

Received February 22, 2021, accepted March 17, 2021, date of publication March 23, 2021, date of current version March 30, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3068075

# Synthesis of Micro-Doppler Signatures of Human Activities From Different Aspect Angles Using Generative Adversarial Networks

**IBRAHIM ALNUJAIM<sup>1</sup>, SHOBHA SUNDAR RAM<sup>2</sup>, (Senior Member, IEEE),  
DAEGUN OH<sup>3</sup>, AND YOUNGWOOK KIM<sup>1</sup>, (Senior Member, IEEE)**

<sup>1</sup>Department of Electrical and Computer Engineering, California State University, Fresno, CA 93740, USA

<sup>2</sup>Department of Electrical and Computer Engineering, Indraprastha Institute of Information Technology Delhi, New Delhi 110020, India

<sup>3</sup>Advanced Radar Research Division, Daegu Gyeongbuk Institute of Science and Technology, Daegu 42988, South Korea

Corresponding author: Youngwook Kim (youngkim@mail.fresnostate.edu)

**ABSTRACT** In this paper, we propose to produce synthesized micro-Doppler signatures from different aspect angles through conditional generative adversarial networks (cGANs). Micro-Doppler signatures of non-rigid human body motions vary considerably as a function of the radar's aspect angle. Because the direction of the human motion can be arbitrary, a large volume of training data across diverse aspects is needed for practical target activity classification through machine learning. As measurements can require significant monetary and labor costs, the synthesis of micro-Doppler signatures can be an alternate solution. Therefore, we investigate the feasibility of data augmentation through synthesizing micro-Doppler signatures of human activities from diverse radar aspect angles with input data from a single aspect angle. For the training data, the micro-Doppler radar signatures of 12 human activities are generated from different angles ranging from 0 to 315 degrees, at 45-degree increments, through simulations. For each angle, cGANs are trained to synthesize the micro-Doppler signatures for that specific angle given micro-Doppler signatures from another angle. The output of each model is evaluated by calculating mean-square errors and structural similarity indexes between the synthesized micro-Doppler signatures and the ground-truth ones obtained from simulations. We test three different scenarios, and report the respective results.

**INDEX TERMS** Generative adversarial networks (GAN), deep learning, micro-Doppler signatures, monostatic radar, multistatic radar, human activity.

## I. INTRODUCTION

Radar based analysis of human activities is increasingly applied in defense, surveillance, and health care scenarios for its capability to operate 24/7, in through-object scenarios, under poor weather conditions, and in situations where privacy is a concern. When analyzing human motion, extraction of signatures from radar returns is a key component in radar signal processing. Among many radar signatures, micro-Doppler signatures in the joint time-frequency domains produced from non-rigid body motions [1] have been extensively studied for purposes of human activity classification [2], [3], gait analysis [4], abnormal gait classification [5], senior fall detection [6], hand gesture detection in indoor scenarios [7], and pedestrian detection in

outdoor automotive radar scenarios [8]. However, because the Doppler information embedded in these signatures is a function of the aspect angle of the monostatic radar, the characteristics of radar signatures can significantly vary depending on the target's trajectory with respect to the radar. Accordingly, deep learning algorithms, known as some of the best for classification, require large data sets of micro-Doppler signatures, captured across diverse radar aspects, to exhibit stable classification accuracy for arbitrary human motions.

Deep learning has played a critical role in solving an array of scientific, engineering, medical, and financial problems, and radar target classification problems have been no exception: deep learning has been a main stream in radar target classification. Various deep learning models, such as auto-encoders, deep convolutional neural networks (DCNNs), and deep recurrent neural networks (DRNNs), have been exploited to recognize radar imagery. DCNN captures

The associate editor coordinating the review of this manuscript and approving it for publication was Shuping He<sup>1</sup>.

features in radar imagery such as spectrograms or range-Doppler diagrams for target recognition [9], while DRNN analyzes time-varying radar features and identifies temporal patterns in them [10]. Because deep learning requires a large number of diverse data to avoid overfitting, radar imagery has often suffered from the data deficiency problem, a critical problem in radar imagery that has been addressed by many trials, including ones using transfer learning [11] and data augmentation through synthesizing fake data by generative adversarial networks (GANs) [12], [13]. Acquiring sufficient radar data is a key component of successful training.

From an information point of view, multistatic radars offer a second important advantage over single-aspect angle data obtained from a monostatic radar. When, for example, a human moves tangentially or perpendicular to the boresight of a monostatic radar antenna, no micro-Doppler appears. However, a second orthogonally placed radar antenna would capture maximal micro-Doppler data in response to the human's radial motion toward it. Accordingly, use of multistatic radar reduces aspect angle deterioration or distortion of radar signatures by offering different radar viewpoints on a target. Multistatic radars may also help in overcoming problems relating to shadowing, particularly when used to classify and reveal in-depth information regarding human activity sufficient to identify personnel, distinguish between potentially armed and unarmed personnel [14], classify human activities, determine target orientation [15], and recognize hand gestures [16]. However, although monostatic radar systems are currently in wide use, upgrading to a multistatic radar system may involve challenges related to budget, labor, and physical space. Furthermore, hardware challenges are related to synchronization of data across multiple radars. By contrast, a software solution offers advantages of manageability and scalability over hardware upgrades.

Synthesis of multistatic spectrograms of human activity using data from a single radar has the potential to increase information regarding human movement from different aspect angles. If the target's motions and radar channel conditions are fully known, the signatures can be directly simulated. However, some target motions may be complicated and difficult to realistically model with a low degree of error. For example, non-analyzable targets, such as humans, animals, birds, and insects, are hard to simulate correctly.

In this paper, we explore the feasibility of using radar spectrograms gathered from a single aspect angle of human activity to synthesize spectrograms from other angles using conditional GAN, a variant of the general GAN used in previous micro-Doppler radar literature [12], [13]. We investigate the possibility of image-to-image translation with conditional adversarial nets to generate spectrograms from different monostatic radar points of view. In 2016, image-to-image translation with conditional adversarial nets (Pix2Pix) was introduced for computer vision [17]. In the radar community, Pix2Pix has been used to despeckle colorized SAR images [18]. The goal of our paper is to synthesize

spectrograms from 45 degrees to 315 degrees, at increments of 45 degrees, using the Pix2Pix model, given an input spectrogram at a 0-degree aspect angle that represents a human subject moving away from the radar. It should be noted that changes in aspect angle do not involve a constant cosine relationship, because the angle between radar radial direction and human movement direction is a function of time. To construct the training data, we simulate micro-Doppler signatures of a human subject from several radar aspect angles using MOCAP data. The simulated data set is divided into a training set and a test set to train and evaluate the performance of Pix2Pix. Once trained, Pix2Pix synthesizes micro-Doppler signatures from a specific angle, and we compare them with ground-truth simulated signatures obtained from MOCAP-based modeling. The contributions of this paper are as follows; i) we proposed to synthesize micro-Doppler signatures seen from different aspect angles, ii) image-to-image translation through deep learning has been applied to micro-Doppler signatures.

The paper structure is as follows. In Section II, we describe the method to simulate micro-Doppler signatures of 12 different human activities at varying aspect angles. In Section III, we present a detailed description of the Pix2Pix and its structure. Section IV discuss the results of synthesized micro-Doppler signatures for the three cases. The conclusions and future research suggestions are provided in Section V.

## II. HUMAN MICRO-DOPPLER SIMULATION

The micro-Doppler signatures of humans can be simulated from MOCAP data as described in [19]. These data describe the articulation of a 3D human skeletal structure across a motion trajectory at a video frame rate of 30 Hz. The animation data are suitably interpolated to radar sampling frequencies. Then each bone,  $b$ , in the skeleton is modeled as an ellipsoid primitive whose radar cross-section ( $\sigma_b$ ) at microwave frequencies is analytically defined as a function of radar carrier frequency,  $f_c$ , and aspect angle,  $\theta_b(t)$ . Furthermore, the dielectric properties of the human skin are incorporated in the RCS expression. Time-varying scattering centers ( $\vec{r}_b(t)$ ) of all primitives are identified at the primitives' center. Then radar returns from a monostatic continuous-wave radar located at  $\vec{r}_n$  are given by

$$x_n(t) = \sum_{b=1}^B A \frac{\sqrt{\sigma_b(t)}}{r_{b,n}^2(t)} e^{j2\pi f_c \frac{2r_{b,n}(t)}{c}} \quad (1)$$

where  $r_{b,n}$  is the Euclidean distance between  $\vec{r}_b(t)$  and  $\vec{r}_n$ , where the latter is the position of the monostatic radar. In the preceding expression,  $A$ , encompasses the transmitted power and the gains of the radar antennas. The foregoing method can be easily modified to a multistatic configuration by changing the radar's position.

In this paper, we consider simulated radar returns from twelve human activities that consist of uniform motions and non-uniform motions; crouching to running, crawling, running, running to jumping to walking, running to hopping to walking, running with boxing, running to crouching,

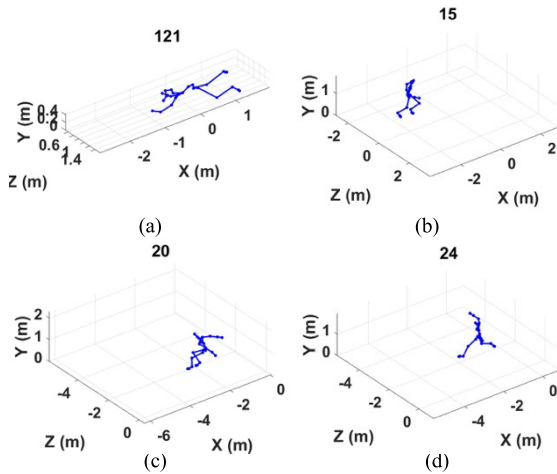


FIGURE 1. Examples of MOCAP data: (a) crawl, (b) run, (c) run to jump to walk, (d) run to hop to walk.

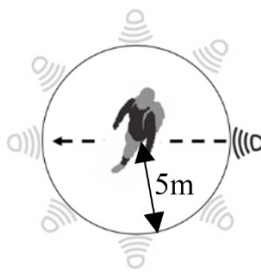


FIGURE 2. Human activity and radar positioning (black radar positioned at 0 degree).

skipping, walking, walking to hopping to walking, walking to leaping to walking, and walking to picking up a box. Fig. 1 shows examples of MOCAP data for different motions, taken from the CMU Graphics Lab Motion Capture Database. We have varied the operating frequency and gain of the antenna in the simulation. The radar carrier frequency used in the simulations ranges from 6.5 GHz to 8.5 GHz at increments of 0.05 GHz, and the isotropic gain of the radar antennas varies between 5 dBi and 15 dBi. This is for the purpose of not only increasing the number of data samples given number of MOCAP data, but including diversity of dataset in terms of the signal-to-noise ratio and a scaling effect in the Doppler frequency.

The simulations are conducted at eight different radar positions with 45-degree angular separation. By changing the center frequency and isotropic gain, 451 spectrograms are generated for a particular angle and activity, for a total of 5,412 data samples. The radar is placed at a radial distance of 5 meters from the center of the simulation space in each case. Fig. 2 shows eight radar locations and simulation configurations.

### III. IMAGE-TO-IMAGE TRANSLATION WITH CONDITIONAL ADVERSARIAL NETWORKS

We use Pix2Pix, a variant model of GAN, to synthesize micro-Doppler signatures. Pix2Pix is designed to carry out image-to-image translation using conditional

GANs (cGANs) [17]. In general, GANs use two neural networks—a generator network and a discriminator network—working against each other. The generator synthesizes a data sample that imitates the training output data from random noise. The discriminator classifies these data as either real or synthetic without any prior knowledge of data type. The main difference between cGANs and general GANs is that the generator in the former neural network takes two inputs, noise and label, to synthesize each data sample. After training, the cGAN generates data based on a given label. We describe the cGAN algorithm in detail hereafter. In Pix2Pix, the label is an input image. The loss function of Pix2Pix is defined as

$$Losses = \operatorname{argmin}_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G) \quad (2)$$

The generator network ( $G$ ) minimizes the loss function to confuse the discriminator network and the discriminator network ( $D$ ) maximizes the classification accuracy. The  $\lambda$  is an integer set to 100 to give importance to the  $l_1$ -norm. In the cGAN, the loss function is defined as

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y} [\log D(x, y)] + \mathbb{E}_{x,z} [\log(1 - D(x, G(x, z)))] \quad (3)$$

where  $x$  is the input data,  $y$  the synthesized data, and  $z$  the noise vector. The  $\mathcal{L}_{L1}(G)$  encourages the generator network to generate output near the ground truth:

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z} [\|y - G(x, z)\|_1] \quad (4)$$

In GANs and cGANs’ loss functions, noise vector is used as an input to the generator network. However, in Pix2Pix, the noise vector is not used as an input. Instead, noises are introduced in the generator network in dropout layers.

The generator is designed to take input spectrograms of 0 aspect angle and produce output spectrograms at another angle designated by the label. The generator network consists of encoder and decoder blocks. The network generates the target output by downsampling the data in the encoder, processing them, and then upsampling them using the U-Net decoder [20], which adds extra connections between the encoder and decoder blocks of the network to preserve low-level information. Fig. 4 presents a simplified illustrative example structure of the U-Net. It should be noted that the more complex version of the network used in our work, which featured more layers, is not shown due to space constraints.

In the U-Net, a single encoder block has a convolution layer with a kernel size of  $4 \times 4$ , stride of 2, batch normalization, and leaky ReLU activation functions. While configuring the U-Net, the number of kernels in each encoder layer must be specified and the batch normalization may be disabled. In our work, the spectrogram is passed through the first encoder layer, with 64 kernels and disabled batch normalization. Then the data are passed through the second and third encoder layers, having 128 and 256 kernels, respectively, with batch normalization enabled. Next the data are passed through the fourth to eighth encoder layers, having 512 kernels each, with

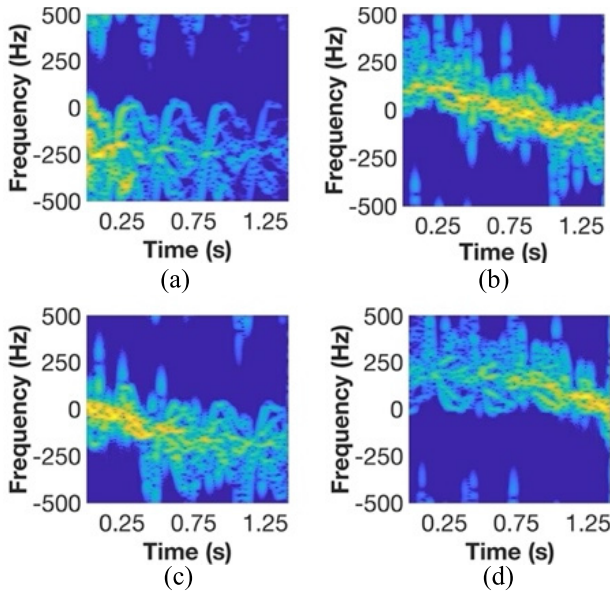


FIGURE 3. Micro-Doppler signatures of the human activity depending on radar positioning. (a) 0 degree, (b) 45 degrees, (c) 90 degrees, and (d) 135 degrees.



FIGURE 4. Simplified U-Net architecture for a generator network.

batch normalization enabled. The output is passed through a series of seven decoders. Each decoder has a transposed convolution layer with a kernel size of  $4 \times 4$ , stride of 2, batch normalization, and dropout rate of 0.5. Again, the number of kernels must be specified in each layer, and the dropout can be activated only for the first three layers. The ninth decoder layer, which consists of 512 kernels, takes the output of the eighth layer as input. The tenth layer concatenates the output of the previous layer with the seventh layer and passes it through a decoder having 512 kernels. The eleventh and twelfth layers take the previous layer and outputs of the sixth and seventh layers, having 512 kernels, and pass it through a decoder. The thirteenth, fourteenth, and fifteenth layers take the previous layer and outputs of the fourth, third, and second layers, having 256, 128, and 64 kernels, respectively, and pass them through a decoder. The last layer has only a single transposed convolution layer, with three kernels of size  $4 \times 4$  and a *tanh* activation layer.

The use of  $l_1$ -norm in the loss function makes the output less sharp. To solve the issue, PatchGAN is used in the discriminator network [17]. PatchGAN segments an image into multiple  $N \times N$  ( $70 \times 70$  pixels, in our case) patches and classifies each patch. The output of the PatchGAN is thus a smaller ( $30 \times 30$ ) matrix with a decision of true or

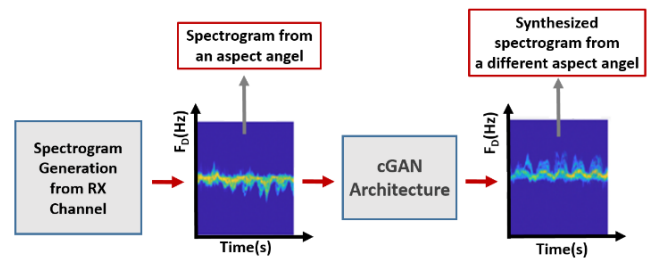


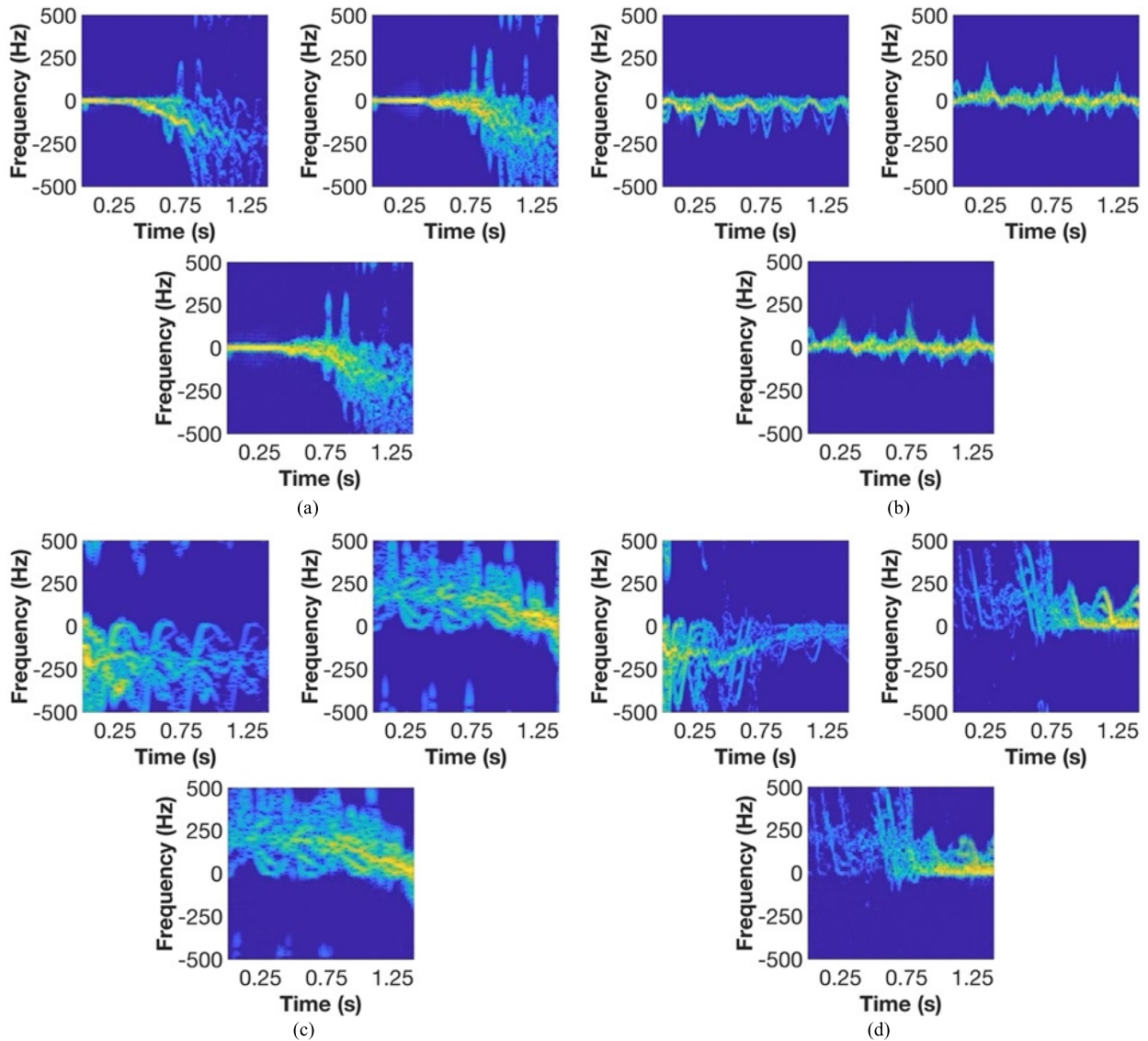
FIGURE 5. System processing steps to generate Pix2Pix output of one angle.

fake for each patch in the corresponding matrix element. The final decision about whether the image is real or fake is based on the average of the decisions across all patches. The discriminator network follows a DCNN structure but takes two inputs instead of one. The two inputs are the Pix2Pix input image and an artificial/real image. These images are concatenated at the first step of the discriminative network. Then the input is passed through a convolution layer having 64 kernels, with stride of 2, followed by leaky ReLU. All convolution layers have a kernel size of  $4 \times 4$ . Afterward, two convolution layers having 128 and 256 kernels, respectively, are followed by batch normalization and leaky ReLU. 2D zero padding is applied to the output of the last layer. This is followed by a convolution layer having 512 kernels and stride of 2, with batch normalization and leaky ReLU. Finally, 2D zero padding and a convolution layer with a single kernel are applied.

The deep learning algorithm was implemented with TensorFlow and Python, using an Intel i9 CPU and an NVIDIA GeForce RTX 2080 GPU. All spectrograms have the same pixel size of  $256 \times 256$ . The dataset is divided into 80% for training and 20% for testing. All transposed convolutions and regular convolutions are initialized using the random normal initializer. The epoch is set to 30. To improve training performance, input spectrograms are randomly jittered and mirrored. The batch size is set to one.

#### IV. GENERATING MULTISTATIC MICRO-DOPPLER SIGNATURES

Using the simulated micro-Doppler signatures, we train the image-to-image translation with cGAN for three scenarios. In the first, we train the Pix2Pix model using single activity data to generate the spectrograms for seven different aspect angles. In the second, we use multiple activities in the training process, and during testing, the trained models generate the corresponding spectrograms at seven different aspect angles given any one of the activities. In the third and most challenging, we use ten activities in the training and use the remaining two activities to evaluate Pix2Pix’s ability to synthesize micro-Doppler signatures at different aspect angles for unknown images of a new motion. In short, the network must generate results for an activity it has not encountered during training. To quantify the performance of the algorithm, we apply two metrics to the synthesized micro-Doppler spectrograms obtained from Pix2Pix and reference ground-truth



**FIGURE 6.** Example results from the first scenario. Three spectrograms (clockwise) are 0-degree input, ground truth at the desirable aspect angle, and output spectrogram of the desired angle: (a) crouch to run at 45°, (b) crawl at 90°, (c) run at 135°, and (d) run to jump to walk at 180°.

spectrograms obtained directly from simulations: normalized mean square error (NMSE) and the structural similarity index (SSIM). Whereas NMSE captures the pixel-by-pixel difference in energy between test output images and ground-truth images, SSIM is useful for comparing visual similarity in two images. In ideal scenarios, when test output images are identical to ground-truth images, NMSE is 0 and SSIM is 1. Fig. 5 summarizes the steps needed to produce the synthesized spectrograms:

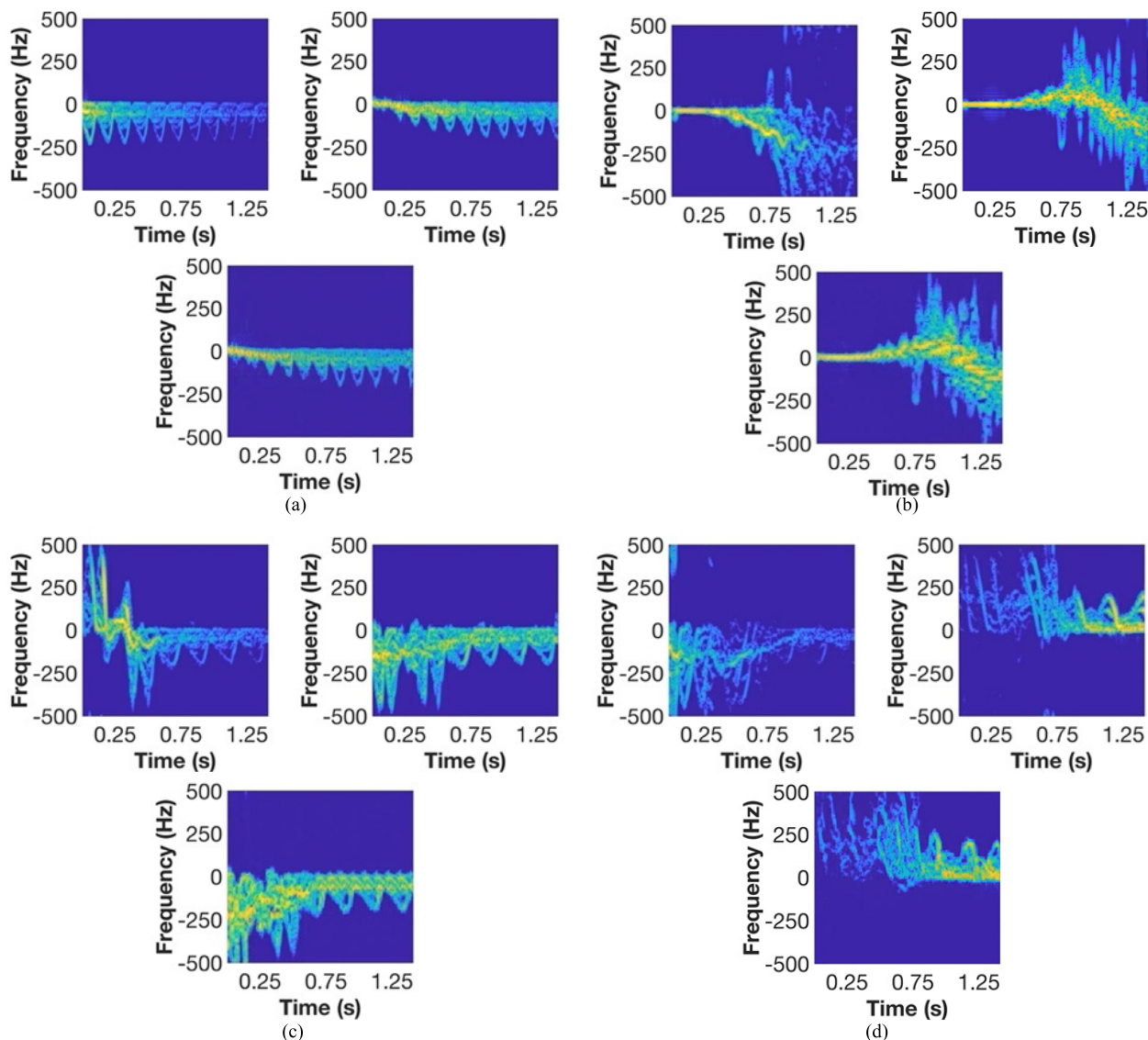
Pix2Pix is trained to generate micro-Doppler signatures of human activities at different aspect angles. Three scenarios are tested. In the first case, a network is trained on a single activity along one aspect angle. During test, the network synthesizes the radar signature of the corresponding activity along the desired aspect angle (different from the input angle). In this scenario, we test a total of twelve different activities one by one. For each activity, seven independent

**TABLE 1.** NMSE and SSIM for each angle for the first scenario.

| Angle | NMSE      | SSIM |
|-------|-----------|------|
| 45°   | 0.68 E-04 | 0.81 |
| 90°   | 0.87 E-04 | 0.81 |
| 135°  | 1.09 E-04 | 0.77 |
| 180°  | 1.30 E-04 | 0.75 |
| 225°  | 1.06 E-04 | 0.77 |
| 270°  | 0.77 E-04 | 0.81 |
| 315°  | 0.50 E-04 | 0.82 |

Pix2Pix models are trained to translate images from 0 aspect angle to 45°, 90°, and up to 315°. As a result, we train 84 independent models in this scenario.

Fig. 6 demonstrates the results of the first scenario. In the figure, there are three images for each case: the input spectrogram at 0 aspect angle, ground truth at the angle of interest, and the synthesized spectrogram at the angle of



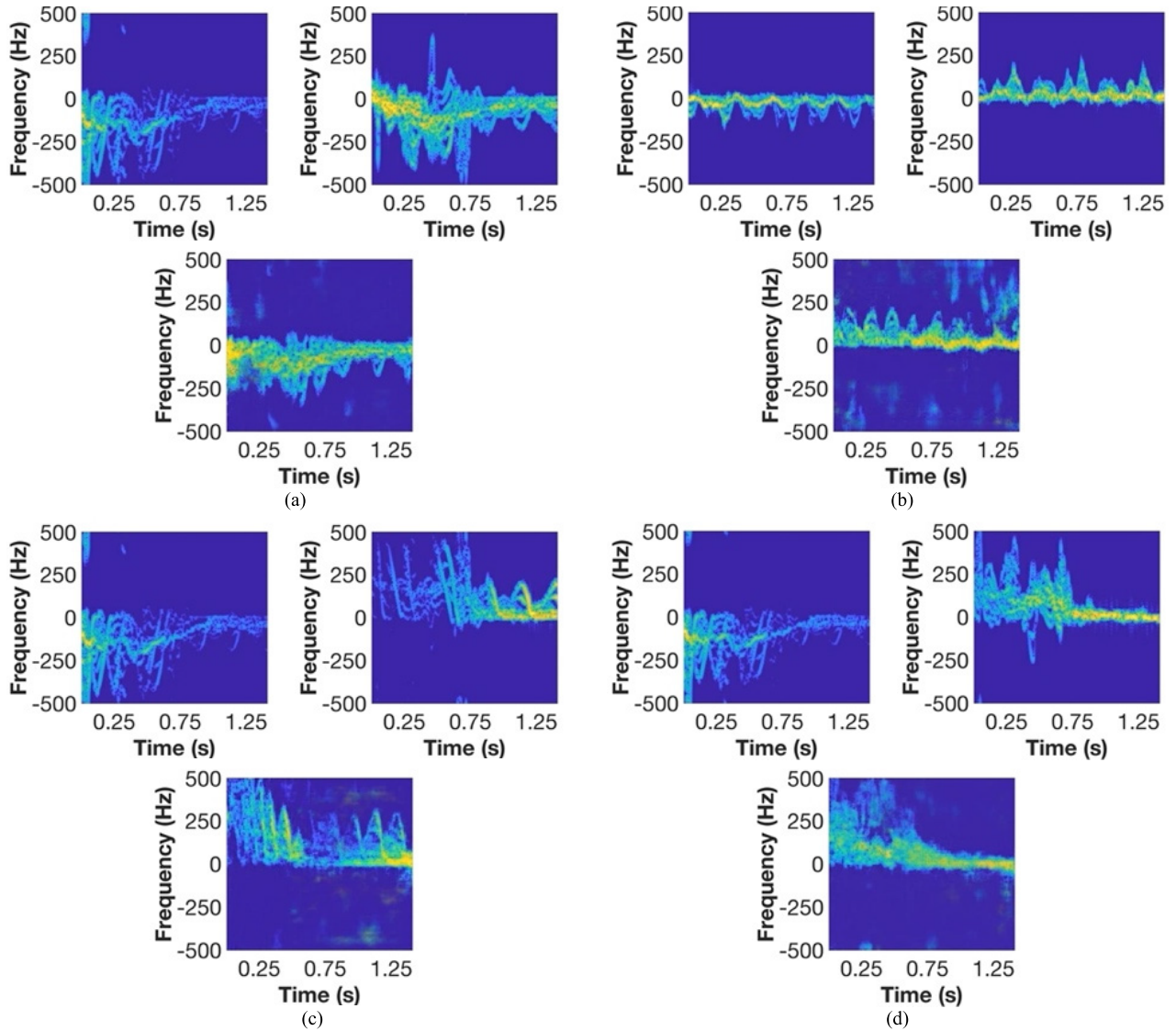
**FIGURE 7.** Example results from the second scenario. Three spectrograms (clockwise) are 0-degree input, ground truth at the desired aspect angle, and output spectrogram of the desirable angle: (a) walk at 45°, (b) crouch to run at 90°, (c) run to hop to walk at 135°, and (d) run to jump to walk at 180°.

interest. Thus, we expect that the second and the third spectrograms will be resemble to each other. Table 1 shows the error value for each angle. For all activities, the average NMSE is 0.9E-4 and the standard deviation is 0.44E-4. The lowest NMSE achieved by *walking to picking up a box* for angle 315° with NMSE of 0.2E-04 and maximum value is 1.85E-04 for *walk to leap to walk* for angle 225°. The average SSIM is 0.79, and the standard deviation is 0.1. For all activities, the lowest value achieved for *running with a box* at angle 180° with a value of 0.5 and maximum SSIM is 0.92 for *walking with a box* at angle 270°. The low value of NMSE and high SSIM show that the proposed algorithm is effective at synthesizing spectrograms at different aspect angles with models trained for a specific activity and input–output angle combinations.

**TABLE 2.** NMSE and SSIM for each angle for the second scenario.

| Angle          | NMSE            | SSIM        |
|----------------|-----------------|-------------|
| 45°            | 1.07E-04        | 0.76        |
| 90°            | 1.31E-04        | 0.77        |
| 135°           | 1.59E-04        | 0.72        |
| 180°           | 1.64E-04        | 0.71        |
| 225°           | 1.42E-04        | 0.74        |
| 270°           | 1.21E-04        | 0.77        |
| 315°           | 0.75E-04        | 0.79        |
| <b>Average</b> | <b>1.28-E04</b> | <b>0.75</b> |

In the second scenario, we train Pix2Pix for each angle, which results in only seven models trained to generate spectrograms by angles, as all twelve activities are used together to train Pix2Pix. In short, the network is customized not



**FIGURE 8.** Example results from the third scenario. Three spectrograms (clockwise) are 0-degree input, ground truth at the desirable aspect angle, and output spectrogram of the desired angle: (a) run to jump to walk at 45°, (b) crawl forward at 135°, (c) run to jump to walk at 180°, and (d) run to jump to walk at 225°.

**TABLE 3.** NMSE and SSIM for each angle for the third scenario.

| Angle          | NMSE            | SSIM        |
|----------------|-----------------|-------------|
| 45°            | 1.56E-04        | 0.76        |
| 90°            | 2.11E-04        | 0.76        |
| 135°           | 2.46E-04        | 0.63        |
| 180°           | 3.27E-04        | 0.59        |
| 225°           | 2.07E-04        | 0.69        |
| 270°           | 2.22E-04        | 0.72        |
| 315°           | 1.94E-04        | 0.73        |
| <b>Average</b> | <b>2.23E-04</b> | <b>0.69</b> |

to a specific activity but rather for a specific combination of input–output aspect angles. Fig. 7 presents the examples of input spectrogram of any kinds of activity, ground-truth spectrograms at the aspect angle of interest, and synthesized spectrograms at the aspect angle of interest. Table 2 shows the averaged error values of all activities for each angle.

If we investigate the number in details, NMSE ranges from 0.75E-04 to 1.64E-4 and SSIM from 0.71 from 0.79. Again, we observe very low NMSE and high SSIM, showing that the models can be trained per the input–output aspect angle but need not be customized for specific types of target activities to synthesize spectrograms at an aspect angle of interest. It should be noted that in this second scenario, all twelve activities were used for training and testing.

In the third scenario, seven models by angles are trained using a subset of ten randomly chosen activities from the twelve activities. During testing, the network is used to generate spectrograms of the remaining two activities—crawling forward and running to jumping to walking—previously not used in training. Fig. 8 shows the synthesized results of the testing activities. Table 3 shows the averaged error values of all activities for each angle. If we investigate the values in details, NMSE ranges from 1.56E-04 to 3.27E-4 and SSIM

from 0.59 from 0.76. We observe higher errors than in the first and second scenarios, but the similarity of synthesized spectrograms.

## V. CONCLUSION

We investigated the feasibility of generating multistatic micro-Doppler signatures using data from a single aspect angle using image-to-image translation with conditional GANs. The training data were obtained by simulations using MOCAP data. We tested three scenarios: (i) translating a specific activity from one angle to a certain aspect angle, (ii) translating spectrograms from all trained activities from one angle to a certain aspect angle, and (iii) translating spectrograms of unknown activities from one angle to another angle. We found that Pix2Pix could generate spectrograms for different aspect angles when a spectrogram at the 0 angle is given. The accuracy of synthesizing spectrograms for the first scenario was higher than for the third scenario, which involved unknown activities. This technique can potentially serve as a data augmentation method to enhance the diversity of the radar data set in terms of aspect angles. The significance of this research is that we augmented the micro-Doppler data set from different aspect angles, increasing the diversity of the dataset. The augmented data set has the potential to improve classification accuracy.

Future research should include collection of large data sets of diverse activities to reduce errors for spectrogram synthesis of unknown activities. In addition, to validate the quality of the synthesized data, it will be necessary to include synthesized data in the training process for the target classification problem to ensure accuracy improvement. Furthermore, this research can be expanded to synthesize radar data from a different carrier frequency as well as to produce inverse synthetic aperture radar imagery from a different angular velocity of target motion.

## REFERENCES

- [1] V. Chen, *The Micro-Doppler Effect in Radar*. Norwood, MA, USA: Artech House, 2019.
- [2] Y. Kim and H. Ling, "Human activity classification based on micro-Doppler signatures using a support vector machine," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1328–1337, May 2009.
- [3] D. P. Fairchild and R. M. Narayanan, "Classification of human motions using empirical mode decomposition of human micro-Doppler signatures," *IET Radar, Sonar Navigat.*, vol. 8, no. 5, pp. 425–434, Jun. 2014.
- [4] A.-K. Seifert, M. G. Amin, and A. M. Zoubir, "Toward unobtrusive in-home gait analysis based on radar micro-Doppler signatures," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 9, pp. 2629–2640, Sep. 2019.
- [5] D. L. Hall, T. D. Ridder, and R. M. Narayanan, "Abnormal gait detection and classification using micro-Doppler radar signatures," in *Proc. 23rd Radar Sensor Technol.*, Baltimore, MD, USA, vol. 11003, Apr. 2019, pp. 110030Q-1–110030Q-11.
- [6] Q. Wu, Y. D. Zhang, W. Tao, and M. G. Amin, "Radar-based fall detection based on Doppler time–frequency signatures for assisted living," *IET Radar, Sonar Navigat.*, vol. 9, no. 2, pp. 164–172, Feb. 2015.
- [7] G. Li, R. Zhang, M. Ritchie, and H. Griffiths, "Sparsity-driven micro-Doppler feature extraction for dynamic hand gesture recognition," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 2, pp. 655–665, Apr. 2018.
- [8] Y. Deep, P. Held, S. S. Ram, D. Steinhäuser, A. Gupta, F. Gruson, A. Koch, and A. Roy, "Radar cross-sections of pedestrians at automotive radar frequencies using ray tracing and point scatterer modelling," *IET Radar Sonar Navigat.*, vol. 14, no. 6, pp. 833–844, 2020.
- [9] Y. Kim and T. Moon, "Human detection and activity classification based on micro-Doppler signatures using deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 1, pp. 2–8, Jan. 2016.
- [10] G. Klarenbeek, R. Harmanny, and L. Cifola, "Multi-target human gait classification using LSTM recurrent neural networks applied to micro-Doppler," in *Proc. Eur. Radar Conf.*, Oct. 2017, pp. 167–170.
- [11] Y. Xu and K. A. Scott, "Sea ice and open water classification of SAR imagery using CNN-based transfer learning," in *Proc. IGARSS*, Fort Worth, TX, USA, 2017, pp. 3262–3265.
- [12] I. Alnujaim, D. Oh, and Y. Kim, "Generative adversarial networks for classification of micro-Doppler signatures of human activity," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 3, pp. 396–400, Mar. 2020.
- [13] B. Erol, S. Z. Gurbuz, and M. G. Amin, "Motion classification using kinematically sifted ACGAN-synthesized radar micro-Doppler signatures," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 4, pp. 3197–3213, Aug. 2020.
- [14] F. Fioranelli, M. Ritchie, and H. Griffiths, "Multistatic human micro-Doppler classification of armed/unarmed personnel," *IET Radar, Sonar Navigat.*, vol. 9, no. 7, pp. 857–865, 2015.
- [15] D. P. Fairchild and R. M. Narayanan, "Multistatic micro-Doppler radar for determining target orientation and activity classification," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 1, pp. 512–521, Feb. 2016.
- [16] Z. Chen, G. Li, F. Fioranelli, and H. Griffiths, "Dynamic hand gesture classification based on multistatic radar micro-Doppler signatures using convolutional neural network," in *Proc. IEEE Radar Conf. (RadarConf)*, Apr. 2019, pp. 1–5.
- [17] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5967–5976.
- [18] Y. Li, R. Fu, X. Meng, W. Jin, and F. Shao, "A SAR-to-optical image translation method based on conditional generation adversarial network (cGAN)," *IEEE Access*, vol. 8, pp. 60338–60343, 2020.
- [19] S. Ram and H. Ling, "Simulation of human micro-Dopplers using computer animation data," in *Proc. IEEE Radar Conf.*, May 2008, pp. 1–6.
- [20] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Med. Image Comput. Comput.-Assist. Intervent.*, 2015, pp. 234–241.

• • •