Received: 27 July 2023 | Revised: 3 October 2023 | Accepted: 20 October 2023

DOI: 10.1049/rsn2.12501

IET Radar, Sonar & Navigation

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

Radar-based human activity recognition using denoising techniques to enhance classification accuracy

Ran Yu[1,2] | Yaxin Du[2] | Jipeng Li[3] | Antonio Napolitano[4] | Julien Le Kernec[1]

[1]James Watt School of Engineering, University of Glasgow, Glasgow, UK

[2]Shenzhen International Graduate School, Tsinghua University, Beijing, China

[3]Shanghai Jiaotong University, Shanghai, China

[4]Department of Engineering, University of Napoli "Parthenope", Napoli, Italy

Correspondence

Julien Le Kernec.

Email: julien.lekernec@glasgow.ac.uk

Abstract

Radar-based human activity recognition is considered as a competitive solution for the elderly care health monitoring problem, compared to alternative techniques such as cameras and wearable devices. However, raw radar signals are often contaminated with noise, clutter, and other artifacts that significantly impact recognition performance, which highlights the importance of prepossessing techniques that enhance radar data quality and improve classification model accuracy. In this study, two different human activity classification models incorporated with pre-processing techniques have been proposed. The authors introduce wavelet denoising methods into a cyclostationarity-based classification model, resulting in a substantial improvement in classification accuracy. To address the limitations of conventional pre-processing techniques, a deep neural network model called Double Phase Cascaded Denoising and Classification Network (DPDCNet) is proposed, which performs end-to-end signal-level classification and achieves state-of-the-art accuracy. The proposed models significantly reduce false detections and would enable robust activity monitoring for older individuals with radar signals, thereby bringing the system closer to a practical implementation for deployment.

KEYWORDS

radar, signal classification, signal denoising

1 | INTRODUCTION

The ageing problem refers to the demographic shift, where the proportion of elderly people in the population increases while the proportion of young people decreases. This phenomenon is a result of several factors, including declining fertility rates, improved healthcare, and advances in technology [1, 2]. Elderly care refers to the services and support provided to older adults to help them maintain their quality of life as they age. Older adults can accidentally fall at home causing serious injuries such as broken bones and head injuries or even be fatal. There exists a number of health-monitoring techniques to solve this problem, including but not limited to camera systems [3], wearable devices [4], and radar sensors [5]. Those health-monitoring techniques could enable people to live independently for longer and keep them in the comfort of their own home for as long as possible. Among those technologies, the radar shows its superiority in many aspects (as shown in Table 1), which is foreseen as a sensing modality of choice for healthcare as it is contactless and privacy-preserving. Some challenges remain for radar including classifying various human activities robustly in a real environment with noise and clutter.

The core of designing a radar-based health-monitoring system is the ability of distinguishing different human activities from the radar signal. There exists a number of relevant research works. For instance, the authors in Ref. [6, 7] discuss human activity recognition (HAR) based on Frequency Modulated Continuous Wave (FMCW) radar and Ultra Wide Band (UWB) radar, respectively.

Relevant research compares the performance of different radar data domains in classification tasks, including raw complex radar data, range-Doppler, range-time, Cadence Velocity Diagrams (CVD), and Cepstrogram [5, 8–10]. Different domains provide different advantages to the classification task. Some recent works classify activities by using cyclostationarity of the radar signal and get promising results [11]. As one

TABLE 1 Comparison between different health-monitoring devices [3, 4, 6].

Device	Longevity	Privacy protection	Lighting independent
Camera	✓	×	×
Wearable devices	×	✓	✓
Radar	✓	✓	✓

statistical characteristic, cyclostationarity is generated by the modulation of the signal caused by the motion of the target in the radar context [12].

Noise exists in the process of radar signal acquisition, impacting the quality and reliability of the extracted information from radar signals [13, 14]. This makes denoising techniques important in the radar signal classification problem. Denoising techniques aid in the removal of unwanted noise and interference. This is particularly crucial in situations where radar systems are subjected to various sources of noise, such as environmental factors, electronic interference, or clutter from the surrounding objects. The signal-to-noise ratio can be enhanced by denoising, thereby improving radar signals' readability. Consequently, this allows for more accurate and reliable classification of targets and facilitates better decision-making in real-time applications.

Various denoising methods have been developed for the radar signal, such as wavelet denoising, match filter, adaptive filter, and high-resolution spectral analysis [15–17], and they are investigated at length in textbooks and lecture notes [18–20]. However, it is still an open question about designing an effective denoising method for resource-constrained platforms for measured data without a clean reference and how denoising contributes to classification accuracy.

In recent years, the field of statistical machine learning and deep learning has rapidly developed and revolutionised many research areas, including computer vision, natural language processing, and autonomous systems. Research has been conducted with statistical-based machine learning algorithms, including support vector machine (SVM), K-Nearest Neighbours (KNN), and random forest (RF) [8, 10]. There is also a large body of work using deep learning-based methods into radar classification tasks, such as using deep neural networks for human activity classification [5, 9, 21–23] and demonstrating the feasibility of classifying with recurrent neural networks (RNN) [24]. However, the importance of signal pre-processing has mostly been forgotten with an over-reliance on machine learning to find patterns for recognition. The pre-processing methods, especially denoising techniques, should be introduced before the classification model for the purpose of improving classification robustness and accuracy.

Many studies related to denoising or restoring digital images have been conducted [25, 26], using both supervised and unsupervised methods. This same idea can be applied to the problem of radar signal denoising. For image denoising, its basic problem is removing the addictive Gaussian white noise from images, which can be realised by using CNN-based structures [27, 28]. More complex noise removal is performed in medical images using self-supervised deep networks [29, 30]. In most cases, the reference clean images are not available for denoising, which means the denoising should learn from noisy samples. This is blind denoising, and many relevant networks have been proposed [25, 31, 32]. The authors in Ref. [25] proposes a self-learning method to remove noise from images without its clean version. This idea can also be transferred into a radar signal denoising problem.

While denoising in machine vision has developed fast [26], there is also a lot of research about signal denoising derived from vision using deep neural networks. For instance, the authors in Ref. [33] utilise the self-learning based denoising for synthetic aperture radar (SAR) images. Model-based denoising approaches for picture restoration mainly attempt to tackle additive white Gaussian noise with architectures such as U-Net [34], residual learning [35], and FFDNet [36]. They however fail when confronted to real noise as opposed to synthetic [37]. The denoising model in Ref. [37] incorporates Gaussian and impulsive noise models for denoising improving on models using only Gaussian models. The DNCNet [38] consists of denoising and classification sub-networks. However, the authors had access to clean labelled samples to train the networks through supervised learning training in 2 phases, the denoiser and the classifier. The literature lacks methods for complex signal (In-Phase and Quadrature—I & Q) denoising as most methods are only suitable for real signals or inputs. Complex signal denoising has been demonstrated in automotive applications based on spectrograms being fed as complex-valued images into a CNN [39], but this is simulated additive white Gaussian noise which is not correlated. In [40], untrained graph neural networks directly estimate the raw signal parameters faster than noise alike one-shot learning showing promise as it has been validated over simulated and real data. This area, however, for complex signals is still undergoing development.

The main contributions of this paper are as follows.

- For cyclostationarity-based classification, we analysed the effect of self-adaptive wavelet filtering to dynamically adapt the noise filtration based on the input signal showing an improvement over threshold-based methods of wavelet denoising and no denoising at all for complex from complex (I & Q) radar signals. This has shown improvement in the classification accuracy with a moderate increase in computational load. The use of wavelet-based denoising and cyclostationary features has shown an outstanding performance for the drinking activity which is still a challenge in state-of-the-art methods, showing the uniqueness of using cyclostationary features as a complementary feature set in fusion methods for improved accuracy overall while the denoising method presented adapts to the input signal and improves the overall accuracy of our proposed method.
- For raw radar signal classification, the limitations of the wavelet-based denoising methods were overcome with a self-supervised deep neural network to denoise complex radar signals corrupted with correlated noise. Denoising with a deep neural network in the literature is limited to the

Gaussian signal on real data or simulated complex signals with independent, identically distributed noise components on each channel.

- We propose a novel network called Double Phase Cascaded Denoising and Classification Network (DPDCNet) that combines our novel self-supervised denoising network with a classification residual network, enabling the denoising and inference in a single model. DPDCNet produces denoised data for complex signals (I and Q) of higher quality compared to what can be achieved through wavelet-based denoising and self-supervised methods used alone, especially dealing with strongly distorted signals. The classification accuracy of DPDCNet outperforms existing state-of-the-art methods and is validated by comparing it with ResNet18 with varying signal-to-noise ratio to demonstrate its robustness.
- Our proposed methods incorporating denoising, preprocessing, and inference are suitable for real-time operation with up to 3.96s for our cyclostationary method with wavelet denoising and 0.26s for DPDCNet for each sample of 10 s.

This paper is organised in the following structure. Section 2 is about the background information including relevant knowledge of cyclostationarity, wavelet denoising, and signal denoising with deep neural network. In Section 3, we introduce the dataset used in this project and the methodology of our simulation procedure, proposed the cyclostationarity-based classification model, and deep learning-based approach. Relevant discussions about this project are covered in Section 5. Finally, our conclusions and future research directions are given in Section 6.

2 | BACKGROUND

2.1 | Cyclostationarity of radar signal

If a process or signal has periodically varying statistical properties with time, it is said to be a cyclostationary process or signal [11, 12]. In other words, cyclostationarity is a property of certain signals or processes that exhibit periodic or cyclical patterns in their statistical characteristics. Cyclostationary signals should be distinguished from both stationary (with constant statistical properties) and non-stationary signals (with irregularly unpredictable properties). As an important signal property, cyclostationarity is widely applied in the fields of signal processing, communication systems, and time–series analysis [12]. For the radar signal of human activity, the received radar signal is cyclostationary, since it is the modulation of the frequency modulated continuous wave (FMCW) signal from the transmitter which is a periodic signal by the human activity performed in the field of view of the radar (see Figure 1). Based on this understanding, cyclostationary or cyclic features of the received signal can be used for human activity recognition (HAR).

In this paper, eight different cyclostationary functions are discussed: cyclic autocovariance function (Cyy), cyclic cross-covariance function (Cyx), 2nd-order cyclic polyspectrum function (Pyy), 2nd-order cyclic cross-polyspectrum function (Pyy), cyclic spectrum function (Syy), cyclic cross-spectrum function (Syx), cyclic autocorrelation function (Ryy), and cyclic cross-correlation function (Ryx), where the definitions of those functions can be found in Ref. [12].

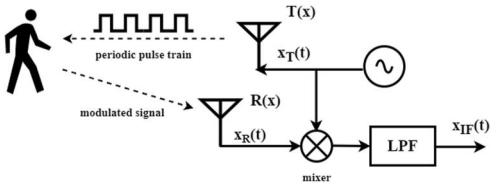

FIGURE 1 Block diagram of the FMCW radar system for human activity detection. The received signal is a modulation between the transmitted impulse train and human activity random phenomenon, which is believed to have cyclostationarity.

2.2 | Wavelet denoising

Denoising is the process of removing noise from a signal to enhance its quality and usefulness for analysis. In the field of signal processing, wavelet denoising is a popular method used to reduce noise from signals with unknown characteristics. This technique is based on the concept of wavelets basis, which are mathematical functions that are used to represent signals and decompose them into their constituent parts [41]. The range of versions of wavelet functions enables the wavelet analysis to cover both the general scope and details of observed signals [20].

Similar to Fourier transform that decomposes a signal into a summation of sinusoidal functions, the wavelet transform (WT) is defined as decomposing the signal into a series of wavelet basis functions, where those basis functions are finite length signals with energy concentration. The WT only decomposes the low-frequency segment of the signal further and does not extend the decomposition to the high-frequency portion, that is, the detailed section of the signal. The intermediate frequency signal from an FMCW radar is relatively stationary with beat frequencies concentrated in the lower frequency band of the receiver while noise is relatively random with higher frequency, therefore the wavelet coefficient amplitude of the noise is relatively small compared to that of the useful signal after wavelet decomposition. This is the theoretical foundation of wavelet denoising. A typical wavelet denoising procedure consists of the following steps [20]: (1) Decompose signal into a multi-level wavelet basis function; (2) Threshold coefficients of different levels with the threshold function; (3) Reconstruct the signal by combining different levels.

2.3 | Deep learning for signal denoising

Deep learning has been developed rapidly in recent years and has achieved a remarkable progress in relevant research areas

[25, 42]. For noise reduction, especially denoising and restoring digital images, much relevant research based on neural networks has been conducted. The most common method is supervised learning, which uses a pair of clean and noisy signals to train a deep neural network that learns the mapping relation from the noisy signal to a clean version. However, the main limitation of the supervised learning method is the difficulty to obtain the noise-free version of received signals in practice.

The self-supervised learning method has first been proposed to remove noise from digital images [25]. The fundamental concept [25] is constructing a U-Net shape deep neural network [43] and compelling the neural network to learn the mapping relation of multiple noisy image pairs with zero-mean Gaussian white noise. When the number of samples is large enough, since random noise is not predictable, from the perspective of minimising the loss function, the neural network tends to output the clean signal itself. This idea is also transferred into radar signal denoising. For example, according to the authors in Ref. [33], a self-supervised method is proposed for synthetic aperture radar (SAR) images denoising problem based on the separation of the real and imaginary parts of single-look complex SAR images. However, they consider the real and imaginary channels to be statistically independent which is not the case for I & Q in our applications.

3 | METHODOLOGY

3.1 | Simulation

Before starting to consider how to remove noise from a real-world captured radar, we first present some analysis on the denoising effect with the simulation radar signal. This is because the effectiveness of the radar signal quality improvement with denoising can be hard to quantify in a real-world signal with no information about its noise-free version. The simulation is conducted by using the HDM05 dataset and based on the method introduced in Ref. [44]. For simplicity, we only simulate the walking activity for human body, which is the most representative activity in our study. Additive Gaussian white noise (AGWN) is added to the ideal signal to test the effectiveness of denoising.

Two performance indicators are used to describe the quality of the simulated radar signal: root mean square error (RMSE) and the signal-to-noise ratio (SNR):

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left[f(n) - \hat{f}(n)\right]^2}, \quad (1)$$

$$SNR = 10\log\frac{\sum_{n=1}^{N} f^2(n)}{\sum_{n=1}^{N}\left[f(n) - \hat{f}(n)\right]^2}, \quad (2)$$

where f is the clean signal and $\hat{f}$ is the noisy signal (or denoised noisy signal). RMSE represents the discrepancy between the denoised signal and the original signal, showing whether distortions exist after denoising. SNR directly reflects how the signal is influenced by the noise component. Generally, the decrease in RMSE and increase in SNR indicate that denoising was successful.

Various factors influence the effect of wavelet denoising, such as decomposition levels, wavelet basis, and threshold values. In this study, the simulation is conducted to analyse how these factors relate to the denoising result, providing guidance for denoising real-world signal datasets. Figure 2 shows the RMSE and SNR of the simulated denoised radar signal by applying different wavelet denoising parameters.

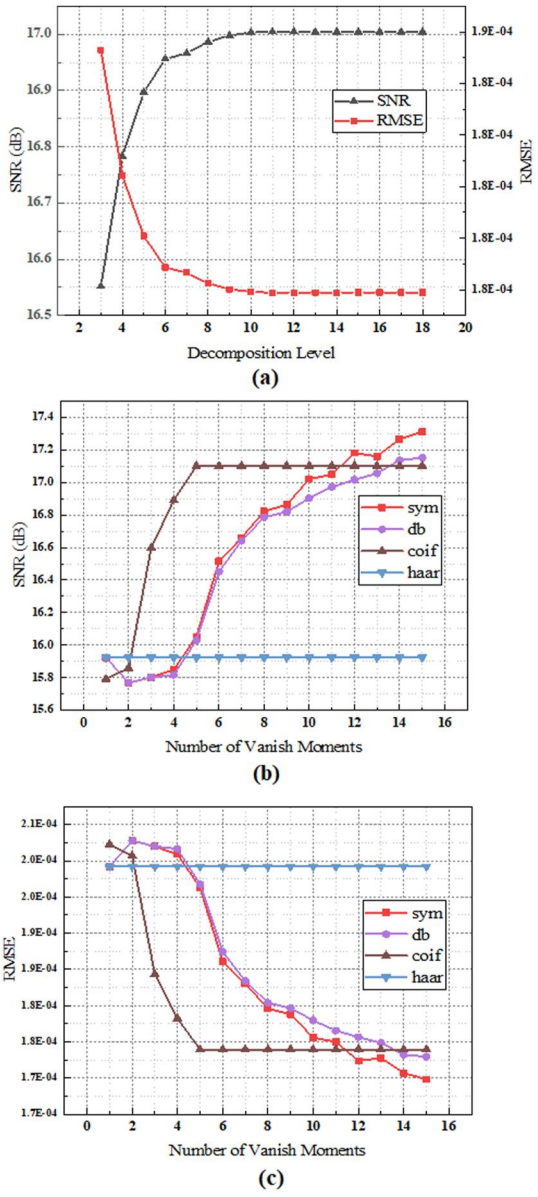

FIGURE 2 Denoising results on the simulation radar signal. (a) The relationship between wavelet decomposition levels and denoising effect; (b) The value of SNR by using a different wavelet basis; (c) The value of RMSE by using different wavelet basis functions.

First, AGWN is added to a noise-free simulated radar signal to generate a noisy version with SNR = 16 dB. Following this, we explore the effects of denoising with different decomposition levels and wavelet basis with a control variate method. Figure 2a shows the relationship between wavelet decomposition levels and the denoising effect. The denoising result improves significantly as the decomposition levels increase from 2 to 12, but further increasing the decomposition level does not contribute to denoising. The increase of computational time is negligible when increasing the decomposition levels. Figures c, d show the relationship between denoising and the applied wavelet basis function. Four different basis functions are discussed: Symlets (sym), Daubechies wavelets (db), Coiflet wavelets (coif), and Haar wavelet (haar). The horizontal axis of the diagrams represents the number of wavelet basis vanish moments. Haar does not have a vanish moment attribute, and Coiflet wavelets' largest vanish moment number is five. Therefore, the results for non-existent vanish moment numbers are substituted with the highest existing value.

Based on the results, Symlets and Daubechies wavelets show better performances compared to other wavelets. The diagrams also show that the number of vanish moments largely affects the denoising result. However, considering that increasing the number of vanish moments puts stress on computational costs, there should be a trade-off between computational efficiency and denoising effect in real practice. Considering the case of using Symlets as a wavelet basis, the computational time of denoising with the vanish moment of 15 is less than 0.1 s while this value increase to 1.5 s as the vanish moment increases to 25. However, the improvement of denoising with an SNR increase <1 dB is costly computationally. As a result, we choose a 14-level decomposition and 15 vanish moments for denoising the measured radar dataset.

3.2 | Dataset

Apart from the HDM05 dataset which is used for simulating the human activity radar signal [45], the main dataset used in this research is the human activity FMCW radar signal dataset from the University of Glasgow [46] for the purpose of monitoring the activity levels and patterns of people which we will refer to as the UoG dataset. The data is gathered by an off-the-shelf FMCW radar (key parameters listed in Table 2). Six different activities are collected in the UoG dataset: (A0) walking, (A1) sitting down, (A2) standing up, (A3) picking up an object, (A4) drinking water, and (A5) falling.

TABLE 2 Parameters of data collecting radar.

Parameter	Value	Definition
T_s	1 ms	Chirp duration
N_s	128	Number of samples
G_i	17 dB	Yagi antennas gain
P_o	18 dBm	Output power gain
f_c	5.8 GHz	Carrier frequency
f_B	400 MHz	Chirp bandwidth

3.3 | Cyclostationarity-based classification model

This section presents a classification procedure that employs cyclostationary functions along with the preprocessing techniques discussed earlier to enhance the data [11]. The flowchart of the procedure is shown in Figure 3. The input data is a one-dimensional complex radar signal, which is first passed through a moving target indicator for clutter rejection. Following that, denoising techniques are applied for data enhancement. The cyclostationary functions of the enhanced radar signal are then generated and used as features for classification.

In the pipeline, we use a single canceller moving target indicator (MTI) to remove the interference of stationary obstacles in the detection region. This is realised by subtracting the current echo of the signal from the previous echo at a certain time T_r, where T_r equals to the chirp duration T_s. Once MTI has been applied for clutter rejection, the complex signals are further enhanced by wavelet denoising to eliminate AGWN in the signal. Two main kinds of denoising methods based on wavelet decomposition have been used.

- The first one is the most conventional wavelet denoising method with thresholding wavelet coefficients on different decomposition levels.
- The second method is denoising by using a self-adaptive filter combined with wavelet denoising.

For wavelet decomposition, noise predominantly exists in the high frequency component with greater values of decomposition levels. The processed signal can be reconstructed without those decomposition levels and used as the reference signal for a self-adaptive filter to remove noise. In this paper, '**D_Threshold**' and '**D_SAF**' are the abbreviations for denoising with the threshold function and denoising with a self-adaptive filter, which are used to represent these two methods respectively.

As one of the most important research objects, cyclostationary functions of the radar signal have been used as features for classification. As mentioned in 2.1, eight different cyclostationary functions are generated for classification. For each function, it has its normal form and conjugated form. We use the small letter 'n' to indicate the normal form and the small letter 'c' to indicate the conjugated form. For example, nRyx and cSyy represent the normal form of the cyclic cross-correlation

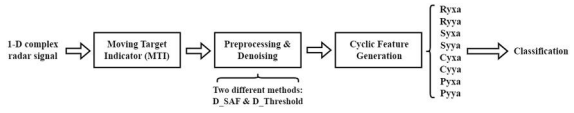

FIGURE 3 Pipeline of cyclostationarity-based classification.

function and the conjugated form of the cyclic spectrum function, respectively. Since the generated cyclostationary functions are in a complex form, we use their real part and imaginary part separately for classification. The prefixes 'real_' and 'imag_' are used to indicate the real part and imaginary part of the generated cyclic function. The eight cyclostationary functions with normal and conjugate forms are calculated, and each of them has real and imaginary parts. The total number of features used in classification is $8 \times 2 \times 2 = 32$, which are defined as the cyclic features for classification.

For the classification method, machine learning algorithms have been used, including support vector machine (SVM), Random Forest (RF), and K-Nearest Neighbours (KNN). According to the performed experimental results of each method, RF has the best classification accuracy compared to the other machine learning methods. For each cyclic feature, we have performed a series of experiments with the same dataset splitting strategy and calculated the average classification accuracy for the final result. Since the focus of the experiment is to analyse how denoising improves classification accuracy, relevant feature selection techniques are not included in the classification. This is to avoid other factors that may influence the experiment.

3.4 | Deep learning based denoising and classification

Considering the complex number characteristic of a radar signal, it naturally provides a pair of noisy signals: real part and imaginary part [33]. We denote the noise complex radar signal as $\hat{x} = \hat{x_r} + j\hat{x_i}$, where the subscript r and i represent the real and imaginary part. For each part, it can be split into a clean signal and additive noise: $\hat{x_r} = x_r + n_r$ and $\hat{x_i} = x_i + n_i$. Assuming the noise on the real and imaginary part is independent and identically distributed (i.i.d.) following a zero mean Gaussian distribution $n_r, n_i \sim \mathrm{N}(\mu = 0, \sigma^2)$, the optimisation target of the signal pair $(\hat{x_r}, \hat{x_i})$ can be represented as [33].

$$\arg\min_\theta \sum \mathrm{L}(f(\hat{x_r}, \theta), \hat{x_i}) = \arg\min_\theta \sum \mathrm{L}(f(\hat{x_r}, \theta), x_i + n_i), \tag{3}$$

where L is the loss function. According to Noise2Noise [25], in the case of infinite data, Equation (3) is equivalent to a supervised learning target that is

$$\arg\min_\theta \sum \mathrm{L}(f(\hat{x_r}, \theta), x_i). \tag{4}$$

For a finite dataset with N samples, the error between Equation (3) and Equation (4) can be represented as

$$\mathbb{E}_{\hat{x_i}}\left[\frac{1}{N}\sum x_i - \frac{1}{N}\sum \hat{x_i}\right]^2 = \frac{1}{N}\left[\frac{1}{N}\sum \mathrm{Var}(x_i)\right]. \tag{5}$$

A larger dataset could decrease this error [25]. Therefore, it is possible to train such a neural network that learns the mapping relation between the real and imaginary part to reduce noise in radar signals with a reasonably large dataset.

Based on the previous analysis, we have constructed a self-supervised neural network for radar signal denoising. As shown in Figure 4, the entire working process of the proposed self-supervised learning method can be divided into two stages: training stage and denoising stage. During the training stage (stage A), one part of the complex signal (real or imaginary) is sent into the network at a time, while the other part is viewed as a target or reference signal. The output is the denoised real or imaginary part of the signal, and the loss is computed between the denoised real (imaginary) signal and the original imaginary (real) signal. The loss function used here is the mean square error (MSE). If we denote the denoised part of the signal as $\hat{x}$ and reference signal part is x, then the loss function is defined as

$$L_r = \frac{1}{N}\sum_{k=1}^{N}(x_k - F(x_k, \theta))^2 = \frac{1}{N}\sum_{k=1}^{N}(x_k - \hat{x}_k)^2 \tag{6}$$

where the subscript of L_r is the abbreviation of reconstruction. Once the network is trained, the denoising stage (stage B) is used to denoise complex signals. It should be noticed that the model shares the same weights for both the real part and imaginary part, since they are both trained in stage A. The output of the two parts is then combined into a denoised complex signal.

The denoising network is constructed using a U-net-like structure, which is a type of convolutional neural network (CNN) architecture that was first proposed in Ref. [43] to solve biomedical image segmentation tasks. Since the radar signal data is one-dimensional complex data which is different from

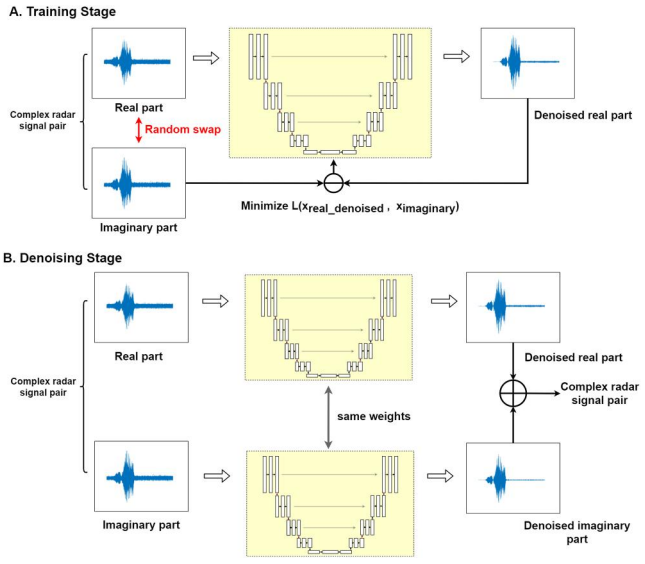

FIGURE 4 Working principle of the radar signal self-supervised denoising network: (a) the training stage optimises the network by using the loss function between the denoised real (imaginary) part and original imaginary (real) part; (b) the denoising stage process between the real and imaginary part of the signal with the same network, and outputs are combined into a denoised complex signal.

two-dimensional images, the U-net-like structure cannot be directly applied. There are two possible solutions to this problem. The first one is to construct a one-dimensional U-net-like network structure to process the signal data. The other consists in reconstructing the data structure and changing it into a data matrix. In this paper, the second method is adopted. Considering the number of samples N_s equals to 128, it is important to avoid truncating chirp signals, which means the length of one row of the matrix should be an integer multiple of 128. Therefore, the data is rearranged into a matrix 728×728, which means one row contains 6 chirps.

The self-supervised learning method is based on the assumption that noise distributes in the real and imaginary parts independently and identically (i.i.d.). Otherwise, the self-supervised network tends to learn the mapping relationship between the real and imaginary part noise, which means the noise cannot be completely removed. However, in the case of the UoG dataset, the noise in the signal's real and imaginary parts is correlated.

To overcome this problem, a new network structure, DPDCNet (Double Phase Cascaded Denoising and Classification Network), is proposed in this section. In DPDCNet, the denoising network is connected with a classification network, as shown in Figure 5. The DPDCNet is developed for two reasons. First, we want to realise the entire process of data enhancement and signal-level classification in one model. Second, since the self-supervised denoising network's results are limited by the correlated noise in the signal's real and imaginary parts, we want to use the high-level task (classification) to assist in training the denoising network [42]. Instead of only optimising the reconstruction loss between the denoised signal and the reference signal in the training process, the loss of classification is also taken into consideration, representing the quality of the data in high-level task performance.

To train DPDCNet, the denoising network and classification network should first be trained separately.

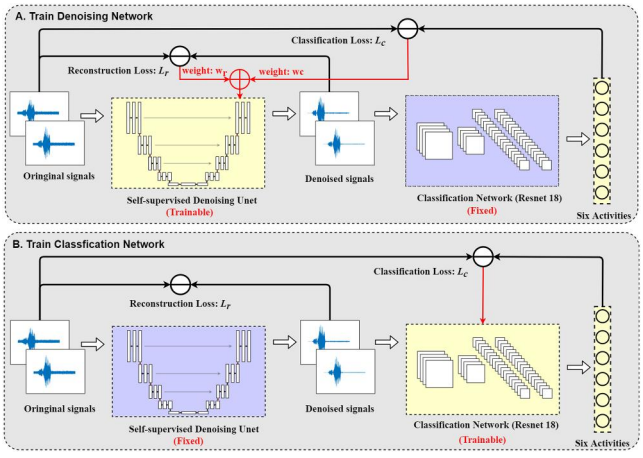

FIGURE 5 Training Procedure of DPDCNet. Two stages are trained separately. The denoising network is first trained by using a loss from both the denoising stage and the classification stage. The classification ResNet is then retrained with the fixed denoising network.

- The denoising network is first initialised with the Kaiming initialisation aiming to mitigate the problems of vanishing or exploding gradients during the training of deep networks [47]. It is trained by sending the input complex raw radar signals from the training set of the UoG radar dataset. The mean square error between real and imaginary parts was set as the loss function (Equation (6)). The number of training epochs of the denoising network is 50, with a learning rate of 10^{-3}.
- Then, the classification network is pre-trained by directly sending the original non-denoised signal data into the network. The one-dimensional complex signal is rearranged into two-dimensional square matrices of size 728×728. The real and imaginary parts are sent into the classification model together through two input channels. The number of training epochs and the learning rate are similar to what is used for training the denoising network.

Once the two networks have been successfully pre-trained, DPDCNet will be trained in combination, as shown in Figure 5. The combinational training can be divided into two stages. In the first stage, the denoising network is trained, and the classification network is fixed. The loss function is constructed using both reconstruction loss and classification loss. The reconstruction loss, as presented in Equation (6), indicates the level of resemblance between the denoised and the reference signal. The classification loss is in the form of cross-entropy. The cross-entropy loss is widely used in deep learning supervised learning tasks. The human activity classification belongs to a multi-classification problem. Each signal is labelled with an index corresponding to the activity it represents, and this index is represented as a one-hot vector p with six dimensions (as there are six activities in total). For instance, the first activity (walking) can be represented as $p = [1, 0, 0, 0, 0, 0]^T$. The output of the classification network is another six-dimensional vector q, where q_i represents the probability of the classified signal belonging to the ith activity and $\sum_{i=0}^{5} q_i = 1$. The cross-entropy loss can be expressed as

$$L_c = H(p, q) = -\mathbb{E}_{x \sim p} \log(q_i) = -\sum_{i=0}^{5} p_i \log(q_i), \qquad (7)$$

where the subscript of L_c is the abbreviation of classification. Therefore, the loss used in the first stage training of DPDCNet is

$$L_1 = w_r L_r + w_c L_c, \qquad (8)$$

where w_r and w_c are the adjustable weights of two losses representing the importance of them. In our experiments, both w_r and w_c are set to be identical, showing equal importance.

The second training stage of DPDCNet retrains the classification network while keeping the denoising network fixed. In this stage, the input data of the classification network is the denoised radar signal, which is formed as a square matrix. Similar to what has been done in the pretraining stage, the real and

complex parts are combined and sent to the network together. The classification network is optimised using only the classification loss L_r. The classification network uses a ResNet18 structure [48]. ResNet was introduced to combat the issue of vanishing gradients that arise in very deep neural networks, as this problem can hinder the effective training of these networks. ResNet18 is a smaller and shallower version of the network [48], with 18 layers, compared to the 152 layers of the original ResNet. We use this structure for its simplicity and its computational efficiency. The architecture of DPDCNet, for both the denoising part and the classification part, is shown in Table A1 and Table A2, respectively. For the details of DPDCNet training, we have listed the training hyper-parameters and relevant information including training epochs and learning rate in Table A3.

4 | RESULTS

4.1 | Cyclostationarity-based classification results

As mentioned in Section 3.3, for the proposed cyclostationarity-based classification method, 32 different cyclic features are used for classification and compared with each other. The results are listed in Table 3. Considering some cyclic features in Table 3 that has relatively poor performance in classification, we mainly focus cyclic features with a satisfying classification accuracy (with accuracy over 80% which is highlighted in Table 3). For those features, the average classification accuracy of Baseline, D_SAF, and D_Threshold is 82.5%, 84.9%, and 82.5%, respectively. Figure 6a shows the performance of the top five highest accuracy cyclic features (imag_nRyy, imag_nSyy, real_nRyy, imag_nCyy, and real_nCyy). It can be seen from the results that the self-adaptive filter-based denoising method can remarkably improve the classification accuracy compared with without denoising. Among those five cyclic features, the self-adaptive filter method achieves the highest average classification accuracy of 86.9% compared to baseline 82.9% and threshold value method 82.3%. The increase of accuracy compared between the self-adaptive filter method and baseline ranges from 0% min in imag_nSyy to 7.3% max in imag_nCyy.

In Table 4, the individual activity classification accuracy of cyclic feature real_nSyy is compared between the baseline and denoising methods with a fixed threshold and a self-adaptive filter. The results show that both denoising methods improve the accuracy among six activities. The self-adaptive filter denoising method significantly improves the accuracy of sitting down (A1) (by over 16% compared to the baseline and threshold methods), standing (A2) (24% better than threshold and 40% than the baseline), and picking up an object activity (A3) (11.1% better than the threshold value and 16.7% better than the baseline). This method keeps the same accuracy compared to the baseline method in drinking (A4) and falling (A5), which are 100% and 85%, respectively. For walking (A0), the self-adaptive filter method experiences a significant drop in performance (11.7% decrease compared to the baseline), while the threshold value method achieves the highest classification accuracy of 92.3% (3.4% increase compared to the baseline). We also compared our results with some existing methods with different radar domains, which used statistical learning-based

TABLE 3 Classification accuracy of each cyclic feature with different denoising methods.

Cyclic feature	Baseline	D_SAF	D_Threshold	Cyclic feature	Baseline	D_SAF	D_Threshold
real_cCyy	0.788	0.679	0.803	imag_cCyy	0.818	0.679	0.781
real_cCyx	0.394	0.292	0.358	imag_cCyx	0.409	0.234	0.307
real_cPyy	0.664	0.701	0.679	imag_cPyy	0.657	0.759	0.65
real_cPyx	0.482	0.394	0.599	imag_cPyx	0.496	0.401	0.599
real_cSyy	0.672	0.788	0.686	imag_cSyy	0.759	0.774	0.679
real_cSyx	0.518	0.453	0.657	imag_cSyx	0.504	0.46	0.613
real_cRyy	**0.818**	**0.803**	**0.796**	**imag_cRyy**	**0.818**	**0.803**	**0.803**
real_cRyx	0.504	0.401	0.547	imag_cRyx	0.504	0.438	0.489
real_nCyy	**0.832**	**0.869**	**0.796**	**imag_nCyy**	**0.803**	**0.876**	**0.818**
real_nCyx	0.380	0.277	0.336	imag_nCyx	0.365	0.336	0.343
real_nPyy	0.708	0.818	0.708	imag_nPyy	0.788	0.781	0.766
real_nPyx	0.511	0.445	0.642	imag_nPyx	0.482	0.453	0.562
real_nSyy	**0.737**	**0.854**	**0.782**	**imag_nSyy**	**0.839**	**0.839**	**0.861**
real_nSyx	0.496	0.504	0.679	imag_nSyx	0.496	0.467	0.708
real_nRyy	**0.832**	**0.876**	**0.803**	**imag_nRyy**	**0.839**	**0.883**	**0.839**
real_nRyx	0.467	0.401	0.577	imag_nRyx	0.504	0.453	0.562

Note: Bold values indicate the text analysis and are the best performances out of the tabe.

methods. As shown in Table 4, the authors in Ref. [49] introduced an adaptive thresholding method into the radar spectrum and phase domains to enhance accuracy. The results show that cyclostationarity-based classification has better classification accuracy when introducing denoising methods, which indicates our method can enhance the value of cyclostationary data domain. In Ref. [50], the authors combined different radar data domains, and it yielded the best overall performance among all the methods. The methods in Refs [52, 53] and the baseline method for cyclostationarity did not employ denoising techniques. However, we have shown that using denoising improves the performance using cyclostationarity, and it outperforms other methods for activity A4 which has been difficult to classify as demonstrated in Ref. [5]. The data domain can be combined with other domains to further improve the accuracy.

The robustness of the proposed denoising methods is further discussed by subjecting the classification task to ten-fold cross-validation [51] (Figure 7). Five representative cyclic features are selected for comparison, including real_nCyy, imag_nCyy, imag_nSyy, real_nRyy, and imag_nRyy. Similar to what is observed in Table 3, the method D_SAF achieves better classification performance in all these five cyclic features compared with the baseline in terms of the mean value (the increase in accuracy ranges from 1.5% min to 3.3% max) and median value (the increase in accuracy ranges from 0.1% min to 4.1% max). The boxplot shows that the classification accuracy of the baseline method is more sensitive to the splitting of training and validation datasets. The accuracy of the baseline method is more volatile, while the accuracy of the two denoising methods, especially D_SAF, is more stable. This further demonstrates the benefits of the proposed denoising methods to improve classification accuracy. However, these methods have shown a shortcoming on classifying walking A0 accurately with D_SAF and difficulty generally with drinking A3 with all methods. Because walking is an essential activity to be classified for further processing to evaluate the risk of falling, another method is being explored in the next section with the DPDCNet to improve the classification accuracy and robustness of our algorithm.

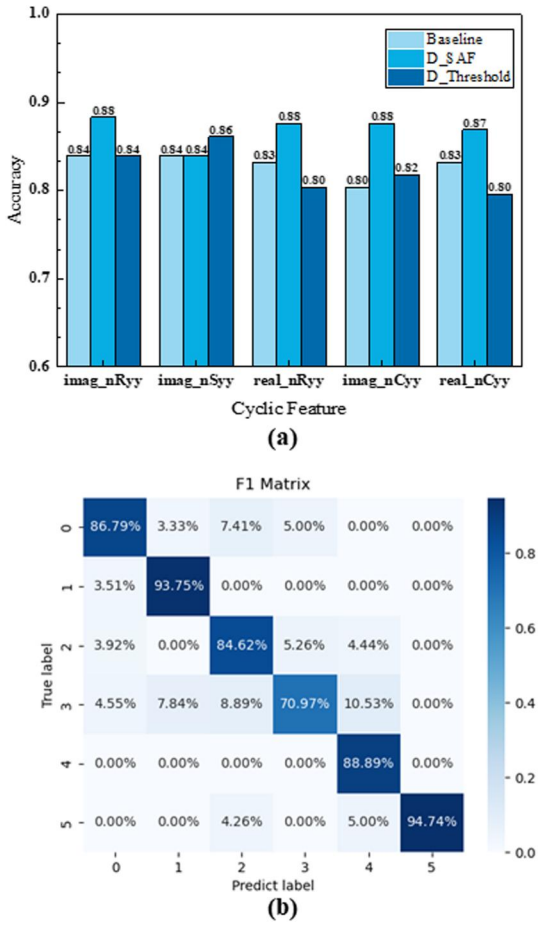

FIGURE 6 (a) Top five cyclic features ranking based on classification accuracy; (b) F1 score matrix of real_nRyy classification results with self-adaptive filter denoising (D_SAF). The overall accuracy reaches 88%.

4.2 | DPDCNet results

DPDCNet can perform both denoising and classification tasks in one model. Figure 8 shows the noisy signal (blue) with its denoised version (red). Figure 8b shows a section with no action and therefore just noise, our method can reduce and suppress the noise significantly. For the segments containing information, the denoised result removes noise without distorting the radar signal as shown in Figures 8c,d,e.

For DPDCNet's performance in human activity radar signal classification, the experiment is carried out by splitting the UoG dataset into the training set (80% of total data) and validation set (20% of total data). Figure 9 is the evaluation results presenting in the form of the average classification

TABLE 4 Individual activity accuracy between different methods with real_nSyy.

Method	A0	A1	A2	A3	A4	A5	Overall
Baseline	88.9%	80.65%	48.0%	44.4%	**100%**	85%	73.7%
D_SAF	77.8%	**96.8%**	88.0%	61.1%	**100%**	85%	85.4%
D_Threshold	92.3%	80.7%	64.0%	50.0%	93.8%	85%	78.2%
Mask data [49]	97.8%	90.7%	89.3%	62.5%	76.7%	92.4%	84.9%
Mask phase [49]	99.4%	86.6%	87.4%	68.0%	71.6%	91.9%	84.1%
Multi-domain [50]	**100.0%**	95.5%	**95.2%**	**76.9%**	84.6%	**100.0%**	**92.0%**

Note: Bold values indicate the text analysis and are the best performances out of the tabe.

accuracy, precision, recall, and F1 score. The overall classification accuracy is 97.4%. The average values of precision, recall, and F1 score are 0.97, 0.98, and 0.97, respectively.

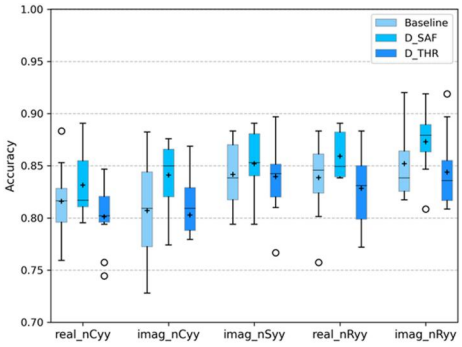

FIGURE 7 Ten-fold cross validation results of five representative cyclic features: real_nCyy, imag_nCyy, imag_nSyy, real_nRyy, and imag_nRyy.

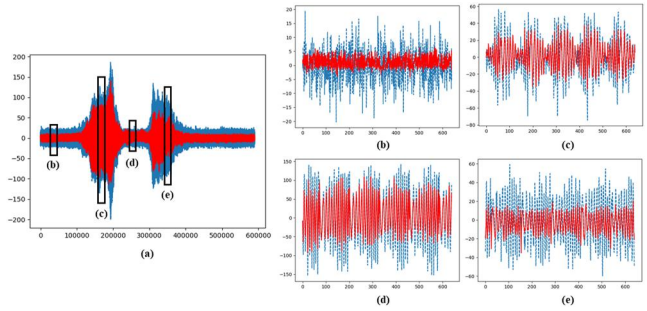

FIGURE 8 Comparison of denoised signal (red) and noisy signal (blue). (a) General view of the signal; (b), (c), (d), and (e) Partial enlarged details of signal.

Comparing it with results shown in Figure 6 with the overall accuracy of 88%, DPDCNet outperforms than the cyclostationarity-based classification method in all six activities. The authors in Ref. [11] proposes a cyclostationarity-based classification model with the feature selection method and Random Forest method for classification which reaches a 95.4% overall accuracy. The proposed method outperforms than the method in Ref. [11] and improves the overall performance by 2%.

To further investigate our method, we conducted a ten-fold cross-validation experiment of the proposed DPDCNet and compared its result with existing state-of-the-art approaches [5]. Five different methods were chosen to compare with the proposed DPDCNet. In Ref. [24], the authors introduce a data augmentation method and a recurrent neural network for human activity classification. The authors in Ref. [52] proposes a lightweight network model based on the Mobilenet v2 model pre-trained on the ImageNet dataset. Refs. [8, 50] are based on SVM with different kernels such as cubic and quadratic for classification. The uthors in Ref. [53] proposed a novel PointNet phase CNN structure and introduced a fusion voting method for combining different models. The overall classification results and individual activity performances are shown in Table 5. The DPDCNet shows the best overall performance among all those methods with an overall accuracy of 95.7%. Compared to Ref. [52], our method has a 0.3% improvement in accuracy. Considering the individual activities, DPDCNet reaches 100% accuracy in walking (A0) and standing up (A2). For drinking (A4), it provides the highest accuracy (94.7%) among all methods and outperforms [8] by 0.3%. For sitting down (A1), picking up an object (A3), and falling (A5), the gaps between our method and the best-performing algorithms are 2.3%, 9%, and 0.5%, respectively.

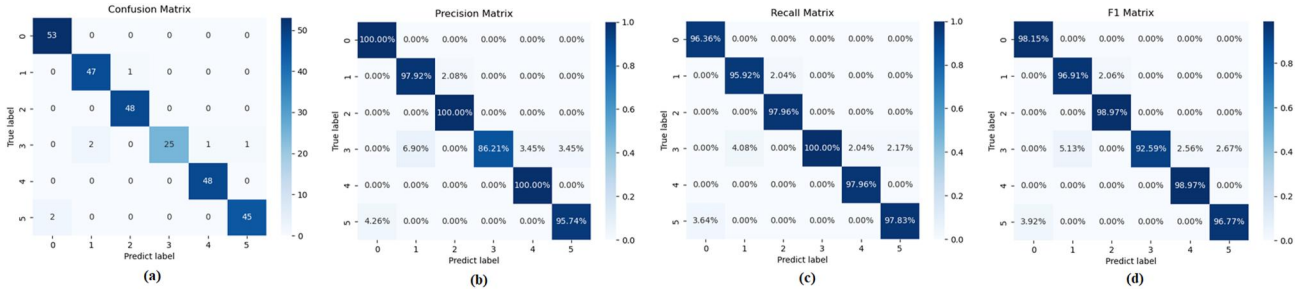

FIGURE 9 (a) Confusion matrix of the DPDCNet classification result (average accuracy is 97.4%); (b) Precision matrix of the DPDCNet classification result (average precision score is 0.97); (c) Recall matrix of the DPDCNet classification result (average recall score is 0.98); (d) F1 matrix of the DPDCNet classification result (average F1 score is 0.97).

TABLE 5 The classification results of existing methods and the proposed method.

Method	A0	A1	A2	A3	A4	A5	Overall
GRU [24]	99%	**98.7%**	97.4%	87.1%	86.4%	99.5%	94.3%
Mobile net v2 [52]	**100%**	94.0%	**100%**	83.0%	87.0%	**100%**	95.4%
SVM quadratic [50]	**100%**	95.5%	95.2%	76.9%	84.6%	**100%**	92.0%
SVM cubic kernel [8]	**100%**	98.4%	98.5%	**95.2%**	94.4%	90.2%	95.4%
PointNet CNN [53]	**100%**	97.1%	88.9%	90.9%	91.4%	**100%**	94.3%
Our method (DPDCNet)	**100%**	96.4%	**100%**	86.2%	**94.7%**	99.5%	**95.7%**

Note: Bold values indicate the text analysis and are the best performances out of the tabe.

5 | DISCUSSION

Classification based on the cyclostationarity of radar signals, and how to improve it by introducing denoising techniques for complex signals, was presented in this paper. Different cyclic functions have been generated and classified using machine learning algorithms. Classification results show that classification accuracy varies a lot between different cyclostationary functions. Generally, auto-cyclostationary functions provide better quality data for classification than cross-cyclostationary functions. One possible interpretation of the varying performance between different cyclic features is that cyclic features extract and capture different information of the moving target. For instance, the cyclic autocorrelation function R_{yy} provides the highest classification accuracy (84%) since it reveals the frequency and phase of the periodic modulation, which can be used to estimate the Doppler shift and rotation rate of the target. As introduced in Ref. [11], the accuracy can reach up to 95.4% when all the cyclic frequencies from −3 to 3 are combined together and could reach even higher if we used feature selection methods to only keep salient features. However, this is beyond the scope of this article.

To improve the classification accuracy, wavelet denoising techniques have been implemented to the radar signals. The first fixed threshold method, D_Threshold, denoises the signal by thresholding different wavelet decomposition levels and requires extensive manual tuning which is not practical to implement as it would need to be returned for new classes. The second method (D_SAF) combines wavelet decomposition with a self-adaptive filter. The advantage is that this method tunes itself depending on the input; however, it has shown a reduction in performance with the walking activity A1 that is not acceptable in the context of assisted living. From the experimental results, it can be concluded that the classification accuracy of the cyclostationarity-based method can be enhanced through denoising and careful tuning of the whole preprocessing chain as shown in Ref. [49]. The maximum accuracy increases to 88.3% by the self-adaptive filter method compared to 83.9% without denoising. We provide an intuitive interpretation of the accuracy improvement.

Considering cyclic feature Real_nRyy, we plot its amplitude before and after denoising in Figures 10a,b and normalised them to make the largest amplitude value to be 1 for easier comparison. From the diagram, it is hard to distinguish between (A2) standing and (A3) picking up an object from an amplitude perspective by using the raw radar signal without denoising, as shown in Figure 10a, where a strong overlap can be observed. Figure 10b shows the result by applying wavelet denoising, where the difference in amplitudes becomes apparent and easier to distinguish.

Table 6 summarises the various denoising methods used for cyclostationarity-based classification, and suggests that the denoising step adds only about 1 s to the overall classification pipeline processing time, which is acceptable. The classification method used in this study is Random Forest, which can be run on a CPU with an inference time of less than 0.1 s per sample. Overall, the denoising methods employed here improved classification accuracy, thereby increasing computational demand by 35% for the D_SAF method and 41% for the D_Threshold method. Although the resulting accuracy does not reach state-of-the-art performance, it still provides a valuable improvement for real-time classification with 4.4% and 2.2% improvement in accuracy.

One of the main difficulties in radar signal denoising is the lack of a noise-free reference signal, which is essential for both analysis and training of denoising methods. To solve this problem, we simulated human activity radar signals based on the HDM05 dataset and constructed a self-supervised denoising neural network for denoising. With the help of simulation, the appropriate parameters for wavelet denoising were tested. The simulation results show that a higher order moment wavelet basis and relatively large decomposition levels are preferred in this problem, considering the complexity of the radar signal. Additionally, the soft threshold function (D_SAF) generally outperformed the hard threshold function (D_Threshold). Self-supervised learning can be used for denoising when the reference clean signal is not observable. For complex radar signals,

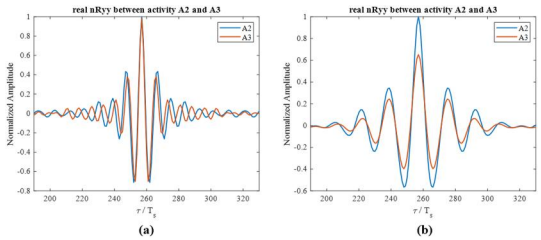

FIGURE 10 Normalised value of Real_nRyy of (A2) standing and (A3) picking up an object. (a) Real_nRyy of raw radar signal; (b) Real_nRyy of denoised radar signal.

TABLE 6 Summary table of different denoising methods.

	Baseline	D_SAF	D_Threshold
Preprocessing time (per sample)	2.7s	2.7s	2.7s
Denoising time (per sample)	None	0.98s	1.16s
Total training time	2.4s	2.4s	2.4s
Inference time (per sample)	<0.1s	<0.1s	<0.1s
Denoising + Pre-processing + Inference time (per sample)	<2.8s	<3.78s	<3.96s
Best overall accuracy	83.9%	88.3%	86.1%

the real and imaginary parts are used as mutual references used for training. However, it is observed that noise in real and imaginary parts of the signal is correlated, which limits the denoising effect.

To address this limitation, the proposed DPDCNet employs a high-level task loss function to aid in the training of the denoising network, which serves as its core component. DPDCNet is a cascaded model that combines denoising pre-processing with signal-level classification, achieving competitive performance compared to existing methods. Its superiority can be further demonstrated in stronger noise environments where conventional classification models face challenges. To simulate this situation, the Gaussian addictive white noise is added to radar signals in the UoG dataset and retrained the network for SNR levels ranging from −14 to 0 dB in steps of 2 dB. Figure 11 compares the classification accuracy between DPDCNet and ResNet18 without the denoising network. The classification accuracy difference between DPDCNet and ResNet18 is less than 0.5%. As the SNR increases to −14 dB, the difference between DPDCNet and ResNet18 increases to 9.5%. The results show that DPDCNet outperforms ResNet18 in all SNR levels less than zero, and the accuracy gap between DPDCNet and ResNet18 is more significant in lower SNR levels. It should be noticed that the ResNet 18 is the same as the classification sub-network structure in our proposed DPDCNet, and the improvement of classification accuracy shows the effectiveness of the denoising part. This experiment further demonstrates the value of our work. Existing deep neural network methods for human radar signal classification problems could reach the same performance in an environment with good SNR; however, their performance drops rapidly when the SNR is degrading. Figure A1 shows the difference in accuracy when training Resnet alone with an SNR of −4 dB and the use of DPDCNet after pre-training the denoising network. The loss and accuracy show that DPDCNet is able to generalise much better than ResNet18 alone.

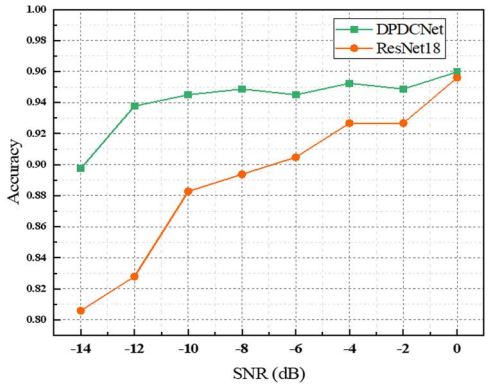

FIGURE 11 Comparison of classification accuracy between DPDCNet and ResNet18 in a simulated strong noise environment (SNR <0 dB).

We summarise the proposed DPDCNet method in Table 7. All the experiments of DPDCNet are performed by using a single NVIDIA GeForce RTX 3090. The training time of DPDCNet is 134.61 s/epoch, and the inference time is 0.26 s/sample. The robustness of DPDCNet comes at the cost of 2.71 times the computation time compared to DPDCNet. A denoising time and inference in 0.26 s for a 10 s recording are sufficient for real-time operation. It is worth mentioning that most of the existing classification methods tend to focus on feature extraction from spectrograms using statistical or deep learning methods. The authors in Ref. [54] implemented the signal pre-processing of a generating micro-Doppler spectrum by squaring the output matrix after the ShortTime Fourier Transform (STFT) and requires 2.77 s to perform the pre-processing only before classification with a dual cores Arm Cortex-A78AE platform. The whole process of spectrograms generation and classification can be accelerated by using GPU

TABLE 7 Summary table of DPDCNet.

	Denoising network	**Classification network**
Trainable parameters	718801	11176454
Training time (per epoch)	92.24s	42.37s
Function	Denoising	Inference
Time (per sample)	0.19s	0.07s
Overall accuracy	None	Similar as DPDCNet
	DPDCNet	**ResNet-18**
Trainable parameters	11985255	11176454
Training time (per epoch)	134.61s	42.37s
Function	Denoising + Inference	Inference
Time (per sample)	0.26s	0.07s
Overall accuracy	97.4% (UoG dataset)	97.4% (UoG dataset)
	95.2% (UoG dataset with SNR = −4 dB)	92.6% (UoG dataset with SNR = −4 dB)
	94.5% (UoG dataset with SNR = −10 dB)	88.2% (UoG dataset with SNR = −10 dB)

(Nvidia Ampere, 2048 NVIDIA CUDA Cores), which reduces the computation time into 1.31 s per sample [54]. This results shows that comparing with classification based on spectrograms, the proposed method shows advantages in faster pre-processing and inference time.

This demonstrates that with a well-designed input data structure, raw radar data can be used directly for high-level machine learning tasks with a powerful deep neural network. DPDCNet is believed to have advantages in flexibility and adaptability as a truly end-to-end radar signal classification model. However, it is important to note the potential limitations of this end-to-end model. Compared with other methods that use spectrograms or other feature extraction techniques for classification tasks, the input patch size of DPDCNet is relatively large, which may affect memory usage and throughput. Although efforts have been made to improve computational efficiency and reduce network complexity, it may still be challenging to deploy a U-net-like denoising network combined with a ResNet18 into platforms like FPGA for resource-constrained embedded platforms for in-home radar healthcare systems.

6 | CONCLUSION AND FURTHER WORK

This study presents a comprehensive evaluation of 3 techniques for human activity radar signals in combination with different classification algorithms. The UoG dataset, containing measurements of activities, is used for evaluation. All techniques employed a moving target indicator to remove static clutter before denoising with wavelet decomposition with a fixed and soft threshold as well as DPDCNet. The use of denoising techniques results in a significant increase in classification accuracy, with a maximum accuracy of 88.3% compared to 83.9% without denoising.

Furthermore, this study introduces a novel deep learning method for radar signal denoising and classification, aimed at overcoming the limitations of traditional denoising techniques. The proposed approach, DPDCNet, performs an end-to-end data enhancement and classification process. Unlike conventional denoising methods, DPDCNet employs a self-supervised learning approach in the denoising stage, without requiring prior knowledge of noise statistical characteristics or clean signal information. The denoising performance of DPDCNet is further enhanced by incorporating a loss function from the classification stage to eliminate correlated noise between signal's real and imaginary parts. According to the experiment results, DPDCNet achieves a classification accuracy of 97.4%, demonstrating its competitive performance compared to previous methods. It is also robust against higher noise levels with a 6% degradation from 0 dB to −14 dB SNR, whereas ResNet18 drops by 15% in the same interval.

The denoising network could be improved further by adding residual blocks to the original U-net structure. While DPDCNet has shown satisfactory performance in the classification task using raw radar data, it can also be used to classify radar signal spectrograms with some modifications. This approach can significantly reduce the size of the input data, leading to faster training and improved memory efficiency. This combined with efficient implementations as shown in Ref. [54] could boost spectrogram-based techniques significantly as demonstrated in this paper for raw data. Therefore, exploring different input data structures and their effects on classification accuracy is another potential avenue for future research. Moreover, evaluating the generalisation ability of DPDCNet on different radar datasets collected under varying conditions could provide further insight into the feasibility of the proposed method in different conditions and classes to verify its generalisation capabilities. The cyclostationariy performances shown in Ref. [11] combined with pre-processing fine tuning shown in Ref. [49] could significantly improve the performances of the classification techniques based on cyclic features by determining the most salient features and therefore calculating only the relevant cyclic frequencies to streamline processing further.

AUTHOR CONTRIBUTIONS

Ran Yu: Formal analysis; Investigation; Software; Validation; Writing – original draft. **Yaxin Du**: Methodology; Software; Visualisation. **Jipeng Li**: Methodology; Software; Visualisation. **Antonio Napolitano**: Conceptualisation; Methodology; Resources; Software; Validation. **Julien Le Kernec**: Conceptualisation; Methodology; Project administration; Supervision; Writing – review & editing.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

https://researchdata.gla.ac.uk/848/.

ORCID

Julien Le Kernec https://orcid.org/0000-0003-2124-6803

REFERENCES

1. Boccardi, V.: Population Ageing: The Need for a Care Revolution in a World 2.0 (2019)
2. Knickman, J.R., Snell, E.K.: The 2030 problem: caring for aging baby boomers. Health Serv. Res. 37(4), 849–884 (2002). https://doi.org/10.1034/j.1600-0560.2002.56.x
3. Shoaib, M., et al.: Altcare: safe living for elderly people. In: 2010 4th International Conference on Pervasive Computing Technologies for Healthcare, pp. 1–4. IEEE (2010)
4. Degen, T., et al.: Speedy: a fall detector in a wrist watch. In: ISWC, vol. 3, pp. 184–187 (2003)
5. Yang, S., et al.: The human activity radar challenge: benchmarking based on the 'radar signatures of human activities' dataset from glasgow university. IEEE J. Biomed Health Inf. 27(4), 1813–1824 (2023). https://doi.org/10.1109/jbhi.2023.3240895
6. Li, H., et al.: A multisensory approach for remote health monitoring of older people. IEEE J. Electromagn, RF and Microw Med Biol 2(2), 102–108 (2018). https://doi.org/10.1109/jerm.2018.2827099
7. Ding, C., et al.: Non-contact human motion recognition based on uwb radar. IEEE J. Emerg. Sel. Top. Circ. Syst. 8(2), 306–315 (2018). https://doi.org/10.1109/jetcas.2018.2797313

8. Li, X., et al.: Hierarchical radar data analysis for activity and personnel recognition. Rem. Sens. 12(14), 2237 (2020). https://doi.org/10.3390/rs12142237
9. Le Kernec, J., et al.: Radar signal processing for sensing in assisted living: the challenges associated with real-time implementation of emerging algorithms. IEEE Signal Process. Mag. 36(4), 29–41 (2019). https://doi.org/10.1109/msp.2019.2903715
10. Taylor, W., et al.: Radar sensing for activity classification in elderly people exploiting micro-Doppler signatures using machine learning. Sensors 21(11), 3881 (2021). https://doi.org/10.3390/s21113881
11. Du, Y., et al.: Radar-based Human Activity Classification with Cyclostationarity (2021)
12. Napolitano, A.: Cyclostationary Processes and Time Series: Theory, Applications, and Generalizations. Academic Press (2019)
13. Thangaraj, A., Kramer, G., Böcherer, G.: Capacity bounds for discrete-time, amplitude-constrained, additive white Gaussian noise channels. IEEE Trans. Inf. Theor. 63(7), 4172–4182 (2017). https://doi.org/10.1109/tit.2017.2692214
14. Wen, F., et al.: Angle estimation for bistatic mimo radar in the presence of spatial colored noise. Signal Process. 134, 261–267 (2017). https://doi.org/10.1016/j.sigpro.2016.12.017
15. Islam, M.S., Chong, U.: Noise reduction of continuous wave radar and pulse radar using matched filter and wavelets. EURASIP J. Image Video Process. 2014(1), 1–9 (2014). https://doi.org/10.1186/1687-5281-2014-43
16. Huether, B.M., Gustafson, S.C., Broussard, R.P.: Wavelet preprocessing for high range resolution radar classification. IEEE Trans. Aero. Electron. Syst. 37(4), 1321–1332 (2001). https://doi.org/10.1109/7.976968
17. Joy, J., Peter, S., John, N.: Denoising using soft thresholding. Int. J. Adv. Res. Electr. Electron. Instrum. Eng. 2(3), 1027–1032 (2013)
18. Lyons, R.G.: Understanding Digital Signal Processing, 3rd ed. Pearson (2010)
19. Cetin, A.E., Tofighi, M.: Projection-based wavelet denoising [lecture notes]. IEEE Signal Process. Mag. 32(5), 120–124 (2015). https://doi.org/10.1109/msp.2015.2440051
20. Walnut, D.F.: An Introduction to Wavelet Analysis. Springer Science and Business Media (2002)
21. Erol, B., et al.: Wideband radar based fall motion detection for a generic elderly. In: 2016 50th Asilomar Conference on Signals, Systems and Computers, pp. 1768–1772. IEEE (2016)
22. Gurbuz, S.Z., Amin, M.G.: Radar-based human-motion recognition with deep learning: promising applications for indoor monitoring. IEEE Signal Process. Mag. 36(4), 16–28 (2019). https://doi.org/10.1109/msp.2018.2890128
23. Li, X., He, Y., Jing, X.: A survey of deep learning-based human activity recognition in radar. Rem. Sens. 11(9), 1068 (2019). https://doi.org/10.3390/rs11091068
24. Jiang, H., et al.: Human Activity Classification Using Radar Signal and Rnn Networks (2021)
25. Lehtinen, J., et al.: Noise2noise: Learning Image Restoration without Clean Data (2018). arXiv preprint arXiv:1803.04189
26. Tian, C., et al.: Deep learning on image denoising: an overview. Neural Network. 131, 251–275 (2020). https://doi.org/10.1016/j.neunet.2020.07.025
27. Zhang, K., et al.: Learning deep cnn denoiser prior for image restoration. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017)
28. Lucas, A., et al.: Using deep neural networks for inverse problems in imaging: beyond analytical methods. IEEE Signal Process. Mag. 35(1), 20–36 (2018). https://doi.org/10.1109/msp.2017.2760358
29. Wu, D., Ren, H., Li, Q.: Self-supervised dynamic ct perfusion image denoising with deep neural networks. IEEE Trans. Radiat. Plasma Med. Sci. 5(3), 350–361 (2020). https://doi.org/10.1109/trpms.2020.2996566
30. Liu, P., et al.: Neural network evolution using expedited genetic algorithm for medical image denoising. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I, pp. 12–20. Springer (2018)
31. Zhao, H., et al.: A simple and robust deep convolutional approach to blind image denoising. In: Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (2019)
32. Yue, Z., et al.: Variational denoising network: toward blind noise modeling and removal. Adv. Neural Inf. Process. Syst. 32 (2019)
33. Dalsasso, E., Denis, L., Tupin, F.: As if by magic: self-supervised training of deep despeckling networks with merlin. IEEE Trans. Geosci. Rem. Sens. 60, 1–13 (2021). https://doi.org/10.1109/tgrs.2021.3128621
34. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Navab, N., et al. (eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, pp. 234–241. Springer International Publishing (2015)
35. Zhang, K., et al.: Beyond a Gaussian denoiser: residual learning of deep cnn for image denoising. IEEE Trans. Image Process. 26(7), 3142–3155 (2017). https://doi.org/10.1109/tip.2017.2662206
36. Zhang, K., Zuo, W., Zhang, L.: Ffdnet: toward a fast and flexible solution for cnn-based image denoising. IEEE Trans. Image Process. 27(9), 4608–4622 (2018). https://doi.org/10.1109/tip.2018.2839891
37. Zhou, Y., et al.: When Awgn-Based Denoiser Meets Real Noises (2019)
38. Du, M., et al.: Dncnet: deep radar signal denoising and recognition. IEEE Trans. Aero. Electron. Syst. 58(4), 3549–3562 (2022). https://doi.org/10.1109/taes.2022.3153756
39. Rock, J., et al.: Complex signal denoising and interference mitigation for automotive radar using convolutional neural networks. Fusion 2019 (2019)
40. Rey, S., et al.: Untrained graph neural networks for denoising. IEEE Trans. Signal Process. 70, 5708–5723 (2022). https://doi.org/10.1109/tsp.2022.3223552
41. Pan, Q., et al.: Two denoising methods by wavelet transform. IEEE Trans. Signal Process. 47(12), 3401–3406 (1999). https://doi.org/10.1109/78.806084
42. Liu, D., et al.: Connecting image denoising and high-level vision tasks via deep learning. IEEE Trans. Image Process. 29, 3695–3706 (2020). https://doi.org/10.1109/tip.2020.2964518
43. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, pp. 234–241. Springer (2015)
44. Chen, V.C.: The Micro-doppler Effect in Radar. Artech house (2019)
45. Müller, M., et al.: Mocap database hdm05. Institut für Informatik II, Universität Bonn 2(7) (2007)
46. Fioranelli, F., Le Kernec, J., Shah, S.A.: Radar for health care: recognizing human activities and monitoring vital signs. IEEE Potentials 38(4), 16–23 (2019). https://doi.org/10.1109/mpot.2019.2906977
47. He, K., et al.: Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1026–1034 (2015)
48. He, K., et al.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
49. Li, Z., et al.: Radar-based human activity recognition with adaptive thresholding towards resource constrained platforms. Sci. Rep. 13(1), 3473 (2023). https://doi.org/10.1038/s41598-023-30631-x
50. Li, Z., et al.: Multi-domains based human activity classification in radar. IET International Radar Conference (IET IRC 2020), vol. 2020, pp. 1744–1749. IET (2020)
51. Zhou, Z.-H.: Machine Learning. Springer Nature (2021)
52. Xiaolong, Z., Tian, J., Hao, D.: A Lightweight Network Model for Human Activity Classifiction Based on Pre-trained Mobilenetv2 (2021)

53. Guo, J., et al.: Complex field-based fusion network for human activities classification with radar. In: IET International Radar Conference (IET IRC 2020), vol. 2020, pp. 68–73. IET (2020)
54. Béranger, C., et al.: Radar-based human activity acquisition, classification and recognition towards elderly fall prediction. Euromicro DSD2023 (2023)
55. Kingma, D.P., Ba, J.: Adam: A Method for Stochastic Optimization (2014). arXiv preprint arXiv:1412.6980

How to cite this article: Yu, R., et al.: Radar-based human activity recognition using denoising techniques to enhance classification accuracy. IET Radar Sonar Navig. 18(2), 277–293 (2024). https://doi.org/10.1049/rsn2.12501

APPENDIX A

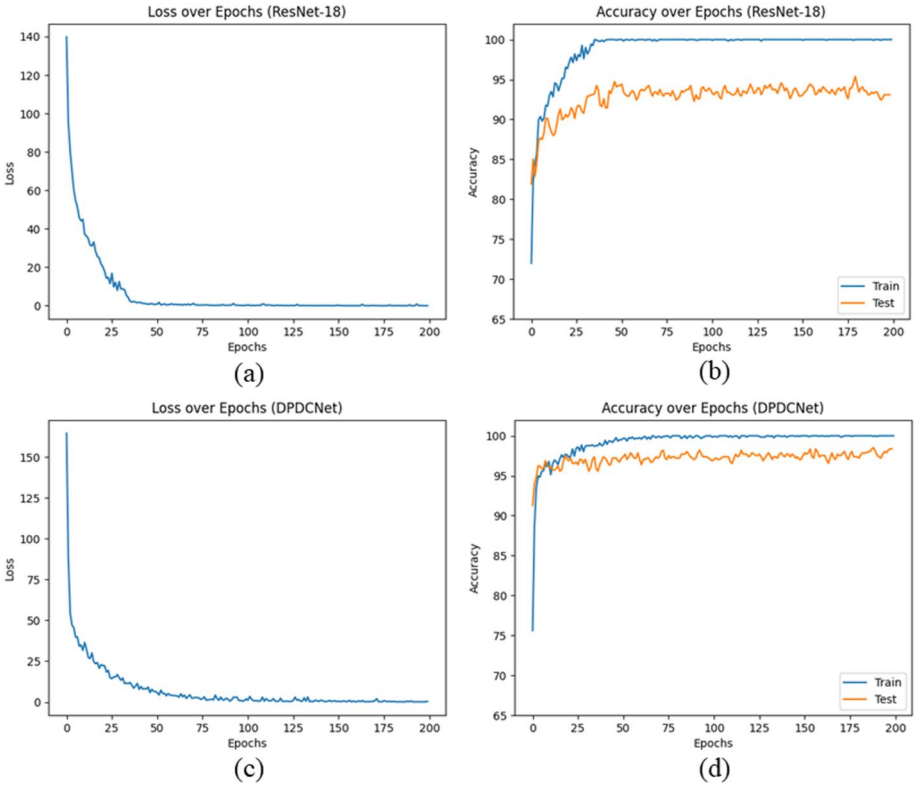

FIGURE A1 Loss and accuracy as a function of epochs of the DPDCNet training procedure when the SNR is −4 dB. (a) Loss over epochs of pretraining ResNet-18. (a) Accuracy over epochs of pretraining ResNet-18. (c) Loss over epochs of retraining ResNet-18 with a fixed denoising network. (d) Accuracy over epochs of retraining ResNet-18 with a fixed denoising network. The improvement in classification performance shows the effectiveness of the denoising.

TABLE A1 Denoising network architecture.

Layer (type)	Output shape
Input	[1, 768, 768]
Encoder block 1	
Conv2d-1	[48, 768, 768]
ReLU	[48, 768, 768]
Conv2d-2	[48, 768, 768]
ReLU	[48, 768, 768]
MaxPool2d	[48, 384, 384]
Repeat blocks for encoder	
Conv2d-3	[48, 384, 384]
ReLU	[48, 384, 384]
MaxPool2d	[48, 192, 192]
Decoder blocks	
Conv2d-4	[48, 192, 192]
ReLU	[48, 192, 192]
ConvTranspose2d	[48, 384, 384]
Concat and up blocks	
Conv2d-5	[96, 384, 384]
ReLU	[96, 384, 384]
Conv2d-6	[96, 384, 384]
ReLU	[96, 384, 384]
ConvTranspose2d	[96, 768, 768]
Final block	
Conv2d-7	[64, 768, 768]
ReLU	[64, 768, 768]
Conv2d-8	[32, 768, 768]
ReLU	[32, 768, 768]
Conv2d-9	[1, 768, 768]
LeakyReLU	[1, 768, 768]

TABLE A2 Classification network architecture.

Layer (type)	Output shape
Input	[2, 768, 768]
Conv2d-1	[64, 384, 384]
BatchNorm2d	[64, 384, 384]
ReLU	[64, 384, 384]
MaxPool2d	[64, 192, 192]
Residual block 1	
Conv2d-2	[64, 192, 192]
BatchNorm2d	[64, 192, 192]
ReLU	[64, 192, 192]
Conv2d-3	[64, 192, 192]
BatchNorm2d	[64, 192, 192]
ReLU (with skip connection)	[64, 192, 192]
Residual block 2	
Conv2d-4	[128, 96, 96]
BatchNorm2d	[128, 96, 96]
ReLU	[128, 96, 96]
Conv2d-5	[128, 96, 96]
BatchNorm2d	[128, 96, 96]
ReLU (with skip connection)	[128, 96, 96]
Residual block 3	
Conv2d-6	[256, 48, 48]
BatchNorm2d	[256, 48, 48]
ReLU	[256, 48, 48]
Conv2d-7	[256, 48, 48]
BatchNorm2d	[256, 48, 48]
ReLU (with skip connection)	[256, 48, 48]
Residual block 4	
Conv2d-8	[512, 24, 24]
BatchNorm2d	[512, 24, 24]
ReLU	[512, 24, 24]
Conv2d-9	[512, 24, 24]
BatchNorm2d	[512, 24, 24]
ReLU (with skip connection)	[512, 24, 24]
AvgPool2d	[512, 1, 1]
Fully connected	[512, 6]
Softmax	[6]

TABLE A3 Description of training parameters of DPDCNet.

	Pretrain classification ResNet	**Pretrain denoising U-net**
Epoch	50	50
Batch size	13	13
Patch size	$2 \times 768 \times 768$	768×768
Learning rate	10^{-3}	10^{-3}
Optimiser	Adam [55]	Adam
Loss type	Cross-entropy	Mean square error (MSE)
	Combinationally train denoising U-net	**Retrain classification ResNet**
Epoch	100	200
Batch size	13	13
Patch size	768×768	$2 \times 768 \times 768$
Learning rate	10^{-3}	10^{-3}
Optimiser	Adam	Adam
Loss type	Cross-entropy + MSE	Cross-entropy