A Novel Micro-Doppler Coherence Loss for Deep Learning Radar Applications

Mikolaj Czerkawski, Christos Ilioudis, Carmine Clemente, Craig Michie, Ivan Andonovic, Christos Tachtatzis

Abstract — **Deep learning techniques are subject to increasing adoption for a wide range of micro-Doppler applications, where predictions need to be made based on time-frequency signal representations. Most, if not all, of the reported applications focus on translating an existing deep learning framework to this new domain with no adjustment made to the objective function. This practice results in a missed opportunity to encourage the model to prioritize features that are particularly relevant for micro-Doppler applications. Thus the paper introduces a micro-Doppler coherence loss, minimized when the normalized power of micro-Doppler oscillatory components between input and output is matched. The experiments conducted on real data show that the application of the introduced loss results in models more resilient to noise.**

Keywords — **Doppler radar, Micro-Doppler, Deep Learning, Radar Classification**

I. Introduction

Advancements in deep learning have motivated diverse research in their application to radar signal processing since many of the relevant tasks can be described as instances of classification or domain translation. The paper reports on the enhanced performance of deep learning models processing Doppler-time radar signals, achieved by changing the content of the objective function.

Although a number of deep learning approaches have been applied to radar signal processing, research to adjust the objective function within this domain has been limited. In [1], one of the early applications of a Convolutional Neural Network (CNN) for the classification of radar signals is reported and although a performance improvement is evident owing to deep learning techniques, the type of loss used is not revealed. The unsupervised approach utilising a stacked auto-encoder architecture presented in [2] utilised a cost function consisting of a reconstruction term, based on the Mean Squared Error (MSE), a weight regularisation term, and a divergence term between a sparsity parameter and the average output of the hidden neurons. A CNN model for human activity classification is introduced in [3], but the exact characteristics of the objective function are not disclosed. A transfer learning approach using a novel DivNet architecture for human motion classification is proposed in [4], without modification of the objective function content. An unsupervised approach to learning relevant features from radar micro-Doppler spectrograms in [5] applies a standard auto-encoder objective function. A significant body of research related to the classification of radar signals has adopted the same methodologies [6]–[13].

Here, a micro-Doppler coherence loss (an additional term within the objective function) is introduced, applicable in a wide range of frameworks operating on radar Doppler signals. The micro-Doppler coherence loss improves results in an unsupervised learning scheme applied for a classification task by promoting aligned periodic characteristics of the reconstructed signal in individual velocity bands to that of the ground truth signal. Results indicate that deep neural networks performing tasks related to micro-Doppler analysis can achieve superior immunity to injected noise when trained using this loss.

II. Problem Formulation

Two representative application contexts are selected to position the scope of the reported results. First, a network executing domain translation where a given time-frequency map is transformed to a different domain - de-noising or interference removal fall into this category; second is classification, one of the common uses of deep learning for radar applications.

In general, each application requires the network to learn relevant features from the input time-frequency map. In the case of domain translation, the features establish a representation used to decode an appropriate output, while in the case of classification, these features form the input to a classification output module (for instance, a relatively shallow stack of fully connected layers). The feature set learned by the encoding module significantly impacts the performance of the network in both cases. Consequently, the challenge of learning the relevant features constitutes the problem.

Furthermore, reliance on the standard losses widely adopted outside of the radar context can advertise features that are not relevant in the goal of interpreting Doppler signals. The most commonly applied objective function J_θ for fully convolutional networks with parameters θ, is the MSE reconstruction loss between the target y and network output $\hat{y}$ (both of size $M \times N$):

$$J_\theta(y, \hat{y}) = \mathcal{L}_{\text{MSE}}(y, \hat{y}) = \frac{1}{N}\frac{1}{M}\sum_{f_D=1}^{N}\sum_{t=1}^{M}(y_{t,f_D} - \hat{y}_{t,f_D})^2 \quad (1)$$

This loss does not prioritise any set of high-level features over others in the sense that the value of each pixel has an equal and direct influence on the error. The choice of prioritised features is challenging, but if the assumption that the relevant information is not uniformly spread over all pixels holds true, then an improvement in performance can be expected through adjustment of the loss terms.
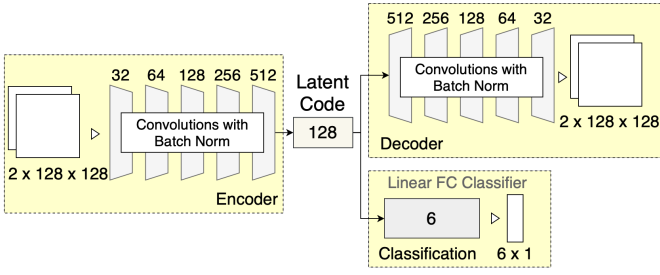
Fig. 1. Diagram of the used hybrid model.

III. PROPOSED SOLUTION

The modulations of Doppler components contain crucial information in micro-Doppler applications. Consequently, uniform contribution of each time-frequency bin to the total error may not be appropriate. A new loss term of micro-Doppler coherence loss $\mathcal{L}_{\mu D}$ with a weight β is proposed to promote the relevant micro-Doppler features. The introduced loss term is designed to promote spectral similarity within each Doppler band. This is achieved by formulating a metric that will be minimized when the normalized spectral distribution of each Doppler band is the same for both compared signals.

$$J_\theta(y, \hat{y}) = \mathcal{L}_{\text{MSE}}(y, \hat{y}) + \beta \cdot \mathcal{L}_{\mu\,\text{D}}(y, \hat{y}) \qquad (2)$$

For $\mathcal{L}_{\mu\,\text{D}}$, both a 2D time-frequency ground truth matrix $y[t, f_D]$ and a corresponding network output $\hat{y}[t, f_D]$ are subject to discrete Fourier transform $\mathcal{F}_t$ applied in the temporal dimension t, transforming it to cadence frequency $t \rightarrow f_c$. This yields two tensors representing Doppler-cadence maps (still of size $M \times N$). The magnitude of the resulting 2D map $\mathcal{F}_t[y]$ is extracted with integral normalized to 1 to obtain the final representation $\mathcal{C}[f_c, f_D]$, constituting a normalized Doppler-cadence magnitude map defined as:

$$\mathcal{C}[f_c, f_D] = \frac{|\mathcal{F}_t(y)]|}{\sum_{f_c=1}^M \sum_{f_D=1}^N |\mathcal{F}_t(y)|_{f_c, f_D}} \qquad (3)$$

An identical derivation applied to $\hat{y}[t, f_D]$ yields $\hat{\mathcal{C}}[f_c, f_D]$.

Since $\mathcal{C}[f_c, f_D]$ and $\hat{\mathcal{C}}[f_c, f_D]$ are normalized 2D representations, the form of the $\mathcal{L}_{\mu\,\text{D}}(y, \hat{y})$ loss term will be similar to (1), since MSE is used to compare them:

$$\mathcal{L}_{\mu\,\text{D}}(y, \hat{y}) = \frac{1}{N} \sum_{f_D=1}^N \underbrace{\sum_{f_c=1}^M \frac{(\mathcal{C}_{f_c, f_D} - \hat{\mathcal{C}}_{f_c, f_D})^2}{M}}_{\mathcal{S}[f_D]} \qquad (4)$$

The new term in $\mathcal{L}_{\mu\,\text{D}}$ in the objective function J_θ in (2) emphasizes features directly related to Micro-Doppler oriented tasks. The initial phase information for each oscillatory mode within a Doppler frequency bin is ignored by applying the magnitude operation. The periodicity of the Doppler components has to match the truth in order to minimise the micro-Doppler coherence loss. Furthermore, the loss is

invariant to small shifts in time of individual micro-Doppler spectral components[1].

IV. EVALUATION

The utility of the micro-Doppler coherence loss is demonstrated using a dataset containing real radar signatures of various human activities. The dataset is publicly available[2], containing 1,752 samples with ground truth [14]. The samples contain signatures from 6 different activities: 1) Walking, 2) Sitting down, 3) Standing up, 4) Object Pick Up, 5) Drinking, 6) Fall. The samples are divided into training, validation and test datasets with ratios of (0.5, 0.25, 0.25), respectively. Since the class imbalances in the dataset are not severe, no balancing countermeasures are applied.

The general learning approach is similar to [5], where classification training is preceded by an unsupervised stage enabling the evaluation in the improvement in both the domain translation as well as the classification contexts.

The structure of the model is shown in Figure 1. The network utilises 128 by 128 spectrogram images with 2 channels to accommodate real and imaginary components. The spectrogram is computed with 128 bins in a 0.2 seconds Blackman window and 0.19 seconds overlap. The resulting spectrogram image is then uniformly sampled in time to obtain 128 spectra, yielding a 128 by 128 complex matrix. The model encoder translates this image to a latent code of size 128, as shown in Figure 1. All convolutional layers in the network use a kernel of size 3 with a stride of 2. The latent code is then input to the decoder module; alternatively, the same code can be fed to the classifier module. The hybrid structure allows for convenient switching between the translation and the classification operation.

The results rely on a comparison between a network where only MSE reconstruction loss $\mathcal{L}_{\text{MSE}}$ is contained in the objective function J_θ and a network where the micro-Doppler loss term $\mathcal{L}_{\mu D}$ is added to the objective with a weight β of 4.

A. Unsupervised Stage

The influence of the proposed loss term can be demonstrated by investigating the loss curves of the trained convolutional auto-encoding model component. In the long term, the decay of the reconstruction loss can be expected to drive the micro-Doppler loss down also. However, the degree to which the two losses are coupled remains to be demonstrated. This stage also provides confirmation on whether using the additional loss term significantly changes the direction of gradients used for backpropagation.

Figure 2 illustrates how both the reconstruction and micro-Doppler losses vary with each weight update. In the case of the MSE-only objective (black), both losses are reduced in the long term; however, the latter decays at a slower rate than in the case of the proposed objective function (red). Thus the reconstruction loss and the micro-Doppler loss gradients

---

[1]However, significant shifts in time of individual components will be penalised by $\mathcal{L}_{\text{MSE}}$ due to the resulting pixel error.

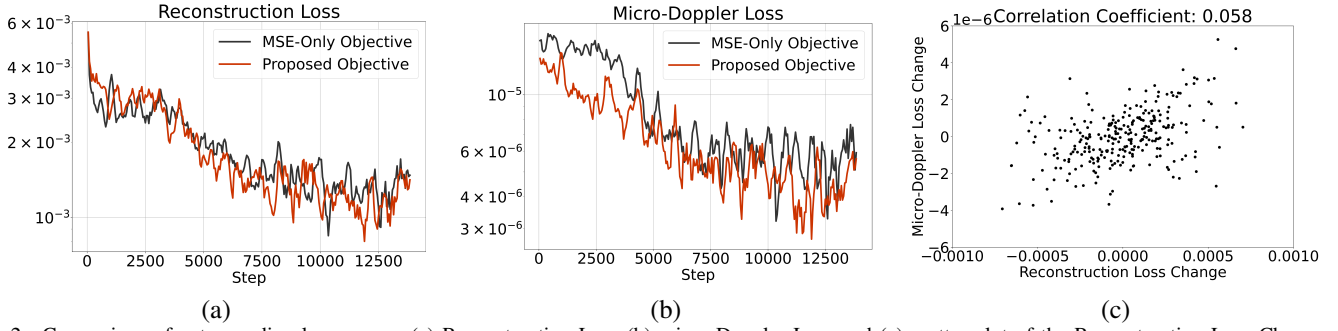[2]Available at http://researchdata.gla.ac.uk/848/

Fig. 2. Comparison of autoencoding loss curves. (a) Reconstruction Loss (b) micro-Doppler Loss and (c) scatter plot of the Reconstruction Loss Change and micro-Doppler Loss Change

are only partially aligned, implying that each loss can be associated with a different set of learned features. Further confirmation can be obtained by investigating the correlation between the change in the MSE-only reconstruction and micro-Doppler losses for the scenario where the objective function contains only the MSE loss (black). The scatter plot in Figure 2(c) illustrates that relationship. The correlation coefficient between the two variables is 0.058 suggesting minimal dependence between the variables and confirming that the gradients propagated from the micro-Doppler loss generally point in a different direction than in the case of an objective based only on the MSE loss.

B. Classification

Two sets of model weights pre-trained in the unsupervised stage have been used subsequently to train a classifier head to discriminate between the three classes contained in the dataset (only cross-entropy loss is included in the objective function at this stage). Figure 3(a) shows the validation curves for the two sets of pre-trained weights. Evident is that the weights obtained using the addition of micro-Doppler loss in the unsupervised stage (red) lead to a smoother validation loss decay than the standard MSE-only objective (black). The dots mark the lowest validation loss achieved in each case; 0.50647 for the standard approach, and 0.52433 for the proposed micro-Doppler coherence loss. The weights from these states have been extracted in order to compare the two approaches.

Using the best-performing weights, the accuracy of the classifier has been tested against varying levels of additive white noise (SNR swept from 10 to -10 dB), as shown in Figure 3(b). Results indicate that the application of the proposed micro-Doppler coherence loss yields a model more robust to noise compared to the conventional approach.

The advantage gained in the context of the classification task is further observed in the confusion matrices for the level of noise where the difference of accuracy is most significant. The confusion matrices presented in Fig. 4 demonstrate the performance achieved by both networks with injected additive input noise of SNR equal to -5 dB. Further, the network output for each sample has been computed for 16 different noise samples in order to obtain a representative example. The proposed micro-Doppler coherence loss results in an increase in accuracy from 0.48 to 0.61 (marked by the two vertically
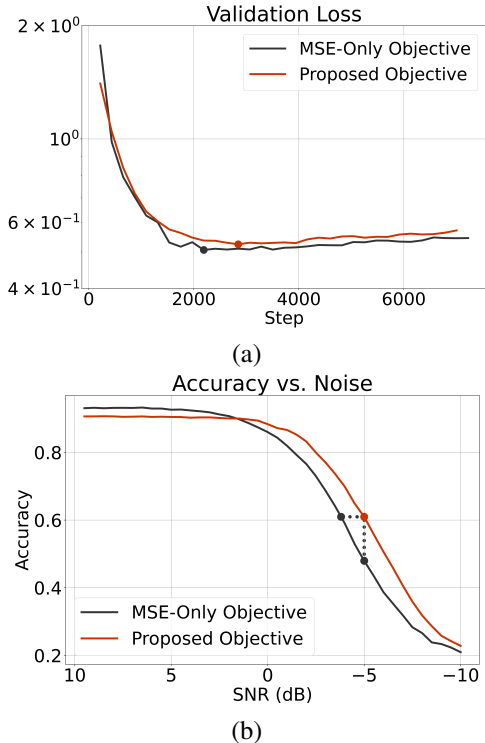

Fig. 3. Comparison of the relationship between accuracy and SNR of tested samples
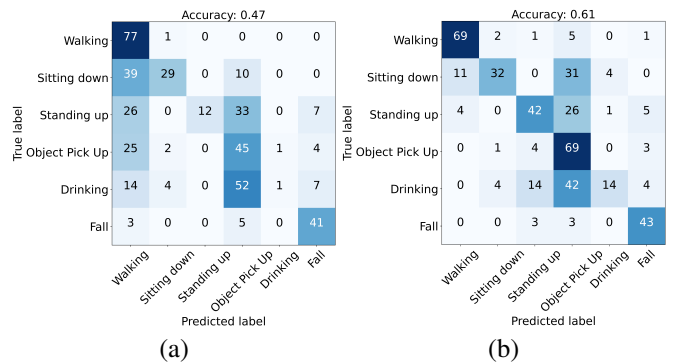

Fig. 4. Comparison of obtained confusion matrix at -5 SNR_{dB} for (a) a scenario with no micro-Doppler loss applied (b) a scenario with micro-Doppler loss backpropagation.

aligned dots). Conversely, for a set accuracy level of 0.61, the proposed approach can accommodate an additional 1.2 dB of added noise with no drop in accuracy. The confusion matrix for the model trained using the micro-Doppler coherence loss is shown in Figure 4(b). The number of correctly classified samples for almost all classes increase significantly compared to the standard approach in Figure 4(a). The numbers of correct predictions for the Walking and the Fall class are lower for the proposed approach, however, the differences are minimal. Nevertheless, the total number of correct predictions is higher for the model trained with micro-Doppler coherence loss over a range of noise level values as illustrated in Figure3(b).

V. Conclusions

A novel coherence loss term has been proposed for training deep learning models operating on Doppler time-frequency representations. Inclusion of the loss term in the objective function provides more appropriate optimization gradients for micro-Doppler applications. Results indicate that this practice can be beneficial not only when the network output target is a time-frequency map but also in a classification framework. The new loss term utilised in the unsupervised pre-training stage leads to a classifier significantly more resilient to noise, making the model accuracy invariant to approximately 1.2 dB of additional noise, or 10 percentage points higher accuracy for the same level of noise is achieved.

References

[1] Y. Kim and T. Moon, "Human detection and activity classification based on micro-doppler signatures using deep convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1, pp. 8–12, 2016, ISSN: 1545598X. DOI: 10.1109/LGRS.2015.2491329.

[2] B. Jokanovic, M. Amin, and F. Ahmad, "Radar fall motion detection using deep learning," *2016 IEEE Radar Conference, RadarConf 2016*, pp. 1–6, 2016. DOI: 10.1109/RADAR.2016.7485147.

[3] Y. Lang, C. Hou, Y. Yang, D. Huang, and Y. He, "Convolutional Neural Network for Human Micro-Doppler Classification," in *European Microwave Conference*, 2017, pp. 497–500, ISBN: 9782874870477.

[4] M. S. Seyfioglu, B. Erol, S. Z. Gurbuz, and M. G. Amin, "DNN Transfer Learning from Diversified Micro-Doppler for Motion Classification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 5, pp. 2164–2180, 2018, ISSN: 15579603. DOI: 10.1109/TAES.2018.2883847. arXiv: 1811.08361.

[5] M. S. Seyfioğlu, A. M. Özbayoğlu, and S. Z. Gürbüz, "Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1709–1723, 2018, ISSN: 00189251. DOI: 10.1109/TAES.2018.2799758.

[6] A. Shrestha, C. Murphy, I. Johnson, A. Anbulselvam, F. Fioranelli, J. Le Kernec, and S. Z. Gurbuz, "Cross-frequency classification of indoor activities with DNN transfer learning," *2019 IEEE Radar Conference, RadarConf 2019*, pp. 1–6, 2019. DOI: 10.1109/RADAR.2019.8835844.

[7] B. Erol, S. Z. Gurbuz, and M. G. Amin, "Frequency-Warped Cepstral Heatmaps for Deep Learning of Human Motion Signatures," *Conference Record - Asilomar Conference on Signals, Systems and Computers*, vol. 2018-Octob, pp. 1234–1238, 2019, ISSN: 10586393. DOI: 10.1109/ACSSC.2018.8645178.

[8] B. Erol, S. Z. Gurbuz, and M. G. Amin, "GAN-based synthetic radar micro-doppler augmentations for improved human activity recognition," *2019 IEEE Radar Conference, RadarConf 2019*, pp. 1–5, 2019. DOI: 10.1109/RADAR.2019.8835589.

[9] I. Alnujaim, D. Oh, and Y. Kim, "Generative Adversarial Networks for Classification of Micro-Doppler Signatures of Human Activity," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 3, pp. 396–400, 2020, ISSN: 15580571. DOI: 10.1109/LGRS.2019.2919770.

[10] L. Wang, J. Tang, and Q. Liao, "A Study on Radar Target Detection Based on Deep Neural Networks," *IEEE Sensors Letters*, vol. 3, no. 3, pp. 1–4, 2019, ISSN: 2475-1472. DOI: 10.1109/lsens.2019.2896072.

[11] A. Huizing, M. Heiligers, B. Dekker, J. De Wit, L. Cifola, and R. Harmanny, "Deep Learning for Classification of Mini-UAVs Using Micro-Doppler Spectrograms in Cognitive Radar," *IEEE Aerospace and Electronic Systems Magazine*, vol. 34, no. 11, pp. 46–56, 2019, ISSN: 1557959X. DOI: 10.1109/MAES.2019.2933972.

[12] 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. DOI: 10.1109/MSP.2018.2890128.

[13] X. Li, Y. He, and X. Jing, "A survey of deep learning-based human activity recognition in radar," *Remote Sensing*, vol. 11, no. 9, 2019, ISSN: 20724292. DOI: 10.3390/rs11091068.

[14] F. Fioranelli, S. A. Shah, H. Li, A. Shrestha, and J. Le Kernec, *Radar signatures of human activities*, en, 2019. DOI: 10.5525/GLA.RESEARCHDATA.848. [Online]. Available: http://researchdata.gla.ac.uk/id/eprint/848.