature computing, data fusion, distribution and transport layers.

A. human-robot collaborative plant

As shown in Fig. 1 the monitoring area is divided into safe and unsafe zones with a robot inside the zone and multiple sensors are deployed in the area. The safe zone is defined as the area in which the worker can move without limit, while in the unsafe zone the worker gets close to the robot and robot movements may cause injuries for the worker. Fig. 5 shows the experimental layout. As shown in the figure, six FMCW radars, sub-THz source and detectors are placed in the monitoring area. We defined 8 postures with body movements at 6 fixed positions marked in the figure. For example, a worker at position number 2 implements two different postures; 2A relates to the worker staying at position 2 and moving in different directions, while activity 2B corresponding to the worker staying at position 2 doing some activities such as taking pieces used for manufacturing in the production line. Each Radar sensor continuously monitors the area and collects samples and sends them to the corresponding micro edge. THz detector also receives signal intensity continuously and sends them to its micro edge node after denoising and normalization pre-processing. Edge node processes data from different pipelines i.e., here 6 radar and THz detectors, transferring to the fusion center extracting features and fusion algorithms, i.e. partial/complete fusion are applied according to the data controller message in the cloud. Data controller could decide about partial or complete fusion according to the third parties' request and specific application. For example, occupancy detection doesn't need to involve multiple data pipelines and presence/absence of the worker inside the area could be obtained from a couple of data pipelines with high accuracy, so in this case data controller decides for partial fusion to speed up the request performance in terms of real-time sensing.

Fused data will be transferred to the cloud through MQTT messages in real time to train data and learn the model. Here, we use LSTM and convolutional neural network in the cloud to train data and discover knowledge for the third parties. In this study, we apply the discovered knowledge to localize the worker position with corresponding activity. In the following section, results are discussed using non-cooperative sensing using individual pipeline and compared with the result of the cooperative sensing via sensor fusion approach.

B. Results and discussion

1) *Non-cooperative sensing*: In this section, we use FMCW radar and sub-THz detector separately to compare their performance for localization and posture recognition.

Fig. 4 on top shows the radar FFT sample image obtained from 6 sensors working at 119 GHz frequency band inside the monitoring area, where nobody is in the unsafe zone (environment) respect to the occupied environment. As shown in the figure, FFT images are sensitive to the worker presence and change clearly in the occupied environment.

Also in Fig. 6, signal intensity obtained from sub-THz camera changes at different positions in layout Fig. 5. Central part of the image shows the received signal intensity focused by camera lenses and changes with different body movements at different positions and gestures. For example, the signal intensity changes more at positions number #2B and #5B with respect to the other marked positions because they are close to the sub-THz detector and any changes in the environment are captured precisely by the camera.

This information is used to classify positions and also activities of the worker inside the plant. LSTM and CNN tools are used for training and position/activity classification. CNN

Fig. 4: Example of multisensory data analytics using hybrid sensors

can capture local features of input radio data because of its spatially located convolutional filters, while LSTM works precisely for time series data sequences. Focusing on THz detector, LSTM uses signal images with 32×32 dimension size as an input and 80 % of data is used for training and 20 % for realtime classification. CNN exploits the same image size (32×32) as LSTM with 18 layers including 4 convolutional layer with 3×3 , and 8, 16, 32, 64 numbers at each layer, 2 pooling layers, 4 batch normalization layers, 4 relu layers, and fully connected and regression layers.

Fig. 7 shows confusion matrix result using LSTM and CNN for non cooperative sensing. Fig.7 (a) and (c) represent sub-THz detector using LSTM and CNN, while Fig.7 (b) and (d) show confusion matrix result radar and LSTM and CNN. Result confirm that CNN outperforms in both accuracy. We also compare the execution time for LSTM and CNN and

TABLE I: Execution time (second)

Method	LSTM	CNN
sub-THz	6.1041e+03	282.8991
Radar	1.0494e+04	1.0374e+02

results are presented in Table 1. With the implemented CNN system the computational burden has been reduced almost 283 second for sub-THz and 1.0374×10^2 seconds for radars

2) *Cooperative sensing via fusion:* In this section, we fused data information from both sub-THz detector and radar sensors. Data normalized based on min-max approach []. LSTM and CNN tools used with the same characteristics of the non cooperative sensing. Fig. 8 shows the confusion matrix result for both LSTM and CNN. Accuracy result in Fig.8 confirms that the proposed sensor fusion approach improves

Fig. 5: Multisensory layout in industrial plant

TABLE II: Execution time (second)

Method	LSTM	CNN
Sensor Fusion	1.0362e+04	1.0444e+03

the accuracy result to 96.9% for LSTM and 94.4% for CNN. Computational time is also computed and result shown in table 2.

C. Worker counting via ambient WiFi signal inspection

Stephan+Sameera.....

VII. CONCLUDING REMARKS AND OPEN PROBLEMS TO BE DONE.....

REFERENCES

- [1] P. A. Lasota, T. Fong, and J. A. Shah, "A survey of methods for safe human-robot interaction," *Found. Trends Robot.*, vol. 5, no. 4, pp. 261–349, 2017.
- [2] F. Vicentini, M. Askarpour, M. G. Rossi and D. Mandrioli, "Safety Assessment of Collaborative Robotics Through Automated Formal Verification," *IEEE Transactions on Robotics*. doi: 10.1109/TRO.2019.2937471, 2019.
- [3] G. Michalos, N. Kousi, P. Karagiannis, et al., "Seamless human robot collaborative assembly – An automotive case study," *Mechatronics*, vol. 55, 2018, pp. 194-211, ISSN 0957-4158, <https://doi.org/10.1016/j.mechatronics.2018.08.006>.
- [4] M. Youssef and F. Kawsar, "Transformative Computing and Communication," *Computer*, vol. 52, no. 7, pp. 12-14, July 2019.
- [5] F. Vicentini, M. Giussani, and L. M. Tosatti, "Trajectory-dependent safe distances in human-robot interaction," in *Proc. IEEE Emerg. Technol. Factory Autom.*, Sep. 2014, pp. 1–4.
- [6] A. Tellaeche, I. Mourtua and A. Ibarguren, "Use of machine vision in collaborative robotics: An industrial case.," *IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, Berlin, 2016, pp. 1-6. doi: 10.1109/ETFA.2016.7733689
- [7] J. Lianghai, et al., "Applying Device-to-Device Communication to Enhance IoT Services," in *IEEE Communications Standards Magazine*, vol. 1, no. 2, pp. 85-91, 2017.
- [8] S. Kianoush, et al., "A Cloud-IoT Platform for Passive Radio Sensing: Challenges and Application Case Studies," in *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3624-3636, Oct. 2018.
- [9] Sigg et al., "RF-sensing of activities from non-cooperative subjects in device-free recognition systems using ambient and local signals," *IEEE TMC*, vol. 13, no. 4, 2014.
- [10] Wilson et al., "See-through walls: Motion tracking using variance-based radio tomography networks," *IEEE TMC*, vol. 10, no. 5, 2011.
- [11] Bartoletti et al., "Device-Free Counting via Wideband Signals," *IEEE J. sel. areas Comm.*, vol. 35, no. 5, 2017.
- [12] Wang et al., "Wifall: Device-free fall detection by wireless networks," *IEEE TMC*, vol. 16, no. 2, 2017.
- [13] Seifeldin et al., "Nuzzer: A large-scale device-free localization system for wireless environments," *IEEE TMC*, vol. 12, no. 7, 2013.
- [14] Deak et al., "Detection of multi-occupancy using device-free passive localisation," *IET Wireless Sensor Systems*, vol. 4, no.3, 2014.
- [15] S. Savazzi, S. Sigg, F. Vicentini, S. Kianoush and R. Findling, "On the Use of Stray Wireless Signals for Sensing: A Look Beyond 5G for the Next Generation of Industry," *Computer*, vol. 52, no. 7, pp. 25-36, July 2019.
- [16] Shi et al. "Accurate Location Tracking from CSI-based Passive Device-

Fig. 6: Image of human body movements captured by Terahertz detector at predefined locations mentioned in layout in Fig.2. Position number at images corresponds to predefined locations, A and B corresponds to human body wo/w body movements.

- free Probabilistic Fingerprinting,” IEEE transaction on vehicular technology, Doi: 10.1109/TVT.2018.2810307, 2018.
- [17] S. Savazzi, et al. “Is someone moving around my cell-phone? Tracing cellular signals for passive motion detection,” in Proc. of IEEE PerCom Workshops, Kona, HI, pp. 10-13, March 2017.
- [18] Yousefi et al., “A survey on behavior recognition using wifi channel state information,” *IEEE Comm. Mag.*, vol. 55, no. 10, 2017.
- [19] S. Kianoush, et al., “Passive detection and discrimination of body movements in the sub-THz band: a case study”, Proc of IEEE Conference on Acoustics, Speech and Signal processing (ICASSP), 2019.
- [20] Rikke Gade and Thomas B. Moeslund, “Thermal cameras and applications: a survey,” *Machine Vision and Applications*, vol. 25, no. 1, pp. 245–262, Jan 2014.
- [21] Ash Tyndall, Rachel Cardell-Oliver, and Adrian Keating, “Occupancy estimation using a low-pixel count thermal imager,” *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3784–3791, May 2016.
- [22] [11] S. Savazzi, V. Rampa, S. Kianoush, A. Minora and L. Costa, “Occupancy Pattern Recognition with Infrared Array Sensors: A Bayesian Approach to Multi-body Tracking,” IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, United Kingdom, pp. 4479-4483, 2019.
- [23] J. Tanaka, H. Imamoto, T. Seki, and M. Oba, “Low power wireless human detector utilizing thermopile infrared array sensor,” Proc. of IEEE Sensors, pp. 461-465, Valencia, Spain, Nov 2014.
- [24] Akyildiz, I. F. et al., “Terahertz band: Next frontier for wireless communications,” *Physical Communication*, 2014, 12, 16 - 32
- [25] T. Nagatsuma, “Generating millimeter and terahertz waves,” in *IEEE Microwave Magazine*, vol. 10, no. 4, pp. 64-74, June 2009.
- [26] V.M. Muravev, et al., “Novel Relativistic Plasma Excitations in a Gated Two-Dimensional Electron System”, *Phys. Rev. Lett.* 114, 106805 ? Published 10 March 2015.
- [27] Shur, M.” Subterahertz and terahertz sensing of biological objects and chemical agents”, *Proc.SPIE*, 2018.
- [28] Mukherjee, M. et al., “GaN IMPATT diode: a photo-sensitive high power terahertz source Semiconductor Science and Technology”, 2007, vol. 22, no.12, pp.1258, 2007
- [29] Savazzi et al., “Device-Free Radio Vision for assisted living: Leveraging wireless channel quality information for human sensing,” *IEEE SP Mag.*, vol. 33, no. 2, 2016.
- [30] Scholz et al., “Sensewaves: Radiowaves for context recognition,” in *Video Proceedings of Pervasive 2011*, Jun. 2011.
- [31] Holl et al., “Holography of Wi-Fi Radiation,” *Phys. Rev. Lett.* 118, 183901, May 2017
- [32] Kaltiokallio et al., “ARTI: An Adaptive Radio Tomographic Imaging System,” *IEEE TVT*, vol. 66, no. 8, 2017.
- [33] Kianoush et al., “Device-Free RF Human Body Fall Detection and Localization in Industrial Workplaces,” *IEEE IoT J.* vol. 4, no. 2, pp. 351-362, April 2017.
- [34] Woznowski et al., “Sphere: A sensor platform for healthcare in a residential environment,” *Designing, Developing, and Facilitating Smart Cities*. Springer, Cham, 2017.
- [35] Wang et al., “Gait recognition using wifi signals,” *ACM Ubicomp*, 2016.
- [36] Di Domenico et al., “Trained-once device-free crowd counting and occupancy estimation using wifi: A doppler spectrum based approach,” *IEEE WiMob*, 2016.
- [37] Liu et al., “Tracking vital signs during sleep leveraging off-the-shelf wifi,” *ACM MobiHoc*, 2015.
- [38] Li et al., “Wifinger: talk to your smart devices with finger-grained gesture,” *ACM Ubicomp*, 2016.
- [39] Wang et al., “We can hear you with wi-fi!” *IEEE TMC*, vol. 15, no. 11, 2016.
- [40] Wang et al., “Human respiration detection with commodity wifi devices: do user location and body orientation matter?” *ACM Ubicomp*, 2016.
- [41] Feng et al., “Mais: Multiple activity identification system using channel state information of wifi signals,” in *WASA*, 2017.
- [42] Zhao et al., “Emotion recognition using wireless signals,” *ACM Mobi-com*, 2016.
- [43] C.Yang et al., “Sensing-Data Curation for the Cloud is Coming: A Promise of Scalable Cloud-Data-Center Mitigation for Next-Generation IoT and Wireless Sensor Networks,” *IEEE Consumer Electronics Magazine*, Vo.6, no.4, pp.48-56, 2017.
- [44] Yasin et al., “Ultra-Low Power, Secure IoT Platform for Predicting Cardiovascular Diseases,” *IEEE TCS*, vol.64, no. 9, 2017.
- [45] Liang et al., “Enhancing QoE-Aware Wireless Edge Caching With Wireless SDN,” *IEEE TWC*, vol. 16, no. 10, 2017.
- [46] Zhu, et. al, “Robust and passive motion detection with COTS WiFi devices,” *IEEE Tsinghua Science and Technology J.*, vol.22, no.4, 2017.

Fig. 7: Confusion matrix result for (a), (c) sub-THz detector and (b) ,(d) Radar technology separately using LSTM and CNN tools for the layout described in Fig.2 when robot moves.8 classes correspond to 8 activities of worker at 6 predefined positions. SPLIT IN TWO FIGURES

Sensor fusion

Fig. 8: Sensor fusion (i.e., radar and sub-THz detector) and confusion matrix result using (a) LSTM and (b) CNN for the layout in Fig.2 when robot moves. 8 classes correspond to 8 activities of worker at 6 predefined positions.

- [47] Savazzi et al., "Wireless Cloud Networks for the Factory of Things: Connectivity modeling and layout design," IEEE IoT J., vol.1, no.2, 2014.
- [48] Alamri, et al., "A Survey on Sensor-Cloud: Architecture, Applications, and Approaches," Int. J. of Distr. Sensor Netw., vol.9, no. 2, 2013.
- [49] Munir et al., "Integrated Fog Cloud IoT: A novel architectural paradigm for future IoT," IEEE Cons. Electr. Mag., vol.6, no. 3, 2017.
- [50] Jayasri et al., "Link Quality Estimation for Adaptive Data Streaming in WSN," Int. J. Wirel. Personal Comm., vol.94, no.3, 2017.
- [51] Chang et al., "FitLoc: Fine-grained and low-cost device-free localization for multiple targets over various areas," IEEE/ACM TN, 2017.
- [52] Pu et al., "Whole-home gesture recognition using wireless signals," Mobicom, 2013.
- [53] Saeed et al., "Ichnaea: A low-overhead robust WLAN device-free passive localization system," IEEE J. Sel. Top. SP, vol. 8, no. 1, 2014.
- [54] Shi et al., "Activity recognition from radio frequency data: Multi-stage recognition and features," in IEEE VTC Fall, 2012.
- [55] Wu et al., "Device-Free WiFi Human Sensing: From Pattern-Based to Model-Based Approaches," IEEE Comm. Mag., vol. 55, no. 10, 2017.
- [56] Fang et al., "Channel State Reconstruction Using Multilevel Discrete Wavelet Transform for Improved Fingerprinting-Based Indoor Localization," IEEE Sensors J., Vol.16, no. 21, 2016.
- [57] Alaya et al., "Toward semantic interoperability in oneM2M architecture," IEEE Comm. Mag., vol. 53, no. 12, 2015.
- [58] Singh et al., "Smart city environmental perception from ambient cellular signals," in International Conference on Algorithms and Architectures for Parallel Processing. Springer, 2017.
- [59] Zou et al., "Grfid: A device-free rfid-based gesture recognition system," IEEE TMC, vol. 16, no. 2, 2017.
- [60] Lien et al., "Soli: Ubiquitous gesture sensing with millimeter wave radar," ACM TOG, vol. 35, no. 4, 2016.
- [61] Adib et al., "Capturing the human figure through a wall," ACM TOG, vol. 34, no. 6, 2015.
- [62] D. Guinard, et al. "Towards the web of things: web mashups for embedded devices," Proc. of the International World Wide Web Conference 2009 (WWW 2009), Madrid, Spain, April 2009.
- [63] He Li, et al., "Learning IoT in Edge: Deep Learning for the Internet of Things with Edge Computing," IEEE Network, vol. 32, no. 1, pp. 96-101, 2018.
- [64] Yuyi Mao, et al., "A Survey on Mobile Edge Computing: The Communication Perspective," IEEE Comm. Surveys and Tutorials, vol. 19, no. 4, 2017.
- [65] L. Mainetti, et. al, "A Software Architecture Enabling the Web of Things," IEEE IoT J., vol. 2, no. 6, pp. 445-454, 2015.
- [66] A. Ferrer, et. al. "Multi-cloud Platform-as-a-service Model, Functionalities and Approaches," Procedia Computer Science. 97. pp. 63-72, October, 2016.
- [67] Wu et al., "Composable 10: A Novel Resource Sharing Platform in Personal Clouds", CloudCom, 2009.
- [68] Cianca et al., "Radios as Sensors," IEEE IoT J., vol. 4, no. 2, 2017.
- [69] M. Tipping, C. Bishop, "Probabilistic principal component analysis," J. of Royal Stat. Society Series, col 21 no. 3, pp. 611-622, 1999.
- [70] R. Vidal, "Subspace Clustering," IEEE Signal Processing Magazine, vol. 28, no. 2, pp. 52-68, March 2011.
- [71] Swetina et al., "Toward a Standardized Common m2m Service Layer Platform: Introduction to onem2m," IEEE Wirel. Comm., vol. 21, no 3, 2014.
- [72] K. Maeda, "Performance evaluation of object serialization libraries in XML, JSON and binary formats," Proc. of Digital Information and Communication Technology and it's Applications (DICTAP), pp. 177-182, 2012.
- [73] Gerla, "Vehicular Cloud Computing", Med-Hoc-Net, 2012.
- [74] Whaiduzzaman et al., "A survey on vehicular cloud computing," J. Netw. & Comp.Applications, vol 40, 2014.
- [75] Burgoon et al., "Nonverbal communication". Allyn & Bacon, Boston, MA, 2016.
- [76] Basseville et al., "Detection of Abrupt Changes", Theory and Applications. Prentice-Hall, 1993.