IET Radar, Sonar & Navigation

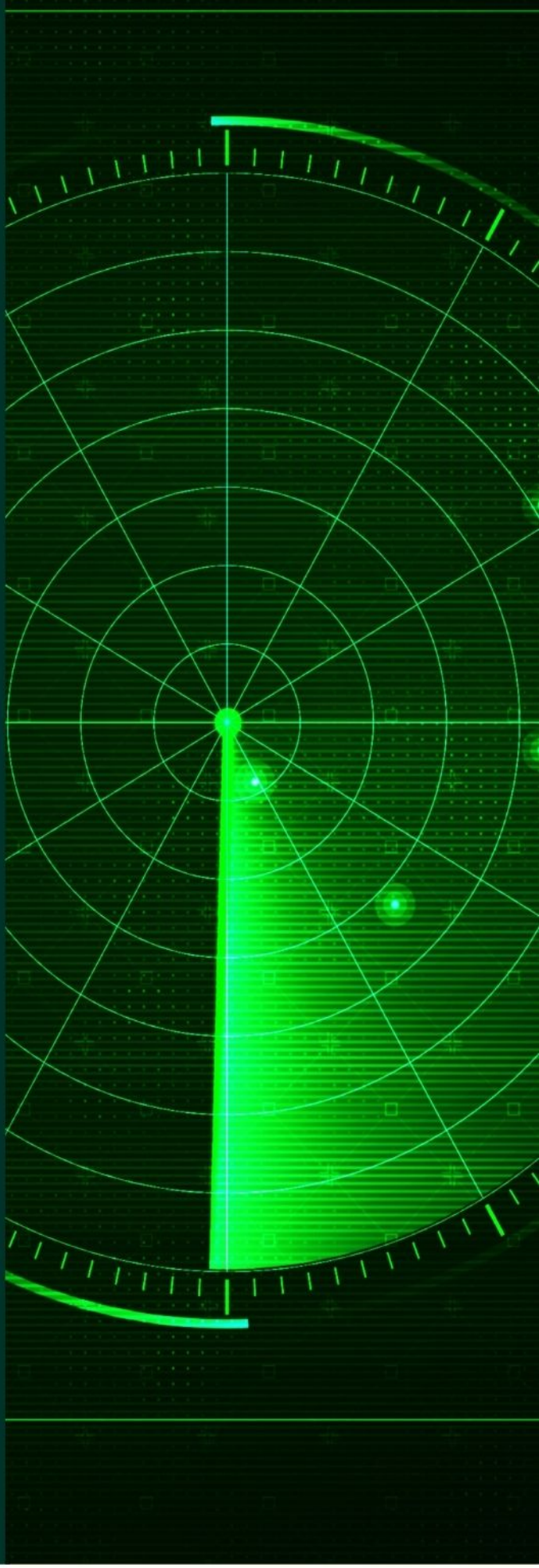
Special Issue Call for Papers

Be Seen. Be Cited.
Submit your work to a new IET special issue

Connect with researchers and experts in your field and share knowledge.

Be part of the latest research trends, faster.

Read more



The Institution of Engineering and Technology

IET Radar, Sonar & Navigation

Research Article

IET Journals
The Institution of Engineering and Technology

Road structure classification through artificial neural network for automotive radar systems

ISSN 1751-8784
Received on 7th December 2018
Revised 1st February 2019
Accepted on 20th February 2019
E-First on 28th March 2019
doi: 10.1049/iet-rsn.2018.5610
www.ietdl.org

Heonkyo Sim[1], Seongwook Lee[2], Byeong-ho Lee[1], Seong-Cheol Kim[1]* ✉

[1]*Department of Electrical and Computer Engineering and Institute of New Media & Communications (INMC), Seoul National University (SNU), 1, Gwanak-ro, Gwanak-gu, Seoul, Republic of Korea*

[2]*Machine Learning Lab, AI & SW Research Center, Samsung Advanced Institute of Technology (SAIT), 130, Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do, Republic of Korea*

**This research was conducted when the author was at SNU*

✉ *E-mail: sckim@maxwell.snu.ac.kr*

Abstract: This study proposes an artificial neural network-based method to classify road structures for automotive radar systems. Generally, road structures generate unwanted echoes called clutter, which degrades the performance of target detection, and may cause critical errors in autonomous driving modes. However, recognition and classification of each individual structure type is very difficult, because there are various types of structures made by different materials which cause ambiguity in the recognition and classification process. To deal with these classification ambiguities, the authors propose an artificial neural network-based recognition and classification method. Road structures are classified by applying artificial neural network to the frequency domain-received signals of automotive radar systems. Once the neural network has been trained with the received signals, it is used to determine the type of road structure with instantaneous received signals. The classification performance was evaluated by experimenting with the number of antenna elements and radar snapshots. In addition, to find the most suitable artificial neural network structure, the authors experimented with the number of nodes and layers. After completing a period of deep learning using actual experimental data, the total classification accuracy was about 95%.

1 Introduction

With the recent growing interests in autonomous vehicles, the importance of various sensors has significantly increased in order to ensure the safety of cars, drivers, and the pedestrians [1]. As currently implemented, autonomous vehicles incorporate cameras, lidars, radars, and ultrasonic sensors. Among these, cameras require light sources to detect a target, and therefore have poor detection performance in shaded areas or at night; however, the detection performance of radar sensors is insensitive to light sources [2]. In addition, radar sensors can detect a target with a relatively long distance compared to ultrasonic sensors [3], and the lidar sensors require higher production cost than the radar sensors. For these reasons, many groups are researching about the automotive radar and its applications [4–7].

The detection performance of radar is affected by the road environments [8–12]. The target detection performance of radar is significantly low in the presence of road structures, such as iron tunnels, sound barriers [13–15]. In particular, if the road structure consists of metal, the intensities of the received signals reflected from the road structure are often stronger than those from the vehicle target, which means that the target vehicle cannot be detected appropriately. Since target detection failure can result in fatal accidents, it is important that the radar systems can detect targets even in harsh road environments to ensure safe autonomous driving. To detect targets in such environments, it is required to develop reliable methods for recognising the environment.

Some studies have proposed methods to detect road structures using lidar systems or camera systems [16–20]; however, the techniques used in these papers are not applicable for radar systems. In [13, 15], the recognition of road structures using radar have been introduced. In [13], a method to recognise iron tunnels using Shannon entropy was investigated. In [15], recognition of guard rails, sound barriers, and iron tunnels was discussed. However, they only recognised metal structures and does not classify the road structures. In [21, 22], a machine learning algorithm was used to classify road structures. They extracted several features and used these features to train a support vector machine (SVM), which is a simple type of machine learning algorithm. However, to use the SVM, it is necessary to extract the representative features well, and the classification performance changes greatly depending on the selected features.

Therefore, we propose a method to classify the types of road structures by using deep learning which can extract unknown features by itself. Road structures are classified based on the fact that the statistical distribution of the received signal strength and fast Fourier transform (FFT) magnitude response differ depending on the road structures. Thus, road structures are classified using FFT magnitude response as the input of artificial neural network (ANN). As the output of ANN, sound barriers, iron tunnels, typical underground tunnels, and open roads are set up. We measured received radar signals for the road structure along the Yongin–Seoul Expressway. The performance of the road structure classification system was evaluated by changing the number of receive antenna elements, the number of snapshots used as input data, and the ANN structure. The training, validation, and test sets comprised 70, 15, and 15% of the used data, respectively. With this approach, it was possible to recognise road structures in real time. Compared to [21] in which road structures are classified by direct extraction of features, the proposed method demonstrated that better classification performance can be obtained by using ANN without extracting features. As a result, once the type of structure has been recognised, it is expected that adaptive target detection algorithms can be used for the according road environment.

The remainder of this paper is organised as follows. A system model of the frequency-modulated continuous wave (FMCW) radar and measurement procedures along the specific road structures are introduced in Section 2. In Section 3, a method of classifying road structures using ANN is presented. Then, the classification results from the ANN method are analyzed in Section 4. Finally, the paper is concluded in Section 5.

2 System model and target detection in automotive FMCW radar

2.1 System model of FMCW radar

Assume a single-input and multiple-output FMCW radar system composed of one transmitting antenna and N identical receiving antenna elements with uniform spacing d. Then, the transmitted signal in the FMCW radar system can be expressed as follows:

$$S(t) = A_{\mathrm{T}}\cos\left(2\pi\left(\left(f_c - \frac{\mathrm{BW}}{2}\right)t + \frac{\mathrm{BW}}{2\Delta T}t^2\right)\right), \tag{1}$$

where A_{T} is the amplitude of the transmitted signal, f_{c} is the center frequency of the transmitted signal, BW is the sweep bandwidth of the transmitted signal, and ΔT is the sweep time of the transmitted signal. The time-domain received signal reflected from multiple targets, $R(t)$, can be expressed as

$$R(t) = \sum_{l=1}^{L} A_{R_l}\cos\left(2\pi\left(\left(f_{\mathrm{c}} - \frac{\mathrm{BW}}{2} + f_{d_l}\right)(t - t_{d_l}) + \frac{\mathrm{BW}}{2\Delta T}(t - t_{d_l})^2\right)\right), \tag{2}$$

where A_{R_l} is the amplitude of the received signal from the lth target, f_{d_l} is the Doppler frequency reflected at the lth target, t_{d_l} is the time delay corresponding to the lth target, and L is the number of targets. The signal received by the antenna elements is demodulated by a mixer and a low-pass filter. The demodulated signal can then be expressed as

$$M(t) = \sum_{l=1}^{L} A_{M_l}\cos\left(2\pi\left(\frac{\mathrm{BW}}{\Delta T}t_{d_l} - f_{d_l}\right)t + 2\pi\left(f_c - \frac{\mathrm{BW}}{2} + f_{d_l}\right)t_{d_l} - \frac{\pi\mathrm{BW}}{\Delta T}t_{d_l}^2\right), \tag{3}$$

where A_{M_l} is the amplitude of the demodulated signal from the lth target. By applying the Fourier transform to the demodulated signal, we obtain a frequency called the beat frequency, from which we can obtain the target information, such as range and velocity. The beat frequency can be expressed as

$$f_{b_l} = \frac{\mathrm{BW}}{\Delta T}t_{d_l} - f_{d_l}. \tag{4}$$

This beat frequency f_{b_l} can be extracted by applying a peak detection algorithm, such as the constant false alarm rate (CFAR) [23]. The sampled time-domain demodulated signal can be expressed as follows:

$$\boldsymbol{x}_i = [x_i(1), x_i(2), \ldots, x_i(N_s)]^T, \quad (i = 1, 2, \ldots, N_{\mathrm{scan}}), \tag{5}$$

where N_s denotes the number of time samples, i is the scan index, and N_{scan} is the total number of scans. The FFT result of $\boldsymbol{x}_i$ is expressed as

$$X_i(k) = \sum_{n=1}^{N_s} x_i(n)\mathrm{e}^{-\mathrm{j}(2\pi/N_{\mathrm{FFT}})(k-1)(n-1)}, \quad (k = 1, 2, \ldots, N_{\mathrm{FFT}}), \tag{6}$$

where N_{FFT} denotes the number of points in FFT. Since $X_i(k)$ is a complex number, the magnitude response is obtained by taking the absolute value for each element, as follows:

$$\boldsymbol{C}_i = [|X_i(1)|, |X_i(2)|, \ldots, |X_i(N_{\mathrm{FFT}})|]^T. \tag{7}$$

2.2 Sensor parameter description

We measured radar signals about the road structures along the Yongin–Seoul Expressway. To judge whether it is applicable directly in an actual road environment, the experiment was conducted in various environments while driving on an actual road rather than an ideal environment. Furthermore, the experiment was performed under various conditions while changing the ego velocity continuously and with varying distances from the target. The measurements were conducted using a long-range radar manufactured by Mando Corporation. This system has a field of view (FoV) from $-10°$ to $10°$ and the maximum detection range of 250 m. The number of transmit and receiving antenna elements were 1 and 4, respectively, the antenna spacing between each receiving antenna element was 1.8λ where λ is the wavelength of the transmitted signal, and FFT length is 2048. In operation, the system transmitted a 76.5 GHz FMCW radar signal with a bandwidth of 500 MHz and a 10 ms sweep time for the up- and down-chirping. One scan takes 100 ms, consisting of up- and down-chirping times of 10 ms and signal processing times of 90 ms.

2.3 Target detection in various road structures

In this section, we discuss the FFT results for different road environments. General elements in driving environments, such as sound barriers, iron tunnels, typical underground tunnels, and open roads, are shown in Fig. 1. A total of 600 scans (i.e. from $\boldsymbol{C}_1$ to $\boldsymbol{C}_{600}$) acquired while driving for 1 min in four different road environments are shown in Fig. 2. The x axis can be converted to time. If the relative velocity between the target and the radar-equipped vehicle is zero, y axis can be converted to distance. Sound barriers appear from about the 300th scan to the 600th scan in Fig. 2*a*, and the horizontally curved lines indicate other vehicles driving in the FoV of our radar. If a target vehicle is faster than the radar-equipped vehicle, the up-slope curve extends from the lower left to the upper right. If the target vehicle is slower than the radar-equipped vehicle, the down-slope curve extends from the upper left to the lower right. Since there are a lot of targets, only a few targets that are noticeable in Fig. 2 are marked. In case of stationary road structures, curves appear as lines having tangential slopes proportional to the velocity of the radar-equipped vehicle. If the radar-equipped vehicle is moving at a constant velocity, lines are straight. In Fig. 2*a*, straight lines appear in the frequency domain when the vehicle enters the section with sound barriers. In Fig. 2*b*, the signals reflected by the metal structures in the iron tunnel are dominant. Thus, the target vehicles could not be accurately detected beyond a certain critical distance. While iron tunnels possess similar characteristics to the sound barrier because they are both made of iron, the received signals reflected from iron tunnels are stronger than those reflected from the sound barriers. In Fig. 2*c*, the 200th to the 600th scan is the inside of a typical underground tunnel, and several unwanted reflections occur due to the tunnel structure. However, since the typical underground tunnel is not made of iron, the intensities of the reflected signals are relatively weak; thus, the preceding vehicle is detected fairly well. In Fig. 2*d*, the effect of the median strip appears periodically, but this is much less likely to act as clutter than the other road structures. The targets are well detected on the open road, unlike environments having metal road structures. As shown in Fig. 2, when 600 scans are obtained, the characteristics of the road structures are distinct enough to tell the difference between them.

Instead, the instantaneous FFT magnitude response is also enough to tell the difference. Instantaneous FFT magnitude responses of a single scan for each road environment are shown in Fig. 3. The red circles in the figure represent the signals that are reflected from the target vehicles, while the red lines indicate the threshold lines using order statistics (OS)-CFAR [24], one of the CFAR algorithms. If an FFT magnitude is greater than the CFAR threshold, the point is recognised as a meaningful target. Thus, in an ideal situation, all of the red circles should be above the red line. Figs. 3*a* and *b* show multiple reflected signals from the road structures, which cause the desired targets to not be properly detected. In addition, the intensities of the reflected signals in the

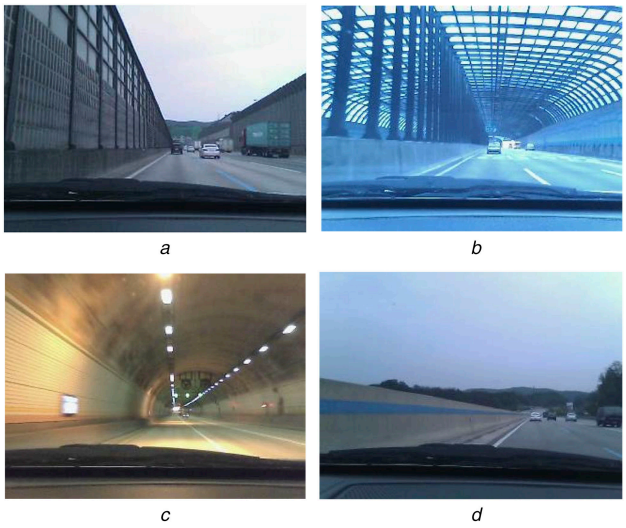


Fig. 1 *Snapshots of*
(a) Sound barrier, ***(b)*** Iron tunnel, ***(c)*** Typical underground tunnel, and ***(d)*** Open road

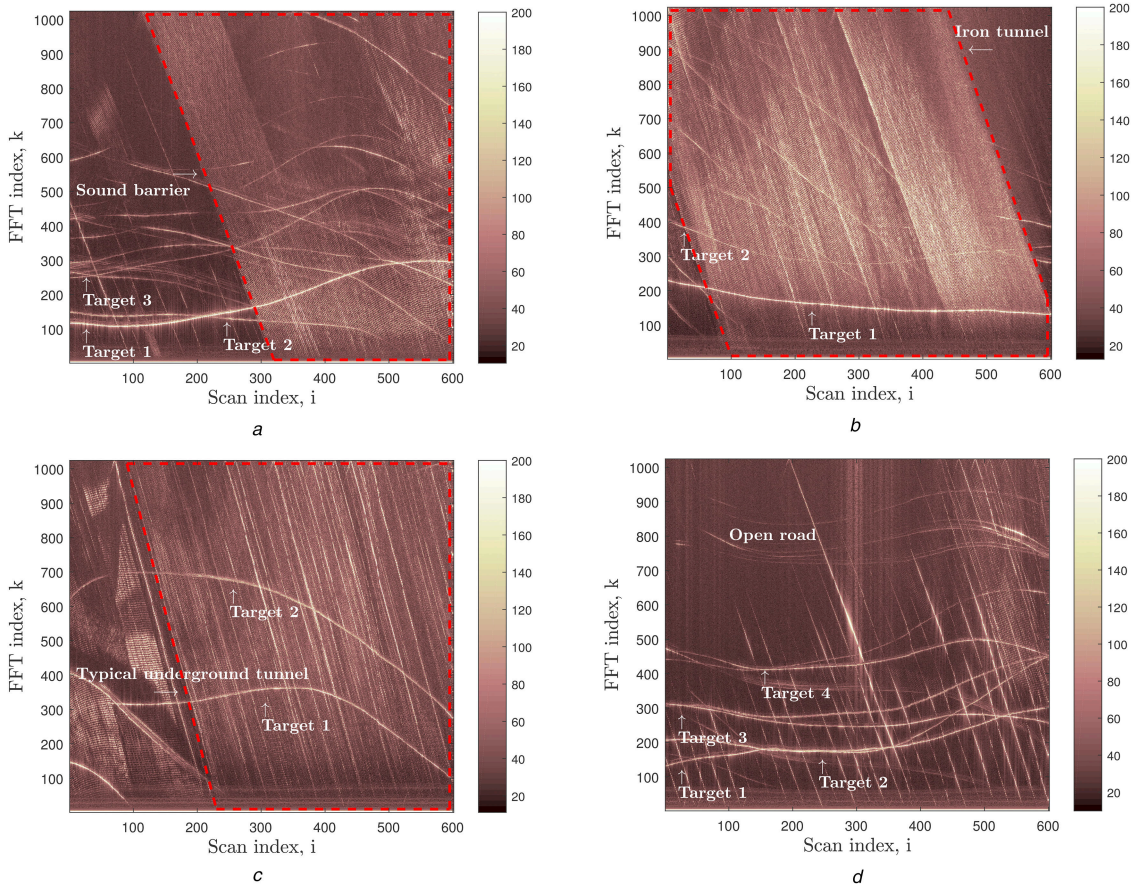


Fig. 2 *Accumulated FFT data for*
(a) Sound barrier, ***(b)*** Iron tunnel, ***(c)*** Typical underground tunnel, and ***(d)*** Open road

iron tunnel are stronger than in other road environments. As shown in Fig. 3*c*, non-target signals are also received while the vehicle is in a typical underground tunnel. However, since a typical underground tunnel is not made of iron structures, the targets are well detected compared to other environments such as iron tunnels or sound barriers. In Fig. 3*d*, unlike metal road structures, the targets are easily detected on the open road. Generally, the number of vehicles does not greatly affect the road structure classification performance, because the reflection signal by the road structure is more intense than the reflection signal by other vehicles. Moreover, we did not consider the influence of the number of vehicles and the experiment was carried out in situations where the number of vehicles was varied. Based on these results, the received signals reflected from each structure have unique characteristics even during one scan. Thus, these characteristics can be used to distinguish between the different structures. Since the different metal density of road structure results in the different intensities of the reflected signals, instantaneous FFT magnitude responses can be used to classify the types of road structures.

3 Road structure classification via ANN

An ANN is a type of machine learning algorithms. While the performance of most machine learning techniques depends on how features are extracted from the received data, deep learning algorithms automatically extract meaningful features through the

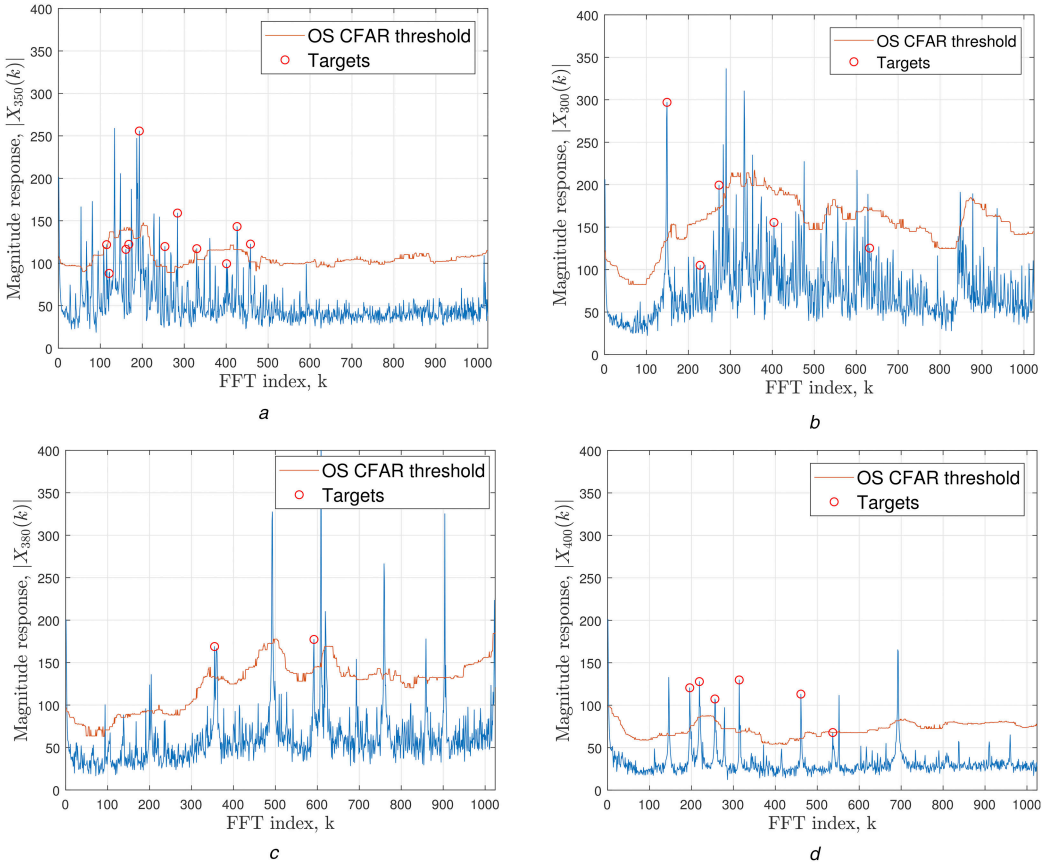


Fig. 3 *Instantaneous FFT magnitude responses for*
(a) Sound barrier, ***(b)*** Iron tunnel, ***(c)*** Typical underground tunnel, and ***(d)*** Open road

training process and are able to identify hidden features that otherwise would not be found.

ANN algorithms used for target classification include convolutional neural network [25, 26] and recurrent neural network [27]. However, they are not suitable for automotive radar that needs real-time signal processing. The computational complexity of CNN is $O\left(\sum_{l=1}^{N_{\text{layer}}} c_l \cdot s_l^2 \cdot n_l \cdot m_l^2\right)$, where l is the index of a CNN layer, N_{layer} is the number of convolutional layers, and c_l is the number of input channels of the lth layer. In addition, s_l is the spatial size of the filter, n_l is the number of filters in the lth layer, and m_l is the spatial size of the output feature map [28]. The computational complexity of RNN is $O(P^2Q)$ per epoch, where P is the number of units in an RNN layer and Q is the total number of time steps [29].

Thus, ANN known as PatternNet, which is relatively simple in network structure and has low complexity, is used to classify and recognise road structures [30–32]. A radar-equipped vehicle is used to gather time domain received signal data across four different road environments, namely, sound barrier, iron tunnel, typical underground tunnel, and open road. Then, the frequency domain signal is obtained by performing FFT on the signal received in the time domain. Next, the magnitude responses of FFT are compared to each other. In this paper, we employ a multilayer neural network called PatternNet. The structure of PatternNet with several hidden layers is shown in Fig. 4. Learning of the multilayer artificial neural network is performed via a backpropagation algorithm that is divided into four stages [33]. Fig. 5 shows the flowchart of the proposed neural network. The specific description of each stage in the flowchart is as follows. The first stage is to initialise the network. This requires each weight and threshold value in each neuron to be assigned within a certain range. The second stage is to determine the activation. The function from the input layer to the hidden layer can be expressed as follows:

$$h_r(p) = f\left(\sum_{q=1}^{N_{\text{in}}} g_q(p) w_{qr}(p) - b_r\right), \quad (r = 1, 2, \ldots, N_{\text{hidden}}) \tag{8}$$

where $h_r(p)$ and $g_q(p)$ denote the values of the rth node of the hidden layer and the qth node of the input layer, respectively. The qth node of the input layer contains $|X_i(q)|$ corresponding to the qth value of the FFT magnitude response. In addition, $w_{qr}(p)$ is the weight between $h_r(p)$ and $g_q(p)$, b_r is the bias of $h_r(p)$, p is the iteration index, N_{in} is the total number of input nodes, N_{hidden} is the total number of hidden nodes, and the function $f(\cdot)$ denotes an activation function. As in (8), the calculation process from the hidden layer to the output layer can be expressed as

$$o_s(p) = f\left(\sum_{r=1}^{N_{\text{hidden}}} h_r(p) w_{rs}(p) - b_s\right), \quad (s = 1, 2, \ldots, N_{\text{out}}) \tag{9}$$

where $o_s(p)$ denotes the value of the sth node of the output layer, $w_{rs}(p)$ is the weight between $o_s(p)$ and $h_r(p)$, b_s is the bias of $o_s(p)$, and N_{out} is the total number of output nodes. Here, we have performed experiments to classify four road structures, so N_{out} is 4. Once the activation has been determined, the weights are learned in the third stage. The goal here is to update the weight to minimise the magnitude of the error, which can be expressed as

$$E(p) = \frac{1}{2} \sum_{s=1}^{N_{\text{out}}} (\hat{o}_s(p) - o_s(p))^2, \tag{10}$$

where $\hat{o}_s(p)$ is the sth actual output value. Here, the value of $w_{rs}(p)$ that minimises $E(p)$ can be expressed as

$$w_{rs}(p+1) = w_{rs}(p) - \eta \frac{\partial E(p)}{\partial w_{rs}(p)} = w_{rs}(p) - \eta \delta_s^o(p) h_r(p), \tag{11}$$

where

$$\delta_s^o(p) = o_s(p)(1 - o_s(p))(\hat{o}_s(p) - o_s(p)), \tag{12}$$

and η is the step size. The term $w_{qr}(p)$ can be updated in the same way. This can be expressed as

$$w_{qr}(p+1) = w_{qr}(p) - \eta \frac{\partial E(p)}{\partial w_{qr}(p)} = w_{qr}(p) - \eta h_r(p)(1 - h_r(p)) \times \sum_{s=1}^{N_{\text{out}}} w_{rs}(p) \delta_s^o(p) g_q(p). \tag{13}$$

Finally, as an iterative stage, p is increased and the activation and weight learning stages are repeated until the error reaches zero in all epochs.

4 Target classification results

This section presents the classification performance when ANN is applied to experimental data obtained using an actual automotive radar system. The experiments were conducted on the Yongin–Seoul Expressway, and the experimental data consist of 13,680 scans to verify the deep learning performance, of which 70, 15, and 15% of the data is reserved for training, validation, and test set, respectively. The FFT magnitude response from (7) is employed as the input to PatternNet. The output is a 4×1 vector since we classify four types of road structures. The output layer is configured as a competitive layer because at least one of the road environments must be recognised in each snapshot. The following results are all based on the FFT magnitude response obtained from four antenna elements. This is because the classification performance is very poor when the phase part of the FFT result or the FFT result itself is directly used as the input.

The performance of deep learning depends on the number of hidden layers, number of nodes, and type of activation function. To ascertain the effects of each factor, the road structure classification process is conducted while varying each factor and monitoring the results. The classification accuracy based on the number of nodes in the hidden layer when one hidden layer is used is shown in Fig. 6. In Figs. 6*a–c* the tanh, sigmoid, and satlin functions are used as the activation functions, respectively, where the tanh function can be expressed as $f(x) = (e^x - e^{-x})/(e^x + e^{-x})$, *sigmoid can be expressed as* $f(x) = 1/(1 + e^{-x})$ [34]. satlin can be expressed as

$$f(x) = \begin{cases} 0, (x < 0) \\ x, (0 \le x < 1) \\ 1, (1 \le x) \end{cases}.$$

Satlin stands for saturating linear function. The results from testing with different number of hidden layer nodes show that, if the number of nodes is <20 or >80, the performance becomes poor. In addition, when the tanh or satlin function is used, the classification performance is the best when the number of nodes is about 50. When the sigmoid function is used, the classification performance is the best when 70 nodes are used.

The effect of the number of layers is shown in Fig. 7. In this test, the number of nodes in the hidden layer is fixed at 50 while the number of hidden layers is increased. In Fig. 7, the best classification performance is achieved when only one hidden layer is used. When the number of hidden layers is increased to two or more, no significant performance change is observed. This is because when the number of hidden layers is increased, it sometimes results in overfitting, which degrades the classification

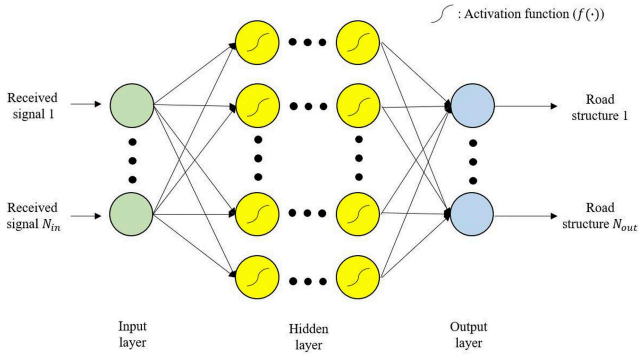


Fig. 4 *Framework of the proposed artificial neural network*

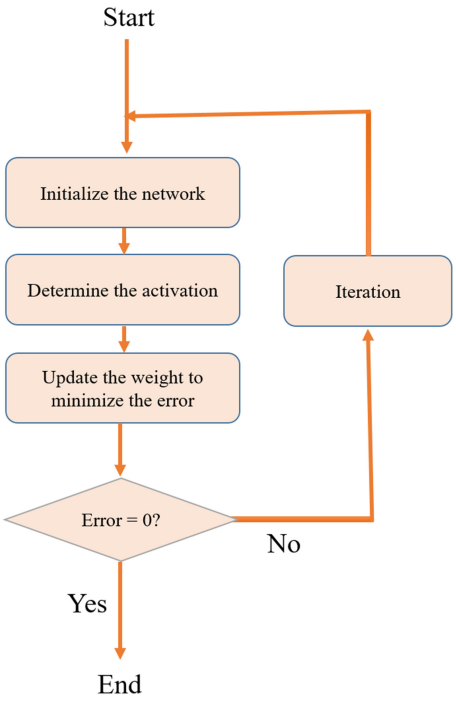


Fig. 5 *Flowchart of the proposed artificial neural network*

performance [35]. In the case of the saturating linear function, classification performance is similar to other activation functions. However, due to the fact that it does not get out of the saturation section, training often fails. For the similar reason, when rectified linear unit was used as an activation function, the training failed frequently and the road structure classification was almost impossible. As shown in Figs. 6 and 7, no significant performance difference between tanh or sigmoid as the activation function is observed. Thus, all following experiments were carried out using tanh as the activation function in the road structure classification process.

Next, the variation in performance is analysed versus the number of receiving antenna elements N. An $N_{\text{FFT}} \times N$ matrix contains the up-chirp FFT magnitude response of the received signals, where $N_{\text{FFT}} = 1024$ and $N = 1, 2, 3$, or 4. The number of hidden layer and the number of node are set based on the previous experimental results. The number of hidden layer is one and the number of nodes in the hidden layer is 50. A tanh function is used as the activation function in the hidden layer. The results of classifying road environments by applying the ANN to actual experimental data are listed in Table 1. Increasing the number of receiving antenna elements also improved the road structure classification performance. Furthermore, the classification accuracy for classifying iron tunnels is almost 100% because they exhibit distinct characteristics compared to other road environments. Sound barriers are not as distinctive as other road structures. Therefore, their classification performance is the lowest among the four road environments.

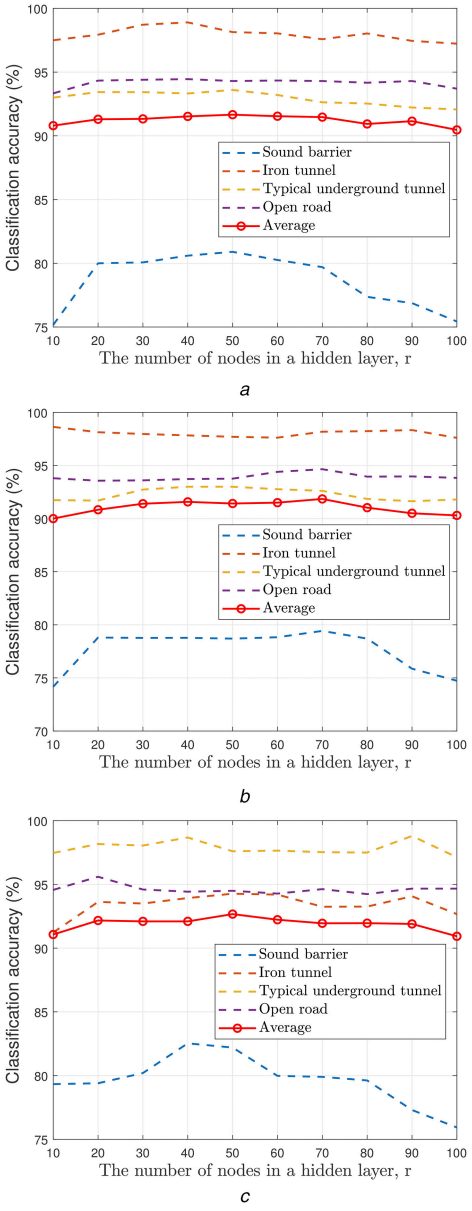


Fig. 6 *Classification performance with an increasing number of nodes in the hidden layer when the activation function is*
(a) tanh, ***(b)*** sigmoid and ***(c)*** satlin

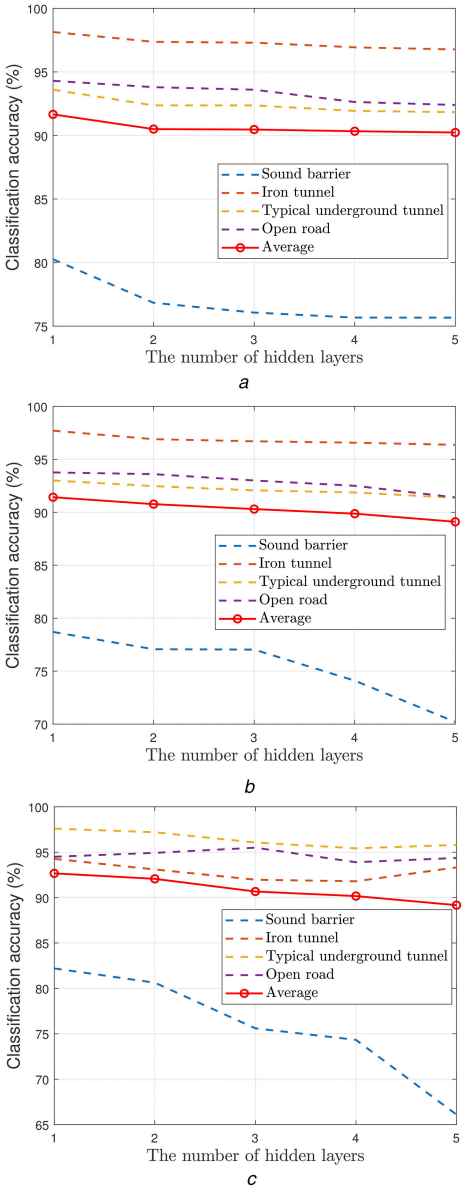


Fig. 7 *Classification performance with an increasing number of hidden layers when the activation function is*
(a) tanh, ***(b)*** sigmoid and ***(c)*** satlin

Table 1 Classification performance according to the number of antenna elements

Input data size	Sound barrier, %	Iron tunnel, %	Typical underground tunnel, %	Open road, %	Average, %
1024 × 1	71.8	96.0	88.6	92.5	87.8
1024 × 2	75.7	96.7	91.6	93.7	90.1
1024 × 3	76.0	98.0	91.9	94.2	90.8
1024 × 4	80.9	98.1	92.6	94.3	91.7

The different types of roads can exhibit similar characteristics when one FFT magnitude response is used. In order to prevent this, the deep learning process is executed with three consecutive magnitude responses (e.g. $\boldsymbol{C}_i$, $\boldsymbol{C}_{i+1}$, and $\boldsymbol{C}_{i+2}$). Thus, input data size is 1024 × 12. Table 2 shows the results of deep learning with three consecutive FFT magnitude responses inserted simultaneously as the input data. In Table 2, the first row represents the actual road structure and the first column represents the road structure estimated from the test result. In these results, the number of hidden layers is one and the number of hidden layer nodes is 50. In addition, a tanh function is used as the activation function. In this test, all of the iron tunnels are properly classified, while, for the sound barriers, nearly 10% of the test data is incorrectly recognised as typical underground tunnels. Consequently, the total classification accuracy is 95%. The average classification performance is enhanced by 3%p when more than one FFT magnitude response is used.

Finally, we compare the classification performance between two input settings; the FFT magnitude response without any feature extraction, and extracted features of the FFT magnitude response. Similar to the previous study [21], the mean, variance, skewness, and kurtosis of the FFT magnitude responses are used as the features. When the features are extracted and used as the input to ANN, the input size of ANN was 4 × 4. The road structure classification results for the deep learning using the above listed features are shown in Table 3. The data used for the classification

Table 2 Confusion matrix when using 3 consecutive scans

Estimated class/actual class	Sound barrier, %	Iron tunnel, %	Typical underground tunnel, %	Open road, %	Average, %
sound barrier	**86.0**	0.0	3.8	1.4	
iron tunnel	0.0	**100.0**	0	0.0	
typical underground tunnel	9.3	0.0	**96.0**	1.9	
open road	4.7	0.0	0.2	**96.7**	
average					**95.0**

The bold values mean the accuracy of classification results. We highlighted them for readability.

Table 3 Confusion matrix when the deep learning is performed using the suggested features in [21]

Estimated class/actual class	Sound barrier, %	Iron tunnel, %	Typical underground tunnel, %	Open road, %	Average, %
sound barrier	**61.0**	1.1	3.9	4.7	
iron tunnel	0.5	**94.6**	1.2	0.0	
typical underground tunnel	34.3	4.3	**94.5**	1.6	
open road	4.2	0.0	0.4	**93.7**	
average					**87.6**

The bold values mean the accuracy of classification results. We highlighted them for readability.

Table 4 Confusion matrix when the support vector machine with a Gaussian kernel is performed using the suggested features in [21]

Estimated class/actual class	Sound barrier, %	Iron tunnel, %	Typical underground tunnel, %	Open road, %	Average, %
sound barrier	**46.8**	5.2	1.9	2.9	
iron tunnel	48.6	**91.2**	5.6	4.1	
typical underground tunnel	1.3	3.5	**92.6**	0	
open road	4.5	0.0	0	**93.1**	
average					**84.0**

The bold values mean the accuracy of classification results. We highlighted them for readability.

of road structures are measured by radar as in [21]. The input of the classifier is the FFT magnitude response of the received time-domain signal, and no further preliminary processing was done on the input, such as OS-CFAR. Next, for the performance comparison with the SVM, the road structure classification was performed using the feature and SVM proposed in [21]. Table 4 shows the classification results using the Gaussian kernel. The mean, variance, coefficient of variance, skewness, and kurtosis of the received signal are used as features. When SVM is used, the classification performance is much lower than ANN algorithm proposed in this paper. Particularly, in the case of the sound barrier, the classification accuracy is <50%. The experiment uses one hidden layer and 50 nodes and tanh is used as the activation function. The detection performance for a typical underground tunnel and open road is similar to that when the FFT magnitude response is used as the input to artificial neural network; however, the detection performance of an iron tunnel and sound barrier is significantly worse. In the case of the sound barrier, the extracted features are not sufficiently distinct compared with those of the other road structures, so they could not be accurately classified. This result shows improvement of classification performance by over 7%p compared with the results using support vector machine.

5 Conclusion

In this paper, we classified road structures by applying an artificial neural network to the frequency domain signals received from various road environments. We evaluated how the classification performance varied with the number of hidden layers, the number of nodes in a hidden layer, and the types of activation function. Using the results obtained from various experiments, we propose which neural network structure suits the road structure classification. In addition, we confirmed how classification performance varies with input data size by changing the number of antenna elements and the number of snapshots. Based on our results, it was demonstrated that our proposed method was able to effectively classify road structures without extracting features from the measured data in advance. When the deep learning method was applied, the four different road environments were accurately recognised with an average accuracy of 95%. The results of this experiment confirm the proposed deep learning method can be used to classify various road structures. Using these results, we can apply a target detection algorithm suitable for each road environment. In future research, we will classify more diverse road structures existing in the urban areas, and additionally attempt to recognise transition regions.

6 Acknowledgments

This work was supported by the Technology Innovation Program (or Industrial Strategic Technology Development Program, No. 10080086, Development of 77/79 GHz Dual Band Radar) funded by the Ministry of Trade, industry & Energy (MI, Korea). This research was conducted when the author B.-H. Lee was at SNU.

7 References

[1] BIS Research: '*Global automotive sensor market demand, supply and opportunities: estimation and forecast of (2015–2022)*' (BIS Research, Fremont, CA, USA, 2015), pp. 1–48

[2] Skolnik, M.I.: '*Introduction to radar systems*' (McGraw-Hill, New York, NY, USA, 2001, 3rd edn.)

[3] Jones, W.D.: 'Keeping cars from crashing', *IEEE Spectr.*, 2001, **38**, (9), pp. 40–45

[4] Eriksson, L., Broden, S.: 'High performance automotive radar', *Microw. J.*, 1996, **49**, pp. 24–238

[5] Rohling, H.: 'A 77 GHz automotive radar system for AICC applications'. Proc. Int. Conf. Microwaves and Radar (MIKON98), Krakow, Poland, 1998

[6] Rohling, H., Meinecke, M.-M.: 'Waveform design principles for automotive radar systems'. Proc. CIE Int. Conf. Radar, Beijing, China, 2001, pp. 1–4

[7] Schneider, M.: 'Automotive radar – status and trends'. Proc. German Microwave Conf., Ulm, Germany, April 2005, pp. 351–354

[8] Nagy, L.L.: 'Electromagnetic reflectivity characteristics of road surfaces', *IEEE Trans. Veh. Technol.*, 1974, **IT-23**, (4), pp. 117–124

[9] Schneider, R., Didascalou, D., Wiesbeck, W.: 'Impact of road surfaces on millimeter-wave propagation', *IEEE Trans. Veh. Technol.*, 2000, **49**, (4), pp. 1314–1320

[10] Pathak, P.H., Burnside, W.D., Marhefka, R.J.: 'A uniform GTD analysis of the diffraction of eletromagnetic waves by a smooth convex surface', *IEEE Trans. Ant. Propag.*, 1980, **AP-28**, (5), pp. 631–642

[11] Matsunami, I., Kajiwara, A.: 'Clutter suppression scheme for vehicle radar'. IEEE Radio and Wireless Symp., LA, USA, January 2010, pp. 320–323

[12] Ma, Y.-Z., Cui, C., Kim, B.-S., *et al.*: 'Road clutter spectrum of BSD FMCW automotive radar'. IEEE European Radar Conf. (EuRAD), Paris, France, September 2015, pp. 109–112
[13] Lee, J.-E., Lim, H.-S., Jeong, S.-H., *et al.*: 'Enhanced iron-tunnel recognition for automotive radars', *IEEE Trans. Veh. Technol.*, 2016, **65**, (6), pp. 4412–4418
[14] Lee, H.-B., Lee, J.-E., Lim, H.-S., *et al.*: 'Clutter suppression method of iron tunnel using cepstral analysis for automotive radars', *IEICE Trans. Commun.*, 2017, **E100-B**, (2), pp. 400–406
[15] Lee, J.-E., Lim, H.-S., Jeong, S.-H., *et al.*: 'Harmonic clutter recognition and suppression for automotive radar sensors', *Int. J. Distrib. Sens. Netw.*, 2017, **13**, (9), pp. 1–11
[16] Takagi, K., Morikawa, K., Ogawa, T., *et al.*: 'Road environment recognition using on-vehicle LIDAR'. IEEE Intelligent Vehicles Symp., Tokyo, Japan, June 2006, pp. 120–125
[17] Broggi, A., Cerri, P., Medici, P., *et al.*: 'Real time road signs recognition'. IEEE Intelligent Vehicles Symp., Istanbul, Turkey, June 2007, pp. 981–986
[18] Zhou, L., Deng, Z.: 'LIDAR and vision-based real-time traffic sign detection and recognition algorithm for intelligent vehicle'. IEEE Int. Conf. on Intelligent Transportation Systems (ITSC), Qingdao, China, October 2014, pp. 578–583
[19] Guan, H., Li, J., Yu, Y., *et al.*: 'Using mobile LiDAR data for rapidly updating road markings', *IEEE Trans. Intell. Transp. Syst.*, 2015, **16**, (5), pp. 2457–2466
[20] Hata, A.Y., Wolf, D.F.: 'Feature detection for vehicle localization in urban environments using a multilayer LIDAR', *IEEE Trans. Intell. Transp. Syst.*, 2016, **17**, (2), pp. 420–429
[21] Lee, S., Lee, B.-H., Lee, J.-E., *et al.*: 'Statistical characteristic-based road structure recognition in automotive FMCW radar systems', *IEEE Trans. Intell. Transp. Syst.*, 2018, pp. 1–12, doi: 10.1109/TITS.2018.2865588
[22] Lee, S., Lee, B.-H., Lee, J.-E., *et al.*: 'Iron tunnel recognition using statistical characteristics of received signals in automotive radar systems'. IEEE Int. Radar. Symp., Bonn, Germany, June 2018
[23] Minkler, G., Minkler, J.: '*CFAR: the principles of automatic radar detection in clutter*' (Magellan, Baltimore, 1990)
[24] Kong, L., Wang, B., Cui, G., *et al.*: 'Performance prediction of OS-CFAR for generalized Swerling-Chi fluctuating targets', *IEEE Trans. Aerosp. Electron. Syst.*, 2016, **52**, (1), pp. 492–500
[25] LeCun, Y., Boser, B., Denker, J.S., *et al.*: 'Handwritten digit recognition with a back-propagation network', *Proc. Adv. Neural Inf. Process. Syst.*, 1990, **2**, pp. 396–404
[26] Krizhevsky, A., Sutskever, I., Hinton, G., *et al.*: 'Imagenet classification with deep convolutional neural networks'. Proc. Adv. Neural Inf. Process. Syst., Stateline, NV, USA, 2012, pp. 1106–1114
[27] Graves, A., Mohamed, A.-R., Hinton, G.: 'Speech recognition with deep recurrent neural networks'. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Vancouver, Canada, May 2013, pp. 6645–6649
[28] He, K., Zhang, X., Ren, S., *et al.*: 'Convolutional neural networks at constrained time cost', IEEE Conf. on Computer Vision and Pattern Recognition, Boston, MA, USA, 2015, pp. 3791–3799
[29] Chauvin, Y., Rumelhart, D.E.: '*Back propagation: theory, architectures, and applications*' (Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1995), pp. 433–486
[30] Cheriet, M.: '*Character recognition systems*' (A John Wiley and Sons, New York, NY, USA, 2007, 1st edn.)
[31] Al-Allaf, O.N.A, Abdalkader, S.A., Tamimi, A.A.: 'Pattern recognition neural network for improving the performance of iris recognition system', *J. Sci. Eng. Res.*, 2013, **4**, (6), pp. 661–667
[32] Ahmed, M., Imtiaz, M.T., Khan, R.: 'Movie recommendation system using clustering and pattern recognition network'. IEEE Computing and Communication Workshop and Conf. (CCWC), Las Vegas, USA, January 2018, pp. 143–147
[33] Rumelhart, D.E., McClelland, J.L., David, E., *et al.*: '*Parallel distributed processing: explorations in the microstructures of cognition*' (MIT Press, Cambridge, MA, 1986, 1st edn.)
[34] Karlik, B., Olgac, A.V.: 'Performance analysis of various activation functions in generalized MLP architectures of neural networks', *Int. J. Artif. Intell. Expert Syst.*, 2011, **1**, pp. 111–122
[35] Wenrui, H., Simon, F.: 'Neural network modeling of salinity variation in Apalachicola river', *Elsevier Signal Process. J.*, 2002, **36**, (1), pp. 356–362