Pedestrian detection using Doppler radar and LSTM neural network

Mussyazwann Azizi Mustafa Azizi[1], Mohammad Nazrin Mohd Noh[2], Idnin Pasya[3], Ahmad Ihsan Mohd Yassin[4], Megat Syahirul Amin Megat Ali[5]
[1,2,4,5]Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam, Malaysia
[3]Microwave Research Institute, Universiti Teknologi MARA, Shah Alam, Malaysia

Article Info

ABSTRACT

Integration of radar systems as primary sensor with deep learning algorithms in driver assist systems is still limited. Its implementation would greatly help in continuous monitoring of visual blind spots from incoming pedestrians. Hence, this study proposes a single-input single-output based Doppler radar and long short-term memory (LSTM) neural network for pedestrian detection. The radar is placed in monostatic configuration at an angle of 45 degree from line of sight. Continuous wave with frequency of 1.9 GHz are continuously transmitted from the antenna. The returning signal from the approaching subjects is characterized by the branching peaks higher than the transmitted frequency. A total of 1108 spectrum traces with Doppler shifts characteristics is acquired from eight volunteers. Another 1108 spectrum traces without Doppler shifts are used for control purposes. The traces are then fed to LSTM neural network for training, validation and testing. Generally, the proposed method was able to detect pedestrian with 88.9% accuracy for training and 87.3% accuracy for testing.

Corresponding Author:

Megat Syahirul Amin Megat Ali,
Microwave Research Institute,
Universiti Teknologi MARA,
40450 Shah Alam, Malaysia.
Email: megatsyahirul@uitm.edu.my

1. INTRODUCTION

Pedestrian detection is among the most vital safety element in an increasingly complex driver assist systems [1-2]. These ensure that the vehicle would detect pedestrians in blind spots and perform evasive maneuvres when required; whether through warning systems [3] or emergency braking mechanisms [4]. Thus far, numerous sensing approaches have been tested which include computer vision [5], laser scanner [6] and automotive radar [7] technologies. Each of these methods has unique capabilities that enable the vehicle to detect pending collisions with pedestrians [8]. For example, radar sensors enable the utilization of Doppler and micro-Doppler information obtained from body movements to identify and discriminate between signals reflected between from pedestrians and other targets [9-10], making them a suitable candidate for this purpose. Hybrid systems were also proposed to improve pedestrian detection capabilities [11-12], however radar sensors remain an attractive choise for their ability to obtain unique signature from reflected signals.

Embedded in these innovative sensing approaches are intelligent algorithms; designed to automatically detect and perform evasive maneuvres [13]. Thus far, the two-dimensional information acquired from imagery information has also been tested using advanced artificial intelligent models such as

the convolutional neural network (CNN) [14]. Despite implementation of LSTM in CNN architectures, the systems still rely on two-dimensional imagery inputs which result in high computational requirements [8]. Two major issues have been identified. 1) The use of radar as the primary sensing element is limited since it relies only on the reflected time-domain signals from targets [15]. Current technology only adopts it as support to the computer vision and laser scanning systems. 2) With radar as the primary input, an LSTM neural network is most suited as the architecture is capable of extracting common features of approaching pedestrians from the sequential information [16-17]. These however, remain untested.

To solve the aforementioned problems, the following objectives are outlines. 1) The study proposes a relatively simple continuous-wave Doppler radar to characterize between approaching pedestrians and controlled condition. 2) The returning pulses deflected off the subjects will used as input to train, validate and test the LSTM recurrent neural network architecture. This paper is structured as the following. Section 2 describes on the data collecton and intelligent classification method used for the study. Subsequently, Section 3 discusses on spectral trace characteristics and subject detection using LSTM neural network. Finally, Section 4 summarizes contribution of the study and its prospective application for driver assist technology.

2.   METHODOLOGY
2.1. Experimental setup and acquisition protocol
Data collection was performed at the Microwave Research Institute, Universiti Teknologi MARA. The equipments used include Agilent MXG Analog Signal Generator, KeySight FieldFox Microwave Analyzer, as well as transmitter and a receiver antenna. As shown in Figure 1, the radar system is placed in a monostatic configuration. An absorber is positioned between transmitter and receiver to minimize spectral leakage. A 10 dBm continuous wave with frequency of 1.9 GHz from the signal generator is transmitted by a Vivaldi antenna. The returning waves deflected from the approaching subjects are captured by the receiver and the information is converted to spectrum traces by the microwave analyzer. Spectral resolution is set to 500 Hz to allow observable Doppler shift signatures as subjects move closer to the radar setup.
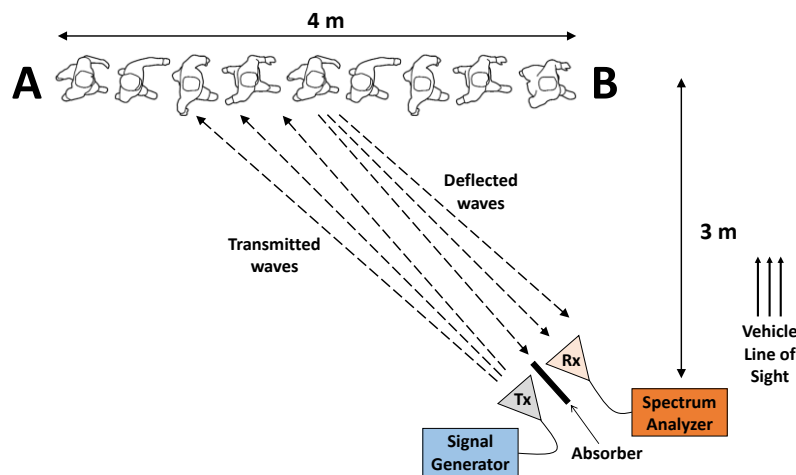

Figure 1. Experimental setup

Eight volunteers have participated in this study. Subjects are required to walk along the specified path from Point A to Point B at moderate pace. The system captures the returning signal of approaching subjects until they pass right in front of the assumed vehicle's line of sight; thus simulating the situation in that would probably result in a collision. Each subject is required to repeat the trials twenty times and every trial will produce between six and eight spectrum traces.

2.2. Pedestrian detection using LSTM neural network
LSTM is an improvement of the recurrent neural network (RNN) used for modelling sequential data. Figure 2 shows the theoretical architecture of RNN with the recurrent layer unfolded into a network [18-19]. U, V and W are hyperparameters of different network layers. x is the input and h is

the hidden state that grants the network memory ability. The different time instances are indicated by $t-1$, t and $t+1$. Through activation function, Γ_1, the output of hidden layer with present information is transferred to the hidden layer of the next time instance as part of the input. The feedback preserves the information of preceding time instance to retain data dependency; thus, improving learning and abstracting from the sequential data [20-21]. The vanishing gradient issues during computation of back-propagation learning however, adversely affect the amount of distant memories to be transferred. Therefore, these restrict the capability of RNN for modelling long-dependency sequential information and not suitable to be implemented this study [22].
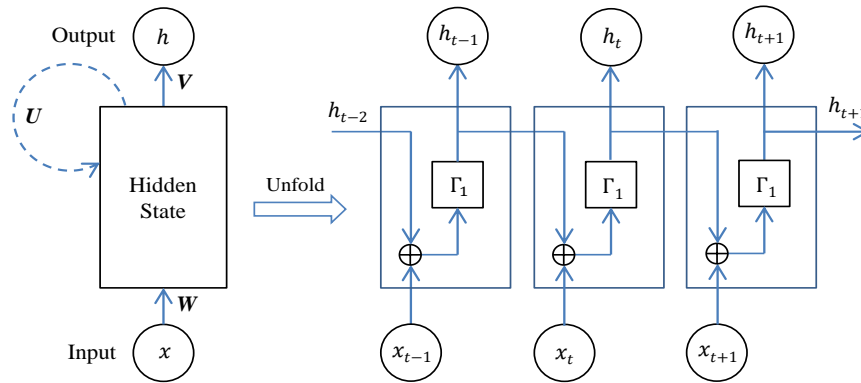

Figure 2. Hidden state of RNN structure [20]

To solve the vanishing gradient issue, LSTM neural network has been proposed. A standard LSTM block shown in Figure 3 is comprised of memory cell state, forget gate, input gate and output gate. The memory state plays a defining role throughout the entire chain in selectively adding or removing relevant information to the cell state through the three-gate system [23].
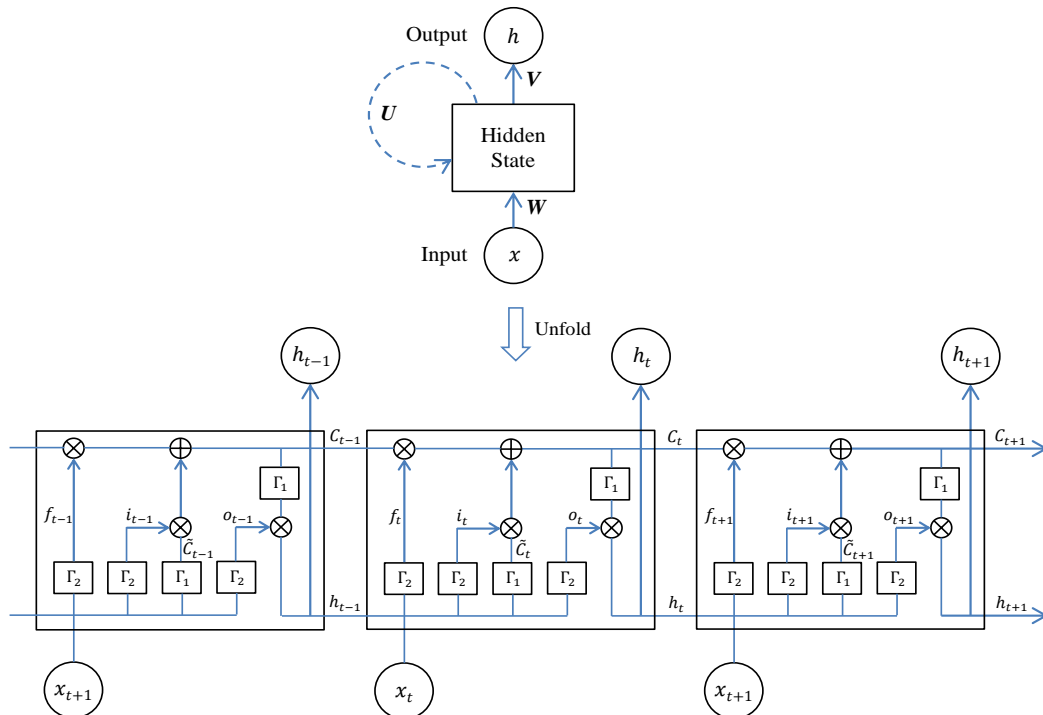

Figure 3. Hidden state of LSTM structure [20]

Initially as shown by (1), cell state, C_t decide on information that should be discarded from previous cell state, C_{t-1} through the forget gate, f_t.

$$f_t = \Gamma_2\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{1}$$

Subsequently as expressed by (2), the input gate, i_t identifies the information from input x_t that should be stored in the cell state, C_t. Input information, i_t and candidate cell, $\tilde{C}_t$ is then updated through (3).

$$i_t = \Gamma_2\left(W_i x_t + U_h h_{t-1} + b_i\right) \tag{2}$$

$$\tilde{C}_t = \Gamma_1\left(W_c x_t + U_c h_{t-1} + b_c\right) \tag{3}$$

Subsequently as shown by (4), the combined candidate memory, $\tilde{C}_t$ and the long-term memory from C_{t-1} is updated for cell state, C_t.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{4}$$

The output at present time instant, h_t is then computed by considering both the output information o_t and cell state, C_t. These are mathematically expressed by (5) and (6).

$$o_t = \Gamma_2\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{5}$$

$$h_t = o_t * \Gamma_2\left(C_t\right) \tag{6}$$

Based on the aforementioned equations, f, i and o each represents the forget gate, input gate, and output gate. W are the input weights, U are the recurrent weights, and b are the biases for the respective gates and cell states. Γ_1 is hyperbolic tangent and Γ_2 sigmoid function. Both activation functions are used to improve non-linearity of the network and can each be expressed by (7) and (8).

$$\Gamma_1(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{7}$$

$$\Gamma_2(x) = \frac{1}{1 + e^{-x}} \tag{8}$$

LSTM architecture that incorporates memory cells and regulated by the gating mechanism provides solution to the vanishing gradient problem of RNN. Thus, the improved network structure is capable to extract historical information and predicts future trend for long-term dependencies of sequential data. In this study, the input to the LSTM neural network is the spectral traces obtained from the spectrum analyzer. The output classes from hidden states are defined as indexes for pedestrian and the controlled condition. 70% of the data is used for training, 15% is used for validation, and the remaining 15% is used for testing [24].

The performance of LSTM recurrent neural network for pedestrian detection is assessed in terms of accuracy (Acc), positive predictivity (Pp), and sensitivity (Se). Acc is described as the ability of the system to correctly differentiate between approaching subjects and control condition. Subsequently, Se is defined as the ability of the system to correctly identify approaching pedestrians. On the other hand, Pp is described as the probability of that following a positive detection, the subject will be within the line of sight of the vehicle. Each of these parameters is expressed by (7-9), where TP is true positive, TN is true negative, FP is false positives and FN is false negative classification [25].

$$Acc = \left(\frac{TP + TN}{TP + TN + FP + FN}\right) \times 100\% \tag{9}$$

$$Se = \left(\frac{TP}{TP + FP}\right) \times 100\% \tag{10}$$

$$Pp = \left(\frac{TP}{TP + FN} \right) \times 100\%$$

(11)

3. RESULTS AND DISCUSSION
3.1. Spectral profiling using doppler shift signature

In theory, the signal deflected off subjects moving farther away from Doppler radar will exhibit longer wavelength than the transmitted signal. These should be reflected in the presence of secondary peak with frequency characteristics lower than 1.9 GHz. In contrast, subjects moving closer towards the Doppler radar will exhibit shorter wavelength than the transmitted signal. These could be be characterized by the presence of secondary peak with frequency higher than 1.9 GHz. Figure 4 shows a sample of the spectrum trace obtained from KeySight FieldFox Microwave Analyzer. Doppler shift can be seen visible at frequency higher than 1.9 GHz. The result is thus valid as the waves deflected conform to the characteristics of an approaching pedestrian.
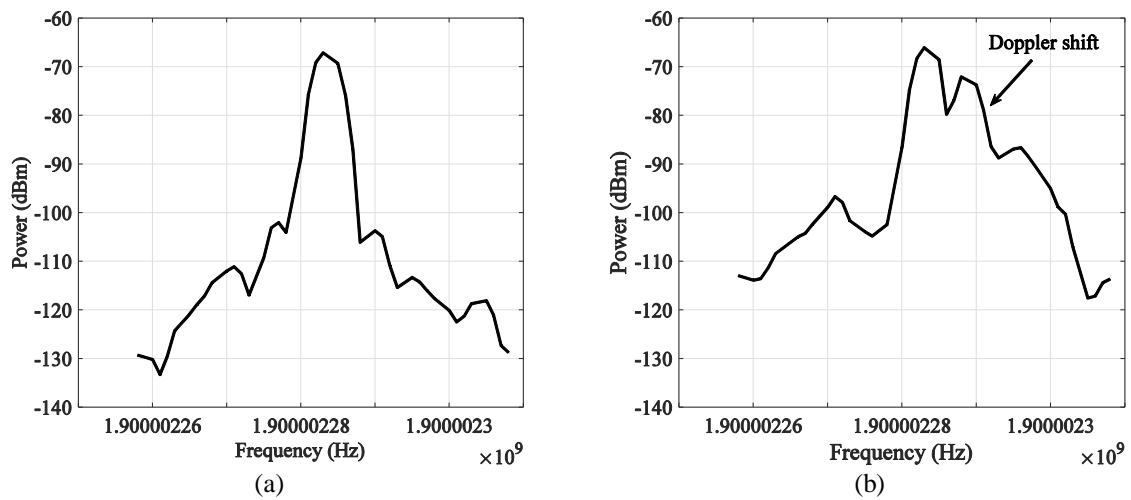

Figure 4. Spectrum traces for (a) controlled condition, and (b) approaching subject

To further confirm the collective pattern of acquired data, results from each sample are combined to form a composite display of spectrum traces. Figure 5 shows the overall spectrum traces for controlled condition. The results show a consistent pattern with a dominant peak at frequency of 1.9 GHz.
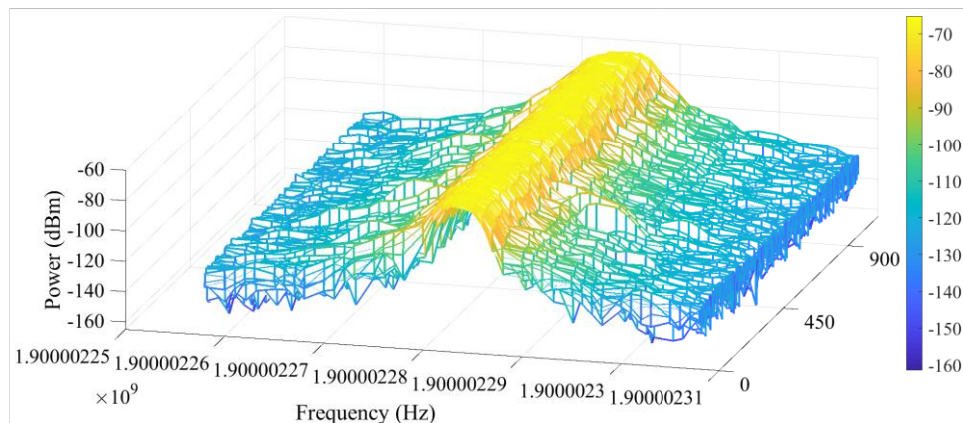

Figure 5. Composite spectrum traces for controlled condition (N = 1108 samples)

The results were also compared with collective pattern of spectrum traces for approaching subjects. As shown in Figure 6, an increase in spectrum activity is detected at frequencies higher than 1.9 GHz. These provide a conclusive proof that the Doppler radar is indeed capturing the correct information from deflected signals.
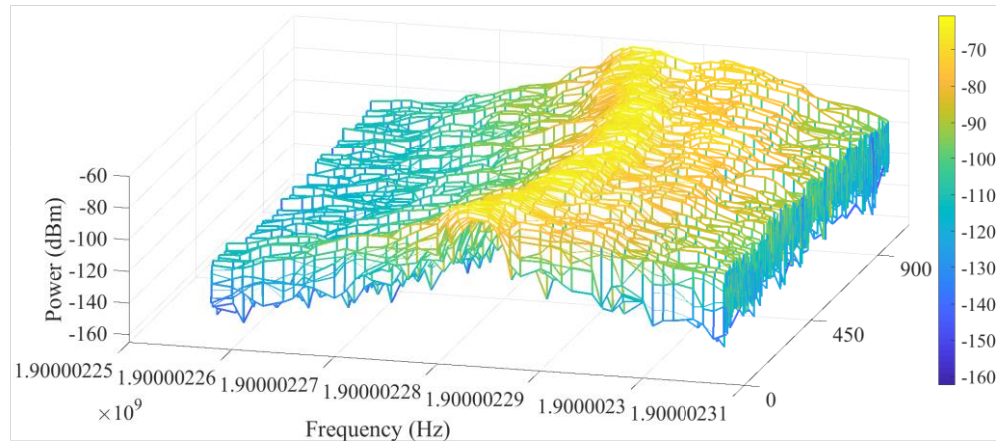

Figure 6. Composite spectrum traces for approaching subjects (N = 1108 samples)

3.2. Pedestrian detection using LSTM recurrent neural network

The spectrum traces which are assumed as sequential information is subsequently fed as input to the LSTM neural network. As shown in Table 1, satisfactory results have been obtained with 88.9% *Acc* for training, 88.9% *Acc* for validation, and 87.3% *Acc* for testing. It is also worth noting that both *Se* and *Pp* measures range between 76.8% to 99.4% when detecting between subject or controlled condition. These indicate that the model is capable of detecting approaching subjects by extracting the Doppler shift information from the respective spectrum traces.

Table 1. Performance of LSTM neural network for pedestrian detection

Parameters		Pedestrian	Control	*Acc*
Training	*Se*	78.4%	99.0%	88.9%
	Pp	98.7%	82.7%	
Validation	*Se*	80.0%	98.2%	88.9%
	Pp	97.8%	82.5%	
Testing	*Se*	99.4%	76.8%	87.3%
	Pp	99.3%	79.0%	

4.   CONCLUSION

The study initially sets out to 1) implement Doppler radar as primary sensing element for detecting approaching pedestrian, and 2) assess the performance of LSTM neural network for extracting sequential information from spectrum traces for distinguishing between incoming subject and controlled condition. Through a relatively simple experiment setup, the study was able to produce satisfactory results. First, the adopted radar system was capable of capturing Doppler shift signatures through the spectrum analyzer. Second, the LSTM neural network has proven capable of extracting the required information for detecting approaching pedestrians. While the overall detection accuracy is satisfactory, there is still opportunity for improvement. Based on the observation of spectrum traces, there are samples in which the Doppler shift is not prominent. Hence, these are presented as outliers that exist within the broad range of samples. Furthermore, the network had to rely on relatively small sample size for capturing relevant information. To overcome these problems, a larger pool of samples is recommended. This is to ensure that the LSTM architecture is capable of extracting the long-term dependency characteristics of the sequential information and successfully generalize the features of incoming pedestrians.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Y. Xu, *et al.*, "Detection of Sudden Pedestrian Crossings for Driving Assistance Systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, pp. 729-739, 2012.

[2]   A. Brunetti, *et al.*, "Computer Vision and Deep Learning Techniques for Pedestrian Detection and Tracking: A Survey," *Neurocomputing*, vol. 300, pp. 17-33, 2018.

[3]   G. De Nicolao, *et al.*, "Onboard Sensor-Based Collision Risk Assessment to Improve Pedestrians' Safety," *IEEE Transactions on Vehicular Technology*, vol. 56, pp. 2405-2413, 2007.

[4]   C. Flores, *et al.*, "A Cooperative Car-Following/Emergency Braking System With Prediction-Based Pedestrian Avoidance Capabilities," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, pp. 1837-1846, 2019.

[5]   N. K. Ragesh and R. Rajesh, "Pedestrian Detection in Automotive Safety: Understanding State-of-the-Art," *IEEE Access*, vol. 7, pp. 47864-47890, 2019.

[6]   B. Kim, *et al.*, "Pedestrian/Vehicle Detection Using a 2.5-D Multi-Layer Laser Scanner," *IEEE Sensors Journal*, vol. 16, pp. 400-408, 2016.

[7]   W. Zhang, *et al.*, "Enhanced Detection of Doppler-Spread Targets for FMCW Radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, pp. 2066-2078, 2019.

[8]   D. Gerónimo, *et al.*, "Survey of Pedestrian Detection for Advanced Driver Assistance Systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 1239-1258, 2010.

[9]   S. Mohd Basir, *et al.*, "Improvement of Doppler Measurement using Multiple-Input Multiple-Output (MIMO) Concept in Radar-based Automotive Sensor Detecting Pedestrians," *Sensor Review*, vol. 38, pp. 239-247, 2018.

[10]   J. V. B. Severino, *et al.*, "Pedestrian Recognition using Micro Doppler Effects of Radar Signals Based on Machine Learning and Multi-Objective Optimization," *Expert Systems with Applications*, vol. 136, pp. 304-315, 2019.

[11]   D. M. Gavrila, "Sensor-Based Pedestrian Protection," *IEEE Intelligent Systems*, vol. 16, pp. 77-81, 2001.

[12]   F. García, *et al.*, "Context Aided Pedestrian Detection for Danger Estimation Based on Laser Scanner and Computer Vision," *Expert Systems with Applications*, vol. 41, pp. 6646-6661, 2014.

[13]   J. Baek, *et al.*, "Fast and Efficient Pedestrian Detection via the Cascade Implementation of an Additive Kernel Support Vector Machine," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, pp. 902-916, 2017.

[14]   Y. Zhao, *et al.*, "Accurate Pedestrian Detection by Human Pose Regression," *IEEE Transactions on Image Processing*, vol. 29, pp. 1591-1605, 2020.

[15]   V. Lekic and Z. Babic, "Automotive Radar and Camera Fusion using Generative Adversarial Networks," *Computer Vision and Image Understanding*, vol. 184, pp. 1-8, 2019.

[16]   X. Tang, "Large-Scale Computing Systems Workload Prediction Using Parallel Improved LSTM Neural Network," *IEEE Access*, vol. 7, pp. 40525-40533, 2019.

[17]   M. Rhanoui, *et al.*, "Forecasting Financial Budget Time Series: ARIMA Random Walk vs LSTM Neural Network," *IAES International Journal of Artificial Intelligence*, vol. 8, pp. 317-327, 2019.

[18]   S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, pp. 1735-1780, 1997.

[19]   T. Bhandarkar, *et al.*, "Earthquake Trend Prediction using Long Short-Term Memory RNN," *International Journal of Electrical and Computer Engineering*, vol. 9, pp. 1304-1312, 2019.

[20]   X. Song, *et al.*, "Time-Series Well Performance Prediction Based on Long Short-Term Memory (LSTM) Neural Network Model," *Journal of Petroleum Science and Engineering*, vol. 186, pp. 1-11, 2020.

[21]   C. G. Pachon-Suescun, *et al.*, "Abnormal Gait Detection by Means of LSTM," *International Journal of Electrical and Computer Engineering*, vol. 10, pp. 1495-1506, 2020.

[22]   Y. Bengio, et al., "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, vol. 5, pp. 157-166, 1994.

[23]   A. Sherstinsky, "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network," *Physica D: Nonlinear Phenomena*, vol. 404, pp. 1-28, 2020.

[24]   A. H. Jahidin, *et al.*, "Classification of Intelligence Quotient using EEG Sub-band Power Ratio and ANN During Mental Task," in *2013 IEEE Conference on Systems, Process and Control*, 2013, pp. 204-208.

[25]   A. I. Amiruddin, *et al.*, "Feature Reduction and Arrhythmia Classsification via Hybrid Multilayered Perceptron Network," in *2013 IEEE 3rd International Conference on System Engineering and Technology*, pp. 290-294, 2013.

BIOGRAPHIES OF AUTHORS

Mussyazwann Azizi Mustaffa Azizi is currently completing his final year studies in B. Eng (Electronics) at the Faculty of Electrical Engineering, Universiti Teknologi MARA, Malaysia. His research interests are in intelligent radar systems and its applications.

Mohammad Nazrin Mohd Noh is currently completing his final year studies in B. Eng (Electrical) at the Faculty of Electrical Engineering, Universiti Teknologi MARA, Malaysia. His research interests are in intelligent radar systems and its applications.

Dr. Idnin Pasya Ibrahim received his B.Eng. and M.Eng. in Information and Communication Engineering, as well as Ph.D. in Information, Communication and Media Design Engineering from Tokyo Denki University, Japan. He is a senior lecturer at the Faculty of Electrical Engineering, Universiti Teknologi MARA, Malaysia. Prior to joining the academia, Dr. Idnin has worked as an engineer at Toshiba PC & Network, Tokyo, Japan. His research interests are in wideband communication systems, MIMO radar and its applications. He is currently the Deputy Director of Microwave Research Institute, Universiti Teknologi MARA.

Ir. Dr. Ahmad Ihsan Mohd Yassin received his B.Sc. in Electrical Engineering from Universiti Tun Hussein Onn, Malaysia, M.Sc. as well as Ph.D. in Electrical Engineering from Universiti Teknologi MARA, Malaysia. Dr. Ihsan is currently a senior lecturer at the Faculty of Electrical Engineering, Universiti Teknologi MARA. His research interests are in deep learning, block-chain technology, optimization, and system identification. He is also a Professional Engineer; registered with the Board of Engineers Malaysia.

Dr. Megat Syahirul Amin Megat Ali received his B.Eng. (Biomedical) from University of Malaya, Malaysia, M.Sc. in Biomedical Engineering from University of Surrey, United Kingdom, and Ph.D. in Electrical Engineering from Universiti Teknologi MARA, Malaysia. He is currently a senior lecturer at the Faculty of Electrical Engineering, Universiti Teknologi MARA. His research interests include biomedical signal processing and artificial intelligence. Dr. Megat is also a research fellow at the Microwave Research Institute, Universiti Teknologi MARA.