Accessing the Effects of GANs with Different Loss Functions in Synthesizing Radar Micro-Doppler Signatures for Human Gait Recognition

Group: K

Abstract: Generative adversarial networks (GANs) have been recently proposed for the synthesis of RF micro-Doppler signatures to mitigate the problem of low sample support and enable the training of deeper neural networks (DNNs) for improved RF signal classification. However, when applied to human micro-Doppler (MD) signatures for gait analysis, GANs suffer from systemic kinematic discrepancies that degrade performance [11]. In this class project, we evaluated the kinematic validity of synthetic signatures generated by various GANs with varying loss functions. Particularly, we examined the synthetic samples from DCGAN with Binary Cross Entropy, WGAN with Wasserstein loss and a kinematically enhanced Multi-Branch Discriminator GAN. Kinematic validity was evaluated based on (1) Dynamic Time Wrapping distances between envelopes of Real and synthetic samples; (2) feature space comparison and (3) Classification accuracy comparison.

I. Introduction

Recently, deep learning has enabled a wide variety of new indoor monitoring applications [1], [2], ranging from home security and smart environments, to remote health monitoring and gesture-controlled non-contact user interfaces. These techniques are heavily reliant, however, on the availability of sufficient data for the training of deep neural networks (DNNs). A key challenge in the application of deep learning to RF datasets has been the greater difficulty, in comparison with computer vision, to acquire datasets of sufficient sample size and statistical variance to train DNNs with high accuracy and generalization. The generation of synthetic data is an attractive solution to this challenge as it also allows for training networks for scenarios that may, in practice, be difficult to acquire. Two approaches for synthetic data generation have been proposed: 1) model-based synthesis, and 2) synthesis with generative adversarial networks (GANs) [3]. Model-based approaches [4], [5] leverage video motion capture data to track the time-varying position of various points on the human body. From these trajectories, the expected RF micro-Doppler signature for the motion can be generated. Transformations that scale time and body size, as well as perturb the trajectories of the centroid of each body part, may be applied to generate many more statistically independent samples than the number of motion capture recordings acquired [6]. While this approach ensures a high degree of consistency in the data with how a human actually moves, it does not take into account backscattering from the entire human body, environmental sources of interference, sensor artifacts, or multipath, which may alter the received radar signal.

In this regard, generative adversarial networks (GANs) have own great promise, enabling applications such as style transfer in computer vision. Similarly, GANs have been proposed as a tool for synthesizing RF data [7]– [10]. However, in prior work [11], it has been shown that the RF micro-Doppler signatures synthesized by GANs possess systemic kinematic errors. In some cases, this makes the synthesized signature resemble other activity classes, such as a progressive fall (rather than a hard fall), walking in the reverse direction, or stopping for a moment before walking again. In other cases, the errors are more catastrophic and add artifacts that would never be observed in typical human motion, such as disjoint motion components. When these types of deviant signatures were tossed out of the synthetic training dataset, a performance increase of about 10% was observed for an 8-class activity recognition problem.

In this class project we evaluated the kinematic validity of synthetic signatures generated by various GANs with varying loss functions. Once we generate a reasonable amount of samples par class, we evaluated the quality of the signatures statistically through kinematic metric. Specially we investigated the envelope

matching techniques like Dynamic Time Wrapping. Accurate representation of the envelope for a motion is of particular importance because the envelope represents the maximum velocities of a body part during the motion. Synthetic points that fall outside this envelope will thus misrepresent the motion trajectory and physically attainable speeds. The significance of the envelope is underscored by recent work which demonstrates hand [12] and arm recognition [13] based solely on the micro-Doppler signature envelope.

By evaluating the synthetic signatures through envelope matching, we were able to quantify the number of erroneous samples each of the GAN generated. Thus, we make a conclusion about the best GAN architecture for Micro-Doppler synthesis.

II. Radar Data collection and Micro-Doppler Signatures Generation

The RF Data were collected with a software-defined radar, SDRKIT 2500B, developed by Ancortek, Inc. The radar was set to transmit a frequency modulated continuous waveform (FMCW) centered at 25 GHz with 1500 MHz bandwidth. The sensor was placed on top of a table of 1-meter height from the ground, with the test subject moving about 0.5-3 meters from the sensor. Six participants of various ages, heights and weights were involved in this study. Five different walking styles were acquired: namely, walking towards the radar (WTowards), short step walking (SHSTEP), Walking on toes (WTOES), Scissors gait walking (SCSSR) and walking with a Cane (WCANE). Each participant conducted 10 repetitions of each activity, resulting in a total of 60 samples per class. All types of walking were conducted along the radar line-of-sight.

The radar transmits LMF chirps and received the backscattered signal as complex I/Q time series, from which line-of-sight distance and radial velocity maybe computed. The amplitude and phase of the complex data is related to the electromagnetic scattering and kinematics of the target being observed. So, the time-varying positions of points on the body result in unique patterns in the time-frequency representation of the RF data. These are known as micro- Doppler frequencies, and are captured by the micro-Doppler signature, or, spectrogram, of the data. The spectrogram is computed as the square modulus of the Short-Time Fourier Transform (STFT) of the received signal. In this study, the STFT was computed using Henning windows with 50% overlap to reduce sidelobes in the frequency domain. Backscatter from static objects was removed using a moving target indicator (MTI) filter, while sensor noise and artifacts were mitigated using a thresholding algorithm.

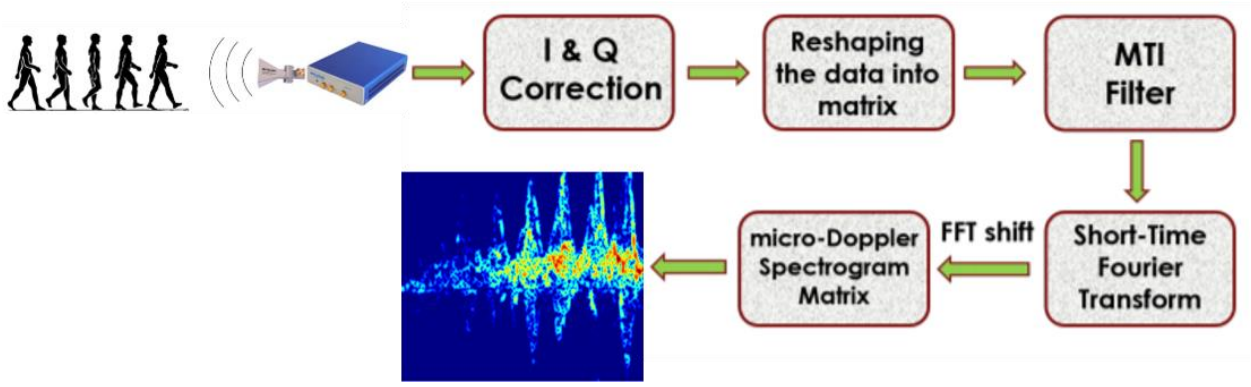


Figure 1: Micro-Doppler signatures generation from Raw Radar data.

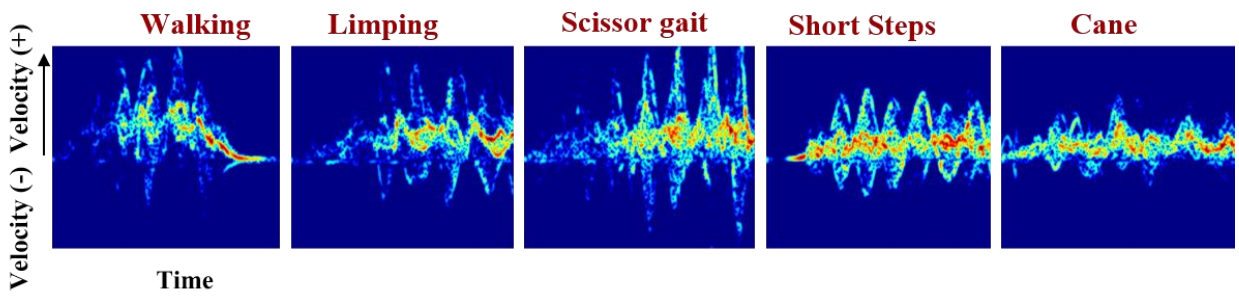


Figure 2: Micro-Doppler Spectrograms

III: Synthetic Data Generation With GAN

A. Deep Convolutional GAN (DCGAN):

DCGANs are the 'image version' of the most basic implementation of GANs. This architecture essentially leverages Deep Convolutional Neural Networks to generate images belonging to a given distribution from noisy data using the Generator-Discriminator framework.

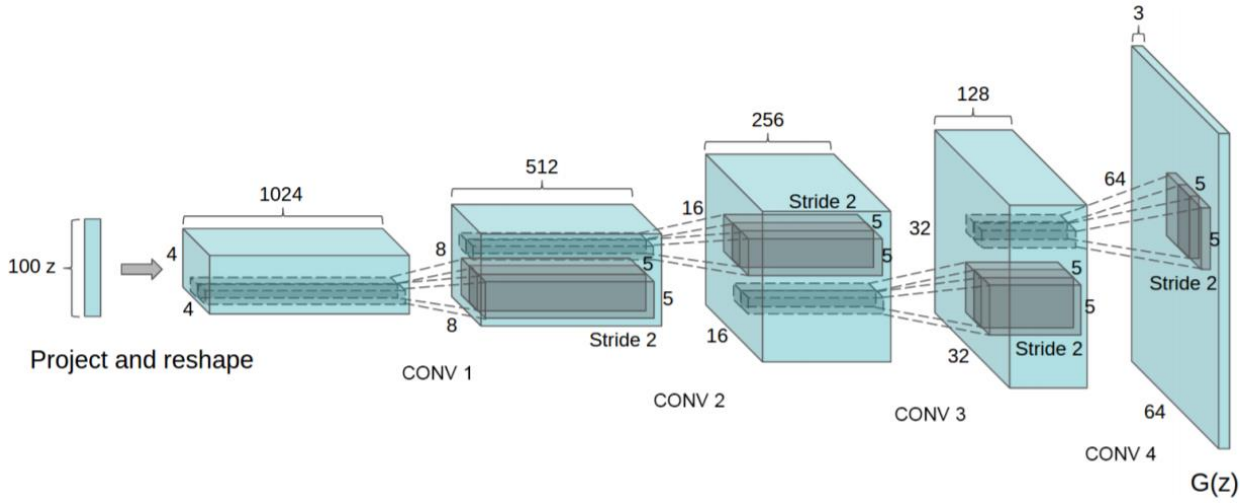


Figure 3: DCGAN generator architecture

The generator model of DCGAN mainly consists of Deconvolution layers or more accurately, Transposed Convolution layers that basically perform the reverse of a Convolution operation [2]. In our DCGAN implementation, we take a 100-dimensional latent space vector and map it to an 8x8x256 dimensional vector using a fully connected layer. This vector is reshaped to (8, 8, 256). These are essentially 256 activation maps of size 8x8. Further, we apply several Deconv layers and finally obtain a "3 channel image of size 128x128".

The discriminator is nothing but a binary classifier which consists of several convolution layers (same as any other image classification task). Finally, the flattened activation maps mapped to a probability distribution to predict if the image is real or fake. Since this is a two class, binary classification problem (real or fake), the ultimate loss function would be Binary Cross entropy

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i . \log\big(p(y_i)\big) + (1 - y_i) . \log(1 - p(y_i))$$

However, this loss is adjusted and applied to both of the networks separately in order to optimize their objective. When we train the DCGAN, we observed that after some number of epochs, all the generated images become the same as can be seen in Figure 2. This issue is also known as "mode collapse".

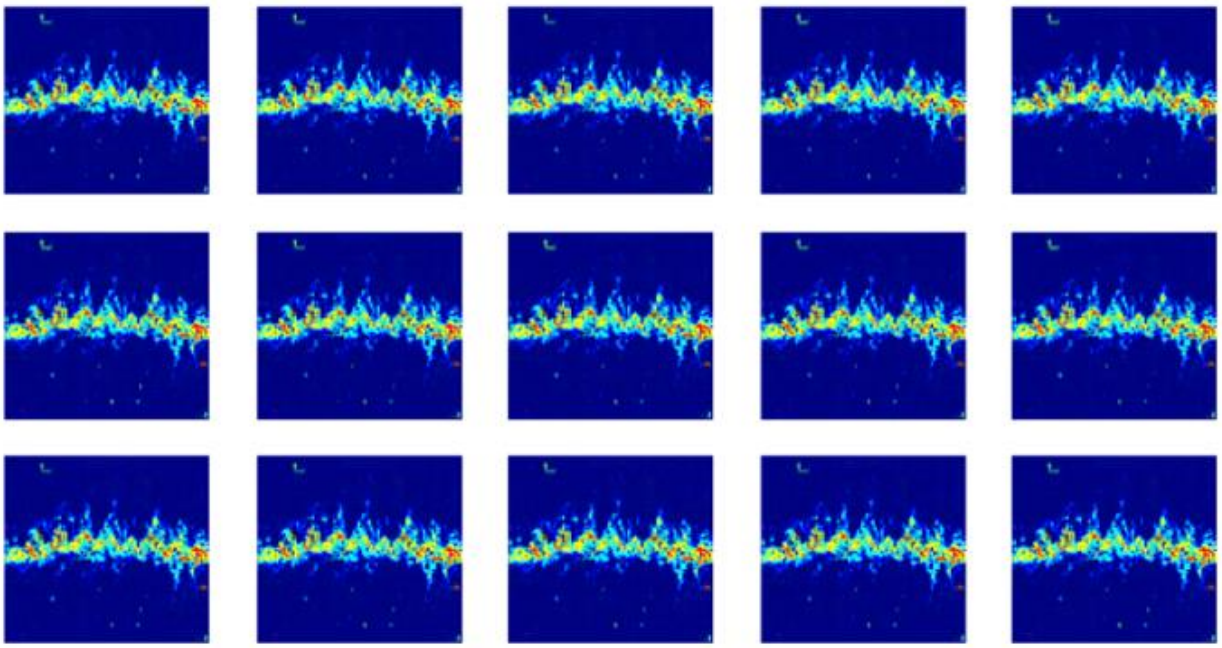


Figure 4: DCGAN generated synthetic images with mode collapse.

We tried different methods to overcome this problem such as with and without normalization, using Batch Normalization, reducing the number of convolutional layers, changing some of the activation functions from ReLU to **tanh**, changing the dimensions of the latent space vector, adding more fully connected layers with dropout etc. but none of these solved the problem. Thus, we didn't use DCGAN results in the final benchmarking stage.

B. Wasserstein GAN (WGAN)

The generator network samples a predefined latent space and up samples via transposed or deconvolutional layers to produce a synthetic image whereas the discriminator network takes that synthetic images as input and attempts to classify them as being real or fake. As the discriminator gets increasingly better, the gradient vanishes, meaning there is no gradients to update the loss during the training process. Wasserstein GAN (WGAN) [14] treats this vanishing gradient problem by applying a gradient penalty (GP) after every gradient update on the discriminator/critic function and hence the name WGAN-GP comes up. To provide a more stable training process, with proven convergence of the loss function, WGAN-GP uses a new loss function derived from the Wasserstein distance; as the asymmetric Kullback–Leibler (KL) divergence causes buggy results when the intention is just to measure the similarity between two equally important distributions and the Jenson–Shannon divergence fails to provide a meaningful value when two distributions are disjointed.

KL Divergence:

$$D_{KL}(p||q) = \sum_{i=1}^{N} p(x_i).(\log \frac{p(x_i)}{q(x_i)})$$

Wasserstein Loss:

$$\min_{g} \max_{c} \mathrm{E}\big(c(x)\big) - \mathrm{E}(c(g(z)))$$

Where, c(x) denotes the critic's output for a real instance, g(z) denotes the generator's output when given noise z and c(g(z)) is the critic's output for a fake instance.

The whole process of WGAN is shown in figure 5.

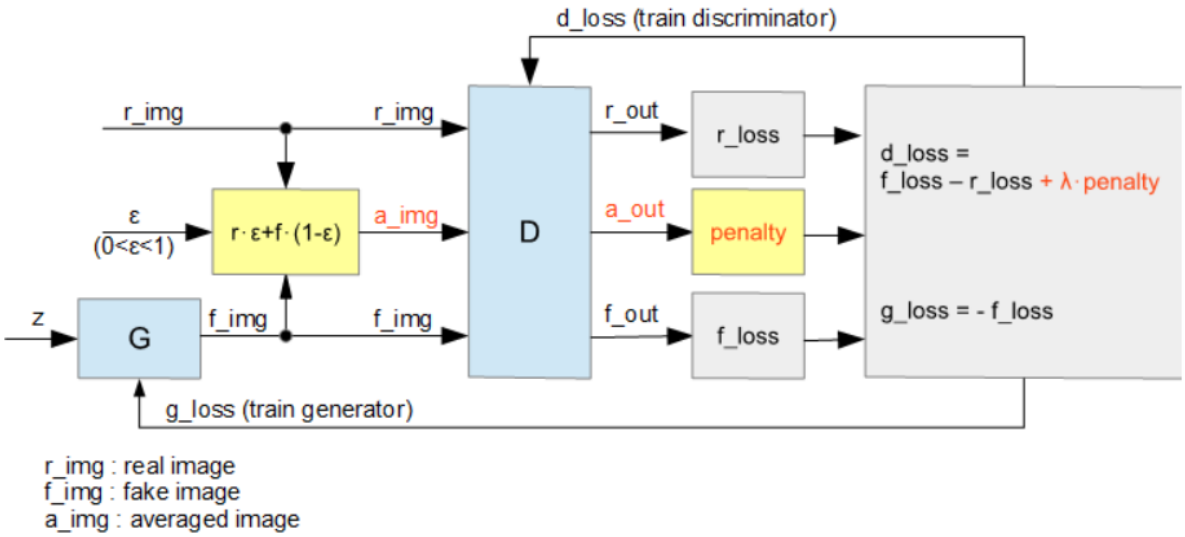


Figure 5: Outline of WGAN

Some of the synthetic images generated by WGAN are shown in figure 6. However, with the WGAN generated signatures, it was observed that there is a lack of kinematic fidelity in a significant percentage of signatures generated. Many of these samples have features that are deviant from the typical properties of micro-Doppler, such as high frequency components disconnected from the low-frequency micro-Doppler, negative micro-Doppler corresponding to motion in the reverse direction etc. Such kinematically inaccurate samples are also shown in figure 6.

A more statistical comparison among real data vs good synthetic data and inaccurate/ outlier's samples can be draws through PCA based feature space plotting as shown on the right side of figure 6. It is evident from the feature space that good synthetic samples are closer to the real samples and hence reflects the presence of the same features in synthetic data as that in real data. However, the outlier's samples are falling apart from the real samples as they carry the inaccurate kinematics. Using these inaccurate signatures in training will degrade the classification performance. So how can we do better in data generation. As most of the errors are because the GANs were not able to follow the motion trajectories accurately, so applying some constraints to impose the motions trajectory will help generating less erroneous spectrograms.

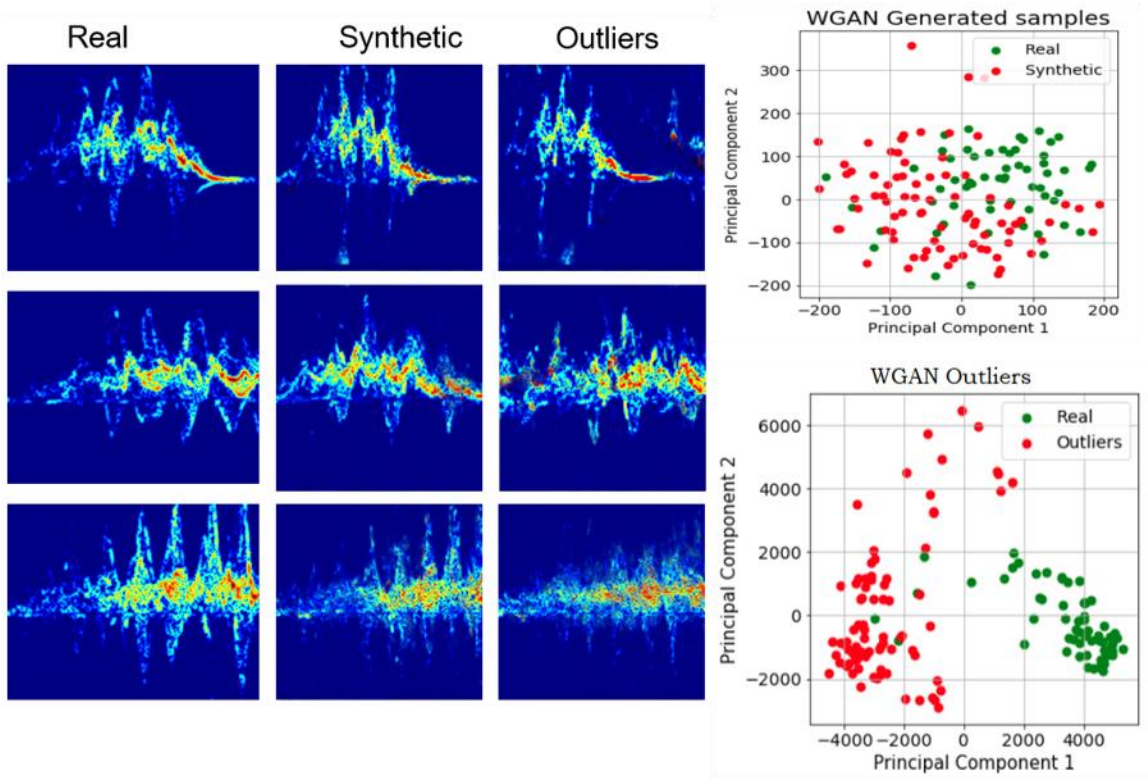


Figure 6: WGAN Generated signatures and feature space using PCA

C. Multi- Branch Discriminator GAN (MBGAN)

To maintain critical kinematic features of the data by better capturing gross properties of the micro-Doppler signature in adversarial learning process, the signature's envelope can be incorporated as an additional, second branch into the discriminator of a WGAN, which we call Multi- Branch Discriminator GAN.

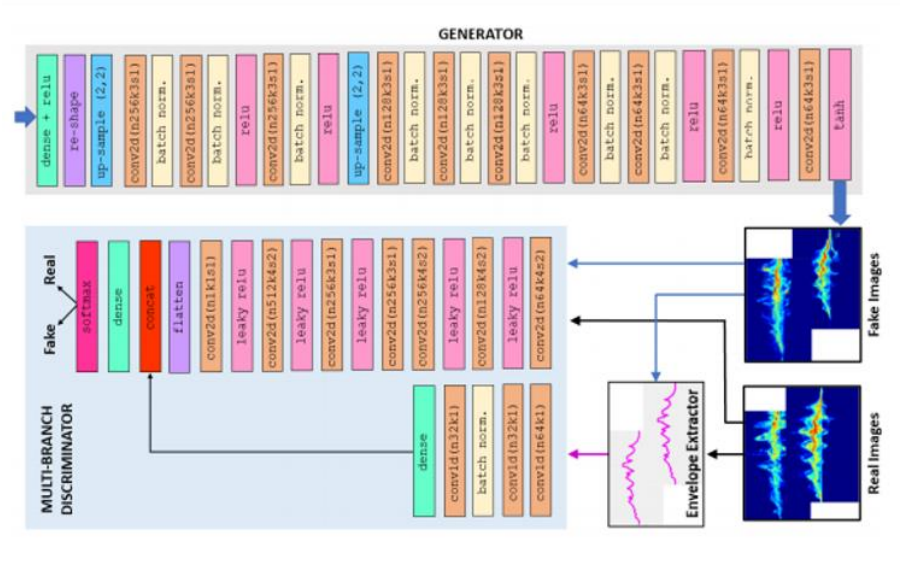


Figure 7: MBGAN Architecture

IV. Results

A. Feature space comparison:

The quality of the WGAN and MBGAN generated samples can be accessed through plotting them with respect to real data in PCA feature space and visually inspect the overlapping between them. From figure 8 we can see that the MBGAN samples overlap more with the real samples compared to the WGAN samples. In the encircled region, WGAN failed to overlap whereas MBGAN perfectly overlap on that region as well.

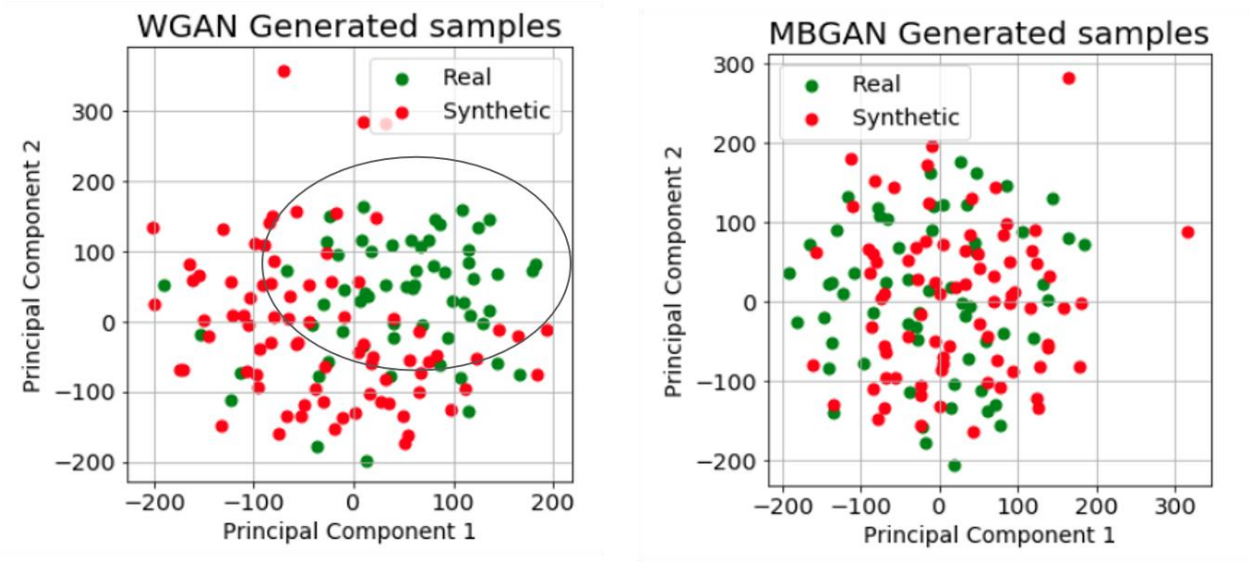


Figure 8: Comparison of Feature space

B. Outlier Detection Through Curve Matching:

Considering the envelope as a time series or a curve, the similarity between curves can be measured in a way that takes into account both the location and ordering of the points along the curve. It is noted that different measures for curve matching appear in several application domains, including time series analysis, shape matching, speech recognition, and signature verification. In time series analysis, Dynamic Time Wrapping is an algorithm for measuring the similarity between two temporal sequences that may vary in time or speed. In our work, we first extract the envelope from Real samples for a particular class. Then we measured the DTW distance among the all combinations of the real spectrogram's envelopes and took an average of them which we used as threshold. We also make an average envelope by taking mean of all the envelopes. Then for each synthetic sample, we first extract the envelope and measured its distance from the average envelope. If the distance is less than or equal to the threshold, then we considered it as a valid/ accurate signature, otherwise we marked it as an outlier. In this way we sifted out all the generated samples and found that out of 2500 samples across 5 classes, WGAN produced 373 (14%) outliers where as MBGAN produces 191 (8%) outliers. Figure 9 shows the envelope extraction and total outliers.

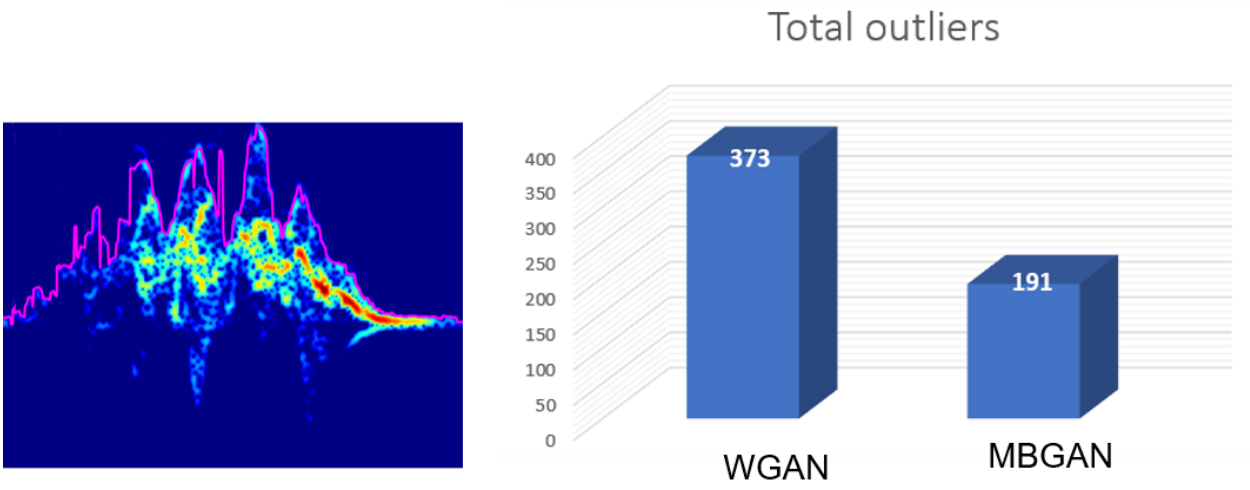


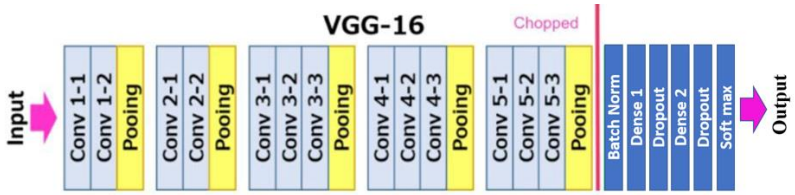
Figure 9: Envelope extraction and total outliers.

C. Classification Performance

Finally, two measure synthetic samples kinematic quality, we've trained VGG16 net as a classifier with the samples generated by WGAN and MBGAN separately. We tased the classifier with the real data. As a baseline, we trained the VGG net with Real samples only and tested with real samples. By comparing with this baseline, we see that training with the large amount of WGAN synthetic samples boost the performance by 12% whereas with MBGAN we get more than 14% performance boost. Clearly, MBGAN performs better than WGAN here. All the hyperparameters are listed in the figure 10.

No.of classes: 5
Training samples:
Real data: 60 samples/class
GAN synthesized data: 500/class
Test samples: 10 samples/class

Classifier: VGG16
Learning Rate: 0.0002
Batchsize: 64
Epoch: 100
Early stopes: Patience= 5
Loss: Categorical Cross-Entropy
Dropout: 0.5



Training set	Testset	Accuracy (%)
Real Data	Real Data	81.25
WGAN Synthesized Data	Real Data	93.70
MBGAN Synthesized Data	Real Data	95.72

Figure 10: VGG16 net with all the hyper parameters and the classification performance.

V. Conclusion

In this class projects, we Evaluated the DCGAN with Binary cross entropy, WGAN with EMD loss function and Multi-Branch Discriminator GAN (MBGAN) for synthesizing Radar micro-Doppler signatures. We found that, MBGAN outperform rest of the GANS in terms of generating a smaller number of outliers, greater overlap with real samples and providing best classification performance. The reason MBGAN

performs so good is because of embedding the envelope kinematics into the multi-branch discriminator, which improved the kinematic fidelity of the generated samples.

REFERENCES

[1] M. Amin, Radar for Indoor Monitoring: Detection, Classification, and Assessment. CRC Press, 2017.

[2] S. Z. Gurbuz and M. G. Amin, "Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring," IEEE Signal Processing Magazine, vol. 36, no. 4, pp. 16–28, 2019.

[3] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in Proc. 27th Int. Conf. Neural Information Proc. Systems - Volume 2, ser. NIPS'14. Cambridge, MA, USA: MIT Press, 2014, p. 2672–2680.

[4] S. Sundar Ram and H. Ling, "Simulation of human microdopplers using computer animation data," in IEEE Radar Conference, 2008, pp. 1–6.

[5] B. Erol, C. Karabacak, and S. Z. Gurbuz, "A kinect-based human microdoppler simulator," IEEE Aerospace and Electronic Systems Magazine, vol. 30, no. 5, pp. 6–17, 2015.

[6] M. S. Seyfioglu, B. Erol, S. Z. Gurbuz, and M. G. Amin, "Dnn transfer learning from diversified micro-doppler for motion classification," IEEE Trans. Aerosp. Electronic Sys., vol. 55, no. 5, pp. 2164–2180, 2019. [7] J. Guo, B. Lei, C. Ding, and Y. Zhang, "Synthetic aperture radar image synthesis by using generative adversarial nets," IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 7, pp. 1111–1115, 2017.

[8] B. Erol, S. Z. Gurbuz, and M. G. Amin, "Gan-based synthetic radar micro-doppler augmentations for improved human activity recognition," in 2019 IEEE Radar Conference (RadarConf), 2019, pp. 1–5.

[9] I. Alnujaim, D. Oh, and Y. Kim, "Generative adversarial networks to augment micro-doppler signatures for the classification of human activity," in IEEE Int. Geosci. Rem. Sens. Symp., 2019, pp. 9459–9461.

[10] H. G. Doherty, L. Cifola, R. I. A. Harmanny, and F. Fioranelli, "Unsupervised learning using generative adversarial networks on microdoppler spectrograms," in 16th Eur. Radar Conf., 2019, pp. 197–200.

[11] B. Erol, S. Z. Gurbuz, and M. G. Amin, "Motion classification using kinematically sifted acgan-synthesized radar micro-doppler signatures," pp. 3197–3213, 2020.

[12] M. G. Amin, Z. Zeng, and T. Shan, "Hand gesture recognition based on radar micro-doppler signature envelopes," in 2019 IEEE Radar Conference (RadarConf), 2019, pp. 1–6.

[13] Z. Zeng, M. G. Amin, and T. Shan, "Automatic arm motion recognition based on radar micro-doppler signature envelopes," IEEE Sensors Journal, vol. 20, no. 22, pp. 13 523–13 532, 2020.

[14]T. hang, Z. Li, Q. Zhu, and D. Zhang, "Improved procedures for training primal wasserstein gans," in 2019 IEEE Smart World, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (Smart World/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), 2019, pp. 1601–1607.