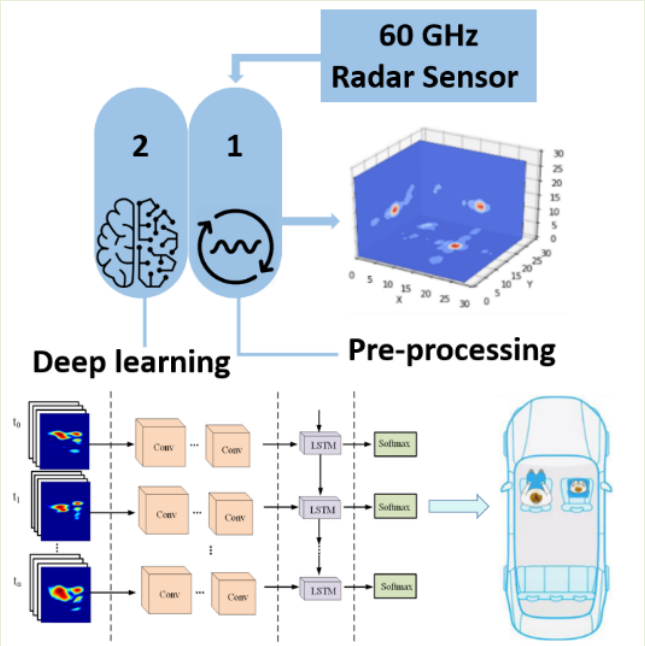
Deep Learning-Based In-Cabin Monitoring and Vehicle Safety System Using a 4-D Imaging Radar Sensor

Hajar Abedi, *Graduate Student Member, IEEE*, Martin Ma, James He, Jennifer Yu, Ahmad Ansariyan, and George Shaker, *Senior Member, IEEE*

***Abstract*— In this article, we use a multiple-input–multiple-output (MIMO) radar for in-vehicle passenger detection. We propose a 2-D convolutional neural network-long short-term memory (CNN-LSTM) to accurately detect, count, and classify passengers inside five-seater vehicles. Our deep learning model first extracts the feature using a CNN model, and then, a series of frames will be delivered to a time series model (LSTM) to predict new scenarios. In addition, we provide the outcomes of various deep learning models and show that temporal deep learning models perform better in our radar datasets. Furthermore, we provide reliable session-dependent datasets collected from different car models with various passengers (including infants/children and adults). The results show that our proposed 2-D CNN-LSTM model can detect unattended infants/children in vehicles with more than 95% accuracy, count passengers and identify their occupied seats with an accuracy of 89%, and classify passengers with more than 74% accuracy. Since our model is evaluated in a new car with new passengers, it ensures the generality of our proposed method to be deployed in any five-seater vehicle.**

***Index Terms*— Contactless passenger monitoring, in-cabin sensing, multiple-input–multiple-output radar (MIMO), passenger classification and counting.**



Manuscript received 15 January 2023; revised 15 March 2023; accepted 28 March 2023. Date of publication 28 April 2023; date of current version 31 May 2023. This work was supported in part by the National Sciences and Engineering Research Council of Canada and in part by the Ontario Center of Innovation. The associate editor coordinating the review of this article and approving it for publication was Dr. Sevgi Gürbüz. *(Corresponding author: Hajar Abedi.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Clinical Research Ethics Committee under Project 31235 (27 August 2018).

Hajar Abedi is with the Department of Systems Design Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: habedifi@uwaterloo.ca).

Martin Ma is with the School of Engineering and Applied Science, Harvard University, Cambridge, MA 02138 USA (e-mail: martinma@g.harvard.edu).

James He is with Qualcomm, Waterloo, ON N2L 3G1, Canada (e-mail: james.he@uwaterloo.ca).

Jennifer Yu is with the Department of Computer Science, University of Toronto, Toronto, ON M5S 2E4, Canada.

Ahmad Ansariyan and George Shaker are with the Electrical Engineering Department, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: ahmad.ansariyan@uwaterloo.ca; gshaker@uwaterloo.ca).

Digital Object Identifier 10.1109/JSEN.2023.3270043

I. Introduction

In THE last decade, there has been significant technological advancement in modern vehicles, such as adaptive cruise control, hands-free highway driving, and other advanced driver assistance systems (ADASs). There is a growing need for in-cabin monitoring systems for infant/pet presence detection to prevent hot car death, enhance seat belt reminders, and alert emergency services in the event of a crash that can report how many passengers are in the vehicle [1].

Accordingly, in the European New Car Assessment Programme (EuroNCAP) 2025 roadmap, child presence detection has been added as a safety requirement [2]. In addition, a rear occupant alert system in cars is required in USA [3]. Therefore, it is clear that passenger occupancy detection in vehicles has become increasingly important.

Various in-vehicle occupancy detection solutions already exist, such as the pressure sensor [4], [5], capacitive sensors [6], [7], and carbon dioxide sensors [8]. However, the high false alarm rate is the major issue these technologies suffer from [9]. While vision-based systems could be used for

in-cabin occupancy monitoring [10], [11], [12], [13], they are sensitive to illumination and invade occupants' privacy. Note that in automotive industries, cameras are mainly being used to monitor the surroundings of the car [14].

Radar-based systems, on the other hand, have recently been introduced and integrated into automotive safety technologies [15], [16], [17], [18], [19], [20], [21], [22], [23]. Typical applications tend to be applied to control passenger-side airbags, safety belts, warning devices, and commuter transport logistics [1]. Unlike vision-based sensors, radar-based sensors offer a solution that is robust to lighting conditions and preserves privacy [24]. Furthermore, millimeter-wave (mm-wave) radars have small form factors, resulting from integrated silicon and antenna-in package, which can be easily mounted [25].

Radar-based human detection facilitates people counting and density estimation [26], [27], [28], [29], [30], [31]. For example, in [31], a people counting method based on bi-motion-model framework was proposed. Features, such as activation index, connected regions, and energy of frames, were extracted and fed to support vector machines (SVM) to classify people's motion states. CNN was then used to count the number of people in an indoor environment when people were standing far away from each other, and radar was installed at a height of 2.2 m. In [28], preprocessing pipelines were proposed to prepare range–amplitude signals for a deep learning model. An unsupervised pretraining process and a defined loss function were proposed to stabilize network convergence for indoor people counting. Most currently available people counting methods using radar sensors have been applied in a large environment where a sensor was installed on a stand with a height of more than 2 m and tested with people relatively far from each other (compared with an in-vehicle environment). In addition, these methods only estimated the number of people, while the locations of people were unknown. In summary, common techniques that we use today for estimating the number of people in a given large area using radars are often inaccurate and require very complex signal processing methods leading to high computational costs.

Currently, the radar systems used inside vehicles are mainly centered on detecting the presence or absence of a living body to save children and pets left in vehicles to prevent death in extremely hot or cold weather [32]. In [33], a sensor system based on the electromagnetic coupling between a transmitter and a receiver patch antenna was developed for in-vehicle occupancy detection. In addition to a high false alarm rate, another limitation of this system is that it just detects one seat, while various sensors are required to monitor all occupied seats and the number of passengers. Most of the available techniques are based on the micro-Doppler effects created due to the breathing cycles of an alive subject inside a car [32] without considering the occupied seats and the number of passengers. For example, in [34], a seat-occupancy detection system based on a pulsed coherent radar was present based on breathing cycles. The breathing rate was also estimated from the amplitude peaks of the signals. In our previous work [32], we proposed a presence or absence detection algorithm to detect an unattended child or pet inside a car. The primary purpose was to propose a fast, easy-to-implement solution to prevent hot car death. However, the number of passengers and their occupied seats was not obtained. This issue refers to multitarget detection and positioning via radar inside vehicles. To detect passengers inside a vehicle, a passenger's vital signs were quantified [19]. In addition, in [23], multipassenger occupancy detection inside a vehicle was performed using a single-channel frequency modulated continuous wave (FMCW) radar. The physiological characteristics acquired from the breathing and heartbeat of the radar signal were analyzed in a time–frequency spectrum to obtain features for multipassenger occupancy detection [23].

However, detection approaches based on the Doppler effects produce false alarms with other moving objects. Not to mention that detection based on the breathing rate would be challenging in the presence of multiple people inside a car, especially when passengers move their bodies, such as hands or torsos. Moreover, these studies were primarily focused on sensing a single passenger or two. In contrast, multiple-people counting, namely, how many people are present as well as their location, is crucial information [24].

Instead of relying on micro-Doppler, range and angle information of the car environment or passenger point cloud were used for in-cabin occupancy detection in [9], [15], [17], [18], [35], [36], [37], and [38]. In [35], a 77-GHz mm-wave radar-based rear occupant detection system was proposed with the main focus on detecting forgotten children in a vehicle's rear seats. Two-dimensional fast Fourier transform (FFT) was applied to build a two-dimensional heat map. Then, a constant false alarm detection (CFAR) algorithm was used for object detection, and thus, a 2-D cloud points image was obtained to identify an occupied versus empty car. Experiments were performed in a room (to mimic the rear-car seat) to detect a maximum of three passengers. This experimental setup was flawed that there was a huge space between three seats, while passengers sat at almost zero distance in the second or third row with three seats. In [38], 2-D imaging was obtained based on digital beamforming using a multiple-input–multiple-output (MIMO) FMCW radar. A logistic regression presence–absence detection method was used to count the number of passengers and identify the occupied seat in a real car. However, there was no reported accuracy in counting and localization, and the number of passengers occupying a car was only two. A 2-D point cloud was generated to count passengers in a four-seater car with a 77-GHz MIMO FMCW radar. Although these methods were not based on the micro-Doppler, they rely on the point cloud of passengers inside a vehicle while windowing the results according to the car seat. The windowing and point cloud methods have several limitations, such as scalability and accuracy. For instance, since these methods were based on the distance of each seat to the radar, meaning that the relative distance must be measured for every car. Moreover, the number of detected points significantly varies passenger by passenger (depending on the size, type, gender, weight, and so on). For instance, if a passenger moves more than the other, his/her motion will be the dominant signal so that other passengers will be concealed. Therefore, with the point cloud and windowing method, we could not obtain a generalized model to be sued in any car.

In [9], we showed that a machine learning algorithm is a promising solution to distinguish occupants with a zero distance and count the number of occupants. Although a good result was obtained for passenger counting and localization, the datasets were obtained from only one car. All collected samples, collected from one car, were combined, shuffled, and then divided into train, validation, and test sets (k-fold method [39]) to evaluate the system performance. Although this approach could provide very high accuracy, the evaluation method is not trustable for a completely new scenario, especially in a new car. Part of the reason is that the radar frame rate is high; thus, mixing all samples and taking some of them for the test set would not ensure that our samples are completely unseen. Generally, in the case of radar-based datasets for machine learning/deep learning models, the evaluation method plays a crucial part. We believe that the appropriate evaluation method is to use a separate sample of recorded data for the test set (i.e., session independent) to report a model performance reliably. This ensures the generality of machine learning applicability in radar sensors. In addition to the issues mentioned above, in previous studies, there was no report on identifying the type of occupant in each seat (adult versus child) using radar sensors.

In this article, we aim to address four identified and major challenges in the field of radar-based in-cabin sensing to reach: 1) a good and large and reliable dataset comprising different cars and various scenarios with different passengers, including real infants/children, instead of a phantom mimicking their chest motion; 2) a proper method of model performance analysis; 3) occupancy-type classification (adult versus child); and 4) a generalized method to be working in every five-seater car.

To propose a more generalized model that could be used for new cases, we collected data from a wide range of scenarios in four cars occupied by different passengers (including infants/children and adults) over several months. To create session-independent datasets, we trained our model based on the samples collected from the first three cars and evaluated the model based on the datasets collected from the fourth car. Classical machine learning methods, such as SVM, performed poorly in predicting new scenarios in the test set. This is because SVM is unable to extract the necessary features and handle large and complicated datasets [40]. However, we developed a deep learning model coupled with radar signal processing for: 1) occupancy counting and localization dedicated for each seat, whether occupied or not and counts passengers (location and the total number of passengers); 2) classification differentiates between two kinds of occupants: an adult and a child; and 3) lock car enables an alarm when a single passenger is left alone in a locked vehicle. We analyzed different deep learning models and compared the performance of the models. Our results demonstrated that time series deep learning algorithms perform better in radar-based in-vehicle occupancy detection. Based on the results, we proposed a temporal convolution network based on 2-D convolutional neural network-long short-term memory (2-D CNN-LSTM) algorithm that is trained based on samples over time. Our proposed model accurately detects passengers, identifies occupied seats, counts the number of passengers, and classifies them. Therefore, in this article, we address four main challenges in previous works for in-vehicle occupancy detection and provide a generalized algorithm for passenger monitoring in five-seat cars. Our main contributions are given as follows.

1) Deep learning models for in-cabin sensing trained and tested on a huge dataset captured from several different cars occupied by different people.
2) Deep learning-based occupancy counting and localization based on data from adults and children/infants of different sizes and ages.
3) Deep learning-based occupancy type identification trained and tested based on actual children and infants at different ages.
4) Deep learning-based occupancy presence/absence detection with 100% accuracy.

II. Methodology

We have been using mm-wave radars to perform the sensing functionalities in various applications [41], [42], [43], [44], [45], [46], [47], [48], [49]. Many of those are available commercially off-the-shelf, such as the ones from Analog Devices [50], Infineon [51], Texas Instruments [52], NXP [53], and Vayyar [54]. As in our past works [9], we chose an MIMO radar, based on its antenna configuration, since it could provide azimuth information in addition to elevation information of the car environment as well as the range of passengers. It should be mentioned that all the results provided in this article are radar-agnostic. In an MIMO radar, for each transmitting (Tx) antenna, there are multiple receiving (Rx) antennas to record the received signals leading to increased angular resolution [9]. Due to the geometry of the radar sensor antennas, we can generate a 3-D representation (range, azimuth, and elevation) by measuring the strength of the reflected signal from an in-car environment. Since the radar we have been using transmits radio frequency (RF) signals continuously, it provides the micro-Doppler, leading to a 4-D imaging system.

A. RF Image Construction

Depending on the antenna array geometry, the number of elements, and the application, various methods could be used to compute the range, azimuth, and elevation information of the environment (referred to as the raw radar images in this article). In our work, to generate raw radar images, we used a Capon beamformer [9], [45] and delay and sum (DAS) imaging method [55]. More details of the Capon beamformer can be found in our previous works [9]. Since all the RF images were constructed based on DAS in this article, we provide more details on the DAS imaging method in this section. Having the frequency (a series of frequencies with a frequency step of Δf) and spatial samples of reflected signals from a car environment, a 3-D image of the car could be constructed. A transceiver transmits a wideband RF signal and collects the return signal over a rectangular planar aperture ($L_{z\min}, L_{z\max}, L_{y\min}, L_{y\max}$). The return signal at the receiver is then coherently processed to construct a 3-D image of the environment. Under the point target model, which ignores the multiple scattering effects, the received signal can

be written as [55]

$$E_s\left(r_{rm}, r_{tm}, k_p\right) = \int G(r_{rm}, r, k_n) G\left(r, r_{tm}, k_n\right) \sigma\left(r\right) dr \quad (1)$$

where $E_s\left(r_{rm}, r_t, k_n\right)$ is the received scattered field at the mth receiver location due to the illumination of the nth transmitter, σ is the reflectivity of the target, r_{rn}, r_{tm}, and r are the position vectors of the nth transmitter, the mth receiver, and targets, i.e., $r_{tm} = (x_{tm}, y_{tm}, z_{tm})$, $r_{rn} = (x_{rn}, y_{rn}, z_{rn})$, and $r = (x, y, z)$, k_p is the free-space wavenumber of the pth operating frequency, and $G(r_{rm}, r, k_p)$ and $G\left(r, r_{tn}, k_p\right)$ are the free-space Green's functions that relate the wave propagation process from the transmitter to the target and from the target to the receiver.

Then, the in-cabin image can be reconstructed through the adjoint operation [55]

$$I\left(r\right) = \int_{k_{\min}}^{k_{\max}} dk_n \int_{L_{z\min}}^{L_{z\max}} \int_{L_{y\min}}^{L_{y\max}} dr_{rm} E_s\left(r_{rm}, r_{tm}, k_p\right) G^*\left(r_{rm}, r, k_p\right) G^*\left(r, r_{tn}, k_p\right) \quad (2)$$

where subscript * represents complex conjugation.

From the concept of time reversal imaging, the complex conjugation in the frequency domain is equivalent to the time reversal in the time domain. The backpropagation of the time-reversed field will focus on the target. In the above equation, the inner two integrands correspond to the beamforming over the cross range, and the outer integrand is the coherent summation over all the operating frequencies.

Since the target (a passenger) is in the far field, this then leads to an approximate expression for Green's function as [55]

$$G\left(r, r_{tn}, k_p\right) = \frac{e^{-jk_p r_{tn}}}{4\pi r_{tn}}, \quad G\left(r_{rm}, r, k_p\right) = \frac{e^{-jk_p r_{rm}}}{4\pi r_{rm}}. \quad (3)$$

The value of each voxel is calculated by compensating for the travel time of each signal to the voxel location and summing them.

B. Data Collection

Our dataset was recorded with a commercially available MIMO radar system [54] with the parameters listed in Table I. For each Tx antenna, there are 20 Rx antennas for collecting and recording the received RF signals. Each Rx antenna and its associated Tx antenna serve as an antenna pair. Fig. 1 shows the experimental setup, while the car was occupied by five passengers. All experiments were carried out in stationary cars, while the engine and car air conditioner were on. Since the primary purpose of this work was to showcase the radar capability in occupant counting and localization, identifying occupied seats, and recognizing the type of occupants, the impact of vibration and jarks on our proposed methods was not analyzed. These analyses would be our future work.

In total, 780 scenarios with a variety of seating situations, occupancy types, and occupancy numbers, with each lasting roughly 30 s, were recorded with a radar frame rate of 5. From each frame, a 3-D matrix of shape 29 × 29 × 24 (X samples, Y samples, and range samples) was generated, containing information about the range, angles, and amplitude of reflected signals. In contrast to previous works [9], where only one car was used, this dataset is collected in four different cars. Due to the high degree of similarity among frames from the same recording session, care was taken to ensure that frames from a particular session ended up in either the test set or the train set to prevent overfitting. Unlike previous works in that dolls were designed to mimic infants, one of the distinctive advantages of our datasets is that we included infants and children of different ages in our data collection campaign. Fig. 2 shows the occupant type distribution of our datasets. Note that Fig. 3 shows the seat nomenclature used in this article. As illustrated, except for the driver seats, for the other four seats, the number of samples collected from children is more than from adults because our main purpose was to detect children for safety concerns. In addition, if our system detects children accurately, it could detect adults easier since children (especially infants) usually sleep in the car, which might lead to missed detection.

Fig. 1. Photograph of a vehicle occupied by five passengers.

TABLE I
CONFIGURATION AND THE RADAR PARAMETERS

Characteristic	*Specification*
Number of Tx	20
Number of Rx	20
Frequency Band	62 - 69 GHz
Frame rate	5
Range resolution	7.5 cm
Angular resolution	6.7 *deg*
Number of frames	300
Number of frequencies to scan	192

C. Preprocessing

Each raw radar image, denoted as $\widehat{O}(x, y, z)$, is a complex 3-D matrix. The image values represent the reflective properties of the targets in space $o(x, y, z)$, calculated on a grid of 3-D coordinates. The absolute value of $\widehat{O}(x, y, z)$, $|\widehat{O}(x, y, z)|$, represents the reflected power, which is related to the radar cross section (RCS) and the range of the targets.

Since transmitters and receivers are almost collocated in most commercially available MIMO radars, there is a strong leakage between them, dominating other reflected signals. Fig. 4(a)–(c) represents X-R, Y-R, and X-Y samples of the

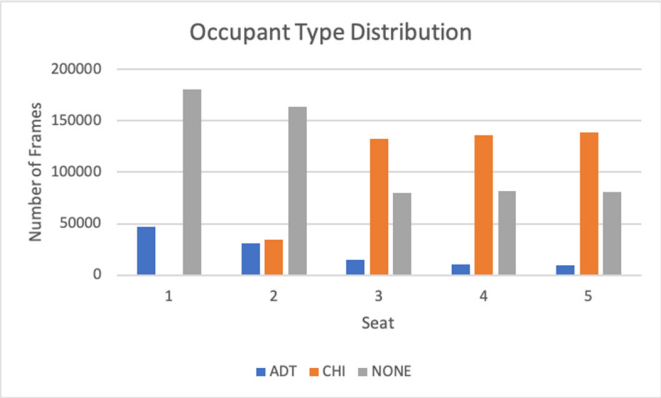


Fig. 2. Visualization of occupant-type distribution of the datasets collected for this study. ADT: adult. CHI: children. NONE: empty cars.

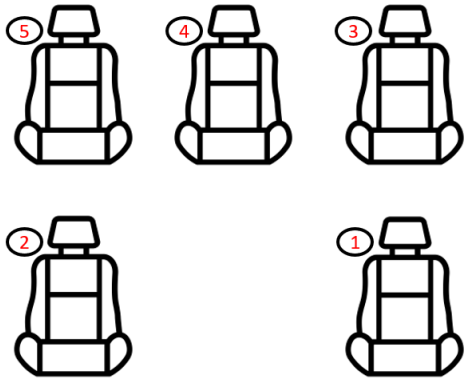


Fig. 3. Inside the car and the seat nomenclature.

case when an infant was in the rear middle seat (seat 4), respectively. As seen, there is a strong line in close range (shown by a red rectangle) to the radar that dominates other signals, and the signals coming from an infant are concealed. This is also visible in a 3-D representation of the same scenario in Fig. 5(a). As seen, the raw radar image is noisy, and the passenger position is not visible. This is because the leakage between transmitters to receivers and the reflections from static objects (stationary clutter) in the car is dominant. For example, the metallic doors and the ceiling are more reflective (higher RCS) compared with the reflection coming from the passenger.

To clean the signal and remove the effects of the static objects, clutter removal and a mutual coupling reduction algorithm are required.

One of the most common clutter removal methods is to compute the average value of the signal of each range bin and subtract it from the aggregated signals in each frame [45]. Another method of clutter removal is background subtraction, where the passive clutter is removed by subtracting the image obtained from a completely vacant vehicle, both in terms of passengers and cargo, of the same model [55]. This method clearly suffers from poor scalability.

In this article, to mitigate the reflections from the stationary clutter and remove the leakage, we propose a derivative-based method. We add another variable to our 3-D images (variation

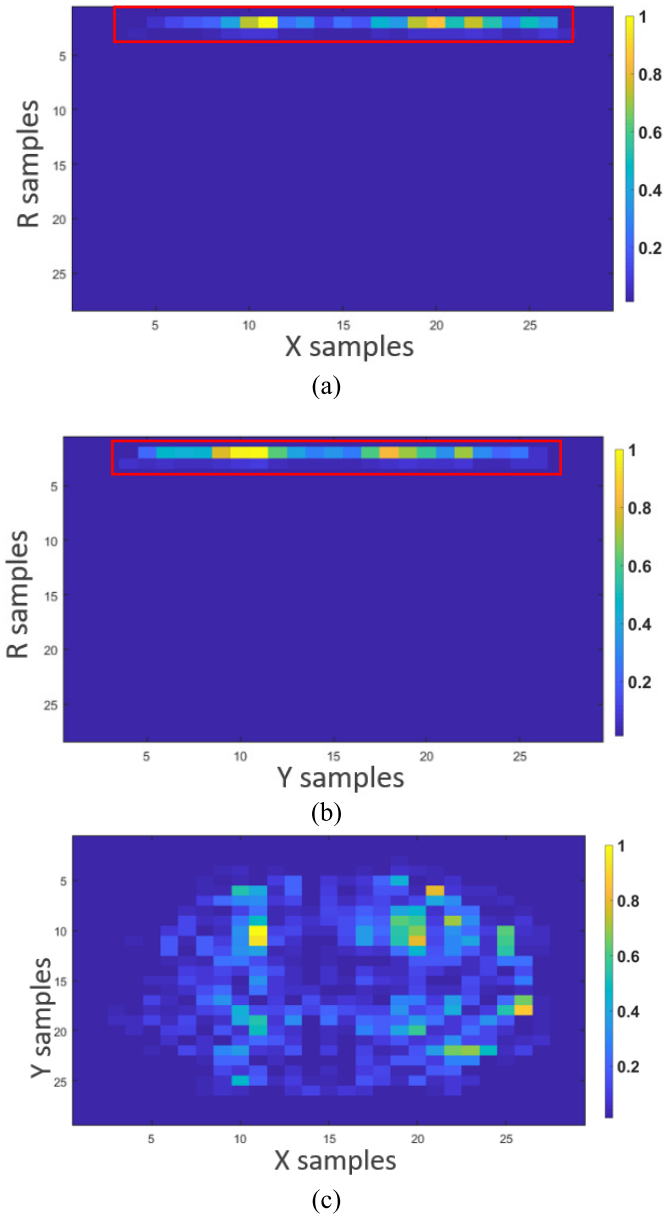


Fig. 4. Two-dimensional visualization of a normalized radar image before removing leakage between transmitters and receivers when seat 4 was occupied by an infant representation of (a) *X-R* samples and (b) *Y-R* samples.

over time), which could be interpreted as velocity. To remove the stationary signals, we apply derivatives on the raw radar image over two raw radar images with M frames interval, which is written as

$$\widehat{O} = I + jQ \tag{4}$$

$$\frac{\mathrm{d}\widehat{O}}{\mathrm{d}t} = \frac{\widehat{O}_i - \widehat{O}_{i-M}}{t_i - t_{i-M}} = I' + jQ' \tag{5}$$

$$Z = \sqrt{I'^2 + Q'^2} \tag{6}$$

where I and Q are the real and imaginary parts of the complex value of $\widehat{O}$, $(\mathrm{d}\widehat{O}/\mathrm{d}t)$ respectively, Z is the derivative function, and Z is the amplitude of the output of the derivative

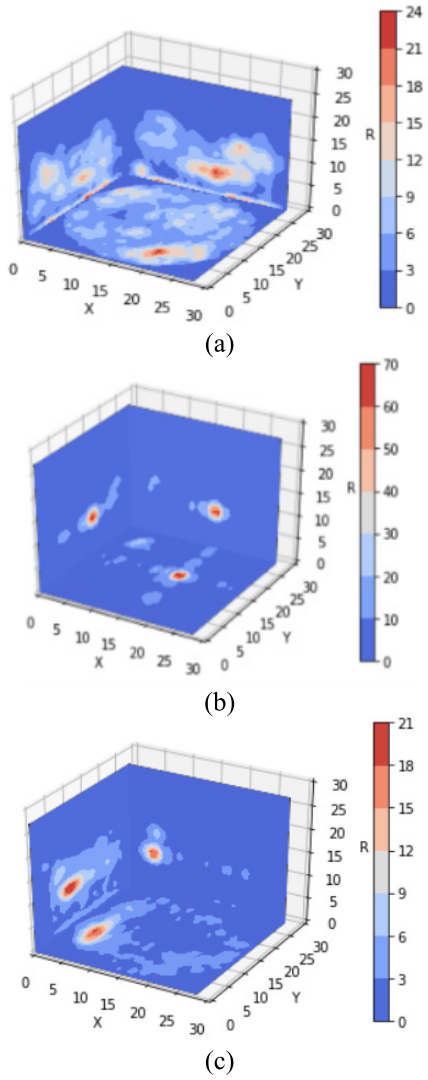


Fig. 5. Three-dimensional visualization of (a) raw radar image and processed radar image, (b) seat 4 was occupied by an infant, and (c) seat 1 was occupied by an adult. The color bar represents signal intensity.

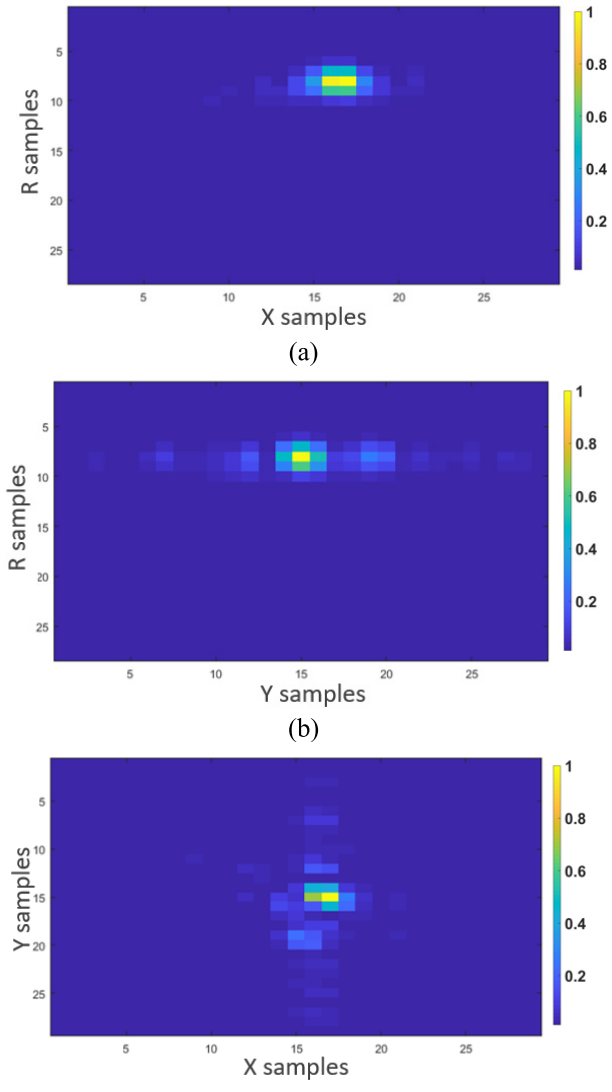


Fig. 6. Two-dimensional visualization of a normalized radar image when seat 4 is occupied by an infant representation of (a) *X*-*R* samples, (b) *Y*-*R* samples, and (c) *X*-*Y* samples.

function, representing the heatmap of the car environment (radar image). The intensity of each pixel denotes the reflected power received from each target in the car at that point [9]. Fig. 5(b) shows the processed radar image over $M = 3$ intervals. Compared with Fig. 5(a), Fig. 5(b) is cleaner and more interpretable. Moreover, Fig. 5(c) shows the case when seat 1 (the driver seat) was occupied by an adult. As seen, although we can count only one passenger for both cases, visually, there is no distinguishable factor that could be used to differentiate the infant and the adult. As noted, the maximum magnitude of the image in Fig. 5(b) is more than that in Fig. 5(c), meaning that we are not able to base our classification on the signal intensity. The signal intensity is affected by many factors, such as RCS, distance to the radar, level of movement, relative angle toward the radar, being concealed by a metallic part of the car seats, and so on. To show the details, Fig. 6 shows the 2-D raw radar images after performing the postprocessing method for the case when an infant was in the rear middle seat. To show more complex scenarios, Fig. 7(a) shows the car occupied by three passengers after applying the proposed preprocessing method [seats 1 (an adult), 3 (a child), and 5 (a child)].

As seen, neither the classification of the type of occupants nor counting their number/occupied seats is tractable based on the RF images provided above. Even with the processed images, passenger counting and classification is a complex problem featuring nonlinear characteristics. The traditional signal processing method requires multiple detection thresholds to be tuned using a set of handcrafted rules. This approach is time-consuming and would not be generalized well to different car models. Thus, there is a pressing need for an algorithm that can detect and classify passengers, count the number of occupants, and identify occupied seats.

As such, we propose a deep learning-based model to count the number of passengers and identify their occupied seats inside a car.

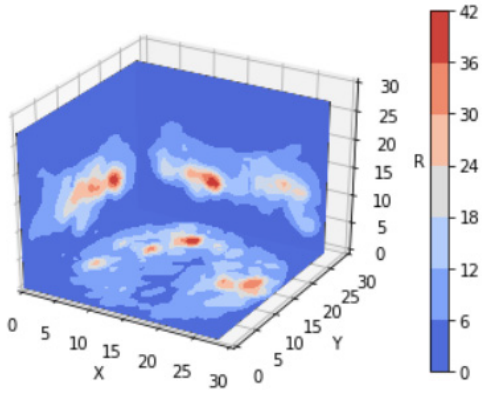


Fig. 7. Three-dimensional visualization of processed radar image seats 1, 3, and 5 that are occupied. The color bar represents signal intensity.

D. Proposed 2-D CNN-LSTM Architecture

Since traditional radar signal processing methods have shown some limitations, particularly in matters of target classification [56], we deployed machine learning in this article. The diagram in Fig. 8 illustrates the processing flow of our proposed in-cabin occupancy detection. As shown, the aforementioned preprocessing method is applied to radar received data to provide a 3-D heatmap of the car environment. The heatmap is then delivered to deep learning models for classification, counting, and detection.

Machine learning-based radar signal processing has been explored in several studies, such as target detection [57], [58], radar image processing and image denoising [59], automatic target recognition [60], and activity recognition [42], [48]. These machine learning algorithms include traditional machine learning (e.g., SVM, decision tree, and random forest) and deep learning [e.g., deep belief networks, autoencoders, CNNs, recurrent neural networks (RNNs), and generative adversarial networks (GANs)]. The raw data are commonly converted into a 2-D image while being treated as an optical image. The corresponding architectures, such as 2-D convolutional neural networks (2-D-CNNs), were used in these systems. However, since a human body motion consists of a series of associated postures through time and the radar heatmaps were collected over successive periods of time (frames by frames), they could be characterized as a time series. For this reason, only CNNs are not a good algorithm for such time-varying data. Given that we have time series data, it is natural to consider using an RNN to keep track of the time series data in terms of memories of what the system has observed so far. In such cases, an appealing method is to use long short-term memory (LSTM) model, an RNN architecture [61]. In the LSTM model, the previous hidden state is passed to the next step of the sequence [48].

Thus, the network uses a set of previously seen samples to make a decision. Our basic approach is based on the idea of LSTM, but since the purpose is to count passengers and classify them, spatial information also plays a key role. A CNN architecture was shown to be the best approach for extracting spatial information. CNN models can learn patterns in a hierarchical manner, and the learned patterns are translation invariant [62]. Our proposed method for counting and classification is to apply the combination of LSTM with CNN and deep neural networks (DNNs) [63] (cascaded networks). Note that this approach is different from the regular LSTM or CNN problem. It is a recurrent layer, but internal matrix multiplications are replaced by convolution operations [62].

Consequently, the data flowing through the 2-D CNN-LSTM cells keep the input dimension (3-D matrix in our work) instead of being flattened as a 1-D vector with features. Therefore, our architecture uses CNN to extract the 3-D spatial features from the radar image, then passes them to LSTM to perform temporal modeling, and finally outputs them to DNN to produce separable feature representations. In other words, we take advantage of the structure in the sequence of 3-D representations of the radar reflections in the car environment to learn valuable patterns from the heatmap data (which is very difficult using 1-D CNN, LSTM, or DNN).

III. Results

In this section, the results of various machine learning and deep learning models are discussed. Data augmentation methods, such as random translation and rotation [64], are applied in both tasks leading to a significant increase in accuracy. Standardization is chosen, which transforms data to have a mean of 0 and a standard deviation of 1. It is worth mentioning that our networks were implemented, trained, validated, and tested in PyTorch in the CMC cloud server [65]. As shown in Fig. 9, N frames of range–azimuth–elevation heatmaps ($29 \times 29 \times 24$) are the inputs to be fed to the models. For our models, a wide range of N was used, and $N = 10$ was found as an optimum number of frames. To count the passengers and identify their occupied seats, and classify them, we followed two approaches explained in Sections III-A and III-B.

A. Counting and Localization

To monitor and count passengers in a five-seater car, each seat is represented as a binary encoding (i.e., 1 if the seat is occupied; 0 otherwise). Therefore, there are 32 scenarios (i.e., 2^5 for five seats) that each index referring to a respective seat number: 1 for an occupied seat (by living subject) and 0 for a nonoccupied seat). A prediction is considered correct if and only if all five seats are correctly classified. We deploy four models to be trained, namely, SVM, CNN, LSTM, and 2-D CNN-LSTM.

To find a proper model for our in-vehicle occupancy detection system, we first shuffled our data sets and split the datasets to train, validate, and test sets as 70%, 20%, and 10%, respectively. We started with the conventional SVM algorithm (the same method used in [9]). However, the accuracy obtained from SVM was 26%. The reason for this poor performance was that the SVM algorithm was not able to handle large datasets and extract features from flattened images. Then, a CNN model was developed, which resulted in 99% accuracy. In order to validate and verify the CNN performance and provide a generalized model for every five-seat car, we deployed the CNN model for our session-independent

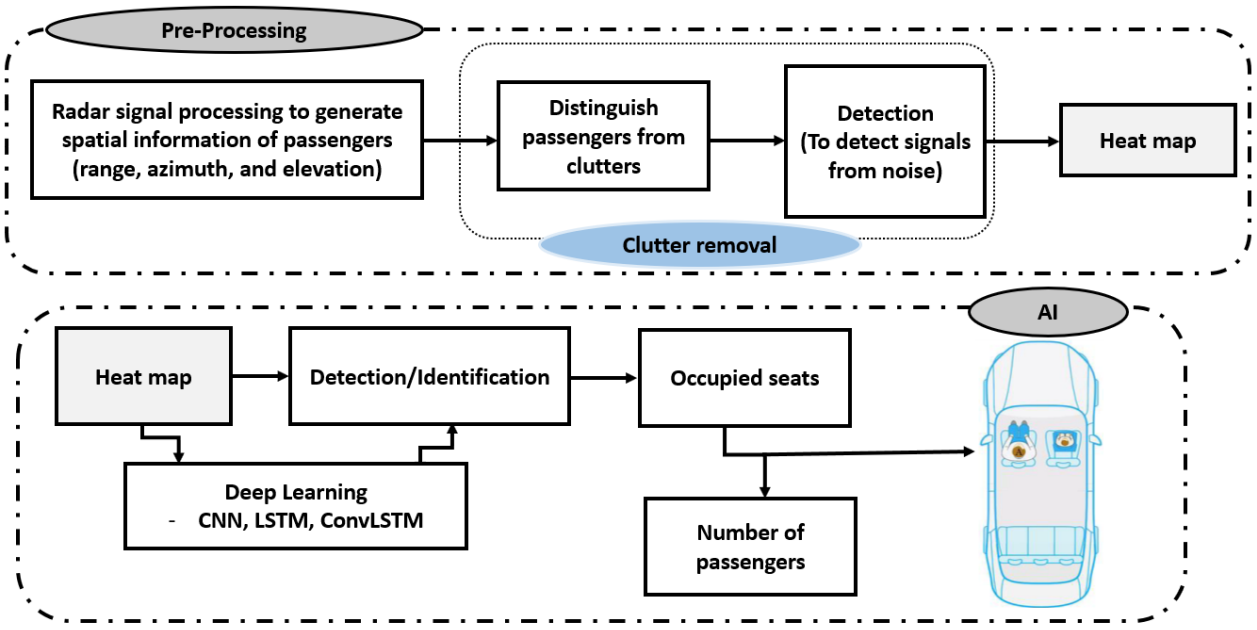


Fig. 8. Proposed in-vehicle occupancy detection algorithm. Preprocessing is applied to radar received data to provide the heatmap. The heatmap is then delivered to deep learning models to identify occupied seats and count the passengers.

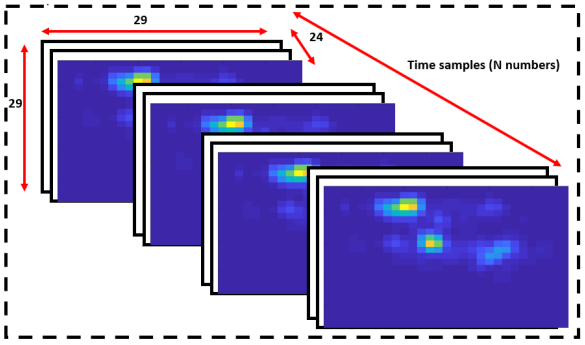


Fig. 9. Time-dependent input for temporal deep learning models. The input is a tensor of 29 × 29 × 24.

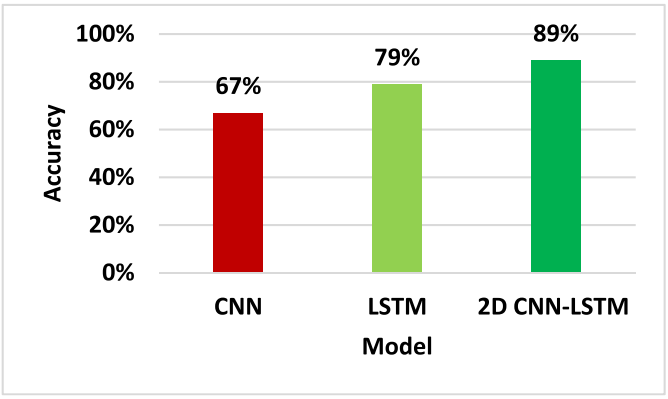


Fig. 10. Accuracy obtained from CNN, LSTM, and 2-D CNN-LSTM.

datasets. The CNN model results in a lower accuracy (67%) because passengers were freely moving and doing different types of activities, such as painting and talking with their phones; their random movements caused the varying heatmaps. Therefore, to overcome this issue and propose a model that performs well in almost all scenarios and situations, we added time as another variable. Hence, the results of all models are shown in Fig. 10. As shown, the proposed 2-D CNN-LSTM achieves 89% classification accuracy on the test set when cross entropy is used as the loss function, while the LSTM is 10% less accurate.

This finding implies that our model learns the underlying mathematical correlation between radar signal and passenger arrangement in the car. This is because 2-D CNN-LSTM is based on spatial and temporal features of our datasets, while CNN and LSTM alone are based on only spatial and temporal features, respectively. Comparing LSTM with CNN, the importance of adding extra information as time in radar datasets is noticeable.

To provide more details of each scenario, a confusion matrix of passenger monitoring and counting is shown in Fig. 11. As seen, the model detects if a person is present in the car with a precision = 100% and recall = 99.6%,

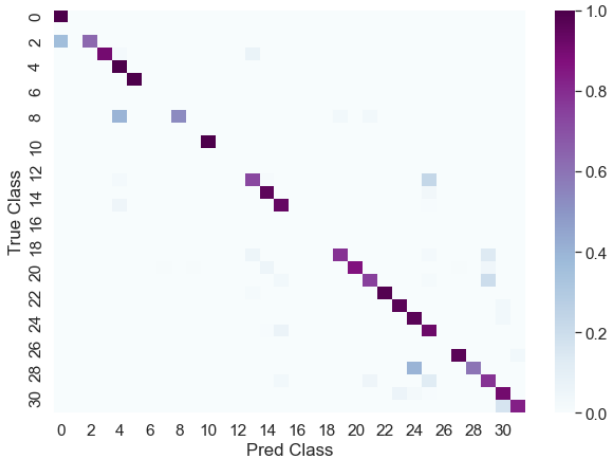


Fig. 11. Confusion matrix of passenger counting and localization. Pred: predicted classes.

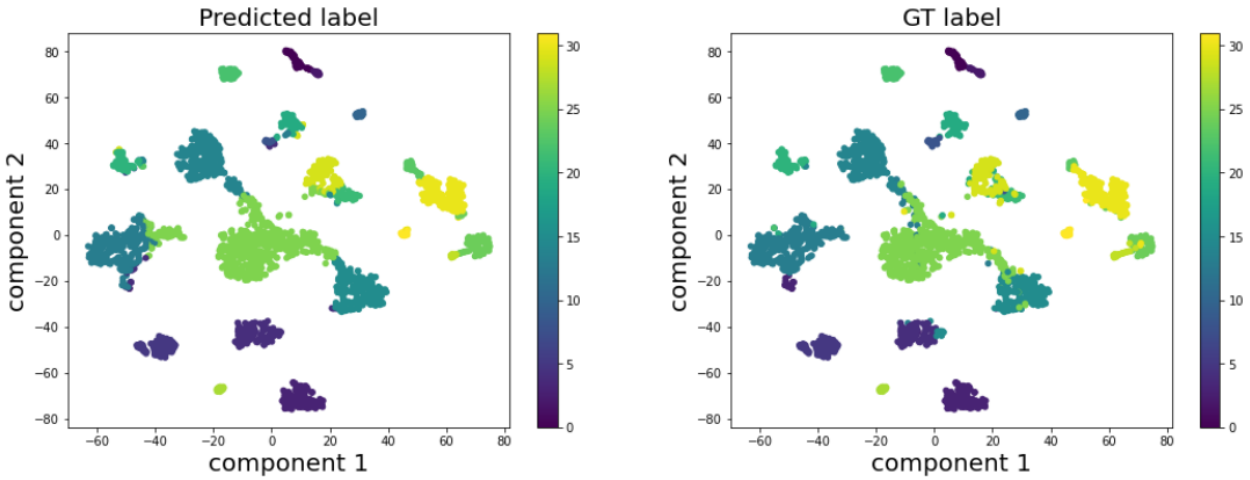


Fig. 12. Clustering plot of the last layer before classifying the test dataset. The color bar represents the ground-truth class. GT is the ground-truth label.

meaning that the probability of missing a passenger inside a car and identify an occupied as a vacant is 0.4%. For example, case 2 [01000] (when a passenger was sitting on seat 2) is misclassified as class 0 [00000] (empty or missed detection). There are some other mispredictions, such as case 8 [10010], misclassified as case 4 [10] or case 28 [11011], and misclassified as case 24 [01011]. In contrast with the "black-box" concept in machine learning, we aim to explain how the model arrives at a specific decision. As each test sample passes through the model, the last hidden layer (i.e., a 512-dimension vector) is extracted and visualized in Fig. 12 after dimensionality reduction via t-distributed stochastic neighbor embedding (t-SNE) [66]. Samples form clusters, and each cluster is associated with a predominant label. In the output layer, the model uses those clusters to assign each sample to a label.

B. Occupant Type

Occupant-type classification is a more challenging task because the model must distinguish between adults and children. For occupancy type, we performed two different methods of analysis: 1) to identify a child/infant inside a car and 2) to distinguish a child/infant from an adult. For the first one, the datasets compose all scenarios of vacant cars (coded as 0) in addition to the cases when children/infants were present in the car (coded as 1). The confusion matrix is constructed in Fig. 13 to summarize the performance of the model in identifying a child occupant. The "Child" class is of particular interest—a desirable in-car safety sensor should alarm the caregiver when a child is left in an unattended vehicle. Our model shows high precision (0.90) and recall (0.95) when used to detect unattended children in vehicles.

For the latter one, each seat has three possible scenarios ("empty," "adult," and "child"). The performance of the proposed 2-D CNN-LSTM model for passenger classification loss function and accuracy is plotted in Fig. 14(a) and (b), respectively. The accuracy of the training set is around 90% (the blue curve), and the validation set is 76% accurate (the brown curve). These results are obtained after hyperparameter

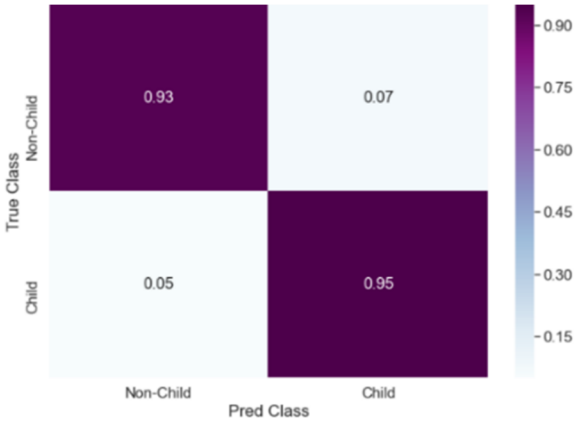


Fig. 13. Test confusion matrix of occupant types.

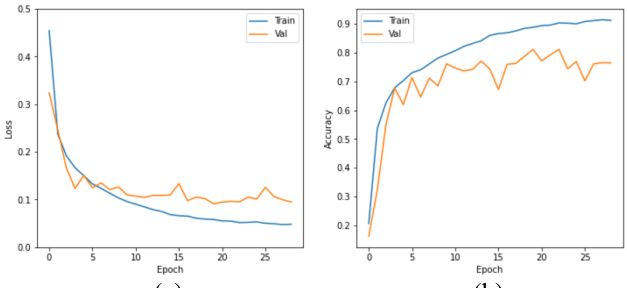


Fig. 14. Performance of the proposed 2-D CNN-LSTM model for passenger classification (a) loss function and (b) accuracy.

tuning. The model was 74% accurate in testing new scenarios in a new car. Based on our criteria, if the occupant type is misclassified at any seat, the output was considered misclassified for the entire sample.

Comparing the accuracy obtained for passenger counting with classification, it is noticed that we can count the number of passengers more accurately. We assume that one of the major reasons is that all infants and children of different ages are considered "child" classes. However, apparently, the size and shape of an infant lying on a rear seat (all seat types

were defined in [32]) are different from an older child sitting on a booster seat. The reason why we regarded all of them as "child" was the lack of enough datasets for children and infants to cover a wide range of ages. Collecting sufficient data from children and infants at different ages inside a car was challenging over a long period of time, especially during the COVID-19 restrictions.

IV. Discussion

Our findings suggest that RF-based continuous passenger monitoring within a vehicle is feasible and has the potential to improve passenger safety. The proposed monitoring system operates noninvasively, analyzing reflections of radio signals to record the presence or absence of a passenger, identify the type of passengers (adults versus kids), count them, and identify their occupied seats (location). This information can be transmitted to a guardian (e.g., parents) upon triggering events, particularly when a child is left in an unattended vehicle. This has significant implications, particularly in the context of the requirement of child presence detection in Europe and a rear occupant alert system in the USA.

The results obtained in this article show a prodigious role for machine learning to play in in-vehicle sensors, not only as an essential component of passenger detection but also as safety precautions such as unattended children alarms. Deep learning-based algorithms were deployed to a new radar research domain to tackle traditional and new challenges from a novel perspective. It was shown that a 3-D RF image of an in-vehicle environment is sufficient information for deep learning models. Applying the proposed preprocessing to raw radar images, direct reflections from transmitters to receivers and reflections from stationary clutter were removed. The remaining signals, represented as a 3-D RF image, contain information on the intensity and velocity of reflected signals in range and angles at each voxel. Since the possible options of seat configuration in a five-seater car are limited to 32, supervised machine learning was trained to predict these limited scenarios. Based on the preprocessing method applied to the RF raw images and given that all passengers cannot move/walk to change their position, it is assumed that the effect of the vehicle environment was removed. This was validated by capturing data from different cars to ensure the generality of this model.

There are several limitations to this study. First, all experiments were conducted in five-seater cars with the same configuration. However, the deep learning model might be unable to predict new scenarios in a different seat configuration. Second, impacts of potential obstructions by chairs or people were not addressed, especially for a car with a third row. Finally, the detrimental effects of noise (especially thermal noise) were not explored. Due to the in-vehicle confined environment surrounded by metallic objects, the radar sensor might be influenced by varying noise levels.

Despite the aforementioned limitations, we believe that this research provides important insights and addresses key unmet needs in in-vehicle occupancy domains.

V. Conclusion

In this article, an in-cabin monitoring and radar-based perception system for increased safety inside the vehicle is reported. It was shown that the AI-powered radar-based system is a promising technology for occupancy detection, counting, and classification using a 4-D radar imaging system. In addition, a derivative-based method of passive clutter removal also is used to remove clutter and mitigate leakage between transmitters and receivers. The performance of different deep learning models is analyzed for in-vehicle occupancy monitoring. It is shown that both temporal and spatial features of the car environment play crucial parts. It is demonstrated that the 2-D CNN-LSTM model performs best in predicting new scenarios based on the session-independent datasets as it uses both feature extraction methods from the CNN and temporal features from the LSTM. Consequently, it was demonstrated that a 4-D radar could be used to monitor both the front and rear seat passengers at the same time.

Acknowledgment

This work would have not been carried out without support from Vayyar. The authors wish to acknowledge the support of the National Sciences and Engineering Research Council of Canada, the Ontario Center of Innovation, and the CMC Microsystems Infrastructure for Supporting Cloud and Edge Computing Research.

References

[1] *Vayyar Announces its 4D Imaging Radar Sensor on a Chip for in-Cabin Monitoring & Vehicle Safety Systems*. Accessed: Aug. 29, 2022. [Online]. Available: https://www.futurecar.com/4313/Vayyar-Announces-its-4D-Imaging-Radar-Sensor-on-a-Chip-for-In-Cabin-Monitoring-&-Vehicle-Safety-Systems

[2] *Euro NCAP 2025 Roadmap: In Pursuit of Vision Zero*. Accessed: Aug. 3, 2021. [Online]. Available: https://cdn.euroncap.com/media/30700/euroncaproadmap-2025-v4.pdf

[3] T. A. Dingus, "The national surface transportation safety center for excellence," Virginia Tech College Eng., Blacksburg, VA, USA, Tech. Rep. 21-UT-098, 2021.

[4] K. Kasten, A. Stratmann, M. Munz, K. Dirscherl, and S. Lamers, "iBolt technology—A weight sensing system for advanced passenger safety," *Adv. Microsyst. Automot. Appl.*, vol. 2006, pp. 171–186, 2006, doi: 10.1007/3-540-33410-6_14/COVER.

[5] A. Vergnano and F. Leali, "A methodology for out of position occupant identification from pressure sensors embedded in a vehicle seat," *Hum.-Intell. Syst. Integr.*, vol. 2, nos. 1–4, pp. 35–44, May 2020, doi: 10.1007/S42454-020-00008-W.

[6] B. George, H. Zangl, T. Bretterklieber, and G. Brasseur, "Seat occupancy detection based on capacitive sensing," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 5, pp. 1487–1494, May 2009, doi: 10.1109/TIM.2009.2009411.

[7] A. Satz, D. Hammerschmidt, and D. Tumpold, "Capacitive passenger detection utilizing dielectric dispersion in human tissues," *Sens. Actuators A, Phys.*, vol. 152, no. 1, pp. 1–4, May 2009, doi: 10.1016/J.SNA.2009.03.005.

[8] *CO_2/Alcohol Sensing. Asahikasei Sensair*. Accessed: Aug. 29, 2022. [Online]. https://asahikasei-automotive-us.com/innovations/senseair

[9] H. Abedi, S. Luo, V. Mazumdar, M. M. Y. R. Riad, and G. Shaker, "AI-powered in-vehicle passenger monitoring using low-cost mm-wave radar," *IEEE Access*, vol. 10, pp. 18998–19012, 2022, doi: 10.1109/ACCESS.2021.3138051.

[10] I. Papakis, A. Sarkar, A. Svetovidov, J. S. Hickman, and A. L. Abbott, "Convolutional neural network-based in-vehicle occupant detection and classification method using second strategic highway research program cabin images," *J. Transp. Res. Board*, vol. 2675, no. 8, pp. 443–457, Apr. 2021, doi: 10.1177/0361198121998698.

[11] Y.-L. Hou and G. K. H. Pang, "People counting and human detection in a challenging situation," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 41, no. 1, pp. 24–33, Jan. 2011, doi: 10.1109/TSMCA.2010.2064299.

[12] M. Devy, A. Giralt, and A. Marin-Hernandez, "Detection and classification of passenger seat occupancy using stereovision," in *Proc. IEEE Intell. Vehicles Symp.*, Oct. 2000, pp. 714–719, doi: 10.1109/IVS.2000.898433.

[13] P. Kuchár, R. Pirník, T. Tichý, K. Rástočný, M. Skuba, and T. Tettamanti, "Noninvasive passenger detection comparison using thermal imager and IP cameras," *Sustainability*, vol. 13, no. 22, p. 12928, Nov. 2021, doi: 10.3390/SU132212928.

[14] *Autopilot | Tesla*. Accessed: Aug. 29, 2022. [Online]. Available: https://www.tesla.com/autopilot

[15] H. Abedi, C. Magnier, and G. Shaker, "Passenger monitoring using AI-powered radar," in *Proc. IEEE 19th Int. Symp. Antenna Technol. Appl. Electromagn. (ANTEM)*, Aug. 2021, pp. 1–2, doi: 10.1109/ANTEM51107.2021.9518503.

[16] M. Alizadeh, H. Abedi, and G. Shaker, "Low-cost low-power in-vehicle occupant detection with mm-wave FMCW radar," in *Proc. IEEE Sensors*, Oct. 2019, pp. 1–4, doi: 10.1109/SENSORS43011.2019.8956880.

[17] H. Abedi, S. Luo, and G. Shaker, "On the use of low-cost radars and machine learning for in-vehicle passenger monitoring," in *Proc. IEEE 20th Topical Meeting Silicon Monolithic Integr. Circuits RF Syst. (SiRF)*, Jan. 2020, pp. 63–65, doi: 10.1109/SIRF46766.2020.9040191.

[18] H. Abedi, M. Ma, J. Yu, J. He, A. Ansariyan, and G. Shaker, "On the use of machine learning and deep learning for radar-based passenger monitoring," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (AP-S/URSI)*, Jul. 2022, pp. 902–903, doi: 10.1109/AP-S/USNC-URSI47032.2022.9887034.

[19] H. Song, Y. Yoo, and H.-C. Shin, "In-vehicle passenger detection using FMCW radar," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Jan. 2021, pp. 644–647, doi: 10.1109/ICOIN50884.2021.9334014.

[20] G. Greneker, "Millimeter wave safety warning system for in-vehicle signing," in *Proc. IEEE Nat. Radar Conf.*, May 1997, pp. 182–185, doi: 10.1109/NRC.1997.588277.

[21] S. Lim, J. Jung, S.-C. Kim, and S. Lee, "Deep neural network-based in-vehicle people localization using ultra-wideband radar," *IEEE Access*, vol. 8, pp. 96606–96612, 2020, doi: 10.1109/ACCESS.2020.2997033.

[22] J. H. Huh and S. H. Cho, "Seat belt reminder system in vehicle using IR-UWB radar," in *Proc. Int. Conf. Netw. Infrastruct. Digit. Content (ICNIDC)*, Aug. 2018, pp. 256–259, doi: 10.1109/ICNIDC.2018.8525708.

[23] H. Song and H.-C. Shin, "Single-channel FMCW-radar-based multi-passenger occupancy detection inside vehicle," *Entropy*, vol. 23, no. 11, p. 1472, Nov. 2021, doi: 10.3390/E23111472.

[24] A. Santra, R. V. Ulaganathan, and T. Finke, "Short-range millimetric-wave radar system for occupancy sensing application," *IEEE Sensors Lett.*, vol. 2, no. 3, pp. 1–4, Sep. 2018, doi: 10.1109/LSENS.2018.2852263.

[25] G. Wang, C. Gu, T. Inoue, and C. Li, "A hybrid FMCW-interferometry radar for indoor precise positioning and versatile life activity monitoring," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 11, pp. 2812–2822, Nov. 2014, doi: 10.1109/TMTT.2014.2358572.

[26] L. Deng, "Deep learning: Methods and applications," Now Found. Trends, Tech. Rep., 2014, doi: 10.1561/9781601988157.

[27] J. W. Choi, D. H. Yim, and S. H. Cho, "People counting based on an IR-UWB radar sensor," *IEEE Sensors J.*, vol. 17, no. 17, pp. 5717–5727, Sep. 2017, doi: 10.1109/JSEN.2017.2723766.

[28] J.-H. Choi, J.-E. Kim, and K.-T. Kim, "Deep learning approach for radar-based people counting," *IEEE Internet Things J.*, vol. 9, no. 10, pp. 7715–7730, May 2022, doi: 10.1109/JIOT.2021.3113671.

[29] L. Servadei et al., "Label-aware ranked loss for robust people counting using automotive in-cabin radar," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2022, pp. 3883–3887.

[30] M. Stephan, S. Hazra, A. Santra, R. Weigel, and G. Fischer, "People counting solution using an FMCW radar with knowledge distillation from camera data," in *Proc. IEEE Sensors*, Oct. 2021, pp. 1–4, doi: 10.1109/SENSORS47087.2021.9639798.

[31] Z. Yang, G. Qi, and R. Bao, "Indoor regional people counting method based on Bi-motion-model-framework using UWB radar," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2022.3159590.

[32] H. Abedi, C. Magnier, V. Mazumdar, and G. Shaker, "Improving passenger safety in cars using novel radar signal processing," *Eng. Rep.*, vol. 3, no. 12, Dec. 2021, Art. no. e12413, doi: 10.1002/eng2.12413.

[33] H. Sterner, W. Aichholzer, and M. Haselberger, "Development of an antenna sensor for occupant detection in passenger transportation," *Proc. Eng.*, vol. 47, pp. 178–183, Jan. 2012, doi: 10.1016/J.PROENG.2012.09.113.

[34] A. Lazaro, M. Lazaro, R. Villarino, and D. Girbau, "Seat-occupancy detection system and breathing rate monitoring based on a low-cost mm-wave radar at 60 GHz," *IEEE Access*, vol. 9, pp. 115403–115414, 2021, doi: 10.1109/ACCESS.2021.3105390.

[35] W. Li, Y. Gao, Z. Hu, N. Liu, K. Wang, and S. Niu, "In-vehicle occupant detection system using mm-wave radar," in *Proc. 7th Int. Conf. Commun., Image Signal Process. (CCISP)*, Nov. 2022, pp. 395–399, doi: 10.1109/CCISP55629.2022.9974471.

[36] T. Kitamura, T. Kumagai, T. Takei, I. Matsushima, N. Oishi, and K. Suwa, "Occupant body imaging based on occupancy grid mapping," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jul. 2021, pp. 755–761, doi: 10.1109/IV48863.2021.9575618.

[37] N. Munte, A. Lazaro, R. Villarino, and D. Girbau, "Vehicle occupancy detector based on FMCW mm-wave radar at 77 GHz," *IEEE Sensors J.*, vol. 22, no. 24, pp. 24504–24515, Dec. 2022, doi: 10.1109/JSEN.2022.3218454.

[38] Y. Li, C. Gu, and J. Mao, "An in-vehicle occupant detection technique based on a 60 GHz FMCW MIMO radar," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Aug. 2022, pp. 1–3, doi: 10.1109/IWS55252.2022.9977812.

[39] *A Gentle Introduction to Machine Learning | CSEG RECORDER*. Accessed: Aug. 10, 2022. [Online]. Available: https://csegrecorder.com/articles/view/a-gentle-introduction-to-machine-learning

[40] J. Zabalza, C. Clemente, G. D. Caterina, J. Ren, J. J. Soraghan, and S. Marshall, "Robust PCA micro-Doppler classification using SVM on embedded systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 3, pp. 2304–2312, Jul. 2014, doi: 10.1109/TAES.2014.130082.

[41] H. Abedi and G. Shaker, "Low-cost 3D printed dielectric hyperbolic lens antenna for beam focusing and steering of a 79GHz MIMO radar," in *Proc. IEEE Int. Symp. Antennas Propag. North Amer. Radio Sci. Meeting*, Jul. 2020, pp. 1543–1544, doi: 10.1109/IEEECONF35879.2020.9329969.

[42] H. Abedi, C. Magnier, J. Boger, and A. Wong, "Integration of random forests and MM-wave FMCW radar technology for gait recognition," *J. Comput. Vis. Imag. Syst.*, vol. 5, no. 1, p. 2, 2019, doi: 10.1109/bhi.2018.8333371.

[43] H. Abedi, G. Shaker, J. Boger, P. Morita, and A. Wong, "Use of millimeter wave FMCW radar to capture gait parameters," *Amer. J. Biomed. Sci. Res.*, vol. 6, no. 2, pp. 122–123, 2019, doi: 10.34297/AJBSR.2019.06.001009.

[44] H. Abedi, J. Boger, P. P. Morita, A. Wong, and G. Shaker, "Hallway gait monitoring system using an in-package integrated dielectric lens paired with a mm-wave radar," *Sensors*, vol. 23, no. 1, p. 71, Dec. 2022, doi: 10.3390/S23010071.

[45] H. Abedi, J. Boger, P. P. Morita, A. Wong, and G. Shaker, "Hallway gait monitoring using novel radar signal processing and unsupervised learning," *IEEE Sensors J.*, vol. 22, no. 15, pp. 15133–15145, Aug. 2022, doi: 10.1109/JSEN.2022.3184188.

[46] H. Abedi, P. P. Morita, J. Boger, A. Wong, and G. Shaker, "In-package integrated dielectric lens paired with a MIMO mm-wave radar for corridor gait monitoring," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (APS/URSI)*, Dec. 2021, pp. 1795–1796, doi: 10.1109/APS/URSI47566.2021.9704192.

[47] H. Abedi, A. Ansariyan, P. P. Morita, J. Boger, A. Wong, and G. Shaker, "Sequential deep learning for in-home activity monitoring using mm-wave FMCW radar," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (APS/URSI)*, Dec. 2021, pp. 1499–1500, doi: 10.1109/APS/URSI47566.2021.9704291.

[48] H. Abedi, A. Ansariyan, P. P. Morita, A. Wong, J. Boger, and G. Shaker, "AI-powered non-contact in-home gait monitoring and activity recognition system based on mm-wave FMCW radar and cloud computing," *IEEE Internet Things J.*, early access, Jan. 9, 2023, doi: 10.1109/JIOT.2023.3235268.

[49] H. Abedi et al., "Non-visual and contactless wellness monitoring for long term care facilities using mm-wave radar sensors," in *Proc. IEEE Sensors*, Oct. 2022, pp. 1–4, doi: 10.1109/SENSORS52175.2022.9967327.

[50] *Mixed-Signal and Digital Signal Processing ICs | Analog Devices*. Accessed: Aug. 10, 2022. [Online]. Available: https://www.analog.com/en/index.html

[51] *Semiconductor & System Solutions Infineon Technologies*. Accessed: Aug. 10, 2022. [Online]. Available: https://www.infineon.com/

[52] *Analog | Embedded Processing | Semiconductor Company | TI.com*. Accessed: Aug. 10, 2022. [Online]. Available: https://www.ti.com/
[53] *Automotive Radar Systems | NXP Semiconductors*. Accessed: Aug. 10, 2022. [Online]. Available: https://www.nxp.com/applications/automotive/adas-and-highly-automated-driving/automotive-radar-systems:RADAR-SYSTEMS
[54] *IMAGEVK-74 | 4D Millimeter Wave Imaging Kit*. Accessed: Oct. 2022. [Online]. Available: https://www.minicircuits.com/WebStore/imagevk_74.html
[55] H. Abedi and B. Zakeri, "Through-the-multilayered wall imaging using passive synthetic aperture radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4181–4191, Jul. 2019, doi: 10.1109/TGRS.2018.2890027.
[56] P. Lang et al., "A comprehensive survey of machine learning applied to radar signal processing," 2020, *arXiv:2009.13702*.
[57] Z. Liu, D. K. C. Ho, X. Xu, and J. Yang, "Moving target indication using deep convolutional neural network," *IEEE Access*, vol. 6, pp. 65651–65660, 2018, doi: 10.1109/ACCESS.2018.2877018.
[58] J. Metcalf, S. D. Blunt, and B. Himed, "A machine learning approach to cognitive radar detection," in *Proc. IEEE Radar Conf. (RadarCon)*, Jun. 2015, pp. 1405–1411, doi: 10.1109/RADAR.2015.7131215.
[59] J. Ni, D. Xiang, Z. Lin, C. Lopez-Martinez, W. Hu, and F. Zhang, "DNN-based PolSAR image classification on noisy labels," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 3697–3713, 2022, doi: 10.1109/JSTARS.2022.3168799.
[60] Y. Ma, Y. Liang, W. Zhang, and S. Yan, "SAR target recognition based on transfer learning and data augmentation with LSGANs," in *Proc. Chin. Autom. Congr. (CAC)*, Nov. 2019, pp. 2334–2337, doi: 10.1109/CAC48633.2019.8996717.
[61] I. Nirmal, A. Khamis, M. Hassan, W. Hu, and X. Zhu, "Deep learning for radio-based human sensing: Recent advances and future directions," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 2, pp. 995–1019, 2nd Quart., 2021.
[62] G. Ciaburro and B. Venkateswaran, *Neural Networks With R: Smart Models Using CNN, RNN, Deep Learning, and Artificial Intelligence Principles*. Birmingham, U.K.: Packt Publishing, 2017.
[63] X. Shi, Z. Chen, H. Wang, D. Y. Yeung, W. K. Wong, and W. C. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Proc. Adv. Neural Inf. Process. Syst.*, Jan. 2015, pp. 802–810, doi: 10.48550/arxiv.1506.04214.
[64] *Top Data Augmentation Techniques: Ultimate Guide for 2022*. Accessed: Aug. 10, 2022. [Online]. Available: https://research.aimultiple.com/data-augmentation-techniques/
[65] *CMC Cloud Design Environment CMC Microsystems*. Accessed: Aug. 29, 2022. [Online]. Available: https://www.cmc.ca/cmccloud/
[66] *An Introduction to t-SNE With Python Example | by Andre Violante | Towards Data Science*. Accessed: Aug. 29, 2022. [Online]. Available: https://towardsdatascience.com/an-introduction-to-t-sne-with-python-example-5a3a293108d1

Hajar Abedi (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from the Babol Noshirvani University of Technology, Babol, Iran, in 2014 and 2017, respectively. She is pursuing the Ph.D. degree with the Systems Design Engineering Department, University of Waterloo, Waterloo, ON, Canada.

Her research interests include antenna design, radar sensors, signal processing, and machine learning.

Martin Ma is currently pursuing the master's degree in computational science and engineering with Harvard University, Cambridge, MA, USA.

He is with the Institute of Applied Computational Science. His research interests include machine learning and stochastic optimization.

James He received the bachelor's degree in mechatronics engineering from the University of Waterloo, Waterloo, ON, Canada, in 2022.

He is an AI and DSP Systems Engineer at Qualcomm. His work at Qualcomm focuses on designing hardware architectures and software algorithms for performant and power-efficient edge AI computation.

Jennifer Yu is currently pursuing the master's degree in computer science with the University of Toronto, Toronto, ON, Canada, under the supervision of Dr. Anna Goldenberg.

She is a Machine Learning Researcher at the Vector Institute and the Hospital for Sick Children, Toronto. Her research interests widely span time series, computer vision, machine learning, and applications to understand and improve health. She aims to build responsible and interpretable machine learning systems to solve healthcare-related problems.

Ahmad Ansariyan received the B.Sc. degree in computer engineering and the M.Sc. degree in information technology from the K. N. Toosi University of Technology, Tehran, Iran, in 2008 and 2014, respectively.

He is a full-stack Software Developer with over ten years of experience for developing several programming languages. He advanced in C, C++, C#, Java, Python, and MATLAB for desktop platforms. Also, he is proficient in HTML, CSS, and JavaScript for web-based and Java for Android applications, in addition to SQL and NoSQL programming for database management. He has great experience with the IoT, cloud computing, and raspberry pi. He is currently working at the University of Waterloo, Waterloo, ON, Canada, as a Graduate Research Associate.

George Shaker (Senior Member, IEEE) is the Laboratory Director of the Wireless Sensors and Devices Laboratory, University of Waterloo, Waterloo, ON, Canada. He has authored/coauthored more than 150 publications and more than 30 patents/patent applications.

Prof. Shaker has received multiple recognitions and awards, including the IEEE Antennas and Propagation Society (AP-S) Best Paper Award (the IEEE AP-S Honorable Mention Best Paper Award (four times to date) and the IEEE Antennas and Propagation Graduate Research Award (the IEEE Microwave Theory and Technology Society (MTT-S) Graduate Fellowship, the Electronic Components and Technology Best of Session Paper Award, and the IEEE Sensors Most Popular Paper Award). He has coauthored two papers in IEEE journals that were among the top 25 downloaded papers on IEEEXplore for several consecutive months. He was the Supervisor of the student team winning the Third Best Design Contest at IEEE AP-S 2016, the coauthor of the Association for Computing Machinery (ACM) MobileHCI 2017 Best Workshop Paper Award, and the 2018 Computer Vision Conference Imaging Best Paper Award. He co-received with his students' several research recognitions, including the Natural Sciences and Engineering Research Council of Canada (NSERC) Top Science Research Award in 2019, the IEEE APS HM Paper Award in 2019, the Biotec Top Demo Award in 2019, the arXiv Top Downloaded Paper (medical device category) in 2019, the Velocity Fund in 2020, the National Aeronautics and Space Administration (NASA) Tech Briefs HM Award (medical device category) in 2020, the University of Waterloo (UW) Concept Winner in 2021, the U.K. Dragons Canadian Competition Winner in 2021, the Canadian Microelectronics Corporation (CMC) Nano Winner in 2021, the Google Research Award in 2021, and the COIL CoLab Award in 2022.