Accepted Manuscript

Open-Set Human Activity Recognition Based on Micro-Doppler Signatures

Yang Yang, Chunping Hou, Yue Lang, Dai Guan, Danyang Huang, Jinchen Xu

PII: S0031-3203(18)30270-X
DOI: https://doi.org/10.1016/j.patcog.2018.07.030
Reference: PR 6622

To appear in: *Pattern Recognition*

Received date: 1 December 2017
Revised date: 24 May 2018
Accepted date: 31 July 2018

Please cite this article as: Yang Yang, Chunping Hou, Yue Lang, Dai Guan, Danyang Huang, Jinchen Xu, Open-Set Human Activity Recognition Based on Micro-Doppler Signatures, *Pattern Recognition* (2018), doi: https://doi.org/10.1016/j.patcog.2018.07.030

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Open-Set Human Activity Recognition Based on Micro-Doppler Signatures

Yang Yang[a], Chunping Hou[a], Yue Lang[a,*], Dai Guan[a], Danyang Huang[a], Jinchen Xu[a]

[a]*School of Electrical and Information Engineering, Tianjin University, Tianjin, China*

Abstract

Open-set activity recognition remains as a challenging problem because of complex activity diversity. In previous works, extensive efforts have been paid to construct a negative set or set an optimal threshold for the target set. In this paper, a model based on Generative Adversarial Network (GAN), called 'OpenGAN' is proposed to address the open-set recognition without manual intervention during the training process. The generator produces fake target samples, which serve as an automatic negative set, and the discriminator is redesigned to output multiple categories together with an 'unknown' class. We evaluate the effectiveness of the proposed method on measured micro-Doppler radar dataset and the MOtion CAPture (MOCAP) database from Carnegie Mellon University (CMU). The comparison results with several state-of-the-art methods indicate that OpenGAN provides a promising open-set solution to human activity recognition even under the circumstance with few known classes. Ablation studies are also performed, and it is shown that the proposed architecture outperforms other variants and is robust on both datasets.

Keywords: Open-set Recognition, Generative Adversarial Network (GAN), Human Activity, Micro-Doppler Radar

*Corresponding author

Email addresses: yang_yang@tju.edu.cn (Yang Yang), hcp@tju.edu.cn (Chunping Hou), langyue@tju.edu.cn (Yue Lang), gd3012203096@tju.edu.cn (Dai Guan), hdy137851@gmail.com (Danyang Huang), xujinchen@tju.edu.cn (Jinchen Xu)

ACCEPTED MANUSCRIPT

1. Introduction

In the last decades, human activity recognition has aroused general concern in numerous fields [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12] due to the increasing demand for entertainment, medical monitoring, security, emergency rescue and other applications. Human activity recognition has relied on visual data for decades and develops rapidly with the help of computer vision. Most available studies focus on analyzing human activities based on image sequences captured by optical sensors. Confronting severe interference caused by the weather, object shelter and light condition, optical sensors are unsuitable for recording activities in the severe environment. Compared with optical sensors, radar has some unique properties. It is invariable to the environment and is capable of detecting objects through the wall. What's more, radar could operate long-distance detections regardless of shelters. Radar-based perception has broad application prospects such as survival search, enemy status perception, and terrorist detection.

Human activity recognition using radar data aims at automatically recognizing human motions from radar spectrograms. When radar echoes are modulated by human activities, there will be micro-Doppler signatures motivated by micro-movements. Micro-Doppler frequency varies with the velocity of a moving target so that each movement has its unique micro-Doppler signatures, which can be used for activity recognition. Conventional Doppler-based activity recognition requires well-designed features from micro-Doppler spectrogram to support an effective training process of classifiers [13, 14]. Substantial efforts have been dedicated to developing discriminating visual feature descriptors [15], and deep learning is employed due to its superior ability of automatically learning from data.

Deep learning has achieved significant success in the field of image classification, object detection and semantic segmentation [16, 17, 18, 19, 20, 21, 22]. Among diverse deep neural networks, GANs attract great attention recently and have produced abundant accomplishments such as visually realistic images generation [23], super resolution [24] and style transfer [25]. Typically, a GAN

ACCEPTED MANUSCRIPT

contains a generator and a discriminator. The generator tries to produce synthetic samples, while the discriminator learns to determine whether a sample is derived from the generator or the real dataset [26].

Open-set recognition is a relatively new topic at present. Most available classifiers are for 'closed-set' recognition [27, 28, 29], in which the training set and testing set contain the same categories [30]. In practice, the world consists of countless elements, and it is impractical to enumerate all the targets and label them seriatim. Human activity recognition is confronted with the same problem, not only are there marvelous types of activities to be classified, but also the same activity varies when performed by different people. Therefore, a more realistic description to the classification of the real world is an 'open-set recognition'. A natural idea of solving open-set problems is to construct a negative dataset so that they can be transformed into closed-set issues. Existing open-set approaches usually build the negative set by introducing a large scale dataset that contains the target classes, then the other classes in the dataset, namely complementary set of target samples, are utilized as the negative set to train the open-set classifier. As there have been some published datasets of optical images, such as ImageNet, the method mentioned above achieves satisfying performances. However, the lack of large-scale micro-Doppler dataset makes the method inapplicable to micro-Doppler data. Another idea is to collect the negative samples manually, but it is both time-consuming and labor-intensive. Therefore, the open-set recognition on radar data remains an open problem.

In this paper, we propose an open-set recognition network, which is called 'OpenGAN', based on the scheme of GAN. During the training phase, the fake samples from the generator are used as the negative set, while the labeled target samples on multiple classes compose the target set. After alternative training, the output vectors from the discriminator are mapped to probabilities on each known or unknown class by the softmax function. The well-trained discriminator is employed as the open-set classifier in the testing phase. OpenGAN is an integration of a one-class classifier and an open-set classifier, which could not only distinguish target samples from negative samples but also recognize the

ACCEPTED MANUSCRIPT

category of each target sample at the same time. Experimental results on measurement dataset and CMU MOCAP dataset demonstrate that the proposed framework is superior to several state-of-the-art one-class classifiers and open-set recognition methods. Moreover, to learn more effective features, the network adopts the dense connection architecture. We evaluate the performances of different architectures in Section 6, and the results prove that the network with one sub-layer in a dense block outperforms plain CNN in most cases and is better than the architectures with more sub-layers.

The proposed open-set recognition method has several advantages over the previous algorithms. Firstly, the negative net in our work is constructed automatically, leaving out the artificial selection process. Secondly, as the number of negative samples in this framework keeps increasing, the negative set is dynamically updated during the training process. This could prevent the discriminator from over-fitting, which easily occurs on artificial negative set. Thirdly, negative sets in the previous works are far from target sets, so the boundary of target samples is easy to acquire. In this work, however, the distribution of synthesized samples from the generator is highly similar to the target samples, so the discriminator trained with this these samples could achieve better discriminability than other classifiers. Finally, as OpenGAN follows the alternative training principle in the GAN, the discriminator is trained more sufficiently after multiple iterations.

The remainder of this paper is organized as follows. A brief review of related works and basic concepts are provided in Section 2. Section 3 presents some preliminaries for open-set problems, including the description of datasets and the evaluation protocols. In Section4, we elucidate the proposed OpenGAN architecture in detail. The experiment results and comparisons are presented in Section 5. In Section 6, we display some ablation studies of the proposed method and state a conclusion in Section 7.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

2. Related works

This section provides an overview of literature in the areas related to our work, including Doppler-based human activity recognition, open-set recognition and some research on GAN.

2.1. Human Activity Recognition

Considerable achievements about activity recognition have been published using natural images, which are captured by optical cameras [31, 6]. The studies were promoted by the introducing of depth sensors. Some new consumer technologies, such as Kinect [32, 33, 34] (a low-cost RGB-D camera invented by Microsoft for its X-Box gaming platform), provided researchers with an easy access to depth information, so that more attention has been paid to activity recognition tasks [35, 36, 37, 38, 1, 7, 3]. As these kinds of sensors have some drawbacks mentioned in Section 1, radar-based methods are considered as a promising solution to this area.

'Doppler effect' (or 'Doppler shift') is caused by a constant moving target, for the duration, micro-motions (such as vibration and rotation) of any parts of a target will bring about micro-Doppler effect. Micro-motions of hands, arms, head, torso, legs, and feet introduce unique micro-Doppler signatures of each activity, which can be distinguished from others. The signature can be represented by spectrograms for visualization, which are generated by taking Short Time Fourier Transform (STFT) on echo signals. There have been a group of studies utilizing micro-Doppler signature for activity classification. Earlier studies focused on exploring active handcrafted features from radar spectrograms [39, 40], *e.g.*, in [14], Y. Kim and H. Ling extracted six micro-Doppler features (torso Doppler frequency, total bandwidth (BW) of the Doppler signal, the offset of the total Doppler, *etc.*) from the resulting spectrograms, they then proposed an artificial neural network (ANN) to classify activities of a human subject using these features. In [39] they trained a support vector machine (SVM) to classify the activities with the same kinds of features and achieved an accuracy of more

ACCEPTED MANUSCRIPT

than 90%. Dustin P. Fairchild and Ram M. Narayanan used the empirical mode decomposition to get a unique feature vector from the human micro-Doppler signals and classify the activities with an SVM [13]. Recently, CNN has drawn great attention because of its remarkable performances in various tasks, this is mainly due to its non-linear mapping of images and superior ability to extract latent features. [28, 29, 41] applied CNN to the micro-Doppler spectrograms to classify human activities. These works are all based on closed training and testing set, which are not able to recognize unknown activities from target ones.

2.2. Open-set recognition

Open-set recognition is considered as intractable because training set is usually unable to involve the entire knowledge of the testing set [42, 43]. W. Scheirer *et al.* introduced a 1-vs-Set Machine, which sculpts a decision space from the marginal distances of a 1-class or binary SVM with a linear kernel, making a preliminary investigation to address the open-set recognition problem [30]. Then in [44], they proposed a model called Compact Abating Probability (CAP) and combined CAP with the statistical Extreme Value Theory (EVT) to formulate a novel technique called the Weibull-calibrated SVM (W-SVM), which was shown to improve the result of recognition. Abhijit Bendale and Terrance Boult extended Nearest Class Mean type algorithms (NCM) [45], to a Nearest Non-Outlier (NNO) algorithm that could balance open space risk and accuracy [44]. The combination of deep network and open-set recognition was presented in [46], a new model layer called OpenMax was introduced. In the OpenMax layer, the Meta-Recognition score from the penultimate layer (defined as 'activation vector') was used to reject samples far from known inputs. This model demands pre-trained deep neural network model on ImageNet, which makes it unfeasible for other tasks. In general, the performances of these methods strongly relies on threshold values so it may lack robustness when encounters novelties. What's more, all the open-set algorithms are evaluated on natural image datasets such as ImageNet and Caltech, open-set recognition for human activities, especially on micro-Doppler spectrograms, has not been available so far.

ACCEPTED MANUSCRIPT

2.3. Generative adversarial network

After being proposed in 2014 [26], GAN is becoming an emerging research area and has been used for a variety of applications [47]. Radford *et al.* [23] proposed a CNN-based GAN architecture called DCGAN, which successfully implement unsupervised representation learning. Arjovsky *et al.* [48] optimized the traditional GAN training with an alternative cost function to get rid of mode collapse in GAN and improve the stability of learning. The function was derived from the earth mover distance, and it is constrained to Lipschitz functions. There are some conditional GANs that could take advantage of class conditions for better generation of multi-modal data [49] and some interesting applications of GANs arose quickly [50, 51, 52, 25]. Most of the works make efforts to present meaningful and semantic synthetic images from the generator, managing to generate more visually lifelike samples. The discriminator, however, has been undervalued. The potential of the discriminator of deciding 'real' or 'fake' is utilized in our work to address the open-set activity recognition problem and for the first time, GAN is introduced into the recognition of micro-Doppler radar spectrograms.

3. Preliminaries

3.1. Data description

In this paper, two datasets are employed to evaluate our open-set recognition strategy. One is the measurement dataset collected by real experiments, the other is the public MOCAP dataset.

Measurement database. Measurement human activity data are collected with an ultra-wideband (UWB) radar module named PulsON 440 (P440). The radar operates between 3.1 and 4.8 GHz with two directional antennas.

The experiments are operated in an indoor environment, and the radar is placed at 1.2-meter height, in order to match the human's centers of gravity. Four subjects perform seven kinds of activities towards the radar in line-of-sight

ACCEPTED MANUSCRIPT

direction. Seven referred activities are listed as follows: (a) boxing while standing in place, (b) crawling on the ground, (c) creeping, (d) jumping forward, (e) running, (f) standing (with slight movements) and (g) walking. The measuring process for 'boxing' is illustrated in Fig. 1(a). Each activity is performed two to four times, and the measurement range is between 1.2 meters and 5.4 meters. The duration of each scenario is about seven seconds. The data are collected over a sampling frequency of 16 GHz, the pulse recurrence frequency (PRF) of Doppler radar is 368 Hz, and the coherent pulse interval (CPI) is 0.2s. Subsequently, we apply STFT with 1024 points on the complex echo data to obtain corresponding spectrograms. One-second slow time window with an overlap of 0.9 is employed to the measured data. After these processing, each activity obtains more than 900 images, which are resized to 120×120. Then we choose 400 images as the training set and 100 images as the testing set for each class.

Fig. 1(b) shows the typical sample spectrograms of 'boxing'. The vertical axis of a spectrogram is radial velocity, while the horizontal axis represents time. Pixels in spectrograms indicate the intensities of human backscattering echoes, which implicitly depict the micro-motions of the human body.

MOCAP database. We also evaluate the proposed approach on the MOCAP database from CMU. It consists of 2605 experiment trials in six main categories (namely, human interaction, interaction with the environment, locomotion, physical activities and sports, situations and scenarios, and test activities) and 23 subcategories. The subject wears a jumpsuit with 41 markers fixed on it, and a Vicon motion capture system with twelve infrared MX-40 cameras is used to capture movements of the skeleton. Fig. 1(c) depicts the data collection process of MOCAP database.

To generate micro-Doppler spectrograms from MOCAP data, we employ the ellipsoid-based human backscattering model in [53].The model used 31 joints to describe a human body, and every two adjacent joints define a body segment. The segment is estimated as an ellipsoid, whose centroid is considered as a moving point scatterer. For each scatterer, the radar echo can be simulated by a Sinc function whose width is inversely proportional to the bandwidth of the

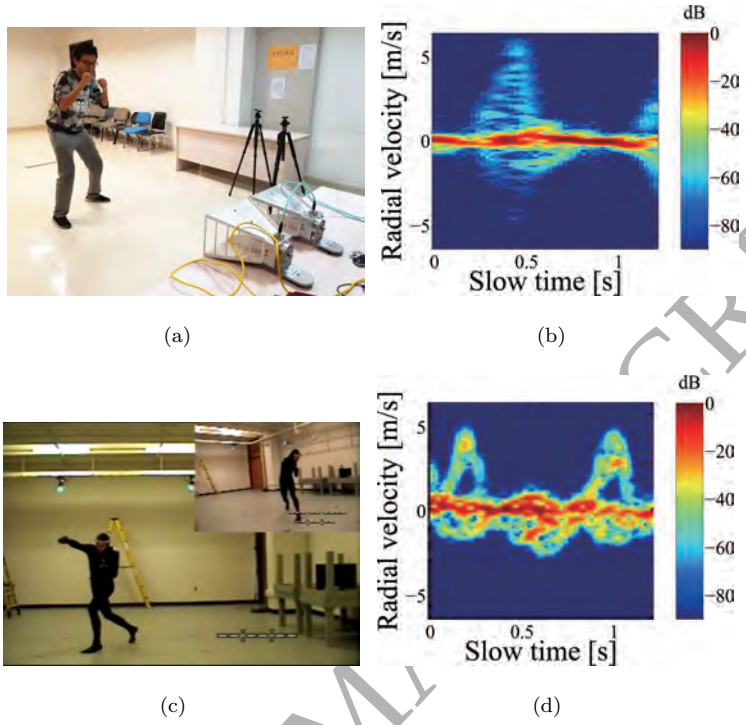


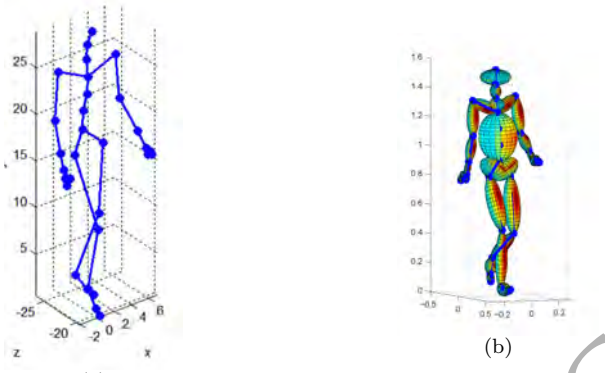
Figure 1: Data collection processes of 'boxing' and corresponding spectrograms. (a) Measurement data collection. (b)Measurement spectrogram. (c) MOCAP data collection. (d) MOCAP spectrogram.

radar system. Then the whole body echo is obtained by coherently summing echoes from multiple scatters. Fig. 2 illustrates the ellipsoid-based human motion model.

We set the frequency band of simulated UWB module as 3-5 GHz, and the PRF is 600 Hz. The CPI is 0.1s. Other post-processing processes are following the same protocol as the measurement data including the dataset partition. The generated spectrogram is shown in Fig. 1(d). For each activity, there are 400 spectrograms in the training set and 100 in the testing set.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT



(a) (b)

Figure 2: The ellipsoid-based human motion model. (a) Joints on the human body. (b) Ellipsoid-based backscattering model.

3.2. Evaluation protocol

3.2.1. Openness

The open-set recognition problems have various grades, which are related to the number of target classes in the training and testing set. Following the research in [30], we use a universal index called 'openness' to measure the grade of an open-set problem. 'Openness' is defined as:

$$Openness = 1 - \sqrt{\frac{2 \times N_{TA}}{N_{TG} + N_{TE}}} \tag{1}$$

where N_{TA} is the number of training classes, N_{TG} is the number of target classes, and N_{TE} denotes the number of testing classes. 'Openness=0' represents a closed-set classification (*i.e.*, the commonest classification task), in which the testing set contains the same classes as the training set. If the testing set is fixed, increasing the number of target classes or decreasing the number of training classes will enlarge the value of openness, which indicates a more 'open' problem. It is notable that in previous works, target classes and training classes may be different since some classifiers require negative training classes. But for the activity recognition task in this work, the training and target classes remain the same as no negative samples are introduced for training. So, when

ACCEPTED MANUSCRIPT

the testing set keeps constant, the openness depends on the number of target activities.

3.2.2. F-measure

For measuring the performance of open-set algorithms, we adopt the commonly used open-set criterion called 'F-measure' [30]. The F-measure is the harmonic mean of precision and recall, which are:

$$precision = \frac{TP}{TP + FP} \tag{2}$$

$$recall = \frac{TP}{TP + FN} \tag{3}$$

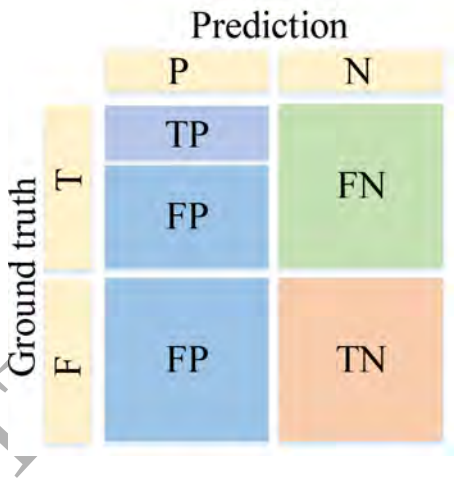


Figure 3: Confusion matrix of open-set problem. T: True; P: Positive; F: False; N: Negative.

Here, TP denotes correctly classified positive samples among all the predicted positive samples (which is slightly different from the one in binary classification problems), and FP includes both misclassified positive samples and miss predicted ones (shown in Fig. 3). So that F-measure evolves into:

$$F\text{-}measure = 2 \times \frac{precision \times recall}{precision + recall} \tag{4}$$

The range of F-measure is [0, 1] and a higher value indicates a more discriminative open-set classifier.

ACCEPTED MANUSCRIPT

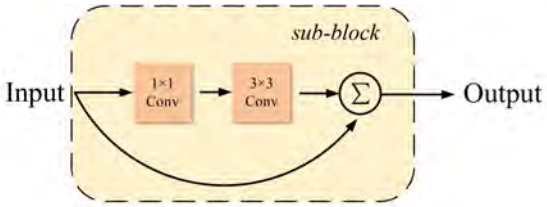


Figure 4: Sub-block structure.

4. OpenGAN model

4.1. Network architecture

The generator and discriminator in OpenGAN are composed of two main parts: (1) dense blocks for feature extraction and combination of multi-level features, (2) transition layers to connect two blocks. The parameters of the network are empirically determined.

Dense block. OpenGAN employs the dense block architecture instead of traditional CNN architecture for both generator and discriminator, which is inspired by [19]. For convenience, we introduce '*sub-block*' to explain the block more clearly. A *sub-block* is the combination of two convolutional layers and a concatenation layer, connecting the feature flow from preceding layers. Each convolutional layer is followed by Batch Normalization (BN) [54] and a Rectified Linear Units (ReLU) [55] (for the generator) or a Leaky ReLU (for the discriminator). Fig. 4 illustrates sub-block structure schematically. After exploration experiments, we apply one sub-block to the dense blocks in our experiments as it reveals better performance.

Transition layer. A transition layer is a stack of layers between two dense blocks. In the generator, the transition layer includes a convolutional layer and a de-convolutional layer. In the discriminator, the transition layer is similar to DenseNet with a convolutional layer and an average pooling layer.

Parameter details. For all experiments in our work, OpenGAN has three dense blocks in the generator and discriminator respectively.

The generator first produces a 1×1×100 small spatial extent tensor, then

ACCEPTED MANUSCRIPT

a de-convolution with 64 outputs is performed on this tensor. The stacked feature maps will be extended to 128 channels by a 1×1 convolutional layer in the dense block, and a 3×3 convolution with 1 padding will subsequently be employed to keep the output size fixed. Between two adjacent dense blocks, a 1×1 convolutional layer and a 2×2 de-convolution layer constitute the transition layer, aiming to double the size of feature maps and expand the size of feature maps gradually. After three dense blocks, the feature maps will be scaled to 32×32 with 512 channels Finally, a 1×1 convolution layer with 3 output channels integrates them into a 32×32 synthetic sample.

The processing in discriminator is basically same as the generator, except that the de-convolution in transition layer is replaced by 2×2 average pooling to operate downsampling. At the end of the third dense block, the feature maps are scaled to a column vector and the linear mapping in the last layer transfers the vector to an n+1 dimension vector, in which n refers to the number of target classes. Then the vector is processed by softmax mapping, and the cost is calculated by Mean Square Error (MSE) loss. Finally, the values in the vector refer to the input sample's probabilities on target categories and unknown class. Fig. 5 and Fig. 6 illustrate the architectures of the generator and discriminator respectively, and the parameters of the network are shown in Table 1 in detail.

4.2. Adversarial training

During the training process, generator and discriminator operate the same as in typical GAN networks. The generator produces synthetic samples that are highly similar to the target samples as the automatic negative set. Then the synthetic samples and target samples are alternately input to the discriminator, enhancing the discernibility of it. At the end of each iteration, the weights of generator and discriminator are stored separately. After training, only the discriminator model is used in the testing phase as an open-set classifier. For movement detection task, each testing sample fed into the discriminator will get a binary judgment of 'target or not', while for activity recognition task, the judgment would be 'unknown' or the class of testing sample.

ACCEPTED MANUSCRIPT

Table 1: Detailed parameters of each layer in OpenGAN. For each convolutional layer, the stride is 1. For each deconvolutional and average pooling layer, the stride is 2.

Model	Generator	Discriminator
Input Layer	4×4 deconv	1×1 conv
Dense Block	1×1 conv	1×1 conv
(1)	3×3 conv, Padding=1	3×3 conv, Padding=1
Transition Layer	1×1 conv	1×1 conv
(1)	2×2 deconv	2×2 Pooling
Dense Block	1×1 conv	1×1 conv
(2)	3×3 conv, Padding=1	3×3 conv, Padding=1
Transition Layer	1×1 conv	1×1 conv
(2)	2×2 deconv	2×2 Pooling
Dense Block	1×1 conv	1×1 conv
(3)	3×3 conv, Padding=1	3×3 conv, Padding=1
Transition Layer	1×1 conv	1×1 conv
(3)	2×2 deconv	2×2 Pooling
Output Layer	1×1 conv	Linear

The objective function in our model is defined as follows:

$$\underset{\hat{x}\sim G(z)}{E}\left(D\left(\hat{x}\right)\right)-\underset{x\sim real}{E}\left(D\left(x\right)\right)+\lambda E\left[\left(||\nabla_t D(t)||_2-1\right)^2\right] \tag{5}$$

where λ is 10 and α is a random value ranges from $\frac{1}{batchsize}$ to 1 and t refers to:

$$t=\alpha x+\left(1-\alpha\right)\widehat{x} \tag{6}$$

We introduce the gradient penalty [56] in our model to prevent the discriminator from fast convergence. When calculating the penalty term gradient in Eq. 5, if the positive samples or negative samples are used for training solely, the second derivative will be biased toward on either class and get out-of-balance. This restrains the optimization process, so the strategy based on gradient penalty should take both kinds of samples into account (shown in Eq. 6).

ACCEPTED MANUSCRIPT

4.3. Implementation details

We train our model on PyTorch, a recently widely used framework provided by Facebook. Training spectrograms are scaled to 32×32. The network parameters are updated with an Adaptive Moment Estimation (Adam) optimizer, and the mini-batch size is 256. At the early training stage, the discriminator convergences faster than the generator. Tto balance the training processes of these two parts of the network, we set the learning rates for the generator and the discriminator to 10^{-2} and 3×10^{-4} respectively. The momentum term remains as the suggested value of 0.9. All experiments are performed on an Intel I7-7700k CPU with 8 GB memory. A GeForce GTX 1080Ti GPU with 11 GB memory is used for training with CUDA for acceleration.

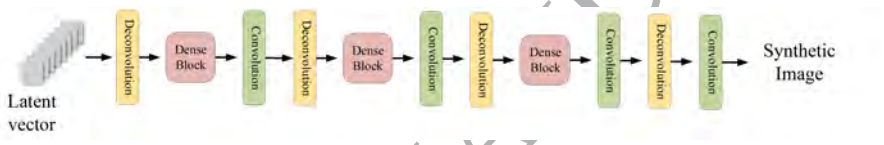


Figure 5: The architecture of generator in OpenGAN.

Figure 6: The architecture of discriminator in OpenGAN.

5. Experiments and results

In this section, the performances of OpenGAN on human movement activity recognition are verified on the measured micro-Doppler radar dataset and compared with several published algorithms. As these conventional methods strongly demand well-selected features, we choose three features for micro-Doppler spectrograms in this work:

ACCEPTED MANUSCRIPT

- Gabor: Following [30], Gabor feature, which has produced impressive results in computer vision tasks, is first employed in this work.

- Spectrograms Envelopes (SE) [57]: Envelopes are unique features of spectrograms. The maximum, minimum, and central velocities of each time slice can be applied to indicate the extremities and torso movement characters. The cumulative amplitude distribution of each column in the spectrogram is calculated as:

$$P(v,t) = \frac{\sum_{v=v_{\min}}^{v} y(v,t)}{\sum_{v=v_{\min}}^{v=v_{\max}} y(v,t)} \tag{7}$$

  where t denotes time, v is the velocity at the time, and y is the cumulative amplitude of each column. The upper, central and lower envelopes are extracted by setting the thresholds of $P(v, t)$ being 0.77, 0.5 and 0.28 respectively. The coordinates of the envelopes are concatenated for training, so each 120×120 spectrogram in this work will produce a 360-dimension feature vector.

- Means of Envelopes (ME) [58]: The means of envelopes are based on the coordinates mentioned above, once the envelopes are obtained, their average values can be calculated and form a three-dimension feature.

The popular Histogram of Oriented Gradients (HOG) [59] descriptor is not applied because it shows poor performances in the preliminary exploring experiments, we assume that this is due to the different properties of radar spectrograms and natural images.

5.1. OpenGAN for human movement detection

As can be seen from Fig. 7, human activity may not continue throughout the measurement scenario, the duration before or after activity will interfere with the performance of open-set recognition of human activity. Human activity

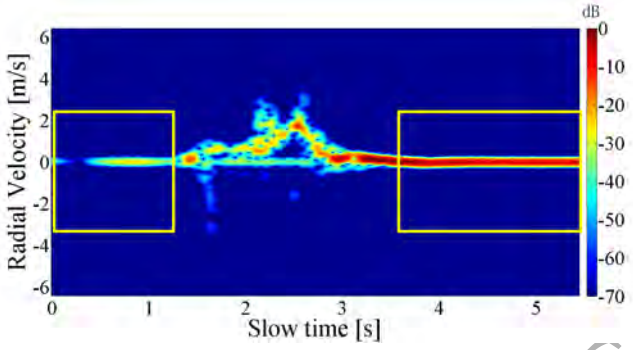
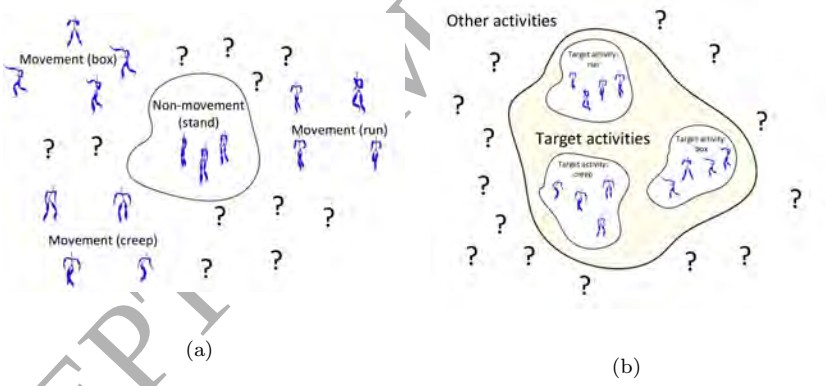


Figure 7: Complete measurement spectrogram. The signatures in the bounding boxes represent 'non-movement'.

detection can provide valuable information regarding an individual's degree of movement and is useful for monitoring the occurrence of the activity. To this end, human movement detection is a pre-processing for activity recognition.



(a) (b)

Figure 8: Illustrations of open-set problem. (a) Human activity detection. (b) Human activity recognition.

The problem can be concretized as a one-class classification problem, which is distinguishing 'standing' from other activities (illustrated in Fig. 8(a)). As Doppler signatures are caused by target movement, the spectrogram of 'standing' has fewer micro-Doppler signatures (such as the trajectories of hands and feet) than other activities. In broad terms, all the still postures generate slight

ACCEPTED MANUSCRIPT

Doppler signatures with a velocity of 0, so it is reasonable to represent 'non-movement' with standing.

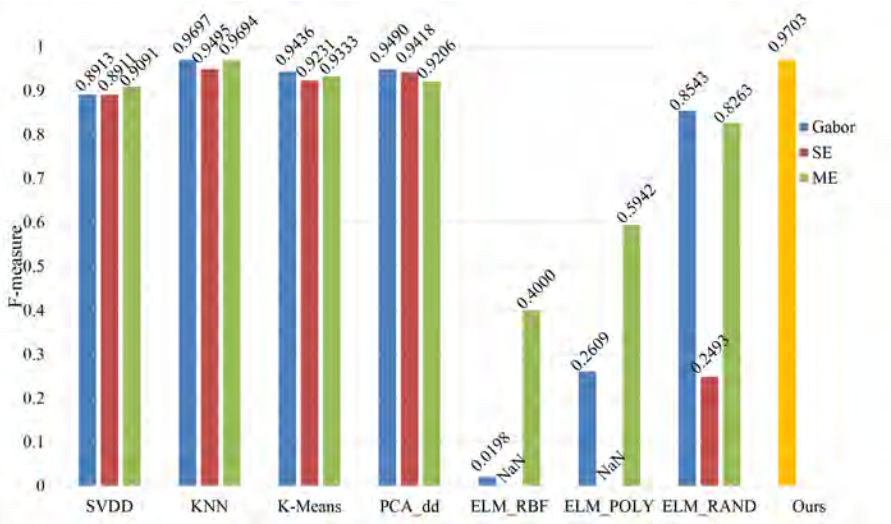
In this context, to tackle human activity detection problem with OpenGAN, we train the model with the spectrograms of 'standing' only, and the test samples of seven activities are collected as the testing set. Then the proposed model is compared with several well-performed one-class classifiers. The competing algorithms in this task include Support Vector Domain Description (SVDD) [60], K-Nearest-Neighbor (KNN) [61], K-means [62, 63], Principal Component Analysis data description (PCA_dd) [64], Extreme Learning Machine (ELM) [65] with Radial Basis Function (RBF) kernel, ELM with polynomial kernel, and ELM with random kernel. All of them are implemented with MATLAB toolbox 'dd_tools' [66] with experimental optimal parameter settings.

The comparison results are summarized in Fig. 9(a) and Fig. 9(b), it can be observed that all the comparison methods obtain lower results on MOCAP database than on measurement database. Among them, KNN achieves better results than other algorithms. Compared with these, OpenGAN outperforms the prior methods under all circumstances both on measurement dataset and MOCAP dataset, showing great robustness and impressive discriminative power with F-measures more than 0.97.

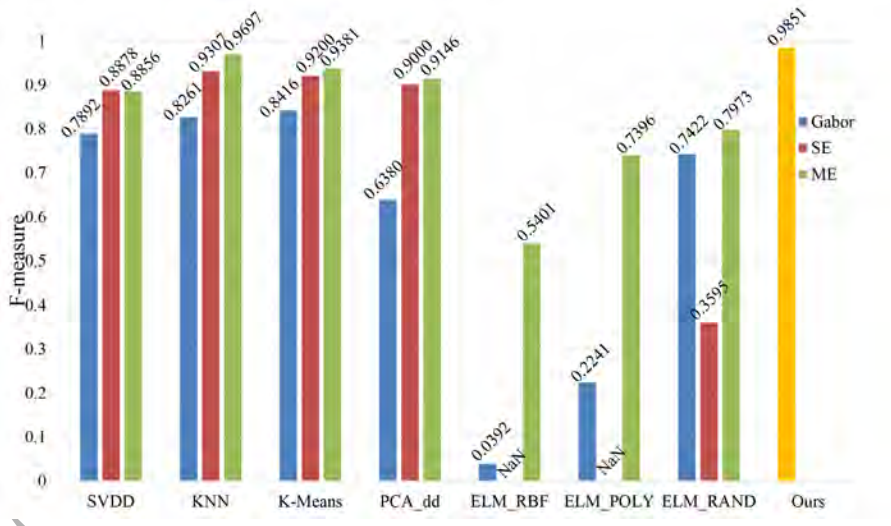
5.2. OpenGAN for human activity recognition

Given a set of micro-Doppler spectrograms, the problem of open-set activity recognition is to find the activities that we are interested in and the categories they belong to (illustrated in Fig. 8(b)). In practical terms, it is hard to enumerate all the activities, which remains open-set recognition an unsolved problem. Fortunately, in most cases we are interested in finite particular activities so that only the target activities samples need to be collected and used as training data, this relieves the open-set problem considerably.

To ensure the extensiveness, we conduct experiments on various openness values by verifying the number of target activities. For each openness value, all the configurations of the training set are enumerated and used to validate



(a)



(b)

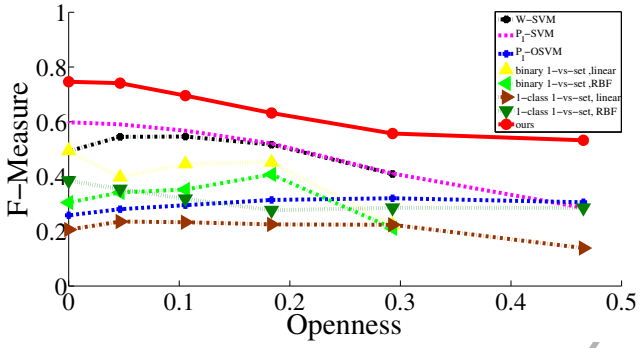
Figure 9: F-measures of open-GAN and comparison methods in human activity detection task. (a) Results on measurement datasets. (b) Results on MOCAP dataset.

ACCEPTED MANUSCRIPT

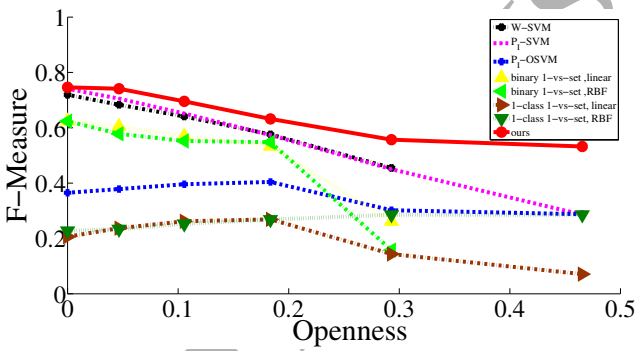
the effectiveness of OpenGAN, then the results are presented in the form of mean value. Also, OpenGAN is compared with several state-of-the-art open-set algorithms, including (1) W-SVM [67], (2) P_I-SVM [68], (3) P_I-OSVM [68], (4) 1-class 1-vs-set machine with linear kernel [30], (5) 1-class 1-vs-set machine with RBF kernel [30] (6) binary 1-vs-set machine with linear kernel [30] (7) binary 1-vs-set machine with RBF kernel [30]. All the competing algorithms are implemented by libsvm-openset package [69].

According to the experimental results shown in Fig. 10 and Fig. 11, the proposed method outperforms competing approaches in terms of F-measure. It can be observed from the figures that, our model achieves about 0.8 F-measure on a complete closed-set classification problem (*i.e.*, openness=0), which is superior to all the other algorithms with Gabor feature and mean envelope feature, only W-SVM and P_I-OSVM using envelope feature achieve comparable results to our model. With the increase of openness, F-measures of all the algorithms decline. Among the competing methods, W-SVM and P_I-OSVM tend to perform better than others, but they are still inferior to OpenGAN. It is worth noting that, W-SVM and binary 1-vs-set machine require negative samples for training, so it is not valid when openness reaches 0.4655, namely only one target activity is in the training set. However, our model retains the F-measure to about 0.5 in this situation, which is 12% higher than other algorithms. These observations imply the effectiveness of OpenGAN in solving open-set problems, even if the problem is highly 'open', the proposed model is still capable of offering optimal performance.

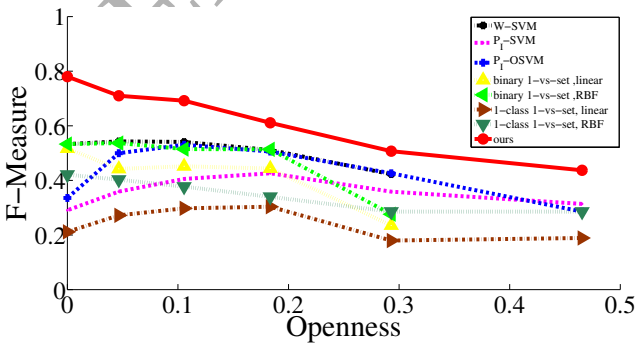
To better explain the efficacy of OpenGAN, we visualize the outputs from the discriminator at different training stages. Each input sample of the discriminator produces a feature vector, then the vectors are reduced to two dimensions with t-SNE [70] for visualization. Fig. 12 shows that during the network training, the mapped target set and negative set gradually become separable. Moreover, samples from the same target class become compact while the distances between different target classes become farther.



(a)

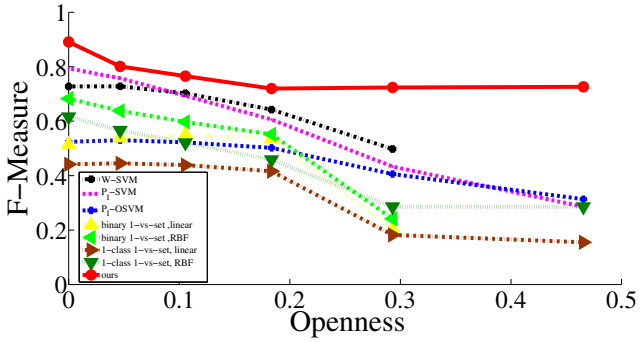


(b)

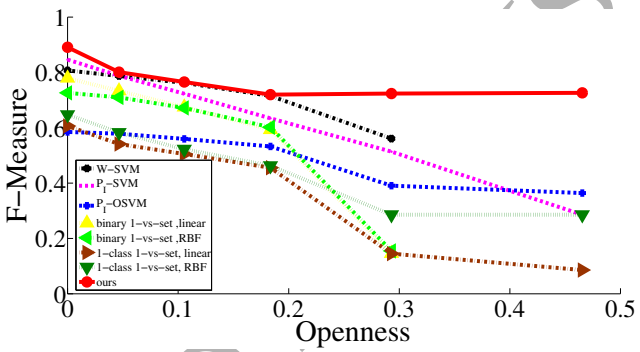


(c)

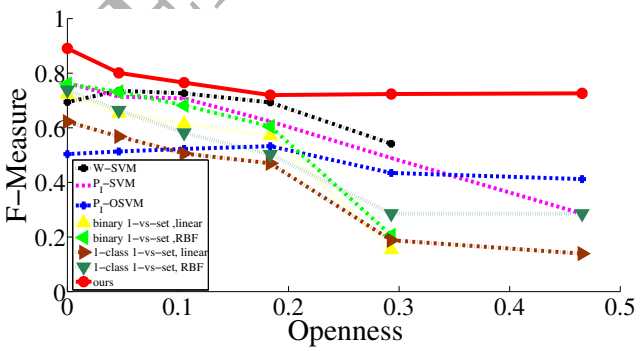
Figure 10: Comparison results on measurement database. (a) With Gabor feature. (b) With spectrograms envelopes feature. (c) With means of envelopes feature.



(a)



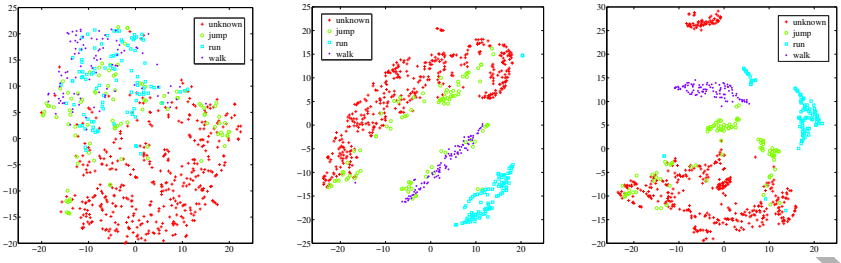
(b)



(c)

Figure 11: Comparison results on MOCAP database. (a) With Gabor feature. (b) With spectrograms envelopes feature. (c) With means of envelopes feature.

ACCEPTED MANUSCRIPT



(a) (b) (c)

Figure 12: Visualization of the feature vectors output from the discriminator. 'Jumping', 'running' and 'walking' are set as the target set, while other activities are regarded as the unknown class. (a) After one iteration. (b) After ten iterations. (c) After 100 iterations.

6. Discussion

To verify the performance of our method versus different hierarchical architectures, we carry out experiments on both activity detection and recognition tasks. The current implementation of the algorithm is based on the dense block with one sub-block in it, to explore the impact of dense connection and hierarchies, we evaluate the performances of a shallower variant without dense connection (i.e., a plain CNN) and three deeper variants with two to four sub-blocks.

Fig. 13 shows the experimental results of activity detection on both datasets. By comparing the results, we can observe that the introduction of dense block boosts the performance of network on both datasets. Nevertheless, with the number of sub-blocks increasing, the F-measure decreases surprisingly.

The activity recognition results are listed in Table 2 and Table 3. It can be seen that the proposed architecture generally outperforms other variants. Although plain CNN achieves two comparable results to the proposed network, the introduction of dense connection is beneficial in most cases. Especially, when openness is 0.0465 (five activities as the target set while one activity as the negative set), the proposed network achieves a 0.2281 improvement on F-measure than a plain CNN. This demonstrates the effectiveness of dense connection.

ACCEPTED MANUSCRIPT

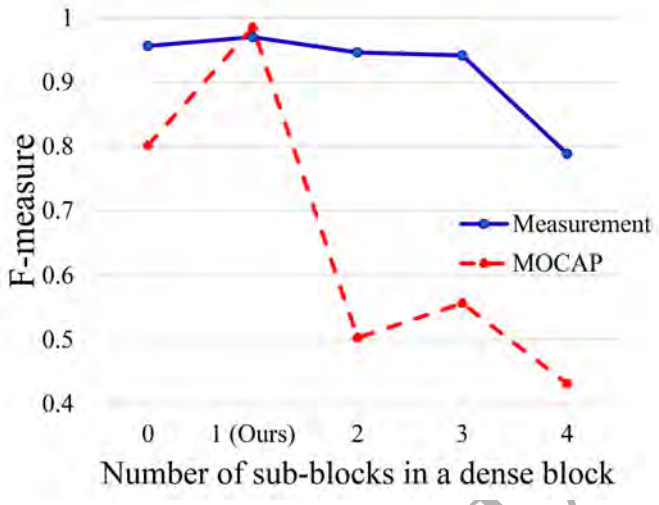


Figure 13: F-measure of human activity detection with different network architectures. '0' refers to a plain CNN network without dense connection.

When the hierarchy gets deeper, just the same as in the detection task, there is a decline in F-measure. Taking the time efficiency and required memory into account, we finally choose the architecture with one sub-block in each dense block.

7. Conclusion and future work

In this paper, we design a novel model called 'OpenGAN' for open-set human activity detection and recognition. The model is inspired by generative adversarial networks and automatically constructs the negative set with synthesized samples from the generator without manual collection. Subsequently, the discriminator is modified with some reasonable modulations to adapt the open-set tasks. Both measurement radar dataset and MOCAP dataset are employed to verify the effectiveness of our model. Extensive experiments show that OpenGAN outperforms several comparison algorithms on the open-set human activity problem. The results suggest the robustness of our method against existing approaches. Exploration experiments about the network structure are also

ACCEPTED MANUSCRIPT

Table 2: F-measure of human activity recognition on measurement dataset with different network architectures.

Sub-block	Openness					
number	0.4655	0.2929	0.1835	0.1056	0.0465	0.0000
0	0.4012	**0.5748**	0.6245	0.5695	0.5130	0.6207
1 (Ours)	**0.5323**	0.5570	**0.6320**	**0.6955**	**0.7411**	**0.7461**
2	0.4169	0.5200	0.5890	0.6255	0.6640	0.7261
3	0.4369	0.5169	0.5842	0.5877	0.6803	0.7461
4	0.4027	0.5117	0.5841	0.6055	0.3974	0.2857

Table 3: F-measure of human activity recognition on MOCAP dataset with different network architectures.

Sub-block	Openness					
number	0.4655	0.2929	0.1835	0.1056	0.0465	0.0000
0	0.6825	0.6767	0.6270	**0.7966**	0.7888	0.8095
1 (Ours)	**0.7262**	**0.7237**	**0.7198**	0.7655	**0.8011**	**0.8909**
2	0.6474	0.6552	0.6957	0.7137	0.8027	0.8083
3	0.3244	0.6431	0.6846	0.7389	0.7669	0.7780
4	0.6360	0.6455	0.7045	0.7518	0.8055	0.8107

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

carried out. The results demonstrate that network with one sub-layer in each dense block achieves the best performances under most circumstances. Future works will include extending the algorithm to other modal data and developing the model to an end-to-end open-set detection and recognition system.

Acknowledgement

This work is supported by the National Natural Science Foundation of China under Grants 61520106002 and 61731003.

References

[1] J. K. Aggarwal, L. Xia, Human activity recognition from 3d data: A review, Pattern Recognition Letters 48 (2014) 70–80.

[2] P. Casale, O. Pujol, P. Radeva, Human activity recognition from accelerometer data using a wearable device, Pattern Recognition and Image Analysis (2011) 289–296.

[3] S. Althloothi, M. H. Mahoor, X. Zhang, R. M. Voyles, Human activity recognition using multi-features and multiple kernel learning, Pattern recognition 47 (5) (2014) 1800–1812.

[4] H. Zhou, M. You, L. Liu, C. Zhuang, Sequential data feature selection for human motion recognition via markov blanket, Pattern Recognition Letters 86 (2017) 18–25.

[5] B. Zhang, L. Wang, Z. Wang, Y. Qiao, H. Wang, Real-time action recognition with enhanced motion vector cnns, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2718–2726.

[6] E. P. Ijjina, K. M. Chalavadi, Human action recognition using genetic algorithms and convolutional neural networks, Pattern Recognition 59 (2016) 199–212.

ACCEPTED MANUSCRIPT

[7] E. P. Ijjina, K. M. Chalavadi, Human action recognition in rgb-d videos using motion sequence information and deep learning, Pattern Recognition 72 (2017) 504–516.

[8] Y. Wang, M. Long, J. Wang, P. S. Yu, Spatiotemporal pyramid network for video action recognition, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017, pp. 2097–2106.

[9] Y. Ji, Y. Yang, X. Xu, H. T. Shen, One-shot learning based pattern transition map for action early recognition, Signal Processing 143 (2018) 364–370.

[10] M. Zhang, Y. Yang, Y. Ji, N. Xie, F. Shen, Recurrent attention network using spatial-temporal relations for action recognition, Signal Processing 145 (2018) 137–145.

[11] A.-A. Liu, Y.-T. Su, W.-Z. Nie, M. Kankanhalli, Hierarchical clustering multi-task learning for joint human action grouping and recognition, IEEE transactions on pattern analysis and machine intelligence 39 (1) (2017) 102–114.

[12] A.-A. Liu, N. Xu, W.-Z. Nie, Y.-T. Su, Y. Wong, M. Kankanhalli, Benchmarking a multimodal and multiview and interactive dataset for human action recognition, IEEE Transactions on cybernetics 47 (7) (2017) 1781–1794.

[13] D. P. Fairchild, R. M. Narayanan, Classification of human motions using empirical mode decomposition of human micro-doppler signatures, IET Radar, Sonar & Navigation 8 (5) (2014) 425–434.

[14] Y. Kim, H. Ling, Human activity classification based on micro-doppler signatures using a support vector machine, IEEE Transactions on Geoscience and Remote Sensing 47 (5) (2009) 1328–1337.

[15] G. Kumar, P. K. Bhatia, A detailed review of feature extraction in image processing systems, in: Advanced Computing & Communication Technolo-

ACCEPTED MANUSCRIPT

gies (ACCT), 2014 Fourth International Conference on, IEEE, 2014, pp. 5–12.

[16] L. Wang, S. Guo, W. Huang, Y. Qiao, Places205-vggnet models for scene recognition, arXiv preprint arXiv:1508.01667.

[17] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, 2012, pp. 1097–1105.

[18] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556.

[19] G. Huang, Z. Liu, K. Q. Weinberger, L. van der Maaten, Densely connected convolutional networks, arXiv preprint arXiv:1608.06993.

[20] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks, in: Advances in neural information processing systems, 2015, pp. 91–99.

[21] J. Dai, Y. Li, K. He, J. Sun, R-fcn: Object detection via region-based fully convolutional networks, in: Advances in neural information processing systems, 2016, pp. 379–387.

[22] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille, Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, arXiv preprint arXiv:1606.00915.

[23] A. Radford, L. Metz, S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, arXiv preprint arXiv:1511.06434.

[24] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al., Photo-realistic single image super-resolution using a generative adversarial network, arXiv preprint arXiv:1609.04802.

ACCEPTED MANUSCRIPT

[25] J.-Y. Zhu, T. Park, P. Isola, A. A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, arXiv preprint arXiv:1703.10593.

[26] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, in: Advances in neural information processing systems, 2014, pp. 2672–2680.

[27] J. Kwon, N. Kwak, Human detection by neural networks using a low-cost short-range doppler radar sensor, in: Radar Conference, 2017, pp. 0755–0760.

[28] Y. Kim, T. Moon, Human detection and activity classification based on micro-doppler signatures using deep convolutional neural networks, IEEE Geoscience and Remote Sensing Letters 13 (1) (2016) 8–12.

[29] Y. Kim, T. Moon, Classification of human activity on water through micro-dopplers using deep convolutional neural networks, in: Radar Sensor Technology, Vol. 9829, International Society for Optics and Photonics, 2016, p. 982917.

[30] W. J. Scheirer, A. de Rezende Rocha, A. Sapkota, T. E. Boult, Toward open set recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (7) (2013) 1757–1772.

[31] I. Laptev, M. Marszalek, C. Schmid, B. Rozenfeld, Learning realistic human actions from movies, in: Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, IEEE, 2008, pp. 1–8.

[32] J. Han, L. Shao, D. Xu, J. Shotton, Enhanced computer vision with microsoft kinect sensor: A review, IEEE transactions on cybernetics 43 (5) (2013) 1318–1334.

[33] H. Chen, G. Wang, J.-H. Xue, L. He, A novel hierarchical framework for human action recognition, Pattern Recognition 55 (2016) 148–159.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

[34] L. Zhou, W. Li, P. Ogunbona, Z. Zhang, Semantic action recognition by learning a pose lexicon, Pattern Recognition 72 (2017) 548–562.

[35] J. Wang, Z. Liu, Y. Wu, J. Yuan, Mining actionlet ensemble for action recognition with depth cameras, in: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, IEEE, 2012, pp. 1290–1297.

[36] W. Li, Z. Zhang, Z. Liu, Action recognition based on a bag of 3d points, in: Computer Vision and Pattern Recognition Workshops, 2010, pp. 9–14.

[37] L. Xia, C. C. Chen, J. K. Aggarwal, View invariant human action recognition using histograms of 3d joints, in: Computer Vision and Pattern Recognition Workshops, 2012, pp. 20–27.

[38] X. Yang, Y. L. Tian, Eigenjoints-based action recognition using naïve-bayes-nearest-neighbor, in: Computer Vision and Pattern Recognition Workshops, 2012, pp. 14–19.

[39] B. Tekeli, S. Z. Gurbuz, M. Yuksel, Information-theoretic feature selection for human micro-doppler signature classification, IEEE Transactions on Geoscience and Remote Sensing 54 (5) (2016) 2749–2762.

[40] M. O. Padar, A. E. Ertan, Ç. ĝatay Candan, Classification of human motion using radar micro-doppler signatures with hidden markov models, in: Radar Conference, 2016, pp. 1–6.

[41] T. S. Jordan, Using convolutional neural networks for human activity classification on micro-doppler radar spectrograms, in: Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security, Defense, and Law Enforcement Applications XV, Vol. 9825, International Society for Optics and Photonics, 2016, p. 982509.

[42] A. Moeini, K. Faez, H. Moeini, A. M. Safai, Open-set face recognition across look-alike faces in real-world scenarios, Image and Vision Computing 57 (2017) 1–14.

ACCEPTED MANUSCRIPT

[43] A. Rozsa, M. Günther, E. M. Rudd, T. E. Boult, Facial attributes: Accuracy and adversarial robustness, Pattern Recognition Letters.

[44] A. Bendale, T. Boult, Towards open world recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1893–1902.

[45] M. Ristin, M. Guillaumin, J. Gall, L. Van Gool, Incremental learning of ncm forests for large-scale image classification, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 3654–3661.

[46] A. Bendale, T. E. Boult, Towards open set deep networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1563–1572.

[47] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, A. A. Bharath, Generative adversarial networks: An overview, arXiv preprint arXiv:1710.07035.

[48] M. Arjovsky, S. Chintala, L. Bottou, Wasserstein gan, arXiv preprint arXiv:1701.07875.

[49] M. Mirza, S. Osindero, Conditional generative adversarial nets, arXiv preprint arXiv:1411.1784.

[50] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, Image-to-image translation with conditional adversarial networks, arXiv preprint arXiv:1611.07004.

[51] T. Kim, M. Cha, H. Kim, J. Lee, J. Kim, Learning to discover cross-domain relations with generative adversarial networks, arXiv preprint arXiv:1703.05192.

[52] Z. Yi, H. Zhang, P. T. Gong, et al., Dualgan: Unsupervised dual learning for image-to-image translation, arXiv preprint arXiv:1704.02510.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

[53] Y. He, P. Molchanov, T. Sakamoto, P. Aubry, F. Le Chevalier, A. Yarovoy, Range-doppler surface: a tool to analyse human target in ultra-wideband radar, IET Radar, Sonar & Navigation 9 (9) (2015) 1240–1250.

[54] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, in: International Conference on Machine Learning, 2015, pp. 448–456.

[55] X. Glorot, A. Bordes, Y. Bengio, Deep sparse rectifier neural networks, in: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, 2011, pp. 315–323.

[56] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. Courville, Improved training of wasserstein gans, arXiv preprint arXiv:1704.00028.

[57] P. Van Dorp, F. Groen, Feature-based human motion parameter estimation with radar, IET Radar, Sonar & Navigation 2 (2) (2008) 135–145.

[58] C. Karabacak, S. Z. Gurbuz, A. C. Gurbuz, M. B. Guldogan, G. Hendeby, F. Gustafsson, Knowledge exploitation for human micro-doppler classification, IEEE Geoscience and Remote Sensing Letters 12 (10) (2015) 2125–2129.

[59] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, Vol. 1, IEEE, 2005, pp. 886–893.

[60] D. M. Tax, R. P. Duin, Support vector domain description, Pattern recognition letters 20 (11) (1999) 1191–1199.

[61] P. E. Hart, An asymptotic analysis of the nearest-neighbor decision rule., Tech. rep. (1966).

[62] J. MacQueen, et al., Some methods for classification and analysis of multivariate observations, in: Proceedings of the fifth Berkeley symposium on

ACCEPTED MANUSCRIPT

mathematical statistics and probability, Vol. 1, Oakland, CA, USA., 1967, pp. 281–297.

[63] G. H. Ball, D. J. Hall, Isodata, a novel method of data analysis and pattern classification, Tech. rep., Stanford research inst Menlo Park CA (1965).

[64] K. Pearson, Liii. on lines and planes of closest fit to systems of points in space, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 2 (11) (1901) 559–572.

[65] G.-B. Huang, Q.-Y. Zhu, C.-K. Siew, Extreme learning machine: theory and applications, Neurocomputing 70 (1) (2006) 489–501.

[66] D. Tax, Ddtools, the data description toolbox for matlab, version 2.1.3 (Jan 2018).

[67] W. J. Scheirer, L. P. Jain, T. E. Boult, Probability models for open set recognition, IEEE transactions on pattern analysis and machine intelligence 36 (11) (2014) 2317–2324.

[68] L. P. Jain, W. J. Scheirer, T. E. Boult, Multi-class open set recognition using probability of inclusion, in: European Conference on Computer Vision, Springer, 2014, pp. 393–409.

[69] C.-C. Chang, C.-J. Lin, Libsvm: a library for support vector machines, ACM transactions on intelligent systems and technology (TIST) 2 (3) (2011) 27.

[70] L. v. d. Maaten, G. Hinton, Visualizing data using t-sne, Journal of machine learning research 9 (Nov) (2008) 2579–2605.

Yang Yang is currently working toward the Ph.D. degree in the School of Electrical and Information Engineering at Tianjin University, China. His research interests include object detection, deep learning, pattern recognition, human activity recognition and micro-Doppler radar.

Chunping Hou received the M.Eng. and Ph.D. degrees in electronic engineering from Tianjin University, Tianjin, China, in 1986 and 1998, respectively.

ACCEPTED MANUSCRIPT

She is currently a Full Professor and the Director of the Broadband Wireless Communications and 3-D Imaging Institute at Tianjin University. Her research interests include deep learning, pattern recognition, 3-D image processing, 3-D display, and wireless communication.

Yue Lang received the B.Eng. degree from the Department of Communication Engineering, Northeastern University, China, in 2016. She is now working toward the M.Eng. degree in the School of Electrical and Information Engineering in Tianjin University, China. Her research interests include deep learning, human activity recognition and micro-Doppler radar.

Guan Dai received the B.Eng. degree from the Department of Electrical Engineering & Automation, Tianjin University, China, in 2016. He is a second year master degree candidate at School of Electrical and Information Engineering in Tianjin University, China. His research interests are in deep learning, human activity recognition and object detection.

DanYang Huang received the B.Eng. degree from the School of Electrical and Information Engineering, Tianjin University, China, in 2017. Now he is working toward the M.Eng. degree in Tianjin University, China. His research interests include deep learning and micro-Doppler radar.

Jinchen Xu received the B.Eng. degree from the School of Electrical and Information Engineering, Tianjin University, China, in 2017. Now he is working toward the M.Eng. degree in Tianjin University, China. His research interests include deep learning and micro-Doppler radar.

ACCEPTED MANUSCRIPT