



Move With the Theremin: Body Posture and Gesture Recognition Using the Theremin in Loose-Garment With Embedded Textile Cables as Antennas

Hymalai Bello^{1*}, Bo Zhou^{1,2}, Sungho Suh^{1,2}, Luis Alfredo Sanchez Marin² and Paul Lukowicz^{1,2}

¹ German Research Centre for Artificial Intelligence, Embedded Intelligence, Kaiserslautern, Germany, ² Department of Computer Science, University of Kaiserslautern, Kaiserslautern, Germany

OPEN ACCESS

Edited by:

Xianta Jiang,
Memorial University of Newfoundland,
Canada

Reviewed by:

Philipp Marcel Scholl,
University of Freiburg, Germany
Junhuai Li,
Xi'an University of Technology, China

*Correspondence:

Hymalai Bello
hymalai.bello@dfki.de

Specialty section:

This article was submitted to
Mobile and Ubiquitous Computing,
a section of the journal
Frontiers in Computer Science

Received: 07 April 2022

Accepted: 25 May 2022

Published: 22 June 2022

Citation:

Bello H, Zhou B, Suh S, Sanchez
Marin LA and Lukowicz P (2022)
Move With the Theremin: Body
Posture and Gesture Recognition
Using the Theremin in Loose-Garment
With Embedded Textile Cables as
Antennas.
Front. Comput. Sci. 4:915280.
doi: 10.3389/fcomp.2022.915280

We present a novel intelligent garment design approach for body posture/gesture detection in the form of a loose-fitting blazer prototype, “the MoCaBlazer.” The design is realized by leveraging conductive textile antennas with the capacitive sensing modality, supported by an open-source electronic theremin system (OpenTheremin). The use of soft textile antennas as the sensing element allows flexible garment design and seamless tech-garment integration for the specific structure of different clothes. Our novel approach is evaluated through two experiments involving defined movements (20 arm/torso gestures and eight dance movements). In cross-validation, the classification model yields up to 97.18% average accuracy and 92% f1-score, respectively. We have also explored real-time inference enabled by a radio frequency identification (RFID) synchronization method, yielding an f1-score of 82%. Our approach opens a new paradigm for designing motion-aware smart garments with soft conductive textiles beyond traditional approaches that rely on tight-fitting flexible sensors or rigid motion sensor accessories.

Keywords: loose garment sensing, theremin, capacitive sensing, activity recognition, posture detection, gesture detection, real-time recognition, RFID

1. INTRODUCTION

Human activity recognition (HAR) is an umbrella term that gives shelter to various specific applications to understand human behavior. An essential piece of HAR is body postures and gestures (BPG) recognition. The popularity of BPG recognition is well earned due to the ability to describe human activities by a sequence of changing postures or by detecting specific gestures (Ding et al., 2020). BPG detection could lead to the generation of emotion and personality profiles (Noroozi et al., 2018; Junior et al., 2019), to understand implicit social interactions (Gaschler et al., 2012; Guedjou et al., 2016), to aid in sign language communication (Enikeev and Mustafina, 2020), and to predict people’s intentions (Sanghvi et al., 2011).

Many wearables sensing applications have found their purpose in BPG, delivering highly developed solutions such as commercial motion capture systems (Schepers et al., 2018). The commercial and research markets for BPG recognition are mainly dominated by inertial

measurement units (IMU) wearable-based techniques (Harms et al., 2008; Sarangi et al., 2015; Butt et al., 2019), and on the textile side by stretch or pressure sensors (Chander et al., 2020).

Most of the current solutions for BPG recognition have a common baseline requirement: the sensors need to be firmly attached to the body using tight garments or dedicated accessories, such as bracelets and straps. Therefore, we could argue that a reliable method for BPG recognition with loose garments remains a largely open problem. The present study explores further the simple and novel method for BPG detection in Bello et al. (2021), which proposes a loose garment solution based on non-contact capacitive sensing with off-the-shelf components. In addition to our previous study in Bello et al. (2021), a set of dance movements is evaluated to demonstrate the potential use case of our design as a sophisticated/elegant game controller. The inspiration came from the Nintendo Wii Rayman Raving Rabbids®: TV Party- ShakeTV (Wii, 2009).

The main component of our system is a modified electronic musical instrument, the theremin (Skeldon et al., 1998) for BPG recognition. The well-known musical instrument usually consists of one or two long metal rod/loop antennas emitting sub-MHz frequencies. As the thereminist moves inside the antennas' range, volume and pitch can be controlled by their hand's position. The theremin antennas are metallic, but any conductive wire/textile can be used as an antenna due to its intrinsic capacitive sensing. We substituted the metal rod with soft wires and integrated them inside a loose-fitting garment.

Our experiment design validates our system within two discrete gesture dictionaries: 20 generic/typical upper-body postures and gestures and a second dictionary with eight dance movements. A distinct aspect of our approach is that the theremin's antennas move with the wearer's body motion, changing the signals. Our main contribution is to expand our wearable and loose-fitting solution for BPG recognition in Bello et al. (2021). The prototype is based on off-the-shelf components, such as a modified electronic musical instrument, the "OpenTheremin" (Gaudenz, 2016), which in conjunction with textile antennas, is embedded into a loose men's jacket. The system shows accuracy above 90% with an evaluation based on several deep neural network models. In this extended study, the "MoCaBlazer" was fused with Radio Frequency Identification (RFID) synchronization for a real-time and wireless recognition for one participant and six classes of a dance movements dictionary, obtaining an f1-score = 82%. Hence, our "MoCaBlazer" could be a promising alternative for an elegant/sophisticated game controller.

Our article structure is as follows; Section 2 introduces related study in the areas of loose-fitting wearables for BPG and capacitive sensing-based solutions. Next, Section 3 provides a detailed description of the electronic prototype, including details about the data collection options. Then, Section 4 describes the experimental design to evaluate our system. Subsequent, Section 5 illustrates the strategy for the evaluation of the system within the two experiment scenarios; a general gestures dictionary and a dance movements dictionary. Next, Sections 6, 7 present the results and discussion of the deep learning models used to verify the feasibility of our method.

Finally, in Section 8, we conclude our study and discuss further ideas.

2. RELATED STUDY

2.1. Loose Fitting Wearables for BPG

Inertial measurement units (IMU) distributed in clothing or accessories for BPG recognition is a widely used technique (Harms et al., 2008; Sarangi et al., 2015; Butt et al., 2019). Another relevant approach for BPG analysis is called kinesiological electromyography (EMG) (Clarys and Cabri, 1993; Zhang et al., 2019). Such approaches are reliable and robust solutions with accuracy above 90%. One of the limitations they share is the need for stable sensor positions to avoid the effect of noise and motion artifacts on the signals. Furthermore, the placement of discrete and rigid sensors around the joints could be uncomfortable for the user. In Loke et al. (2021), the authors employed 100 microchips with memory and temperature sensors interconnected in a flexible fiber on a T-shirt, which is a solution to increase the flexibility and comfort of the user while wearing discrete sensors, a promising idea to explore in the future.

On the other hand, stretchable garments with strain-based or pressure sensing methods have been studied by many researchers (Boyalı et al., 2012; Jung et al., 2015; Zhou et al., 2017; Skach et al., 2018; Mokhlespour Esfahani and Nussbaum, 2019; Lin et al., 2020; Ramalingame et al., 2021; Shin et al., 2021), which demonstrate their value in textile based BPG recognition. Fiber optic embedded in a jacket and pants was proposed in a limited study (one person) (Koyama et al., 2018); the transmitted light changes with the wearer's movements, creating a time series pattern due to the bending of the fiber optics. Wearable optical technology is growing rapidly with multiple hardware designs being proposed by Koyama et al. (2016), Abro et al. (2018), Koyama et al. (2018), Zeng et al. (2018), Leal-Junior et al. (2020), Swaminathan et al. (2020), and Li et al. (2021). A fabric-based triboelectric sleeve is proposed in Kiaghadi et al. (2018). Four Radio Frequency Identification (RFID) tags were proposed on the back, chest, and feet over the persons' clothes and shoes by Wang et al. (2016) to recognize a total of eight activities (standing, sitting, walking, along with others). The piezoelectric effect was employed in Cha et al. (2018), where four flexible piezoelectric sensors were placed on the knee and the hip in slack pants to detect walking, standing, and sitting activities.

Table 1 shows a detailed comparison of state-of-the-art sensing on the garment for activity recognition solutions. At the bottom of the table, our system shows a quick and easy option to integrate e-textile components in loose-fitting garments such as the "MoCaBlazer." The "MoCaBlazer" uses commercial conductive textile parts as the antennas of the modified off-the-shelf theremin (OpenTheremin) based on capacitive sensing.

2.2. Capacitive Sensing

Capacitive sensing is a well-developed technology, available in our everyday life since the invention of the first cellphone with a touch screen (Johnson, 1965). In the cellphone touch screen case, capacitive technology estimates touch or deformation caused by fingers. A capacitance measurement quantifies the electric charge

TABLE 1 | Comparison with state-of-the-art non-camera based sensing on garment methods for human activity recognition.

Studies	Device	Activities	Classification method	Accuracy (%)	Persons
SMASH:Long sleeve shirt (Harms et al., 2008)	Three accelerometers (3D)	Twelve arm movements: Idle, shoulder abduction 40 and 90°, shoulder flexion 90°, shoulder elevation 170°, shoulder abduction 15°, shoulder abduction 90°, shoulder rotation 90° in and out, elbow flexion 90 and 130°, neck-grip, skirt-grip	Nearest Centroid Classifier; window = 1 s	95.00%*	8
Shirt:Digital Electronic (Loke et al., 2021)	Flexible fiber: 100 microchips with temperature sensing	4 motor activities: Sit, Stand, Walk, and Run	CNN; window = 12 s	96.40*	1
Jacket and pant (Koyama et al., 2018)	Hetero-core fiber optics	8 motor activities: Standing, sitting, walking, descending stairs, climbing stairs, lying, eating, and drinking	SVM; window = 4.83 s	98.70*	1
Sleeve (Kiaghadi et al., 2018)	Fabric-based triboelectric joint sensing	4 daily activities: Brushing, eating, walking, idle	SVM; window=** **	91.30*	14
RFID system (Wang et al., 2016)	4 antennas; back, chest, and feet	5 motor + 3 cleaning activities: Sitting, standing, walking, cleaning window, cleaning table, vacuuming, riding bike, going up/downstairs	SVM; window = 5 s	93.60*	4
Sweat jacket (Lin et al., 2020)	Optical-strain sensor	5 motor activities: Standing, sitting, lying, walking, running	CNN-LSTM; window = 4 s	90.90*	12
Elastic sport band (Zhou et al., 2017)	Textile pressure matrix (TPM)	4 gym exercises + 3 non-exercises: Cross trainer, leg press. Seated leg curl. Leg extension. And adjusting machines, pause, and walking as one class	ConfAdaBoost; window = 8 s	93.30*	6
Loose Pants (Cha et al., 2018)	Flexible piezoelectric	5 motor activities + 8 transitions: Walking, standing, sitting, supine, sitting knee extension and 8 transitions between those	Rule-based algorithm (Cha et al., 2017)	93.00*	10
Trousers (3 sizes) (Skach et al., 2018)	Textile pressure sensors	19 Sitting postures/gestures: Sitting postures/gestures: Standing up, sitting down, sitting straight, leaning back, leaning forward, slouching, etc	Random Forest; window = * * *	99.18 [†]	6
Air bladder band (Jung et al., 2015)	Air pressure sensors	6 hand gestures: Flexion and extension of the wrist, flexion, and extension of the fingers, and radial and ulnar deviation of the wrist	Custom Fuzzy Logic	90.00***	6
Stretchable textile tape (Ramalingame et al., 2021)	8 nanocomposite pressure sensors	10 American sign language numbers: numerical gestures (0–9)	ELM; window = * * *	93.00*	10
Glove (Shin et al., 2021)	EGaln-Silicone Soft: React to pressure or stretch	12 Static hand gestures: rest, hand close, and numerical gestures (0–9)	Random Forest; window = 200 ms	97.30*	15
Armband, commercial (Zhang et al., 2019)	Surface EMG	5 hand gestures: Double tap, wave in and out, fingers spread and fist	ANN; window = 400 ms	98.70*	12
Leg/chest band, insole (Haescher et al., 2015)	Capacitive	5 motor activities: Sneaking, walking, fast walking, jogging and walking with weight	Bayesian Classifiers; window = 17 s	88.97**	10
Our approach- General Dictionary	Capacitive	20 posture/gestures (Figure 2)	Conv2D; complete instance ~ 4 s	97.18*, 86.25**	14
Our approach- Dance Dictionary	Capacitive	8 dance movements (Figure 3)	1DConv; complete instance ~ 4 s	92.00*	3

*User dependent solution; **User independent solution; ***Information no available.

storage between two or more conductors, called electrodes. The electrodes are conductive plates that form a chamber, and when they are at different electric potentials (voltages), an electric field is generated. The ratio between the charge (Q) and the differential electric potential is called capacitance. Although the electrodes are usually made of metal, any two plates of conductive material like inks, foils, indium tin oxide (ITO), plastics, textiles, and even the human body can be used to build a capacitor (Grosse-Puppenthal et al., 2017). Capacitance is measured by frequency or duty cycle, which fluctuates when external electrodes disturb

the status quo. Another method is by quantifying the charge balance, or with rising or falling time measurements (Perme, 2007).

In wearable and ubiquitous computing for HAR, capacitive sensing has extensively proven its importance (Braun et al., 2015b; Ye et al., 2020). The applications extend from capacitive furniture (Wimmer et al., 2007; Braun et al., 2015a,b; Liu et al., 2019), to capacitive wristbands (Cohn et al., 2012a; Pouryazdan et al., 2016; Bian et al., 2019a,b), rings (Wilhelm et al., 2015), clothes

(Holleis et al., 2008; Singh et al., 2015), collars (Cheng et al., 2010, 2013) and prosthesis (Zheng and Wang, 2016) up to an entire wall painted as a capacitive array (Zhang et al., 2018) for posture gesture detection.

A textile design was evaluated as an on-body capacitance system by Cheng et al. (2010). The authors validated the technology for eating, head inclination, and arm/leg movements; sensors were placed on the neck, wrist, upper leg, and forearm, though not a loose-fitting solution. In Singh et al. (2015), a flexible textile capacitive matrix was placed on the volunteer's upper leg. The goal was to recognize swipe and hover gestures of paralysis patients. In Cohn et al. (2012b), a capacitive backpack was worn by eight volunteers for posture recognition. The bag works as a receiver of electromagnetic (EM) noise from the power lines and electronic devices inside a room. By measuring the disturbances on the EM field caused by the wearer's movements, the system achieved 93% accuracy for 12 gestures.

The above sensing studies (Cheng et al., 2010; Cohn et al., 2012b; Singh et al., 2015; Zhang et al., 2018) mainly employed tightly coupled or stationary electrodes. In this study, we proposed to use textile theremin antennas in a loose-fitting garment, the "MoCaBlazer" for BPG.

3. ELECTRONICS AND GARMENT PROTOTYPE

The principal component in our electronic garment prototype is an off-the-shelf electronic musical instrument, "The OpenTheremin V3" (Gaudenz, 2016)¹. The theremin produces musical notes based on the frequency fluctuation of its antennas caused by the proximity of a person's hands. In a theremin, we could find two antennas, one for volume (loop antenna) and another for pitch control (rod antenna) (Skeldon et al., 1998). Capacitive sensing is the physical principle governing the behavior of the theremin. The human body could be modeled as a capacitor plate virtually connected to the earth and, in conjunction with the theremin's antennas (second plate), completes a capacitor (Singh et al., 2015). Thus, human proximity changes the effective capacitance of the Clapp LC oscillator in **Figure 1D**, affecting its frequency. Therefore, we could infer that relative differences between body parts and theremin's antennas could be used to distinguish body postures. In the present study, the pitch and volume antennas were embedded in a tailored garment (men's blazer); thus, the person's body moves with the theremin and "makes music" with different postures and gestures (frequency profiles).

To test our approach, we designed a prototype, the "MoCaBlazer," as shown in **Figure 1**. We employed a Tom Tailor®L/52 size blazer (best suited for 184 cm tall persons). In **Figures 1A,C**, the positions and patterns of our four antennas are depicted. The antennas cover the chest, a small part of the shoulders, the arms, and the back, as seen in **Figure 1A**. This setting was appropriate for detecting upper-body postures and

gestures without altering the tailored garment's main structure or hindering the wearer's motion.

The back antennas (standard 28 AWG cables) (AMPHENOL, 2015) in **Figure 1A** (Arm-Left, Arm-Right) start from the side pockets and, following a curving pattern (simulating a volume antenna), pass over the latissimus dorsi muscles toward the deltoids; they then turn sharply to go along the outer sleeve lines and terminate before the cuff buttons. The front antennas (TWC24004B textile cables) (Wear, 2021) in **Figure 1C** (Collar-Left, Collar-Right) were sewn inside the lining without modifying the structural design of the blazer (refer to **Figure 1B**)². The Collar-Left and Collar-Right antennas were arranged to simulate a theremin's pitch antenna as close as possible. Thus, they begin on the side pockets and go to the front-top button, then turn to align with the inner crease of the lapels and reach the notch; consecutively lead out of the crease and climb around the shoulder to the back, and end at the middle edge of the shoulder pad. The antennas' lengths are 80 cm (front) and 100 cm (back) for this particular blazer size (L/52).

Two "OpenTheremin" boards were inside the side pockets of the "MoCaBlazer" (refer to **Figure 1**) to handle four channels. The channel frequencies were modified by changing the capacitor (C2) in the clap-oscillator circuit to minimize cross-talk between them, as depicted in **Figure 1D**. Then, the channels were sampled (frequency-count; Stoffregen, 2014) at 100 Hz by the Teensy®4.1 (Stoffregen, 2020) development board.

Two options are available for the data collection: a UART serial (115,200 Baud rate) as a wired option and a Bluetooth serial (9,600 baud rate) as a wireless option. In the case of the wired alternative, the data is received by the serial port (USB) in a computer. The computer runs a python script with a graphical user interface (GUI) developed using Tkinter (Lundh, 1999), as depicted in **Figure 1E** upper element. For the wireless option, the data of the four channels is sent using the Huzzah-ESP32 Bluetooth serial protocol (Fried, 2022) (in the upper pocket) to a smartphone. The smartphone runs an android application, developed using the Flutter framework (Napoli, 2019), as shown in **Figure 1E** lower element.

4. EXPERIMENT DESIGN

Two experiments were conducted with our garment prototype, the "MoCaBlazer." The experiments were carried out in an office without user calibration, i.e., without tuning the antennas' base frequencies to reduce the impact of different body capacitances. Inside the office, there were few metal objects nearby, which are known to affect capacitive sensing (Osoinach, 2007). All participants signed an agreement following the policies of the university's committee for the protection of human subjects and in accordance with the Declaration of Helsinki. The experiment was video recorded for further confidential analysis. The observer and participant followed an ethical/hygienic protocol following the mandatory public health guidelines at the date of the experiment.

¹The OpenTheremin V3 has been updated to OpenTheremin V4: <https://www.gaudi.ch/OpenTheremin/>.

²TWC24004B textile cables are deprecated, for an alternative option: Interactive Wear <http://www.interactive-wear.com/>.

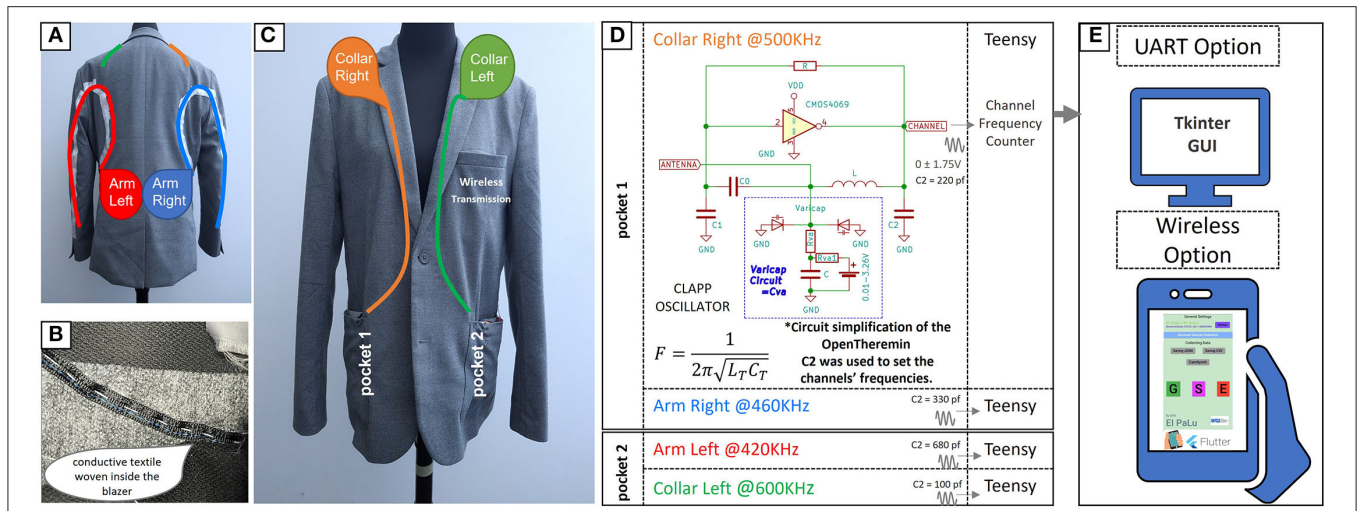


FIGURE 1 | Electronic garment design “the MoCaBlazer,” (A) is the back part of the blazer, (B) is the textile cables sewn inside the garment, (C) is the front part of the blazer, (D) is the circuit simplification design of the Clapp oscillator with the antennas, (E) are the two options for collecting data coming from the blazer; with and UART-wired option, and a Bluetooth based android application (Flutter framework) as the wireless option.

The first experiment scenario was based on a general dictionary of posture and gestures in **Figure 2**. The second one was inspired by dance movements from the Rayman Raving Rabbids: TV Party-Nintendo Wii® as depicted in **Figure 3**.

4.1. General Dictionary Experiment

To study the flexibility of our system to adapt to an abroad type of gestures, a general dictionary of 20 upper-body postures/gestures was defined (see **Figure 2**). Fourteen participants mimicked the postures defined in the dictionary in a random sequence per session while wearing the unbuttoned “MoCaBlazer.” The “MoCaBlazer” is based on a size L/52 blazer (Tom Tailor®), a recommended size for 184 cm tall persons. All participants performed five sessions. One session consisted of four random appearances of each gesture inside the dictionary, giving 400 instances per volunteer. The starting and ending point of a gesture was marked by the null position (standing position). On average, the volunteer’s resting period was at least 20 min (without wearing the blazer) in between sessions. For some volunteers, the experiment was completed in 2 days. The volunteers were seven women, 24–64 years old and 157–183 cm in height; seven men, 25–34 years old, and 178–183 cm in height.

4.2. Dance Movements Experiment

As an application-specific experiment, a dance movements dictionary containing the eight postures depicted in **Figure 3** was defined. It is essential to highlight that the data transmission from the “MoCaBlazer” for this experiment was wireless. Therefore, the capacitive channels were floating (not connected to the ground). The dance movements were selected from the game Rayman Raving Rabbids: TV Party-Nintendo Wii® in order to test the feasibility of using the system as a sophisticated game controller. Three volunteers were asked to imitate the eight movements using the buttoned “MoCaBlazer.” Three sessions

were recorded per volunteer; each session contained five random appearances per gesture inside the dictionary for a total of 120 instances per participant. The volunteers were asked to rest (without wearing the blazer) for at least 10 min in between sessions. The participants were two men and one woman, 26–30 years old and 160–183 cm in height.

5. EVALUATION

As shown in **Figure 1**, the Clapp oscillators generated four data channels. The wearer’s movements alter the channels’ fundamental frequency. The channels’ data is processed as a time sequence. The granularity of the evaluation was a complete gesture/instance. An instance was completed when it included a change from the standing position (starting point) and a return to the standing position (ending point). Furthermore, the impact of common and subtle disturbances on the four capacitive channels was reduced by normalizing the gesture/posture. The digital signal processing was slightly different for the two types of experiments. The videos of both experiments were used as ground truth in a manual labeling procedure.

5.1. General Dictionary Experiment Evaluation

The fundamental frequencies of the channels could be seen as a bias difference between the four channels. A normalization procedure was performed to remove these biases and reduce the capacitive sensing modality reliance on the ground. The normalization consisted of subtracting the average of the gesture’s first (starting point) and last values (ending point). Then, the normalized four channels’ time sequences of each posture/gesture were fed to a fourth-order Butterworth band-pass filter with pass frequencies between 1 and 10 Hz. The duration of gestures performed was not constant, which led to variations in the

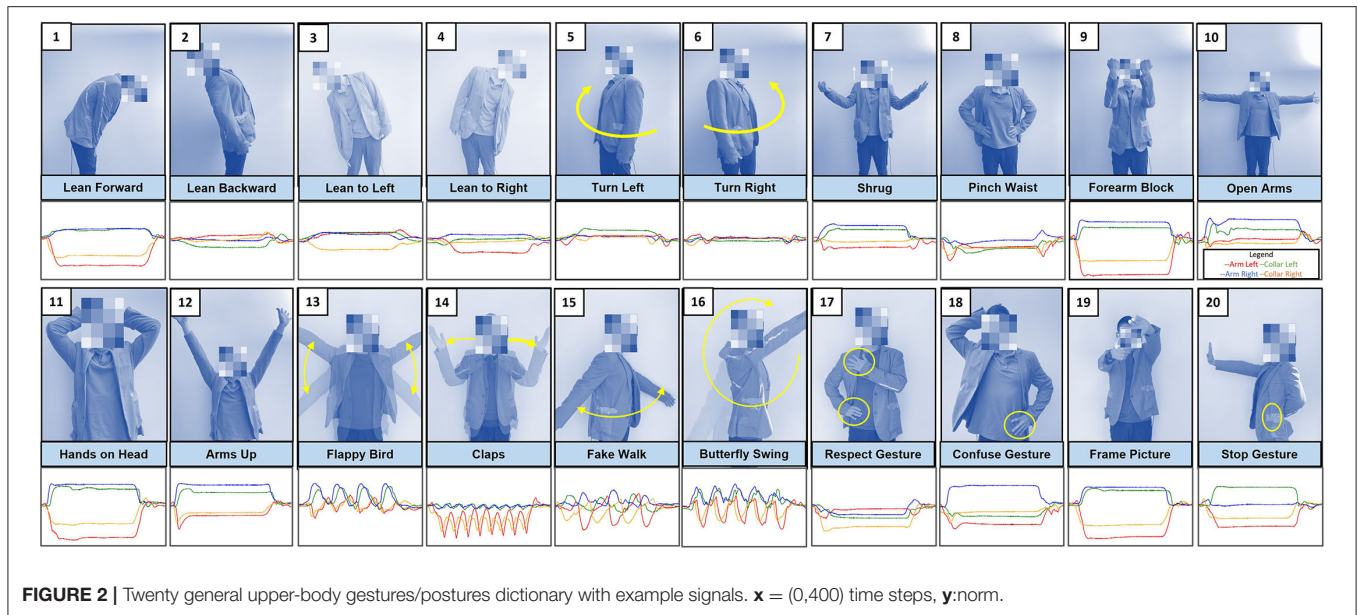


FIGURE 2 | Twenty general upper-body gestures/postures dictionary with example signals. $x = (0,400)$ time steps, y :norm.

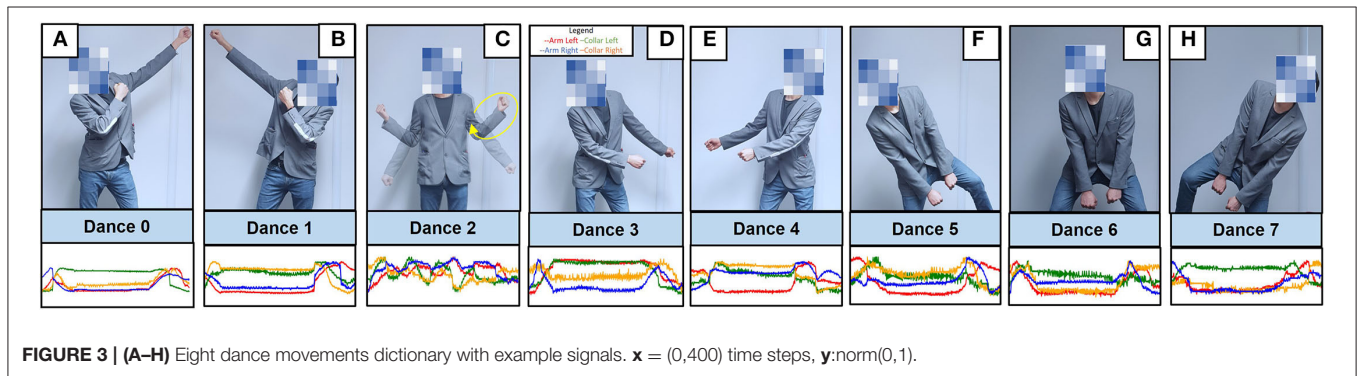


FIGURE 3 | (A–H) Eight dance movements dictionary with example signals. $x = (0,400)$ time steps, y :norm(0,1).

number of samples per instance. The average duration of a gesture was around 2 s (200 samples at 100 Hz). A window of 4 s (400 samples at 100 Hz) was selected to guarantee the activity’s capture. The signals were dilated or contracted depending on whether the gesture contained less or more than 400 samples. Due to the dynamic nature of the applied resampling procedure (dilation or contraction), this is called time-warping (Goldenstein and Gomes, 1999). The signals dynamically resampled (upsampled or downsampled) to 400-time steps provided a fixed size input for the neural network.

The time-warping process was based on the Fourier method (Laird et al., 2004) implemented in the SciPy library (Virtanen et al., 2020). The normalization procedure forced the gesture to start and end circa the same value. Hence, the Fourier method was employed without a window function, which is a method customarily used to avoid ringing artifacts.

A total data of 5,600 gestures/instances of the dictionary in **Figure 2** (14 participants) were processed.

5.1.1. Deep Learning Model

Deep learning models such as 1D-LeNet5 (LeCun et al., 1998; Sornam et al., 2017), DeepConvLSTM (Ordóñez and Roggen,

TABLE 2 | Comparison results for the general 20 body postures and gestures dictionary (in %) with various models.

Method	Accuracy (LRO)	Accuracy (LPO)	Parameters	Training time
1D-LeNet5	96.86 ± 0.46	85.34 ± 7.83	152,880	1.00x
DeepConvLSTM	94.11 ± 0.82	85.42 ± 5.84	440,852	2.32x
Conv2D	97.18 ± 0.70	86.25 ± 8.09	584,800	0.86x

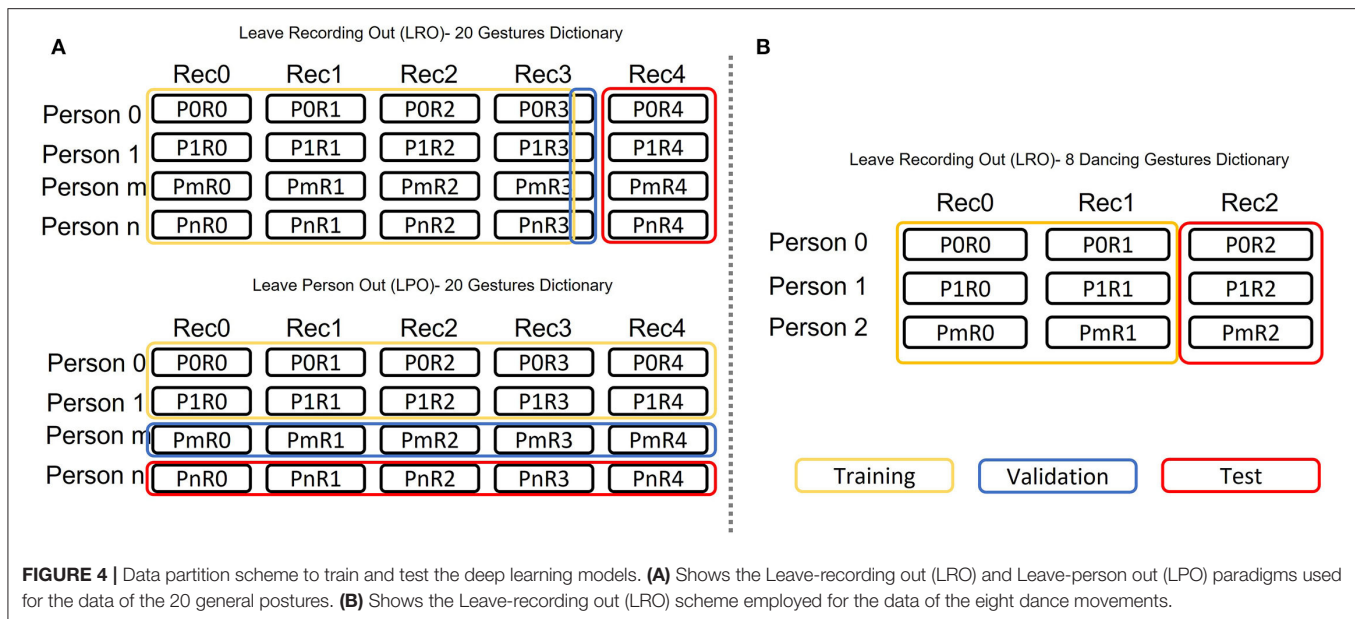
^aLRO, leave recording out; LPO, leave person out.

^bThe accuracy numbers are represented as mean ± std, the standard deviation is from within each complete cross-validation.

^c1.00x Training time of 50 minutes as baseline of complete LRO on NVidia RTX A6000 with the Tensorflow framework.

2016), and Conv2D (Khan et al., 2018; Shiranthika et al., 2020) were evaluated. The best trade-off between performance, parameters, and training time was obtained from a modified 1D-LeNet5 model (refer to **Table 2**). The modified 1D-LeNet5 was defined as a convolution (conv)—max pooling (maxpool)-conv-maxpool-conv—fully connected (fc)-fc-softmax layers with batch normalization (Ioffe and Szegedy, 2015) and dropout (Srivastava et al., 2014) on the convolution layers.

Leave-recording out (LRO) and Leave-person out (LPO) schemes were used as depicted in **Figure 4A**. The LRO paradigm



studies the method's performance for a known group of people, while LPO evaluates the model's performance in the case of unknown persons. We ran all the person's permutations or recordings combinations within each run and summarized the confusion matrix together. That means a complete run of LRO has 5 and LPO has 14×13 train-valid-test cycles. The number of epochs used was 500, stopping when there were signs of overfitting. The three convolution layers are used with a kernel size of 41 and the activation function of ReLU. For max pooling, the pool size was (40, 40) for the first convolution (400, 40) and (4, 40) for the second convolution (40, 40). The third convolution was of size (4, 40) without pooling. A flattening layer of 160 was followed by a fully connected layer of 100. The twenty outputs for the different activities in **Figure 2** are then converted into probabilities by a fully connected layer and softmax function. The categorical cross-entropy loss function and Adam optimizer (Kingma and Ba, 2017) were used in the optimization of the neural network.

5.2. Dance Movements Experiment Evaluation

In this experiment, the time sequences of the four channels were resampled/time-warped to 400-time steps using the same methodology described above in Section 5.1. The signals were normalized between 0 and 1, as $x_{norm} = \frac{x - \min(X)}{\max(X) - \min(X)}$. Where x is a one-time step, X is a sequence of 400-time steps, and x_{norm} is the normalized time step.

In total four deep learning models were generated³. Three individual models per volunteer were trained; two sessions from the same person were used as training, and the third session was

for testing. Moreover, a fourth model was developed using two sessions from each participant (three in total) as training and the third session for testing as shown in **Figure 4B**. A total data of 360 gestures of the dictionary in **Figure 3** (three participants) were fed into a one-dimension convolutional neural network as shown in **Figure 5**. The neural network's input layer was a time series of 400 samples per four channels/antennas (400,4,1). Two convolutional layers followed this with a max-pooling of 10, batch normalization, and dropout of 20%. A third convolutional layer was added but without max pooling. Next, a flattening layer of 160 was followed by a fully connected layer of 100. The eight outputs from the different activities in **Figure 3** were converted to probabilities using a fully connected layer and a softmax function. The training consisted of 500 epochs for all the models. The optimization of the neural network used the categorical cross-entropy loss function and stochastic gradient descent (SGD) (Ruder, 2016) optimizer with learning-rate = 0.005 and momentum = 0.001.

5.2.1. Real-Time Recognition With RFID Synchronization

Following the training and testing paradigms in **Figure 4B** a group model was built for the three participants in the dance experiment. The resulted model considered an entire gesture when the person follows the sequence; standing-gesture-standing. Thus, this sequence needs to be matched to do a real-time evaluation. We proposed to use Radio Frequency Identification (RFID) as a synchronization technique to signal the starting and ending point of the gesture. The RFID synchronization was employed in the calibration of atmospheric pressure sensors to estimate the vertical position of the hand in Bello et al. (2019). The RFID system comprehends two parts; the reader and the tag. The most commonly used extension of RFID is the near field communication (NFC), which is available

³The deep learning framework was TensorFlow version 2.8.0 (Abadi et al., 2015, 2016) and Keras version 2.8.0 (Chollet and Others, 2015) in Google Colab environment (Bisong, 2019).

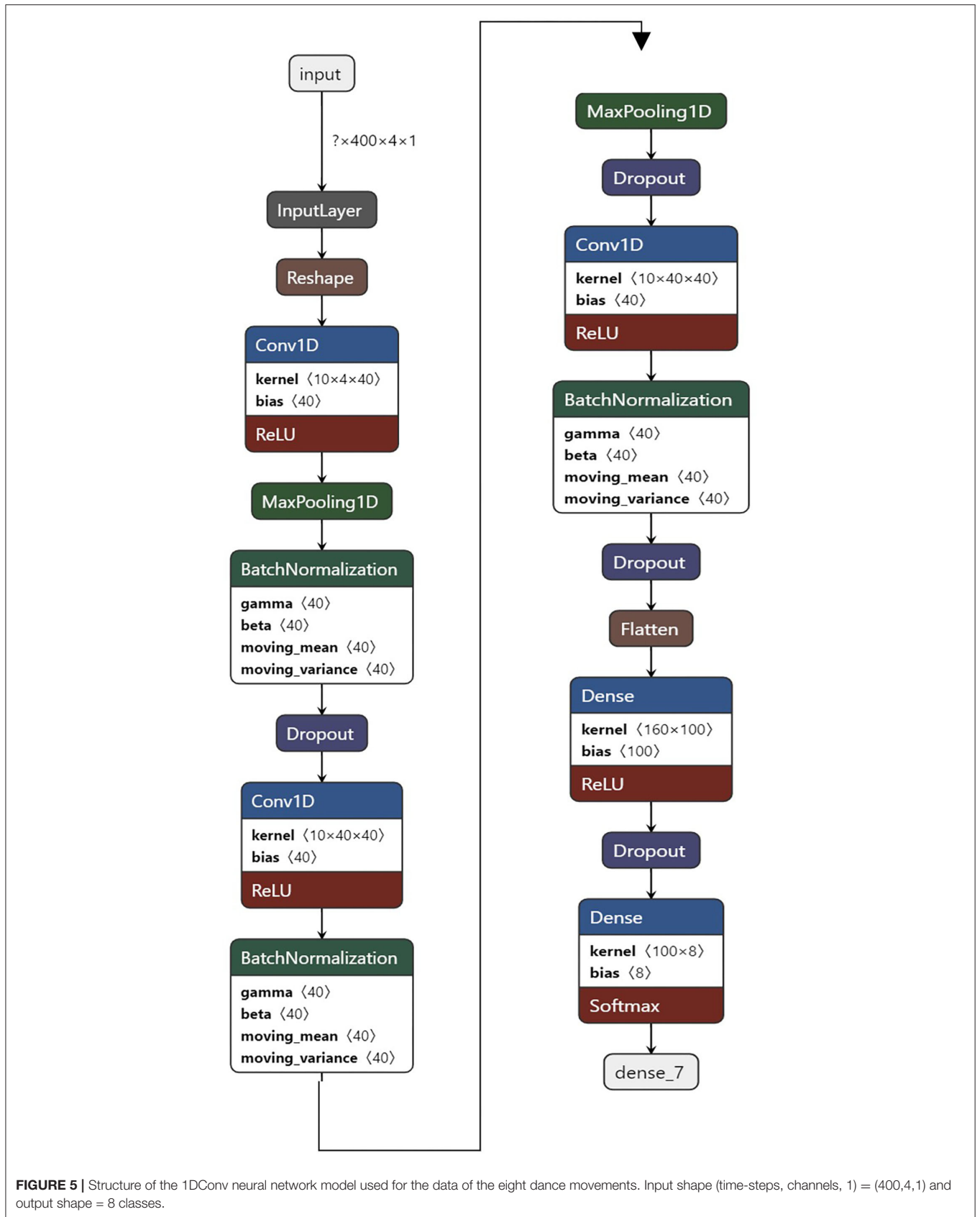


FIGURE 5 | Structure of the 1DConv neural network model used for the data of the eight dance movements. Input shape (time-steps, channels, 1) = (400,4,1) and output shape = 8 classes.

in most smartphones to make over-the-air payments. In Bello et al. (2019), the reader was on the wrist, and the tag was around the pocket to simulate the NFC systems.

It should be noted that there is already an NFC system in our smartphones and that the pocket is a common position to carry our phone. In addition, RFID stickers are nowadays a commonly used solution for tracking merchandise in stores in a ubiquitous and unobstructed manner. Hence, we propose a setting for the real-time evaluation as the one shown in **Figure 6A**. The wrist was the selected position for the reader, and the side pocket of the “MoCaBlazer” was the position for the RFID tag (Mifare Classic 13.56 MHz). **Figure 6B** shows a volunteer wearing the synchronization system. The RFID signal and the “MoCaBlazer” four-channel outputs were sent using Bluetooth serial (wireless) to a python script running the TensorFlow model. The python script follows the flow diagram in **Figure 6C**. The real-time evaluation was performed with participant number two of the three participants pool. The participant was asked to do five repetitions per dance gesture (40 motions).

It is worth mentioning that the real-time recognition with RFID synchronization did not include any pre-training stage with the RFID signal. The model used here was generated from the offline data without RFID. The input data to the offline model was manually labeled with a granularity of 50 fps (recorded video).

6. RESULTS

6.1. General Dictionary Experiment Results

In **Table 2**, the results for the three models; 1D-LeNet5, DeepConvLSTM, and Conv2D are compared. There is not a remarkable variation across the models. The confusion matrices using Conv2D for the Leave-recording out (LRO) and Leave-person out (LPO) are depicted in **Figure 7**. The results confirmed a robust recognition of the 20 postures/gestures dictionary. The LRO or user-dependent case gave an average accuracy of 95%, refer to **Figure 7A**. There was a decrease of around 10% for the LPO or user-independent case, shown in **Figure 7B**. Furthermore, we achieved an average accuracy of 86.25%, with nine classes out of the 20 returning above 95% accuracy. Hence, we could conclude that these results are good enough to consider that our model will perform well for the stranger case; people not included in its training phase.

6.2. Dance Movements Experiment Results

Four models were generated using the neural network structure in **Figure 5**. The results for the three individual models are shown in the confusion matrices in **Figure 8**. **Figure 8A** presents the recognition for the first model trained (2 sessions) and tested (1 session) with the data from volunteer number one. The first participant obtained the lowest performance, f1-score = 93%. **Figure 8B** is the result for Leave-one recording out (LRO) of the second volunteer, showing an f1-score = 100%. For the third participant, the results are only 5% less than the perfect f1-score. With this performance, our design successfully recognized the gesture dictionary in **Figure 8**.

The data partition (train and test) of the fourth model is in **Figure 4B**, and the result is illustrated in **Figure 9A** with

an f1-score = 92%. The fourth model was tested in real-time in conjunction with RFID synchronization and gave an f1-score = 82% as shown in **Figure 9B** for six classes. In the confusion matrix in **Figure 9B**, the classes 4-5 and the classes 6-7 were merged, which gives a total of six classes. In the case of merged classes 4-5, the fourth class was completely confused, with half of its instances being recognized in class number 1 and the other half in class number 5. Moreover, in the case of the merged classes 6-7, the seventh class was recognized consistently as class number 6. The above indicates that the dance movements 4 and 7 in **Figure 3** could not be recognized correctly with the combination of the fourth deep learning model (offline) and the RFID synchronization (online). Despite the negative cases of classes 4 and 7, the real-time recognition with RFID synchronization shows decent performance for the merged classes (six classes in total).

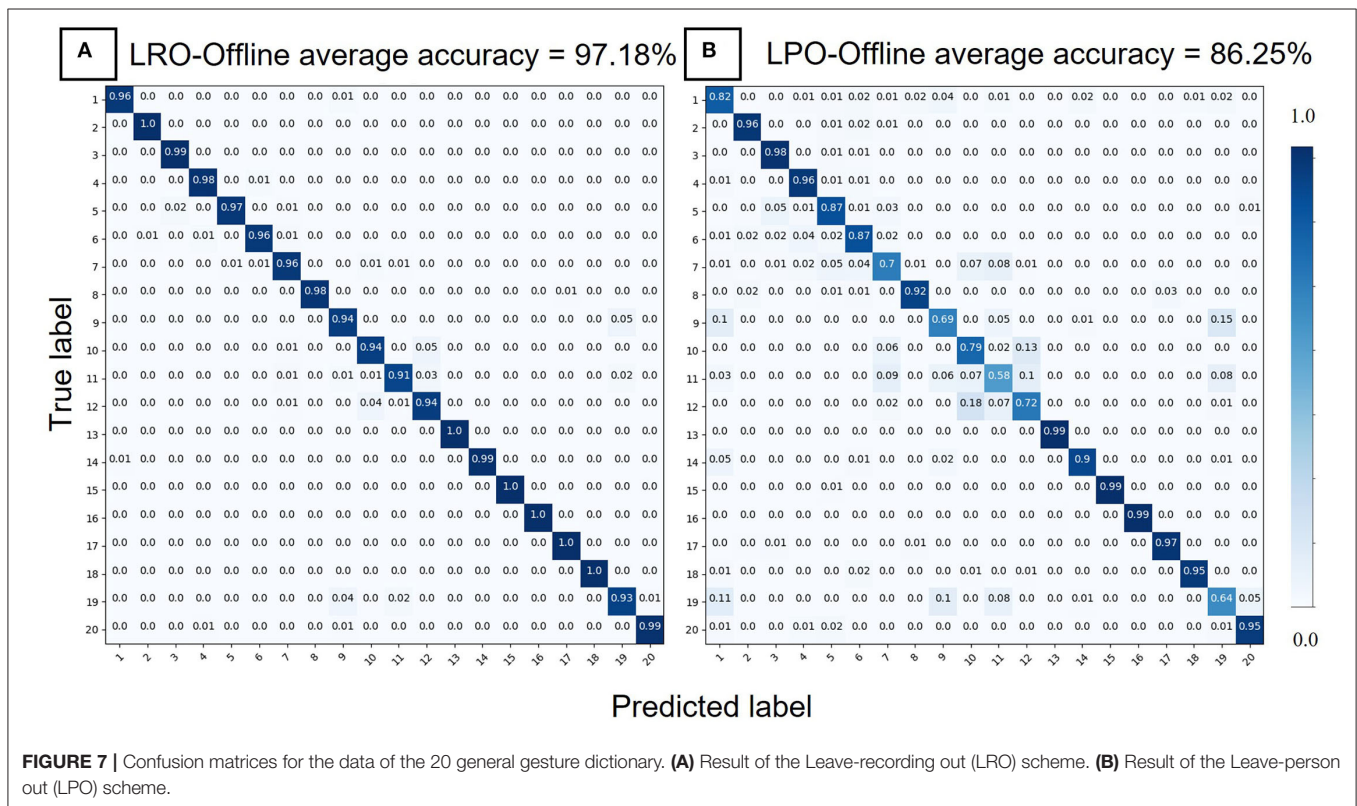
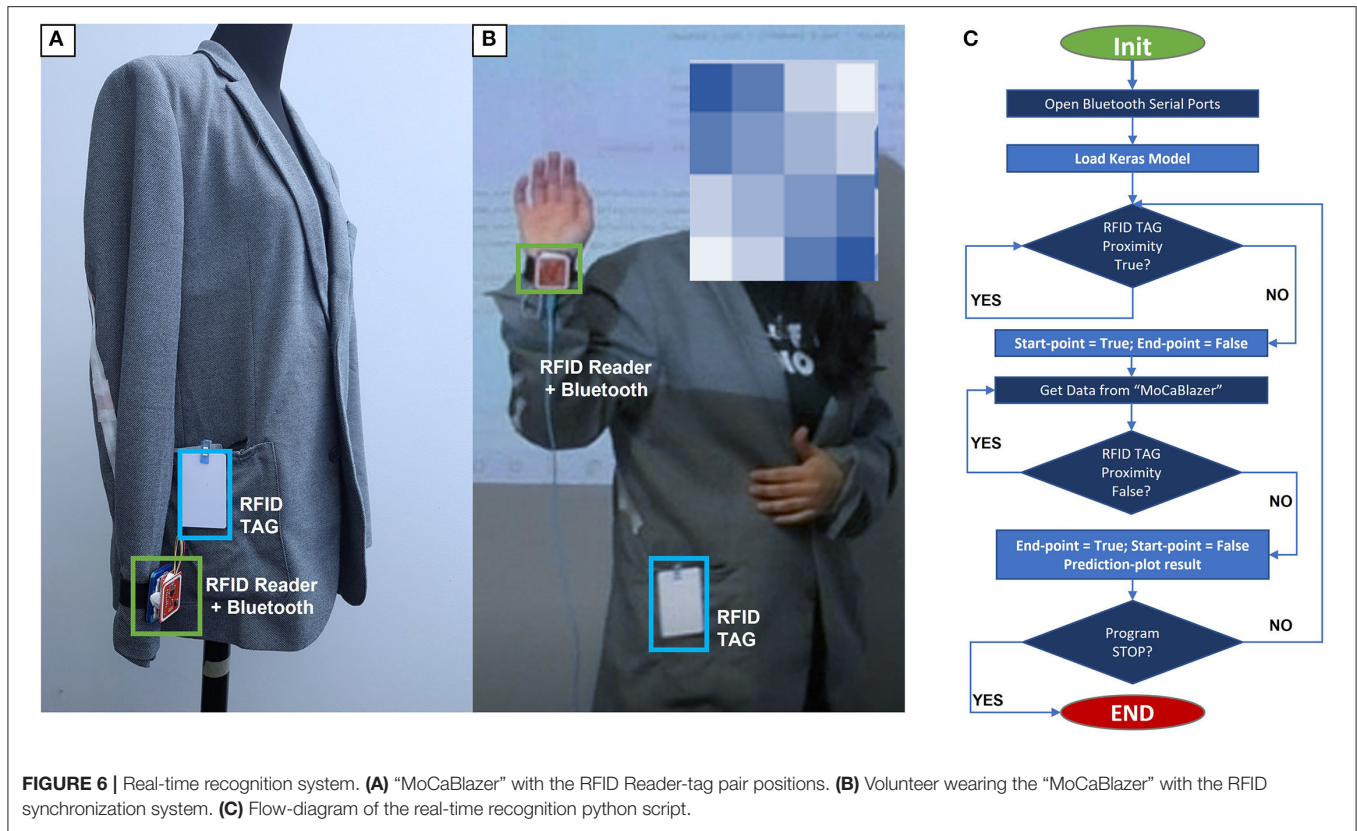
7. DISCUSSION

7.1. General Dictionary Experiment Discussion

To discuss our results, the confusion matrices in **Figure 7** and the 20 gesture/posture dictionary in **Figure 7** will be referenced as a duo. In the case of the Leave-recording out results in **Figure 7A**, the accuracy was above 90% for the 20 classes. On the other hand, in **Figure 7B**, the result for the Leave-person out scheme is depicted, and we could observe several pairs of false recognition. For the pairs of arms-up (Gesture 12)/open-arms (10) and forearms-block (9)/frame-picture (19), the arm motions and directions are physically similar. For the case of lean-forward (1)/frame-picture (19), the similarity is seen in the signals in **Figure 2**; we believe it is a negative effect of participants of different body shapes wearing the same size blazer L/52, which leads to misclassification of 11%. Nonetheless, for forearms-block (9)/hands-on-head(11) pair with similar signal patterns and elbow flexion, the misclassification is only 5%. It is worth noticing that the activities with shoulder motion, such as shrug (7), forearms-block (9), hands-on head (11), arms-up(12), and frame-picture(19), have a reduction in accuracy in the Leave-person out (LPO) result compared to the Leave-recording out (LRO) case. The confusion could be due to the lack of antennas to cover the shoulders of the “MoCaBlazer” and that all fourteen volunteers (of different body shapes) were wearing the same one-size blazer.

7.2. Dance Movements Experiment Discussion

The result of the individual model of participant number one shows some misclassification (refer to **Figure 8A**). For the classes/dance movements 1 and 4, 20% of the gestures are confused; these two gestures have in common that the arms move to the same side of the body trunk but at a different height. The same happens to participant number three as seen in **Figure 8C**. The similarity between these two participants is that they are both men and have a difference in height of 8 cm. In the triplet consisting of dance movements 5, 6, and 7, the seventh and fifth gestures were falsely identified as number six for the case



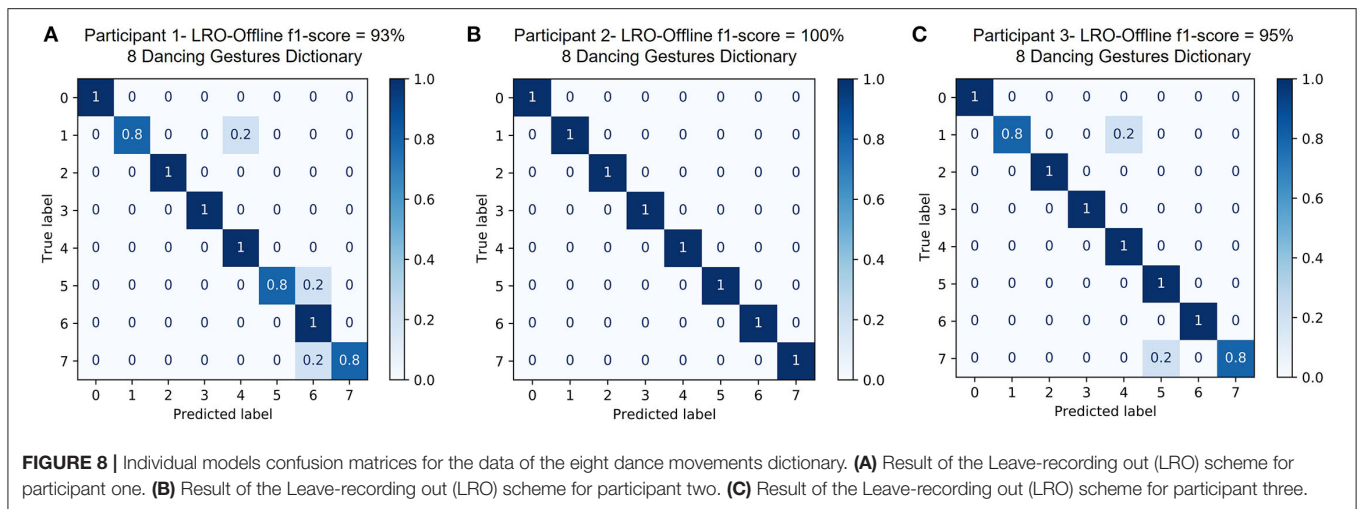


FIGURE 8 | Individual models confusion matrices for the data of the eight dance movements dictionary. **(A)** Result of the Leave-recording out (LRO) scheme for participant one. **(B)** Result of the Leave-recording out (LRO) scheme for participant two. **(C)** Result of the Leave-recording out (LRO) scheme for participant three.

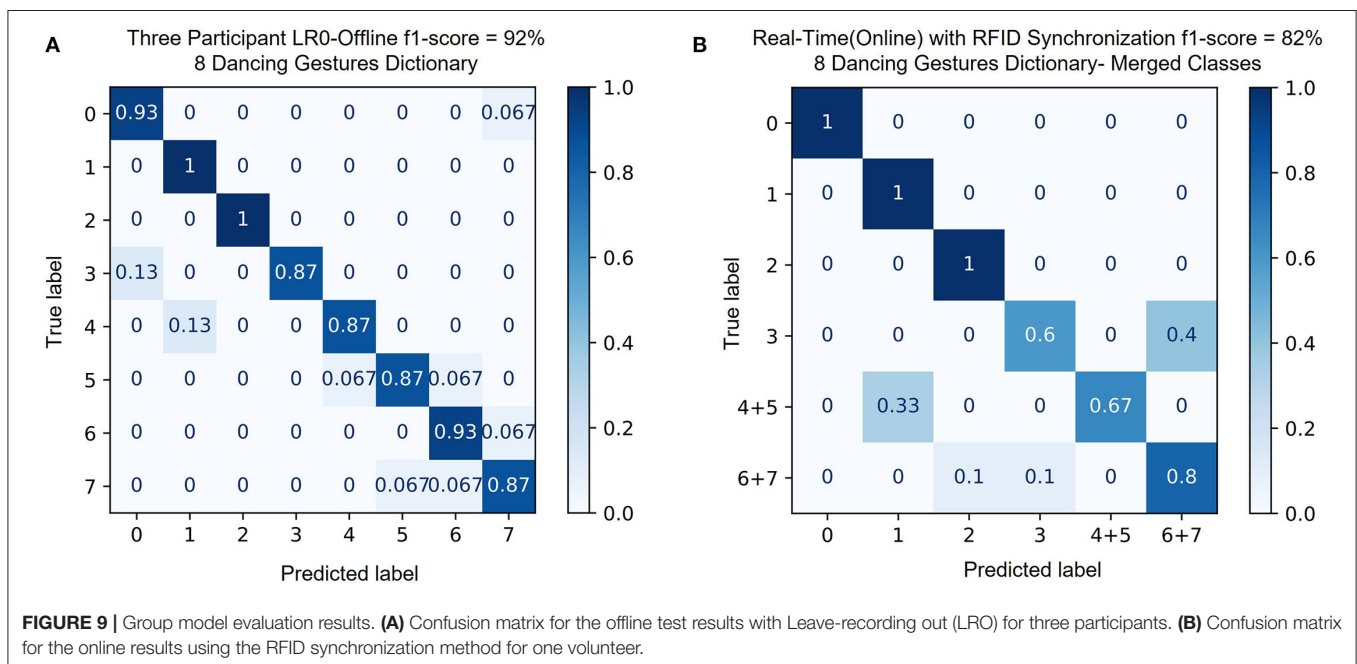


FIGURE 9 | Group model evaluation results. **(A)** Confusion matrix for the offline test results with Leave-recording out (LRO) for three participants. **(B)** Confusion matrix for the online results using the RFID synchronization method for one volunteer.

of participant number one. In the case of the third participant, movement number seven has 20% of its instances confused with the fifth movement. Such gestures include moving both arms in between the legs. A significant difference in the activities is how the legs move; left/right leg in the air or both feet on the ground with the knee bent, and how the shoulders move. The lack of antennas on the shoulder blades and not antennas on the lower part of the body could be the sources of the misclassification. For the second participant, an f1-score = 100% was achieved. This volunteer is a woman with a height of 160 cm. The “MoCaBlazer” was looser for the second participant, which indicates the blazer has more flexibility and could be interpreted as more wrinkles on the garment while doing the movements.

The fourth model was developed using the LRO scheme depicted in **Figure 4B**. With this model, two tests were

performed; LRO-Offline with the three participants and confusion matrix in **Figure 9A**, and the second test was a real-time (online) with RFID synchronization which performance is in **Figure 9B**.

For the first test of the fourth model (offline), the highest recognition error was observed for two pairs of classes, 4/1 and 3/0, with 13% of the instances being wrongly recognized. These two pairs of classes consisted of both arms moving from the standing position (starting point) to the right/left, with the main difference in how much height the arms reach, including a visually distinctive shoulder movement. As seen in the individual models in **Figure 8**, the classes number 5, 6, and 7 are confused between each other, which also occurs in the group-model/fourth-model, so it was a foreseen situation. An f1-score = 92% for the recognition of the gestures in the dance movements

dictionary makes our system a good solution for a sophisticated and elegant dance game controller.

The second test result, the real-time with RFID synchronization in **Figure 9B**, shows perfect recognition for the dance movements 0, 1, and 2. This is not the case for movement number 3, with 40% of its instances being confused with the merged class 6–7. The merged-class 6–7 could be considered as activity number 6 in **Figure 3G**, due to the consistent recognition of dance movement number 7 as dance movement number 6. Therefore, the comparison between dance gesture number 3 in **Figure 3D** and gesture number 6 in **Figure 3G** applies. We suspect two reasons for the 40% wrongly recognized instances; the first could be the slight height difference in the arms' positions and the non-presence of antennas on the shoulders or around the legs. Second, it is essential to remark that this confusion is not present in the offline results, which concludes that our solution depends highly on excellent labeling to mark the gesture's starting point and ending point.

The offline results were obtained using labeling/marking the starting point and ending point with high accuracy in 50 fps/camera. The RFID labeling/marking of the starting point and ending point has an intrinsic error of a slight hand movement (location of RFID reader) to get close enough and detect the RFID tag (on the side pocket). In addition, the RFID solution has a granularity of seconds instead of milliseconds (video based labeling /offline-case).

The merged-class 4–5 has a 33% misclassification with class number 1, and this confusion can also be observed in the offline result of the three volunteers model. Despite the far from perfect RFID synchronization to signal a gesture sequence “standing-gesture-standing” in comparison with the offline version (in the order of milliseconds), we could consider it a promising technique for real-time recognition. A solution to improve the RFID fusion results could be to train the model with data synchronized through the RFID *in-situ* labeling.

8. CONCLUSION AND OUTLOOK

This article has explored a method for posture and body gesture recognition based on a commercially available electronic theremin, the “OpenTheremin,” which, together with conductive textile antennas, was embedded in a loose-fitting garment, the “MoCaBlazer.” Our solution can be deployed and integrated in a fashion and fast manner into loose garments. The “MoCaBlazer” was evaluated with fourteen participants (gender-balanced) mimicking a general dictionary of 20 upper-body movements. Additionally, as an application-specific evaluation, a pool of three volunteers participated in mimicking an eight dance movements dictionary inspired by the Rayman Raving Rabbids: TV Party-Nintendo Wii® game.

For the 20 gestures dictionary, different deep learning models were selected, such as 1D-LeNet5, DeepConvLSTM, and Conv2D. For the case of the eight dance movements dictionary, a one-dimension convolutional neural network was selected. In both evaluations, the system has offered competitive performance compared to state of the art in loose garments for

BPG detection. In the experiment design, a repeated wearing of the “MoCaBlazer” was enforced (per session) to make the results robust against disturbances of re-wearing.

With our chosen sensing modality, the non-contact capacitive method, we use the advantages of being independent of muscular strength/pressure and, therefore, no need for tight or elastic garments. In addition, it is relatively not sensitive to sweat or skin dryness (Zheng and Wang, 2016). A limitation of the capacitive sensing modality is that it is sensitive to conductors, which includes persons/objects in close range with different dielectric properties compared to the antennas (Osoinach, 2007). To avoid the effect of environmental disturbances as much as possible, we normalized our data per gesture window, removing the dependency on absolute values, and built our system upon the relative differences between capacitive channels.

The “MoCaBlazer” data collection for the dance gesture experiment was wirelessly transmitted to an android phone application. With an f1-score = 92% for eight classes with wirelessly collected data, our design demonstrated robustness against capacitive channel drifting values due to floating ground conditions (typical case in wearables). Moreover, a real-time test with RFID synchronization was done (wireless-online) for one volunteer with an f1-score = 82% for six classes.

Our “MoCaBlazer” evaluation has shown promising results in loose garments as a body posture detection method. Hence, we would continue developing elaborated garment integration; with miniaturized sensing modules, more channels, stretchable antennas, and different antenna pattern designs. In the future, the fusion with other sensors such as IMU for continuous posture detection will be an exciting field to explore, in addition to real-time system deployment/evaluation at the edge (embedded devices).

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://github.com/HymalaiDFKI/MoveWithTheTheremin>.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Kaiserslautern and DFKI. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

HB, BZ, and PL: conceptualization. HB and LS: data curation. HB and BZ: formal analysis, methodology, validation, and visualization. PL: funding acquisition and supervision. HB: investigation, data collection software/hardware, and writing original draft. HB, BZ, and SS: data analysis. HB, BZ, SS, and PL: writing review and editing. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work has been partially supported by BMBF (German Federal Ministry of Education and Research) in the project SocialWear.

REFERENCES

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. (2015). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Available online at: <http://download.tensorflow.org/paper/whitepaper2015.pdf>
- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., et al. (2016). "Tensorflow: a system for large-scale machine learning," in *12th \$USENIX\$ Symposium on Operating Systems Design and Implementation (\$OSDI\$ 16)*, 265–283.
- Abro, Z. A., Zhang, Y.-F., Hong, C.-Y., Lakho, R. A., and Chen, N.-L. (2018). Development of a smart garment for monitoring body postures based on FBG and flex sensing technologies. *Sensors Actuat. A Phys.* 272, 153–160. doi: 10.1016/j.sna.2018.01.052
- AMPHENOL (2015). *Flat Ribbon Cable*. Available online at: <https://at.farnell.com/amphenol-spectra-strip/191-2801-150/kabel-flachb-grau-rast1-27mm-50adr/dp/1170229> (accessed April 5, 2022).
- Bello, H., Rodriguez, J., and Lukowicz, P. (2019). "Vertical hand position estimation with wearable differential barometry supported by RFID synchronization," in *EAI International Conference on Body Area Networks* (Florence: Springer), 24–33. doi: 10.1007/978-3-030-34833-5_3
- Bello, H., Zhou, B., Suh, S., and Lukowicz, P. (2021). "Mocapaci: posture and gesture detection in loose garments using textile cables as capacitive antennas," in *2021 International Symposium on Wearable Computers* (Virtual), 78–83. doi: 10.1145/3460421.3480418
- Bian, S., Rey, V. F., Hevesi, P., and Lukowicz, P. (2019a). "Passive capacitive based approach for full body gym workout recognition and counting," in *2019 IEEE International Conference on Pervasive Computing and Communications (PerCom)* (Kyoto), 1–10. doi: 10.1109/PERCOM.2019.8767393
- Bian, S., Rey, V. F., Younas, J., and Lukowicz, P. (2019b). "Wrist-worn capacitive sensor for activity and physical collaboration recognition," in *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)* (Kyoto), 261–266. doi: 10.1109/PERCOMW.2019.8730581
- Bisong, E. (2019). *Google Colaboratory*. Berkeley, CA: Apress, 59–64. doi: 10.1007/978-1-4842-4470-8_7
- Boyalı, A., Kavaklı, M., et al. (2012). "A robust and fast gesture recognition method for wearable sensing garments," in *Proceedings of the International Conference on Advance Multimedia* [Chamonix; Mont Blanc: International Academy, Research, and Industry Association (IARIA)], 142–147.
- Braun, A., Frank, S., Majewski, M., and Wang, X. (2015a). "Capseat: capacitive proximity sensing for automotive activity recognition," in *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Nottingham: Association for Computing Machinery), 225–232. doi: 10.1145/2799250.2799263
- Braun, A., Wichert, R., Kuijper, A., and Fellner, D. W. (2015b). Capacitive proximity sensing in smart environments. *J. Ambient Intell. Smart Environ.* 7, 483–510. doi: 10.3233/AIS-150324
- Butt, H. T., Pancholi, M., Musahl, M., Murthy, P., Sanchez, M. A., and Stricker, D. (2019). "Inertial motion capture using adaptive sensor fusion and joint angle drift correction," in *2019 22th International Conference on Information Fusion (FUSION)* (Ottawa, ON), 1–8.
- Cha, Y., Kim, H., and Kim, D. (2018). Flexible piezoelectric sensor-based gait recognition. *Sensors* 18:468. doi: 10.3390/s18020468
- Cha, Y., Nam, K., and Kim, D. (2017). Patient posture monitoring system based on flexible sensors. *Sensors* 17:584. doi: 10.3390/s17030584
- Chander, H., Burch, R. F., Talegaonkar, P., Saucier, D., Luczak, T., Ball, J. E., et al. (2020). Wearable stretch sensors for human movement monitoring and fall detection in ergonomics. *Int. J. Environ. Res. Publ. Health* 17:3554. doi: 10.3390/ijerph17103554
- Cheng, J., Amft, O., Bahle, G., and Lukowicz, P. (2013). Designing sensitive wearable capacitive sensors for activity recognition. *IEEE Sensors J.* 13, 3935–3947. doi: 10.1109/JSEN.2013.2259693
- Cheng, J., Amft, O., and Lukowicz, P. (2010). "Active capacitive sensing: exploring a new wearable sensing modality for activity recognition," in *International Conference on Pervasive Computing* (Berlin; Heidelberg: Springer), 319–336. doi: 10.1007/978-3-642-12654-3_19
- Chollet, F., and Others. (2015). *Keras*. Available online at: <https://keras.io>
- Clarys, J. P., and Cabri, J. (1993). Electromyography and the study of sports movements: a review. *J. Sports Sci.* 11, 379–448. doi: 10.1080/02640419308730010
- Cohn, G., Gupta, S., Lee, T.-J., Morris, D., Smith, J. R., Reynolds, M. S., et al. (2012a). "An ultra-low-power human body motion sensor using static electric field sensing," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (Pittsburgh, PA: Association for Computing Machinery), 99–102. doi: 10.1145/2370216.2370233
- Cohn, G., Morris, D., Patel, S., and Tan, D. (2012b). "Humantenna: using the body as an antenna for real-time whole-body interaction," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, TX: Association for Computing Machinery), 1901–1910. doi: 10.1145/2207676.2208330
- Ding, H., Guo, L., Zhao, C., Wang, F., Wang, G., Jiang, Z., et al. (2020). RFNET: automatic gesture recognition and human identification using time series RFID signals. *Mobile Netw. Appl.* 25, 2240–2253. doi: 10.1007/s11036-020-01659-4
- Enikeev, D., and Mustafina, S. (2020). "Recognition of sign language using leap motion controller data," in *2020 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA)* (Lipetsk), 393–397. doi: 10.1109/SUMMA50634.2020.9280795
- Fried, L. (2022). *HuzzahESP32*. Available online at: <https://learn.adafruit.com/adafruit-huzzah32-esp32-feather> (accessed April 02, 2022).
- Gaschler, A., Jentzsch, S., Giuliani, M., Huth, K., de Ruiter, J., and Knoll, A. (2012). "Social behavior recognition using body posture and head pose for human-robot interaction," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Vilamoura; Algarve), 2128–2133. doi: 10.1109/IROS.2012.6385460
- Gaudenz, U. (2016). *Opentheremin-Gaudilabs*. Available online at: <https://gaudishop.ch/index.php/product-category/opentheremin> (accessed April 05, 2022).
- Goldenstein, S., and Gomes, J. (1999). "Time warping of audio signals," in *Computer Graphics International Conference* (IEEE Computer Society), 52. doi: 10.1109/CGI.1999.777905
- Grosse-Puppenthal, T., Holz, C., Cohn, G., Wimmer, R., Bechtold, O., Hodges, S., et al. (2017). "Finding common ground: a survey of capacitive sensing in human-computer interaction," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 3293–3315. doi: 10.1145/3025453.3025808
- Guedjou, H., Boucenna, S., and Chetouani, M. (2016). "Posture recognition analysis during human-robot imitation learning," in *2016 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* (Cergy-Pontoise), 193–194. doi: 10.1109/DEVLRN.2016.7846817
- Haescher, M., Matthies, D. J., Bieber, G., and Urban, B. (2015). "Capwalk: a capacitive recognition of walking-based activities as a wearable assistive technology," in *Proceedings of the 8th ACM International Conference on Pervasive Technologies Related to Assistive Environments* (Corfu: Association for Computing Machinery), 1–8. doi: 10.1145/2769493.2769500

ACKNOWLEDGMENTS

We thank Juan Felipe Vargas Colorado, for his corrections and making the workplace a fantastic place to work.

- Harms, H., Amft, O., Tröster, G., and Roggen, D. (2008). "Smash: a distributed sensing and processing garment for the classification of upper body postures," in *Proceedings of the ICST 3rd International Conference on Body Area Networks* (Tempe, AZ: Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 1–8. doi: 10.4108/ICST.BODYNETS2008.2955
- Holleis, P., Schmidt, A., Paasovaara, S., Puikkonen, A., and Häkkinen, J. (2008). "Evaluating capacitive touch input on clothes," in *Proceedings of the 10th International Conference on Human Computer Interaction With Mobile Devices and Services* (Amsterdam: Association for Computing Machinery), 81–90. doi: 10.1145/1409240.1409250
- Ioffe, S., and Szegedy, C. (2015). "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *International Conference on Machine Learning* (Lille: PMLR), 448–456.
- Johnson, E. A. (1965). Touch display—a novel input/output device for computers. *Electron. Lett.* 1:219. doi: 10.1049/el:19650200
- Jung, P.-G., Lim, G., Kim, S., and Kong, K. (2015). A wearable gesture recognition device for detecting muscular activities based on air-pressure sensors. *IEEE Trans. Indus. Inform.* 11, 485–494. doi: 10.1109/TII.2015.2405413
- Junior, J. C. S. J., Güçlütürk, Y., Pérez, M., Güçlü, U., Andujar, C., Baró, X., et al. (2019). First impressions: a survey on vision-based apparent personality trait analysis. *IEEE Trans. Affect. Comput.* 1:1. doi: 10.1109/TAFFC.2019.2930058
- Khan, M. U., Abbas, A., Ali, M., Jawad, M., and Khan, S. U. (2018). "Convolutional neural networks as means to identify apposite sensor combination for human activity recognition," in *2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)* (Washington, DC), 45–50. doi: 10.1145/3278576.3278594
- Kiaghadi, A., Baima, M., Gummesson, J., Andrew, T., and Ganesan, D. (2018). "Fabric as a sensor: towards nonobtrusive sensing of human behavior with triboelectric textiles," in *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems* (Shenzhen: Association for Computing Machinery), 199–210. doi: 10.1145/3274783.3274845
- Kingma, D. P., and Adam, B. J. (2017). A method for stochastic optimization. *arXiv [Preprint]*. arXiv: 1412.6980. Available online at: <https://arxiv.org/pdf/1412.6980.pdf>
- Koyama, Y., Nishiyama, M., and Watanabe, K. (2016). "Gait monitoring for human activity recognition using perceptive shoe based on hetero-core fiber optics," in *2016 IEEE 5th Global Conference on Consumer Electronics (Kyoto)*, 1–2. doi: 10.1109/GCCE.2016.7800431
- Koyama, Y., Nishiyama, M., and Watanabe, K. (2018). "Physical activity recognition using hetero-core optical fiber sensors embedded in a smart clothing," in *2018 IEEE 7th Global Conference on Consumer Electronics (GCCE)* (Nara), 71–72. doi: 10.1109/GCCE.2018.8574738
- Laird, A. R., Rogers, B. P., and Meyerand, M. E. (2004). Comparison of Fourier and wavelet resampling methods. *Magn. Reson. Med.* 51, 418–422. doi: 10.1002/mrm.10671
- Leal-Junior, A. G., Ribeiro, D., Avellar, L. M., Silveira, M., Diaz, C. A. R., Frizzera-Neto, A., et al. (2020). Wearable and fully-portable smart garment for mechanical perturbation detection with nanoparticles optical fibers. *IEEE Sensors J.* 21, 2995–3003. doi: 10.1109/JSEN.2020.3024242
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 2278–2324. doi: 10.1109/5.726791
- Li, J., Wang, Y., Wang, P., Bai, Q., Gao, Y., Zhang, H., and Jin, B. (2021). Pattern recognition for distributed optical fiber vibration sensing: a review. *IEEE Sensors J.* 21, 11983–11998. doi: 10.1109/JSEN.2021.3066037
- Lin, Q., Peng, S., Wu, Y., Liu, J., Hu, W., Hassan, M., et al. (2020). "E-jacket: posture detection with loose-fitting garment using a novel strain sensor," in *2020 19th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)* (Sydney, NSW), 49–60. doi: 10.1109/IPSNS.2020.00-47
- Liu, H., Sanchez, E., Parkerson, J., and Nelson, A. (2019). "Gesture classification with low-cost capacitive sensor array for upper extremity rehabilitation," in *2019 IEEE Sensors* (Montreal, QC), 1–4. doi: 10.1109/SENSOR43011.2019.8956862
- Loke, G., Khudiyev, T., Wang, B., Fu, S., Payra, S., Shaoul, Y., et al. (2021). Digital electronics in fibres enable fabric-based machine-learning inference. *Nat. Commun.* 12, 1–9. doi: 10.1038/s41467-021-23628-5
- Lundh, F. (1999). *An Introduction to tkinter*. Available online at: www.pythonware.com/library/tkinter/introduction/index.htm (accessed June 08, 2022).
- Mokhlespour Esfahani, M. I. and Nussbaum, M. A. (2019). Classifying diverse physical activities using "smart garments". *Sensors* 19:3133. doi: 10.3390/s19143133
- Napoli, M. (2019). *Introducing Flutter and Getting Started*. Indianapolis, IN: John Wiley and Sons. p. 1–23. doi: 10.1002/9781119550860.ch1
- Noroozi, F., Corneanu, C. A., Kamińska, D., Sapiński, T., Escalera, S., and Anbarjafari, G. (2018). Survey on emotional body gesture recognition. *IEEE Trans. Affect. Comput.* 12, 505–523. doi: 10.1109/TAFFC.2018.2874986
- Ordóñez, F. J., and Roggen, D. (2016). Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors* 16:115. doi: 10.3390/s16010115
- Osoinach, B. (2007). Proximity capacitive sensor technology for touch sensing applications. *Freescale White Paper* 12, 1–12. Available online at: <https://www.nxp.com/docs/en/white-paper/PROXIMITYWP.pdf>
- Perme, T. (2007). *Introduction to Capacitive Sensing*. Available online at: <http://www.t-es-t.hu/download/microchip/an1101a.pdf> (accessed April 05, 2022).
- Pouryazdan, A., Prance, R. J., Prance, H., and Roggen, D. (2016). "Wearable electric potential sensing: a new modality sensing hair touch and restless leg movement," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* (Heidelberg: Association for Computing Machinery), 846–850. doi: 10.1145/2968219.2968286
- Ramalingame, R., Bariouli, R., Li, X., Sanseverino, G., Krumm, D., Odenwald, S., and Kanoun, O. (2021). Wearable smart band for American sign language recognition with polymer carbon nanocomposite based pressure sensors. *IEEE Sensors Lett.* 5:6001204. doi: 10.1109/LSENS.2021.3081689
- Ruder, S. (2016). An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*.
- Sanghvi, J., Castellano, G., Leite, I., Pereira, A., McOwan, P. W., and Paiva, A. (2011). "Automatic analysis of affective postures and body motion to detect engagement with a game companion," in *Proceedings of the 6th International Conference on Human-Robot Interaction* (Lausanne), 305–312. doi: 10.1145/1957656.1957781
- Sarangi, S., Sharma, S., and Jagyasi, B. (2015). "Agricultural activity recognition with smart-shirt and crop protocol," in *2015 IEEE Global Humanitarian Technology Conference (GHTC)* (Seattle, WA), 298–305. doi: 10.1109/GHTC.2015.7343988
- Schepers, M., Giuberti, M., Bellusci, G., et al. (2018). Xsens MVN: Consistent tracking of human motion using inertial sensing. *Xsens Technol.* 1. doi: 10.13140/RG.2.2.22099.07205
- Shin, S., Yoon, H. U., and Yoo, B. (2021). Hand gesture recognition using E-gain-silicone soft sensors. *Sensors* 21:3204. doi: 10.3390/s21093204
- Shiranthika, C., Premakumara, N., Chiu, H.-L., Samani, H., Shyalika, C., and Yang, C.-Y. (2020). "Human activity recognition using CNN & LSTM," in *2020 5th International Conference on Information Technology Research (ICITR)*, 1–6. doi: 10.1109/ICITR51448.2020.9310792
- Singh, G., Nelson, A., Robucci, R., Patel, C., and Banerjee, N. (2015). "Inviz: low-power personalized gesture recognition using wearable textile capacitive sensor arrays," in *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)* (St. Louis, MO), 198–206. doi: 10.1109/PERCOM.2015.7146529
- Skach, S., Stewart, R., and Healey, P. G. (2018). "Smart arse: posture classification with textile sensors in trousers," in *Proceedings of the 20th ACM International Conference on Multimodal Interaction* (Boulder, CO: Association for Computing Machinery), 116–124. doi: 10.1145/3242969.3242977
- Skeldon, K. D., Reid, L. M., McNally, V., Dougan, B., and Fulton, C. (1998). Physics of the theremin. *Am. J. Phys.* 66, 945–955. doi: 10.1119/1.19004
- Sornam, M., Muthusubash, K., and Vanitha, V. (2017). "A survey on image classification and activity recognition using deep convolutional neural network architecture," in *2017 Ninth International Conference on Advanced Computing (ICoAC)*, 121–126. doi: 10.1109/ICoAC.2017.8441512
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1929–1958.

- Stoffregen, P. (2014). *Freqcount*. Available online at: https://www.pjrc.com/teensy/td_libs_FreqCount.html (accessed April 05, 2022).
- Stoffregen, P. (2020). *Teensy*. Available online at: <https://www.pjrc.com/store/teensy41.html> (accessed April 05, 2022).
- Swaminathan, S., Fagert, J., Rivera, M., Cao, A., Laput, G., Noh, H. Y., et al. (2020). Optistructures: fabrication of room-scale interactive structures with embedded fiber bragg grating optical sensors and displays. *Proc. ACM Interact. Mobile Wear. Ubiquit. Technol.* 4, 1–21. doi: 10.1145/3397310
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). Scipy 1.0: fundamental algorithms for scientific computing in python. *Nat. Methods* 17, 261–272. doi: 10.1038/s41592-020-0772-5
- Wang, L., Gu, T., Tao, X., and Lu, J. (2016). Toward a wearable RFID system for real-time activity recognition using radio patterns. *IEEE Trans. Mobile Comput.* 16, 228–242. doi: 10.1109/TMC.2016.2538230
- Wear, I. (2021). *Twc24004b*. Available online at: <http://www.interactive-wear.com/solutions/10-textile-wires/45-textile-cables> (accessed April 05, 2022).
- Wii, N. (2009). *Nintendo Wii's Rayman Raving Rabbids: TV Party - ShakeTV*. Available online at: <https://www.youtube.com/watch?v=eoxjA6E1mDs> (accessed August 05, 2011).
- Wilhelm, M., Krakowczyk, D., Trollmann, F., and Albayrak, S. (2015). “ERING: multiple finger gesture recognition with one ring using an electric field,” in *Proceedings of the 2nd international Workshop on Sensor-Based Activity Recognition and Interaction* (Rostock: Association for Computing Machinery), 1–6. doi: 10.1145/2790044.2790047
- Wimmer, R., Kranz, M., Boring, S., and Schmidt, A. (2007). “Captable and capshelf-unobtrusive activity recognition using networked capacitive sensors,” in *2007 Fourth International Conference on Networked Sensing Systems* (Braunschweig), 85–88. doi: 10.1109/INSS.2007.4297395
- Ye, Y., Zhang, C., He, C., Wang, X., Huang, J., and Deng, J. (2020). A review on applications of capacitive displacement sensing for capacitive proximity sensor. *IEEE Access* 8, 45325–45342. doi: 10.1109/ACCESS.2020.2977716
- Zeng, Q., Xu, W., Yu, C., Zhang, N., and Yu, C. (2018). “Fiber-optic activity monitoring with machine learning,” in *Conference on Lasers and Electro-Optics/Pacific Rim* (Hong Kong: Optical Society of America), W4K-5. doi: 10.1364/CLEOPR.2018.W4K.5
- Zhang, Y., Yang, C., Hudson, S. E., Harrison, C., and Sample, A. (2018). “Wall++ room-scale interactive and context-aware sensing,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC: Association for Computing Machinery), 1–15. doi: 10.1145/3173574.3173847
- Zhang, Z., Yang, K., Qian, J., and Zhang, L. (2019). Real-time surface EMG pattern recognition for hand gestures based on an artificial neural network. *Sensors* 19:3170. doi: 10.3390/s19143170
- Zheng, E., and Wang, Q. (2016). Noncontact capacitive sensing-based locomotion transition recognition for amputees with robotic transtibial prostheses. *IEEE Trans. Neural Syst. Rehabil. Eng.* 25, 161–170. doi: 10.1109/TNSRE.2016.2529581
- Zhou, B., Sundholm, M., Cheng, J., Cruz, H., and Lukowicz, P. (2017). Measuring muscle activities during gym exercises with textile pressure mapping sensors. *Pervas. Mobile Comput.* 38, 331–345. doi: 10.1016/j.pmcj.2016.08.015

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Bello, Zhou, Suh, Sanchez Marin and Lukowicz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.