On the Use of Machine Learning and Deep Learning for Radar- Based Passenger Monitoring

Hajar Abedi, Martin Ma, Jennifer Yu, James He, Ahmad Ansariyan, and George Shaker
Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, Canada
{habedifi, z74ma, j336yu, james.he, ahmadansariyan, gshaker}@uwaterloo.ca

***Abstract*—In the past decade, there has been an increasing demand for in-vehicle safety sensors. In this paper, we use a multi-input multi-output (MIMO) frequency modulated continuous wave (FMCW) radar for in-vehicle passenger detection and occupant type classification. We propose a Convolutional Long Short-Term Memory (ConvLSTM) that requires no handcrafted rules and accurately detects passengers and classifies occupant type (empty/adults/children). Our model shows high precision (0.90) and recall (0.95) when used to detect unattended children in the vehicles.**

***Keywords*—radar imaging, in-car passenger detection, occupant type classification, deep learning, temporal convolutional network**

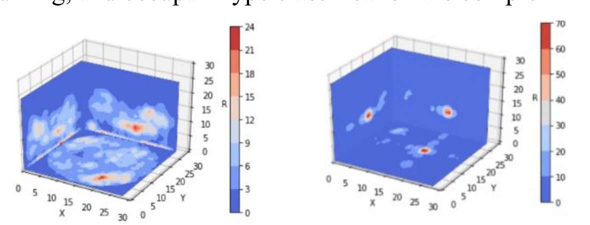
I. Introduction

Radar systems are being increasingly integrated into automotive safety technologies. Typical applications tend to be applied to control passenger-side airbags, safety belts, warning devices and commuter transport logistics. With the integration of machine learning technology, radar sensors can tackle crucial disadvantages of current state-of-the-art monitoring devices, such as the dependency on illumination level, leaking privacy of camera-based sensors and high false alarm rates of mechanical sensors [1]. In our previous work [2], we proposed a novel presence or absence detection algorithm to detect an unattended child or pet inside a car. The main purpose was to propose a fast and easy-to-implement solution to prevent hot car death. However, the number of passengers, their occupied seats as well as the occupancy type were not provided. In [3], we showed that for multiple passenger monitoring, particularly in the third row where they sit at a zero distance from each other (shoulder to shoulder seating), we need a radar with a higher angular resolution. Since the angular resolution increases by increasing the number of transmitters and receivers, implementing a radar system with a large number of transmitters and receivers would result in a higher system cost and more operational complexity. Not to mention that even having a high-resolution radar might not result in an inaccurate in-vehicle occupancy monitoring system. A part of the reason is that reflections from a subject are like a Sinc function [4], and the sidelobe of reflected signals from one passenger leaks to reflections bouncing off from the adjacent passenger. Hence, to distinguish between occupants with a zero distance and count the number of occupants, we need to isolate signals coming from each passenger, remove the leakage from them, and then apply the conventional methods used to identify occupied seats inside a car. However, this process requires information about passengers, such as height and width, which varies from one passenger to another. This would require tuning several parameters and applying more sophisticated methods. To overcome the limitation of the low-resolution radar, we applied machine learning algorithms to the data to detect the occupied seats [4]. To the best of our knowledge, no previous research has investigated the possibility of identifying the type of occupant inside a car, which is critical to safety sensors. In this work, we developed a deep learning model coupled with radar signal processing to accurately detect passengers, count them and classify the type of occupant using a multi-input multi-output (MIMO) frequency modulated continuous wave (FMCW) radar.

II. Method

A. Radar Signal Preprocessing

Many works have demonstrated using radars for in-cabin sensing [2], [3]. Here, we utilize an mm-Wave radar to perform the sensing functionalities. Many of those are available commercially off the shelf as the ones from Analog Devices, Infineon, Texas Instruments, NXP, and Vayyar. Due to the geometry of the radar sensor antennas, we can generate a 3D representation by measuring the strength of the reflected signal from an in-car environment. The intensity of each pixel denotes the reflected energy received at that point (raw radar image). Fig. 1 (a) shows a raw radar image when a passenger is on the rear middle seat. To reduce the effect of noisy signals from reflection, we apply derivatives on the raw radar image over 10 consecutive frames. As shown in Fig. 1 (b), the processed radar image is cleaner and more interpretable. However, even with the processed images, passenger counting, and occupant type classification are complex



(a) (b)

Fig. 1. 3D visualization of (a) raw radar image (b) processed radar image (in this case, rear middle seat is occupied). Colour bar represents signal intensity.

problems 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 does not generalize well to different car models. Therefore, there is a pressing need for an algorithm that not only can detect passengers and count the number of occupants but also identify the type of occupants. As such, we propose a deep learning-based model.

B. *Data Collection*

To construct the dataset to train and test the deep learning model, we collected 780 recording sessions across several days from different passengers. Each recording session is 30 seconds, and radar images are captured at 0.1-second intervals. Unlike our previous works [1], [3] where we used only one car, in this paper, we used four different cars. The dataset is split into train/validation/test datasets in a 70/15/15 ratio. Each dataset is session-independent which means that any recording session will only be used for only tests or training. In other words, we can not use a part of the data of one scenario for the train set and the remaining for the test. This method is important to prevent overfitting and to evaluate the network performance to predict new scenarios reliably. It should be pointed out that our approach in this paper is in contrast to many reported works. For example, in our previous works [1], [3], we combined all collected samples, shuffled them and then divided them into train, validation and tests to report accuracy. Although, this approach could provide very high accuracy, the evaluation method is not trustable for a completely new scenario. A part of the reason is that the radar frame rate is so high, and thus mixing all samples and grabbing some of them for test sets would not ensure that our samples are completely unseen. Generally, in the case of radar-based data sets for machine learning/ deep learning models, the evaluation method plays a crucial part. We believe that in order to report a model performance reliably, the appropriate evaluation method is to use a separate sample of recorded data for the test set.

C. *ConvLSTM Architecture*

The model used in this work is based on ConvLSTM, a combination of Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) and Deep Neural Network (DNN) [4]. The model was originally developed for precipitation nowcasting. Our architecture uses CNN to extract the 3D spatial features from the radar 3-D image, then passes them to LSTM to perform temporal modelling, and finally outputs them to DNN to produce separable feature representations.

III. Results

In this section, the results of the model for 2 different tasks (i.e., passenger counting and occupant type classification) are discussed. Data augmentation methods, such as random translation and rotation, are applied in both tasks leading to a significant increase in accuracy. A prediction is considered correct if and only if all 5 seats are correctly classified.

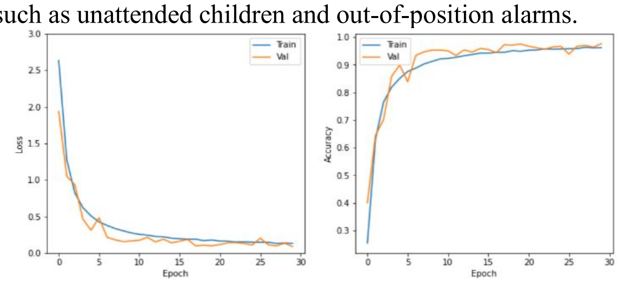
A. *Passenger Counting*

In this task, each seat is represented as a binary encoding (i.e., 1 if the seat is occupied; 0 otherwise). There are 32 scenarios (i.e., 2^5) depending on the presence of people in each seat. As shown in Fig. 2, after training for 25 epochs, the proposed ConvLSTM achieves 89% classification accuracy on the test set when cross-entropy is used as the loss function. This finding implies that our model learns the underlying mathematical correlation between radar signal and passenger arrangement in the car.

B. *Occupant Type Classification*

Occupant type classification is a more challenging task because the model also needs to distinguish between adults and children. The confusion matrix is constructed in Fig. 4 to summarize the performance of the model on 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 the vehicles.

As shown, there is a prodigious role for machine learning to play in in-vehicle sensors, not only as an essential component of passenger detection but as safety precautions such as unattended children and out-of-position alarms.



(a) (b)

Fig. 2. Performance of the proposed CONVLSTM model for passenger counting (a) loss function and (b) accuracy.

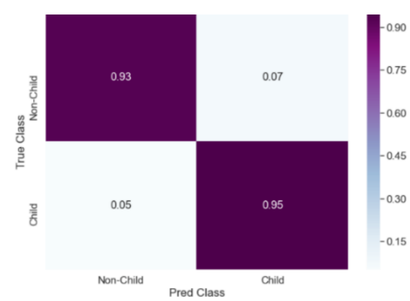


Fig. 3. Confusion matrix of occupant types.

References

[1] M. Alizadeh, H. Abedi and G. Shaker, "Low-Cost Low-Power In-Vehicle Occupant Detection with mm-Wave FMCW Radar,," *IEEE Sensor Conference,* pp. 1-4, 2019.

[2] H. Abedi, C. Magnier, V. Mazumdar and G. Shaker, “Improving Passenger Safety in Cars Using Novel Radar Signal Processing,” Engineering Reports, 2021. Doi: https://doi.org/10.1002/eng2.12413.

[3] H. Abedi, S. Luo, V. Mazumdar, M. Riad and G. Shaker, “AI-Powered In-Vehicle Passenger Monitoring using Low-Cost mm-Wave Radar,” IEEE Access, pp. 1–1, 2021.

[4] X. Shi, et al, "Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting," *Advances in Neural Information Processing Systems*, 2015.