IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 2 March 2024, accepted 24 March 2024, date of publication 29 March 2024, date of current version 5 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3383143

RESEARCH ARTICLE

3D Pain Face Expression Recognition Using a ML-MIMO Radar Profiler

MARÍA-JOSÉ LÓPEZ, (Student Member, IEEE),
CÉSAR PALACIOS-ARIAS, (Student Member, IEEE),
JORDI ROMEU, (Fellow, IEEE), AND
LUIS JOFRE-ROCA, (Life Fellow, IEEE)
Department Signal Theory and Communications, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain

Corresponding author: María-José López (maria.jose.lopez.montero@upc.edu)

This work was supported in part by the Spanish "Comision Interministerial de Ciencia y Tecnologia" (CICYT) under Project PID2019-107885GB-C31, Project PID2022-136869NB-C31, and Project MDM2016-0600; in part by Metropolis under Grant PLEC2021-007609; in part by the Catalan Research Group, Centre Específic de Recerca en Comunicació i Detecció-UPC (CommSensLab-UPC), under Grant 2021 SGR 01415; and in part by the "Secretaría Nacional de Educación Superior, Ciencia, Tecnología e Innovación" (SENESCYT) from the Ecuadorian Government.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the ethics committee of the Universitat Politècnica de Catalunya.

ABSTRACT This study proposes a new method for the detection of facial expressions of pain using a 3D profiler that combines a multiple-input-multiple-output (MIMO) radar system with a machine learning (ML) model (ML-MIMO radar profiler). It offers a solution for pain detection of facial expressions in a non-invasive, non-intrusive, and cost-effective manner. The ML-MIMO radar profiler employs six radars behind a lens to monitor changes in six facial regions and build a 3D facial profile with real-time facial activity information. A dielectric lens was used to ensure an optimal beam size to effectively illuminate each facial region. Signal processing is performed using dynamic time deformation to determine the longitudinal distance and a discrete stationary wavelet transform to filter the signal and improve accuracy. The information from the 3D profiler was compared with the facial action coding system (FACS) to determine actual facial expressions. A machine learning algorithm was trained to learn action units from the FACS and compare them with the information provided by the ML-MIMO radar profiler, thereby performing facial expression classification. In this study, we analyzed four facial expressions: hapiness, sadness, anger, and pain. Identification and classification were performed using a machine-learning model based on multilayer perceptrons. The results revealed 92% accuracy of the system for pain expression, whereas expressions of happiness, sadness, and anger were detected with 88, 86, and 87% accuracy, respectively.

INDEX TERMS 3D radar profiler, contacless, detection, facial expressions, machine learning, MIMO radar, multilayer perceptrons, pain detection, radar, sensing.

I. INTRODUCTION

Facial expressions are a crucial aspect of human communication. They offer a window into a person's emotional state, personality, and intentions [1], [2]. Facial expressions can be particularly useful in cases where communication barriers exist, such as with babies, non-verbal individuals, or those with cognitive disabilities [3], [4]. By analyzing facial expressions, healthcare professionals can gather detailed information about a patient's pain, enhance diagnostic accuracy, and improve treatment outcomes. Furthermore, the study of facial expressions has implications for the

The associate editor coordinating the review of this manuscript and approving it for publication was Cheng Hu.

diagnosis and treatment of psychopathologies such as depression, post-traumatic stress disorder, and autism spectrum disorder [5], [6], [7].

Automatic facial expression recognition techniques can utilize computer vision [8], time-of-flight sensors [9], structured light sensing [10], and infrared [11] for the accurate interpretation and categorization of human facial expressions [12]. Despite their benefits, in this context, the motivation of this study is to introduce a machine-learning multiple-input multiple-output (ML-MIMO) radar profiler that offers high-resolution 3D depth measurements, compactness, light resistance, temperature tolerance and cost-effectiveness.

The proposed ML-MIMO radar profiler aims to capture facial movements from the face profile in a more comfortable and hygienic manner than other methods [13], [14]. As a cost-effective system, the ML-MIMO radar profiler provides 3D spatial high-resolution information and reveals details and movements that may be missed by 2D cameras [15]. ML-MIMO radar profiler performs well under low-light conditions [16] and respects privacy by not capturing facial images [17].

In clinical settings, early and accurate pain assessment is vital for effective management and intervention, enhanced quality of care, and reduced patient distress. Pain assessment typically employs a scale in which patients rate their pain level [18]. Pain detection techniques include self-reporting, observer-based, and physiological measurements [19], [20], [21]. Self-reporting methods, such as visual analog, numerical rating, and verbal rating scales, are common but depend on each patient's subjective pain experience [22], [23], [24]. Observer-based methods involve observing pain-related behaviors, such as facial expressions and body movements, but these depend on inter-rater reliability and observer bias [25]. Physiological measurements, such as EEG, EMG, and MRI, offer objective pain measures but require specialized equipment [26], [27]. Alternatively, the Facial Action Coding System (FACS) provides a potentially more objective, contactless, and non-invasive method based on facial expressions [28], [29], [30]. However, FACS decoding in a clinical environment can be laborious and requires trained personnel, which underscores the need for automated FACS decoding methods [31].

Therefore, this study introduces a compact, cost-effective ML-MIMO radar profiler, specifically engineered for the non-invasive detection of pain-related facial expressions, utilizing FACS. The proposed system strategically places six radars behind a dielectric lens, enabling the tracking of three-dimensional changes across six distinct facial regions. These tracked variations were then used to identify specific action units (AUs) associated with pain. To account for individual differences and improve accuracy over time, a machine learning model was integrated into the system. This model adapts to unique user characteristics, thereby enhancing the precision of the pain detection.

The objectives of this study were as follows: a) to bridge the gap in contactless, compact, and low-cost 3D technology for pain detection from facial expressions. b) Present a MIMO radar system that tracks the longitudinal profile variation of six facial areas linked to specific AUs, aiding in the identification of pain expressions. c) Utilize a machine learning model for the accurate tracking and analysis of AU dynamics, providing insights into changes in facial expressions.

Compared with other state-of-the-art technologies, such as time-of-flight, structured light, computer vision, and infrared systems [32], offers several unique differences.

- Detailed 3D facial profile capturing from face depth measurements.
- Non-contact and Real-time Monitoring make it less intrusive and more comfortable for the patient.
- Incorporation of a machine-learning model to adapt to individual differences enhance its accuracy over time.
- Cost-effectiveness and compactness.

Although some technologies can mitigate privacy concerns, they often require complex setups and are not always cost-effective.

The remainder of this paper is organized as follows. Section II provides a detailed Literature Review related to this study. Section III introduces the proposed materials and methods, including radar-based system architecture, data acquisition and signal processing, and machine learning algorithms for facial expression recognition. Section IV presents the measurement results and performance evaluation metrics of the machine learning model. Finally, Section V concludes the study by summarizing the findings, highlighting their contributions, and outlining potential avenues for future research.

II. THEORETICAL BACKGROUND

The Facial Action Coding System (FACS), developed by Ekman and Friesen in 1978 [33], is a comprehensive, anatomically based system for evaluating human facial expressions. They describe all visually discernible facial movements, termed Action Units (AUs), which correspond to the activation of one or more facial muscles [34]. The system includes 12 AUs for the upper face and 18 AUs for the lower face [35]. FACS is instrumental in coding both basic emotions: anger, disgust, fear, happiness, sadness, and surprise; and complex expressions such as pain [36]. This precise approach allows for detailed study of facial expressions, providing significant insights into human emotions and reactions.

Facial expressions of pain, which have been studied for over three decades, are characterized by specific AUs. LeResche [37] described a pain expression involving a furrowed brow (AU4), tightened skin around the eyes (AU6), horizontally stretched open mouth (AU20), and a deep nasolabial groove (AU9). Craig and Patrick [38] identified six facial actions indicative of pain: cheek lifting (AU7), upper lip raising (AU10), lip corner stretching (AU12), lip separation

(AU20), mouth opening, and eyelid closing (AU6). Further research [39] has associated increased brow-lowering (AU4) with pain intensity. Prkachin et al. studies [40], [41] found significant increases in four facial actions–brow lowering (AU4), orbit tightening (AU6, AU7), upper-lip raising/nose wrinkling (AU9, AU10), and eye closure (AU43)–across four stimuli (cold, pressure, ischemia, and electric shock).

Altough pain can be detected by monitoring four AUs, this preliminary system evaluates at least eight AUs (AU04, AU05, AU06, AU07, AU09, AU10, AU15, and AU43), enabling the identification of other expressions, such as happiness, sadness, and anger, in addition to pain. Table 1 lists the AUs associated with the facial expressions of interest in this study.

TABLE 1. Involved AUs for the studied facial expressions.

Facial Expression	Activated AUs
Anger	4, 7, 17, 24
Happiness	6, 7, 12, 25
Pain	4, (6, 7), (9, 10), 43
Sadnes	1, 4, 15

III. ML-MIMO RADAR SYSTEM METHODOLOGY

This section describes the methodology used in this study to identify pain using the ML-MIMO radar profiler, based on measurements of the longitudinal 3D profile of the face. The involved AUs define six facial regions for individual radar pointing: upper left (UL) and right (UR) encompassing the eyes and eyebrows; middle left (ML) and right (MR) focusing on the cheekbones; and lower left (LL) and right (LR) concentrating on the mouth corners, as depicted in Fig. 1. Each radar measures instantaneous longitudinal variations in a determined region of the face. The six radars collectively determine a facial expression based on a set of activated action units over a single time slot. Signal processing and machine learning configurations are further discussed in this section.

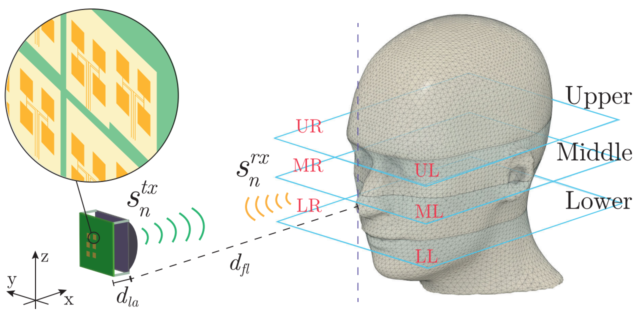


FIGURE 1. ML-MIMO radar profiler configuration. The variable d_{fl} is the face to lens distance and d_{la} represents the lens diameter.

A. RADAR SECTORIAL ILLUMINATION

Six radars, each composed of two antennas (transmitter and receiver), were arranged in a 3×2 matrix positioned behind a focusing lens at a distance $d_{la} = 15$ mm. A dielectric lens was used to define the beam width and the pointing angle to illuminate each of the six facial regions. The pointing angle of each beam was set based on the physical position of the single antenna behind the lens. A displacement in the z-axis points to the upper, middle, and lower regions of the face, whereas the displacement in the y-axis allows the right and left regions of the face to be reached. The activation of each radar is sequential in time-division mode.

The lens provides both the transverse and longitudinal resolutions of the beamwidth. In this regard, each radar illuminates through the lens with a half-power beamwidth (HPBW), whose -3 dB angle is computed as $(\Delta\theta)_{-3dB} = \frac{\lambda_{fc}}{(0.8)d_{ln}}$, where λ_{fc} is the wavelength at the carrier frequency, d_{ln} is the diameter of the lens, and 0.8 is the correction factor owing to the lens geometry. The transversal resolution of the radar is $\delta_{trns} = d_{fl} \cdot (\Delta\theta)_{-3dB}$, where d_{fl} is the observation distance. The ML-MIMO radar profiler was placed at a distance $d_{fl} = 35$ cm from the face, resulting in $\delta_{trns} \approx 20$ mm. The longitudinal resolution, on the other hand, is computed as $\delta_{long} = \frac{c_0}{2BW}$, where c_0 is the speed of light in free space and BW is the bandwidth. In this case, $\delta_{long} = 25$ mm. Under these considerations, the radar-lens configuration maintains the performance of the system in range of 35 cm $\pm$ 10 cm and a variation angle of $\pm 10°$.

A basic numerical study was performed using an iterative spectral algorithm [42] to study the effect of the non-planar shape of the face on incident illumination. Consistent with a previous analytical study, an illuminated region on the order of a circle of 20 mm diameter may be observed, as shown in Fig. 2, resulting in a radar S_{21} parameter propagation radar-face-radar of -50 dB compatible with a radar sensitivity of -90 dBm.

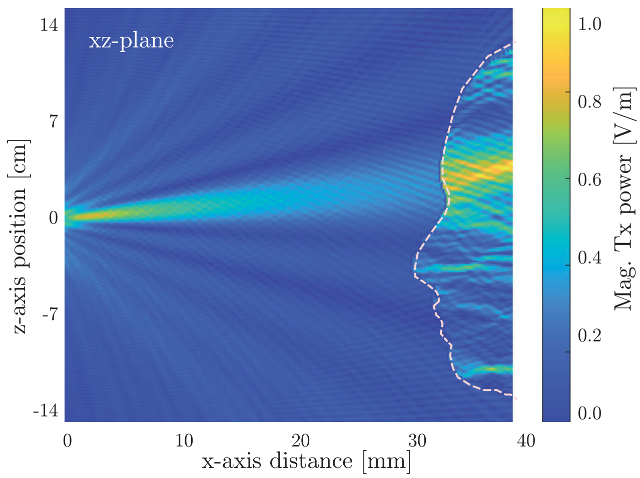


FIGURE 2. Numerical simulation of the transmitted signal in the XZ-plane.

B. MIMO RADAR OPERATION

Each radar operates based on the principle of frequency-modulated continuous wave (FMCW) technology and works within the ISM band at 120 GHz. It exhibits a phase noise of −90 dBc/MHz, enabling minimal frequency fluctuations during high-frequency operations, and consumes a low power of 380 mW. The six radars were integrated into a single chip housing six antennas for transmission and six for reception. A microcontroller interfaces these radars with a computer, facilitating the reception of magnitude and phase.

Each radar sequentially transmits a signal at 120 GHz that sweeps across a 6 GHz bandwidth. The individual trigger signal of each radar controlled the function of the time-division mode. This activates an individual Phase-Locked Loop (PLL), generating a frequency ramp from f_c to f_c + BW. The radar front-end detects the signals across the frequency range of each ramp.

The transmitted S_n^{tx} and received S_n^{rx} signals are then multiplied and passed through a low-pass filter to obtain the intermediate frequency signal:

$$s_n^{IF} = A_{IF} \exp\left[j\left(4\pi \frac{\mathrm{BW}}{t_{rmp}} \frac{(d_0+!d_{fl}(t))}{c} t + 4\pi \frac{(d_0 + d_{fl}(t))}{\lambda_{fc}}\right)\right], \tag{1}$$

where A_{IF} is the amplitude, BW is the bandwidth, d_0 is the initial distance from the radar front end to the face surface, $d_{fl}(t)$ is the face surface displacement due to the activation of the facial muscles during the execution of a facial expression over time, and the subscript n is used to identify each of the six regions.

The longitudinal variation $d_{fl}(t)$ can be calculated from,

$$d_{fl} = \frac{c_0\,|\Delta f\,|}{2(\mathrm{d}f/\mathrm{d}t)}, \tag{2}$$

where Δf is the measured frequency difference and $(\mathrm{d}f/\mathrm{d}t)$ is the frequency shift per unit of time, which is defined for the ramp time as

$$t_{rmp} = t_{smp} * (N_{smp} + 55)/(\mathrm{ADC}_{clk}), \tag{3}$$

where t_{smp} is the sampling time, N_{smp} is the number of samples, 55 samples represent a fixed overhead in the protocol, and ADC_{clk} is the sampling speed of the analog-to-digital converter (ADC), which is equal to 27 MHz.

In this application, longitudinal facial movements are determined from changes in the phase of the received signal using the interferometric distance measurement approach. The resolution of the system was determined by the ADC velocity.

Using Eq. 1, it is possible to calculate the smallest changes in frequency and phase that the ML-MIMO radar profiler can measure. For example, consider a change of 100 µm; the minimum frequency and phase changes that the ML-MIMO radar profiler can detect are 519.9 Hz and 28.8°, respectively. Measurements were performed using standard laboratory equipment. After these measurements were taken, a Fast Fourier Transform (FFT) was applied to each measured instantaneous IF signal. This process creates a *slow-time matrix* that shows the signal spectrum, as illustrated in Fig. 3. This matrix is then processed for feature extraction.

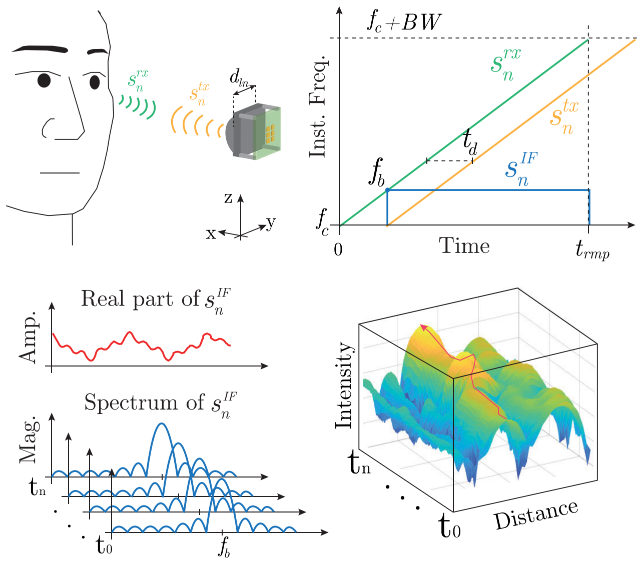


FIGURE 3. **FMCW radar operating principle for longitudinal measurement of facial factions.**

C. SIGNAL PRE-PROCESSING

The radar system continuously measured the instantaneous distances of the six regions of the face and stored them in six individual slow-time matrices. Dynamic Time Warping (DTW) algorithms and wavelet denoising techniques, such as the discrete stationary wavelet transform (DSWT) are used to eliminate noise and enhance the quality and precision of data.

1) FACE MOTION MEASUREMENT BY DTW ADAPTATION

Dynamic Time Warping (DTW) has been employed to assess the temporal variations between multiple instantaneous time-series measurements of facial action units. DTW is a well-established algorithm for time-series analysis, particularly for measuring the similarity between sequences with varying lengths or temporal distortions [43]. By applying DTW to the collected facial action unit data, a quantitative measure of dissimilarity or distance between each pair of time series was obtained. This analysis provides a robust framework for elucidating subtle temporal changes in units, which is essential for a more comprehensive characterization of facial dynamics and expressions. The basic steps involve creating a distance matrix with two consecutive signals, and finding the optimal path through this matrix. Given two consecutive measurements of a single region,

$$s_{n,i}^{IF}(t) = s_{n,1}^{IF}(t), s_{n,2}^{IF}(t), \ldots, s_{n,N_{smp}}^{IF}(t) \tag{4}$$

and

$$s_{n,j}^{IF}(t+\Delta t) = s_{n,1}^{IF}(t+\Delta t), s_{n,2}^{IF}(t+\Delta t), \ldots, s_{n,N_{smp}}^{IF}(t+\Delta t); \tag{5}$$

where n indicates each of the six individual radar signals, N_{smp} is total number of samples; and i and j indicates the actual sample of the respective signal. At this point, the goal is to find the optimal alignment by minimizing the accumulated distance between corresponding elements. The distance matrix $D_{DTW}[i,j]$ represents the local distance or similarity between $s_{n,i}^{IF}(t)$ and $s_{n,j}^{IF}(t+\Delta t)$. D_{DTW} is computed using the sum of squared differences (Euclidean metrics) [44]. Subsequently, an accumulated cost matrix C_{DTW} of size $(N_{smp}+1)\times(N_{smp}+1)$ is initialized and populated using dynamic programming, as shown in Fig. 4. In this study, the possible match number was determined using the Delannoy number, which describes the number of paths from position (1,1) to (N_{smp}, N_{smp}). The computation of each cell involves minimizing the cumulative cost by considering the distances and adjacent cells in C_{DTW} [45], leading to the following expression for C_{DTW}:

$$C_{DTW}[i,j] = D_{DTW}[i,j] + \mathtt{min}(C[i-1,j-1], C[i-1,j], C[i,j-1]) \quad (6)$$

This iterative process fills the matrix C_{DTW} by calculating the cumulative cost for each cell, ensuring optimal alignment. Fig. 4 shows the C_{DTW} generated for two consecutive measurements using a single radar. Finally, the optimal path obtained by backtracking from $C_{DTW}\left[(N_{smp}, N_{smp})\right]$ to $C_{DTW}[0,0]$ reveals the alignment between sequences and determines the DTW distance as the motion variation of two instantaneous distance measurements.

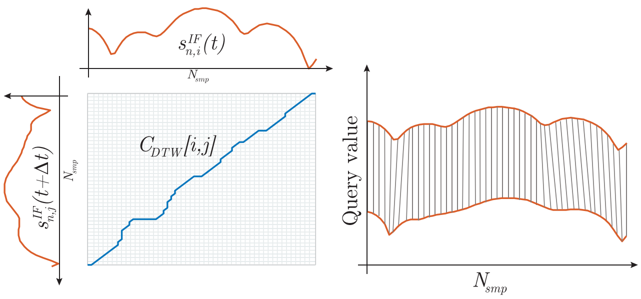


FIGURE 4. a) Dynamic Time Warping (DTW) alignment of two 256-N_{smp} time series of longitudinal distance to the upper right region of the face,the blue line shows the shortest path. b) Distances alignment between two sequential measurements.

Figure 5(a) shows the results of applying the DTW algorithm to radar signals on the upper right and left when a person opens and closes their eyes for 180 s. The peaks in the figure correspond to flashes. The signal was then filtered using a Discrete Stationary Wavelet Transform (DSWT) to isolate the signal corresponding to the facial movements that were being identified.

The Discrete Stationary Wavelet Transform (DSWT) has been used to clear the signal and effectively remove oscillations while preserving the relevant signal information. The DSWT is a robust signal-processing method adapted in this study for feature extraction [46], [47]. The DSWT is a variation of the discrete wave transform that dissects a signal into a collection of wave coefficients at different scales and positions, allowing the analysis of the signal frequency

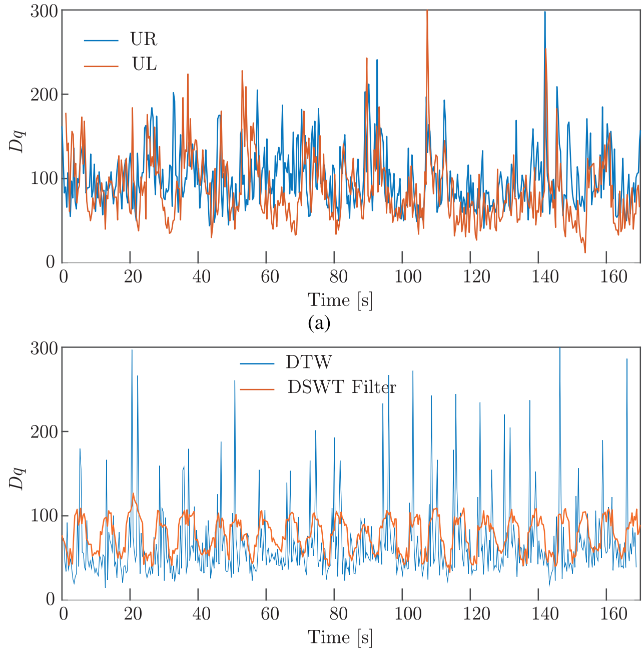


FIGURE 5. (a) Time variation signal of a single region, left side of the upper region (blue) and right side of the upper region (red). (b) DSWT of the original time distance computed using DTW. The blue line correspond to DWT and the red one to the DSWT.

components and extraction of significant characteristics. In this study, the DSWT was implemented using a level 5 stationary wavelet transform with the Haar wavelet [48], [49]. The universal Donoho-Johnstone threshold [50] is then determined based on the first-level detail coefficients, which are used for hard thresholds. Finally, the signal was obtained by inverting the standing wave transform into threshold coefficients. Figure 5(b) shows the DTW and DSWT signals. Each eyeblink is identified in the filtered signal (red line).

Figure 6 shows the flicker measurement results for the six regions of the face. The activity was mainly performed in the upper right and left regions. The middle regions had minimal interaction owing to the possible movements of the cheekbones when closing the eyes. The lower regions remained almost constant because there was almost no motion detection in the mouth.

D. FACIAL EXPRESSION IDENTIFICATION

The pre-processed information from the six radars was analyzed together within the same time slot. Figure 7 illustrates the AU identification process. First, a 1×6 array was formed using the measured facial expression information of the six radars for the same time slot, which is represented by mfe_k, where k is the time slot counter. A binary value was then assigned to identify the activation of a particular action unit. This binary matrix is then correlated with the matrix of Encoded Facial Expressions (EFE) that is created with the data of all the action units of the four facial expressions, which is the aim of this study. The initial

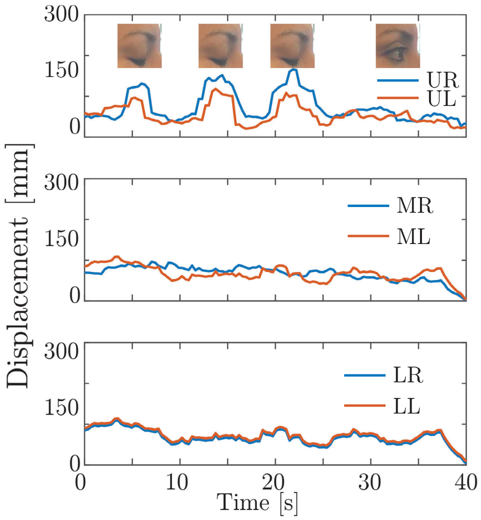


FIGURE 6. **Blinking eye measurement result using ML-MIMO radar profiler.**

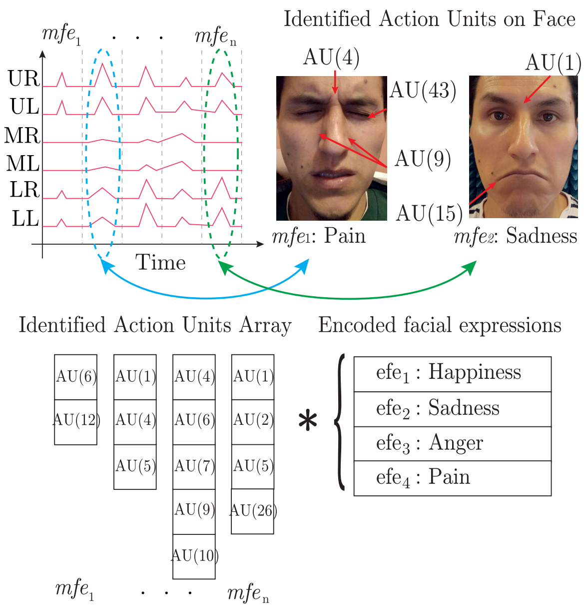


FIGURE 7. **Facial expression identification proccess.**

tests were conducted manually using mathematical software. Subsequently, a machine learning model was trained and tested to identify AUs and classify facial expressions.

E. MACHINE LEARNING MODEL

A machine-learning model was trained to optimize the identification and classification of facial expressions in two stages. Initially, multi-classification identifies the action units, followed by facial expression identification through a joint analysis of these units. The dataset for the multilabel classification stage comprises pre-processed signal values from the ML-MIMO radar profiler, which are used to configure a Multilayer Perceptron (MLP) via the Keras deep learning library.

MLP, a class of artificial neural network, has shown potential for facial expression detection, particularly for smaller datasets because of the lower adjustment requirement [51]. CNNs are commonly used for image processing tasks, including facial expression recognition [52]. SVMs handle high-dimensional data well and avoid overfitting using hyperplanes [53] but may struggle with complex, non-linear patterns common in facial expression detection [54]. Decision trees are interpretable, but can overfit and miss complex patterns. DBNs can model complex, high-dimensional distributions and reduce data redundancy [55]; however, their training can be computationally expensive and slower than that of MLPs [56]. In terms of computational time, CNNs typically consume the most, followed by DBNs, Multilayer Perceptrons MLPs, and finally SVMs consume the least, especially for smaller datasets.

The designed MLP had six inputs for each radar dataset, four outputs for the identified action units, and three densely connected layers. It uses ReLU activation for the hidden layers and Softmax activation for the output layer. The node count of the output layer aligns with the binary value of a particular AU's presence or absence, employing sigmoid activation to predict class probabilities. The model training utilized the binary cross-entropy loss function and Adam optimizer with a batch size of 64 for efficiency and stability.

IV. RESULTS

The testing stages of the ML-MIMO radar profiler included measuring the facial expressions of ten volunteers, five women and five men) who were seated in front of the MIMO radar system. For the preliminary test, they wore an inmovilizer that fixed the face at 35 cm, apart from the lens of the radar system. The volunteers performed facial activities including facial movements and facial expressions. To help each participant feel more comfortable with the requested facial expressions, a screen was placed in front of them to show the facial expressions required in each experiment. Before participating in the experiment, all participants were informed of the study's purpose, procedures, potential risks, and benefits, and signed consent was obtained. This process was followed to ensure that the rights and welfare of participants were protected throughout the study.

Each experiment began by taking the scattering reference level of each subject through face measurements using the ML-MIMO radar profiler, when the subject maintained a neutral expression in a relaxed manner. This approach ensures accuracy and provides a solid foundation for subsequent analysis.

A. BLINKING EYES AND OPEN MOUTH

The first experiment measured the sensitivity of the radar in detecting blinking eyes and mouth opening. Each volunteer was asked to maintain a relaxed position and perform only the blinking action to avoid moving another part of their face. The test consisted of three measurement repetitions in which the person blinks its eyes for one minute with an interval of 3 s between blinks. Before beginning a new repetition, the person takes a 20-second break to avoid data corruption due to muscle exhaustion. Each blink is indicated using a beep. After a 5-minute break, a new test was performed at the same time intervals, but this time the participant was asked to raise his eyebrows. The ML-MIMO radar profiler recorded the activities of the six regions, as shown in Fig. 8, where it is possible to show that the upper right and upper left regions exhibit greater activity. In particular, different patterns of blinking and raising the eyebrows were identified. This allowed us to differentiate between the four action units in the eye region: AU4, AU6, AU7, and AU43.

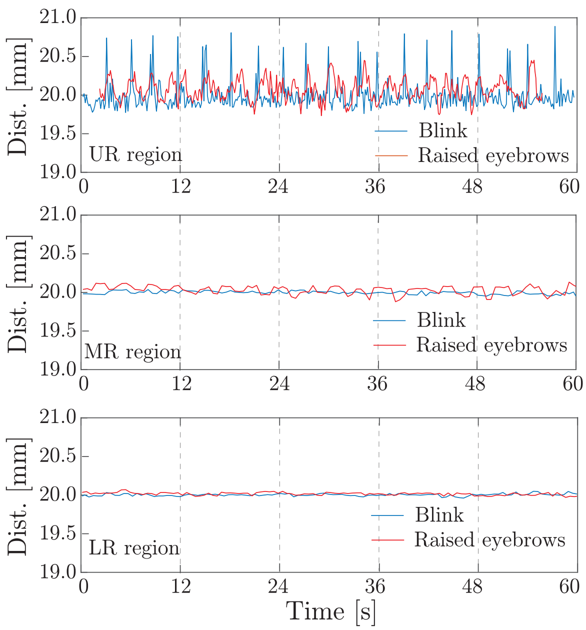


FIGURE 8. Blinking eye and raised eyebrows experiment.

In the second part of the experiment, the focus was on measuring the signals produced by two specific mouth actions: an open mouth and a smile (semi-open mouth and clenched teeth). Fig. 9 shows the results. Once more, the system successfully discerns these two actions, thereby identifying two other units of action: AU9 and AU10.

B. PAIN FACIAL EXPRESSION MEASUREMENTS

This study assessed, four facial expressions: Happiness, Sadness, Anger, and Pain. Participants were asked to emulate the same facial expression for 10 s, with a 10 s break to complete for 3 min. Between each new facial expression, there was a 10 min break. The expressions were evaluated in the following order: Happiness, Anger, Sadness, Pain. Figure 10 shows the measurements of happiness and sadness, with the units of action represented by the participants.

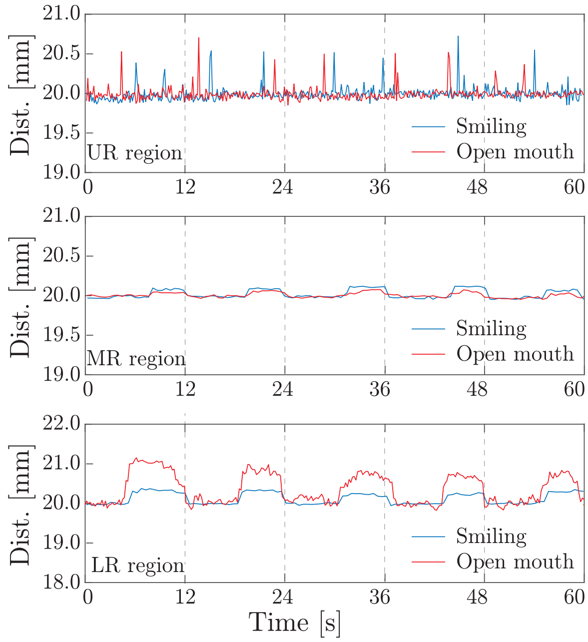


FIGURE 9. Smile and open mouth measurements.

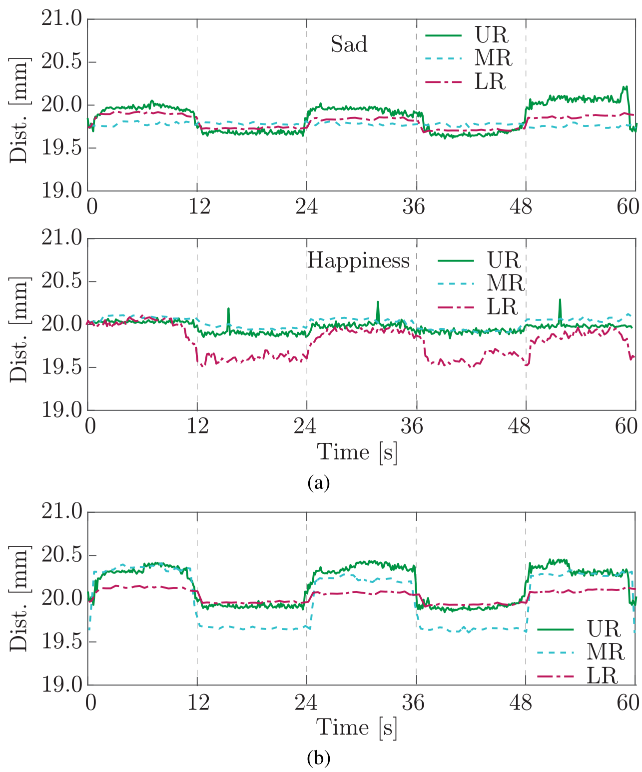


FIGURE 10. (a) Happiness and sadness measurement of the three left regions of the face. (b) Pain facial expression measurement.

A facial expression is formed from the set of signals of the six regions in the same time interval; therefore, if a cut lasting 10 s is made along the x-axis, it is possible to identify the instantaneous facial expressions performed. This set of

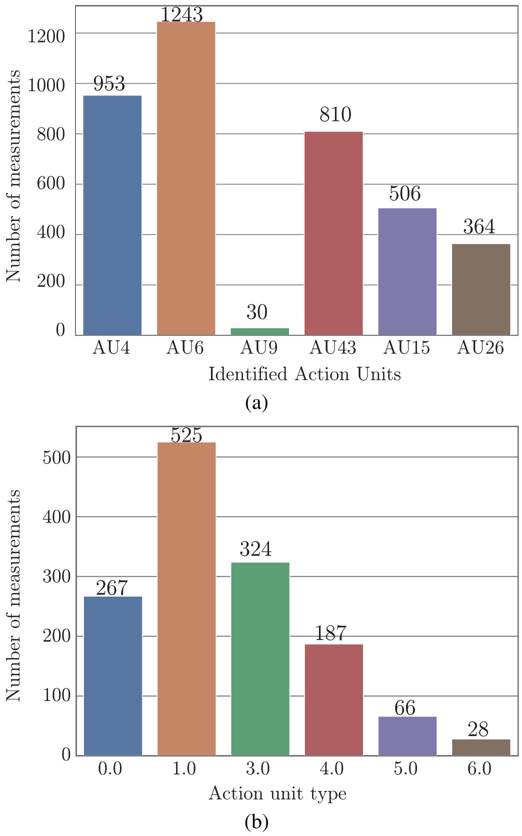


FIGURE 11. **(a) Total number of identified AUs for each category. (b) Total number of measurements including multiple AUs.**

measured AUs was correlated with the database in Table 1 to determine the current facial expressions.

C. MACHINE LEARNING MODEL EVALUATION

The dataset was compiled during two separate data-collection campaigns held on different days. To ensure day-to-day accuracy in the datasets of different campaigns, standardized procedures were used daily for data consistency. The system was calibrated daily for accurate readings and the environmental conditions were controlled for uniformity. Regular quality checks rectify any discrepancies, and statistical methods are used to account for variability, thereby confirming the reliability and validity of the dataset.

The first campaign included measurements of the facial expressions of ten participants in 3 s repetition intervals for 3 min. This resulted in 90 samples for each expression. Four facial expressions, including pain, were considered, resulting in a model training dataset comprising 630 facial expression samples. The second campaign included resampling of the first campaign and a set of random orders of the four expressions at the same time intervals. This resulted in a total of 1890 samples. The model was defined using an MLP network with six input features: measurements of each radar, six output classes that were the AUs identified in each test, and two binary tags that represented the presence or absence

TABLE 2. **Evaluation metrics of the MLP model. Precision, recall, and F-score has been computed for each individual face expressione (happiness, sadness, anger, and pain).**

Classification metrics	Formula	value
Accuracy, *acc* (%)	$(TP + TN)/Total$	88.3%
Precision, *p*	$TP/(TP + FP)$	0.85, 0.90, 0.87, 0.89
Recall, *r*	$(TP/(TP + FN)$	0.87, 0.86, 0.87, 0.92
F-Score	$(2 * p * r)/(p + r)$	0.86, 0.88, 0,87, 0.90

*TP: True Positives, TN: True Negatives, FP: False Positives, FN: False Negatives.

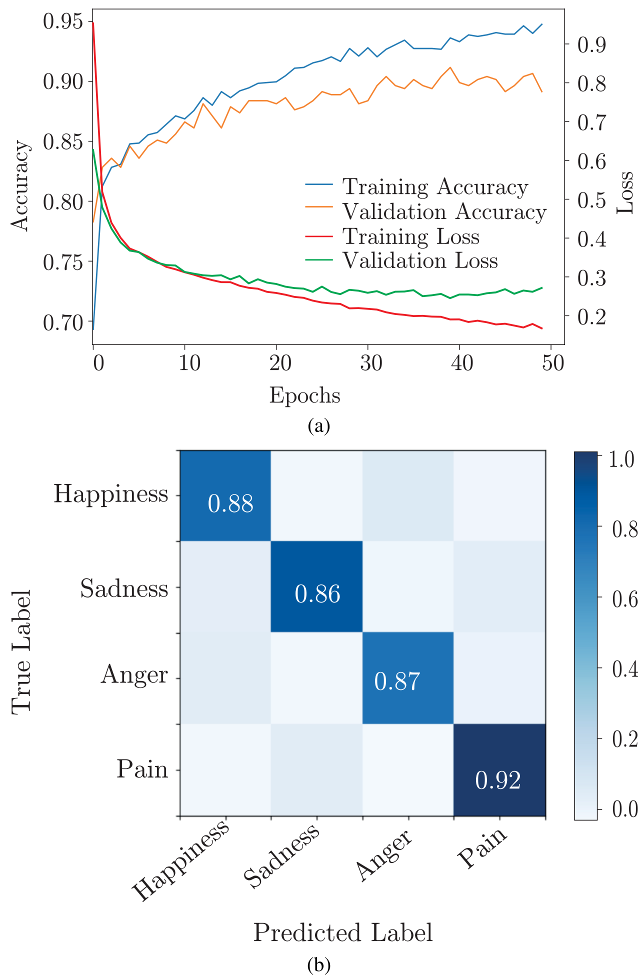


FIGURE 12. **(a) Accuracy and performance of the model. (b) Confusion matrix for the four facial expression tested.**

of the action unit. Figure 11(a) and 11(b) show the number of samples identified for each AU and the number of samples with multiple AUs.

To evaluate the model's performance and robustness comprehensively, a rigorous assessment was conducted using repeated k-fold cross-validation with 10 folds and three repetitions, ensuring reliable estimations of the model's effectiveness. Additionally, a hyperparameter tuning stage was deployed, focusing on optimizing the learning rate and hidden layer sizes to enhance the classification accuracy and generalization capabilities of the model, thereby fine-tuning the model architecture for optimal performance in radar signal classification.

Considering the stochastic nature inherent in the learning algorithm, it is recommended that neural network models be iteratively assessed within the same dataset, particularly with smaller datasets. Reporting the average performance across multiple iterations is advisable in such cases. The evaluation metrics used to assess the performance of the model included a) accuracy, measuring the proportion of correct predictions to total predictions; b) precision, gauging the accuracy of positive predictions; c) recall (sensitivity), indicating the ratio of correctly identified actual positives; and d) F1-score, the harmonic mean of precision and recall. A confusion matrix was also developed to show the true/false positive/negative counts. All computed metrics are reported in Table 2, and the accuracy, loss, and confusion matrices are shown in Fig. 12(a) and 12(b).

V. CONCLUSION

The novelty of the solution presented in this research lies in the combination of a MIMO RADAR system and a machine learning model into a face profiler, enabling high-resolution depth measurement of facial expressions for accurate pain detection in a non-invasive, non-intrusive, and cost-effective manner. The ML-MIMO radar profiler showed promising results in identifying facial expressions of pain with an accuracy of 92%. The system also measured other facial expressions, such as happiness with an accuracy of 88%, sadness with an accuracy of 86%, and anger with an accuracy of 87%. These results demonstrate the potential of MIMO Radar systems and machine learning algorithms for non-invasive pain detection and facial expression recognition. The use of a lens in a MIMO Radar system has improved the focusing and pointing of the system, which has contributed to the high accuracy of the model.

This non-invasive approach offers several advantages over traditional methods that rely on visual imagery or electroencephalography. The radar technology is unaffected by lighting conditions or headwear, making it suitable for use in various settings.

In order to increase the robustness of the system, future studies should focus on improving the accuracy of the MIMO radar system and expanding the dataset to include a wider range of expressions and demographic data. In addition, exploring other machine learning techniques could potentially improve the performance of expression recognition systems.

REFERENCES

[1] A. Awasthi and M. K. Mandal, *Understanding Facial Expressions in Communication: Cross-cultural and Multidisciplinary Perspectives*. Cham, Switzerland: Springer, 2015.

[2] M. Balconi, *The Neuropsychology of Nonverbal Communication: The Facial Expressions of Emotions*. Berlin, Germany: SpringerLink, 2008, pp. 177–202.

[3] E. L. Wanko Keutchafo, J. Kerr, and M. A. Jarvis, "Evidence of nonverbal communication between nurses and older adults: A scoping review," *BMC Nursing*, vol. 19, no. 1, pp. 1–13, Dec. 2020.

[4] M. Bani, S. Russo, S. Ardenghi, G. Rampoldi, V. Wickline, S. Nowicki, and M. G. Strepparava, "Behind the mask: Emotion recognition in healthcare students," *Med. Sci. Educator*, vol. 31, no. 4, pp. 1273–1277, Aug. 2021.

[5] E. L. Rosenberg and P. Ekman, *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. London, U.K.: Oxford Univ. Press, 2020.

[6] B. Martinez and M. F. Valstar, "Advances, challenges, and opportunities in automatic facial expression recognition," in *Advances in Face Detection and Facial Image Analysis*. Cham, Switzerland: Springer, 2016, pp. 63–100.

[7] S. Du, Y. Tao, and A. M. Martinez, "Compound facial expressions of emotion," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 15, pp. E1454–E1462, Apr. 2014.

[8] Z.-Y. Huang, C.-C. Chiang, J.-H. Chen, Y.-C. Chen, H.-L. Chung, Y.-P. Cai, and H.-C. Hsu, "A study on computer vision for facial emotion recognition," *Sci. Rep.*, vol. 13, no. 1, p. 8425, 2023.

[9] S. Bauer, J. Wasza, K. Müller, and J. Hornegger, "4D photogeometric face recognition with time-of-flight sensors," in *Proc. IEEE Workshop Appl. Comput. Vis. (WACV)*, Jan. 2011, pp. 196–203.

[10] Y. Ye, Z. Song, and J. Zhao, "Facial micro-expression analysis via a high speed structured light sensing system," *J. Image Graph.*, vol. 9, no. 1, pp. 15–19, 2021.

[11] A. Bhattacharyya, S. Chatterjee, S. Sen, A. Sinitca, D. Kaplun, and R. Sarkar, "A deep learning model for classifying human facial expressions from infrared thermal images," *Sci. Rep.*, vol. 11, no. 1, Oct. 2021, Art. no. 20696.

[12] R. R. Adyapady and B. Annappa, "A comprehensive review of facial expression recognition techniques," *Multimedia Syst.*, vol. 29, no. 1, pp. 73–103, Feb. 2023.

[13] E. Cardillo and A. Caddemi, "A review on biomedical MIMO radars for vital sign detection and human localization," *Electronics*, vol. 9, no. 9, p. 1497, Sep. 2020.

[14] S. Waqar, M. Muaaz, and M. Pätzold, "Human activity signatures captured under different directions using SISO and MIMO radar systems," *Appl. Sci.*, vol. 12, no. 4, p. 1825, Feb. 2022.

[15] X. Zhang, Y. Zhang, Z. Shi, and T. Gu, "MmFER: Millimetre-wave radar based facial expression recognition for multimedia IoT applications," in *Proc. 29th Annu. Int. Conf. Mobile Comput. Netw.*, Oct. 2023, pp. 1–15.

[16] S. Rao, "Mimo radar," Texas Instrum. Incorporated, Dallas, TX, USA, Tech. Rep. SWRA554A, Jul. 2018. [Online]. Available: https://www.ti.com/lit/an/swra554a/swra554a.pdf

[17] L. Yuan, K. Liang, X. Pu, Y. Zhang, J. Leng, T. Wu, N. Wang, and X. Gao, "PRO-face S: Privacy-preserving reversible obfuscation of face images via secure flow," 2023, *arXiv:2307.09146*.

[18] R. Fink, "Pain assessment: The cornerstone to optimal pain management," *Baylor Univ. Med. Center Proc.*, vol. 13, no. 3, pp. 236–239, Jul. 2000.

[19] R. B. Fillingim, J. D. Loeser, R. Baron, and R. R. Edwards, "Assessment of chronic pain: Domains, methods, and mechanisms," *J. Pain*, vol. 17, no. 9, pp. T10–T20, Sep. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1526590015008652

[20] L. S. Franck, C. S. Greenberg, and B. Stevens, "Pain assessment in infants and children," *Pediatric Clinics North Amer.*, vol. 47, no. 3, pp. 487–512, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031395505702224

[21] P. Werner, D. Lopez-Martinez, S. Walter, A. Al-Hamadi, S. Gruss, and R. W. Picard, "Automatic recognition methods supporting pain assessment: A survey," *IEEE Trans. Affect. Comput.*, vol. 13, no. 1, pp. 530–552, Jan. 2022.

[22] Y. Kang and G. Demiris, "Self-report pain assessment tools for cognitively intact older adults: Integrative review," *Int. J. Older People Nursing*, vol. 13, no. 2, Jun. 2018, Art. no. e12170.

[23] R. H. Bouajram, C. M. Sebat, D. Love, E. L. Louie, M. D. Wilson, and J. J. Duby, "Comparison of self-reported and behavioral pain assessment tools in critically ill patients," *J. Intensive Care Med.*, vol. 35, no. 5, pp. 453–460, May 2020.

[24] D. C. Turk and R. Melzack, *Handbook Pain Assessment*. New York, NY, USA: Guilford Press, 2011.

[25] G. Erogan and S. Celik, "Assessment of postoperative pain by the parent, nurse and an independent observer among 1–7 year old children," *Int. J. Caring Sci.*, vol. 13, no. 2, pp. 1013–1022, 2020.

[26] J. Chen, M. Abbod, and J.-S. Shieh, "Pain and stress detection using wearable sensors and devices—A review," *Sensors*, vol. 21, no. 4, p. 1030, Feb. 2021.

[27] E. Campbell, A. Phinyomark, and E. Scheme, "Feature extraction and selection for pain recognition using peripheral physiological signals," *Frontiers Neurosci.*, vol. 13, p. 437, May 2019.

[28] D. Fabiano, M. Jaishanker, and S. Canavan, "Impact of multiple modalities on emotion recognition: Investigation into 3D facial landmarks, action units, and physiological data," 2020, *arXiv:2005.08341*.

[29] M. Á. Vicente-Querol, A. S. García, P. Fernández-Sotos, R. Rodriguez-Jimenez, and A. Fernández-Caballero, "Development and validation of basic virtual human facial emotion expressions," in *Proc. Int. Work-Conf. Interplay Between Natural Artif. Comput.* Cham, Switzerland: Springer, 2019, pp. 222–231.

[30] J. L. Tracy and R. W. Robins, "The automaticity of emotion recognition," *Emotion*, vol. 8, no. 1, pp. 81–95, 2008.

[31] T. Hadjistavropoulos, "An interdisciplinary expert consensus statement on assessment of pain in older persons," *Clin. J. Pain*, vol. 23, no. 1, pp. S1–S43, 2007.

[32] O. S. Ekundayo and S. Viriri, "Facial expression recognition: A review of trends and techniques," *IEEE Access*, vol. 9, pp. 136944–136973, 2021.

[33] P. Ekman and W. V. Friesen, "Manual for the facial action coding system," in *Consulting*. Psychologists Press, 1978, doi: 10.1037/t27734-000.

[34] M. J. Wieser and T. Brosch, "Faces in context: A review and systematization of contextual influences on affective face processing," *Frontiers Psychol.*, vol. 3, p. 471, Nov. 2012.

[35] Y.-I. Tian, T. Kanade, and J. F. Cohn, "Recognizing action units for facial expression analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 2, pp. 97–115, Feb. 2001.

[36] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," *J. Personality Social Psychol.*, vol. 17, no. 2, pp. 124–129, 1971.

[37] L. LeResche, "Facial expression in pain: A study of candid photographs," *J. Nonverbal Behav.*, vol. 7, no. 1, pp. 46–56, 1982.

[38] K. D. Craig and C. J. Patrick, "Facial expression during induced pain," *J. Personality Social Psychol.*, vol. 48, no. 4, p. 1080, 1985.

[39] C. J. Patrick, K. D. Craig, and K. M. Prkachin, "Observer judgments of acute pain: Facial action determinants," *J. Personality Social Psychol.*, vol. 50, no. 6, pp. 1291–1298, 1986.

[40] K. M. Prkachin, "The consistency of facial expressions of pain: A comparison across modalities," *Pain*, vol. 51, no. 3, pp. 297–306, 1992.

[41] P. Ekman and E. L. Rosenberg, *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. London, U.K.: Oxford Univ. Press, 1997.

[42] J. M. Rius, A. Lozano, L. Jofre, and A. Cardama, "Spectral iterative algorithm for RCS computation in electrically large or intermediate perfectly conducting cavities," *IEEE Trans. Antennas Propag.*, vol. 42, no. 6, pp. 790–797, Jun. 1994.

[43] K. Li, K. Sward, H. Deng, J. Morrison, R. Habre, M. Franklin, Y.-Y. Chiang, J. L. Ambite, J. P. Wilson, and S. P. Eckel, "Using dynamic time warping self-organizing maps to characterize diurnal patterns in environmental exposures," *Sci. Rep.*, vol. 11, no. 1, p. 8674, Dec. 2021.

[44] A. Karlhede, "Classification of Euclidean metrics," *Classical Quantum Gravity*, vol. 3, no. 1, pp. L1–L4, Jan. 1986.

[45] Y. Wang, S.-N. Zheng, and X. Chen, "Analytic aspects of delannoy numbers," *Discrete Math.*, vol. 342, no. 8, pp. 2270–2277, Aug. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0012365X19301141

[46] X. Wang and Y. Dai, "An improved denoising method based on stationary wavelet transform," in *Proc. Int. Symp. Commun. Eng. Comput. Sci. (CECS)*, 2018, pp. 481–485.

[47] E. Stefanutti and F. Bruni, "Signal denoising using the stationary wavelet decomposition," in *Proc. IMEKO TC19 Workshop Metrology Sea*," *MetroSea 2017: Learn. to Measure Sea Health Parameters*, vol. 2017, pp. 104–110, 2017.

[48] R. S. Stanković and B. J. Falkowski, "The Haar wavelet transform: Its status and achievements," *Comput. Electr. Eng.*, vol. 29, no. 1, pp. 25–44, Jan. 2003.

[49] H. Kanagaraj and V. Muneeswaran, "Image compression using Haar discrete wavelet transform," in *Proc. 5th Int. Conf. Devices, Circuits Syst. (ICDCS)*, Mar. 2020, pp. 271–274.

[50] A. Halidou, Y. Mohamadou, A. A. A. Ari, and E. J. G. Zacko, "Review of wavelet denoising algorithms," *Multimedia Tools Appl.*, vol. 82, no. 27, pp. 41539–41569, Nov. 2023.

[51] M.-C. Popescu, V. E. Balas, L. Perescu-Popescu, and N. Mastorakis, "Multilayer perceptron and neural networks," *WSEAS Trans. Circuits Syst.*, vol. 8, no. 7, pp. 579–588, 2009.

[52] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 7553.

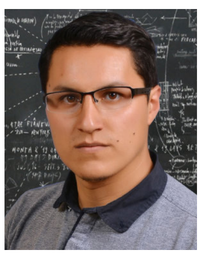
[53] R. Collobert and S. Bengio, "Links between perceptrons, MLPs and SVMs," in *Proc. 21st Int. Conf. Mach. Learn. (ICML)*, 2004, p. 23.

[54] C.-p. Li, X.-y. Zhi, M. Jun, C. Zhuang, Z.-l. Zhu, C. Zhang, and L.-P. Hu, "Performance comparison between logistic regression, decision trees, and multilayer perceptron in predicting peripheral neuropathy in type 2 diabetes mellitus," *Chin. Med. J.*, vol. 125, no. 5, pp. 851–857, 2012.

[55] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.

[56] G. E. Hinton, "Training products of experts by minimizing contrastive divergence," *Neural Comput.*, vol. 14, no. 8, pp. 1771–1800, Aug. 2002.

MARÍA-JOSÉ LÓPEZ (Student Member, IEEE) was born in Ecuador. She received the degree in electronic engineering (telecommunications and networks) from the Higher Polytechnic School of Chimborazo, in 2013, and the master's degree in telecommunications engineering from the University of Calabria, Italy, in 2017. She is currently pursuing the Ph.D. degree with Universitat Politècnica de Catalunya (UPC), Barcelona, Spain. From 2013 to 2019, she taught at the Faculty of Computer Science and Electronics, the Faculty of Sciences, and the Faculty of Livestock Sciences, Polytechnic School of Chimborazo. She is with the AntennaLab Group and the ComSense Lab Group, Department of Signal Theory and Communications (TSC), UPC. She conducted a research stay with the Millimeter-Wave Antennas and Integrated Circuits Laboratory, University of Calabria. In addition, she has completed several complementary training courses in the areas of innovation and vocational training for employment. Since 2021, she has been a part of the M2m Program, UPC, where she is the professional network for women, "Women in Business," and is a volunteer at the Young IT Girl in Spain. Her academic work is focused on the manufacture and use of sensors for physiological measurements in smart cars.



CÉSAR PALACIOS-ARIAS (Student Member, IEEE) received the B.S. degree in telecommunication and electronics engineering from the Private Technical University of Loja (UTPL), Loja, Ecuador, in 2013, and the M.S. degree in electronics engineering from the University of Calabria, Cosenza, Italy, in 2017. He is currently pursuing the Ph.D. degree with the Signal Theory and Communications (TSC) Department, within the research group on microwave interaction with living organisms, CommSensLab, Universitat Politècnica de Catalunya (UPC). He has held positions with ALCATEL-LUCENT, from 2013 to 2015, Corporación Nacional de Telecomunicaciones-CNT, from 2015 to 2015, an External Researcher with the National University of Chimborazo, since 2018, and a Research Support Technician with the Signal Theory and Communications (TSC) Department, UPC, from 2020 to 2021. He is working on micro-systems design and manufacturing for communication with living organisms and sensing at X-wave frequencies.

JORDI ROMEU (Fellow, IEEE) received the Ingeniero de Telecomunicación and the Doctor Ingeniero de Telecomunicación degrees from the Universitat Politècnica de Catalunya (UPC)-BarcelonaTech, Barcelona, Spain, in 1986 and 1991, respectively. In 1985, he joined the AntennaLab, Signal Theory and Communications Department, UPC, where he is currently a Full Professor involved in antennas near-field measurements, electromagnetic scattering and imaging, and system miniaturization for wireless and sensing industrial and bio applications. In 1999, he was a Visiting Scholar with the Antenna Laboratory, University of California at Los Angeles, Los Angeles, CA, USA, on a NATO Scientific Program Scholarship, and the University of California at Irvine, Irvine, CA, USA, in 2004. He holds several patents. He has published 60 refereed articles in international journals and 80 conference proceedings. He was a Grand Winner of European IT Prize, awarded by European Commission for his contributions to the development of fractal antennas, in 1998. He has been involved in the creation of several spin-off companies.

LUIS JOFRE-ROCA (Life Fellow, IEEE) received the M.Sc. (Ing.) and Ph.D. (Doctor Eng.) degrees in electronic engineering (telecommunication engineering) from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 1978 and 1982, respectively. He was a Visiting Professor with École Supérieure d'Electricité Paris, from 1981 to 1982; a Fulbright Scholar with Georgia Institute of Technology, Atlanta, from 1986 to 1987; the Director of the University of California, Irvine, CA, USA, from 2001 to 2002; the Director of the Telecommunication Engineering School, UPC, from 1989 to 1994; the Vice President of the UPC, from 1994 to 2000; and the General Director and a Secretary of Catalan Universities and Research, from 2011 to 2016; the Director of the Catalan Research Foundation, from 2002 to 2004; the Director of the Telefonica Chair on Information Society Future Trends, UPC, since 2003; the Principal Investigator of the Spanish Terahertz Sensing Laboratory Consolider Project, from 2008 to 2013; the Director of the UPC-SEAT Chair on the Future of Automotive; the Research Leader of the CommSensLab Maria de Maeztu Project, from 2017 to 2020; the Academic Director of European Consortium for Future Urban Mobility (Carnet). He has authored more than 200 scientific and technical papers, reports, and chapters in specialized volumes. His research interests include antennas, electromagnetic scattering and imaging, system miniaturization for wireless, and sensing for industrial, scientific, and medical applications. His current work focuses on connected reconfigurable autonomous vehicles for urban mobility, massive MIMO antennas, and microorganism wireless interaction. He is the Chairperson of the EIT-Urban Mobility European Association.