Received: 11 March 2022 | Revised: 28 October 2022 | Accepted: 9 December 2022

IET Radar, Sonar & Navigation

DOI: 10.1049/rsn2.12369

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

Attention-enhanced Alexnet for improved radar micro-Doppler signature classification

Shelly Vishwakarma[1] | **Wenda Li**[2] | **Chong Tang**[2] | **Karl Woodbridge**[3] | **Ravi Raj Adve**[4] | **Kevin Chetty**[2]

[1]Department of Electronic and Computer Science, University of Southampton, Southampton, UK

[2]Department of Security and Crime Science, University College London, London, UK

[3]Department of Electronic and Electrical Engineering, University College London, London, UK

[4]Department of Electrical and Computer Engineering, University of Toronto, Toronto, Ontario, Canada

Correspondence

Shelly Vishwakarma, Department of Electronic and Computer Science, University of Southampton, Southampton, UK.
Email: s.vishwakarma@soton.ac.uk

Funding information

Engineering and Physical Sciences Research Council (EPSRC), Grant/Award Number: EP/R018677/1

Abstract

This work introduces an attention mechanism that can be integrated into any standard convolution neural network to improve model sensitivity and prediction accuracy with minimal computational overhead. The attention mechanism is introduced in a lightweight network – Alexnet and its classification performance for human micro-Doppler signatures is evaluated. The Alexnet model trained with an attention module can implicitly highlight the salient regions in the radar signatures while suppressing the irrelevant background regions and consistently improving network predictions. Network visualizations are provided through class activation mapping, providing better insights into how the predictions are made. The visualizations demonstrate how the attention mechanism focusses on the region of interest in the radar signatures.

KEYWORDS

image classification, micro Doppler, radar target recognition

1 | INTRODUCTION

In recent years, deep convolutional neural networks (DCNNs) have become the state-of-the-art method for classifying human micro-Doppler signatures [1–4]. DCNNs can jointly learn informative features and classification boundaries, resulting in them being an order of magnitude faster than traditional approaches that use additional feature extraction algorithms. The success of DCNNs is attributed to the ever-increasing processing speeds of computers, greater availability of digitally recorded data, and almost unlimited memory capacity.

Unlike the vision community, radar researchers are constrained by the limited availability of open radar databases. Therefore, researchers have used different DCNN initialization methods for micro-Doppler classification with low training sample support. One such method is using a transfer learning technique where pre-trained networks from optical imagery (such as AlexNet, VGGNet, and GoogleNet) are trained with a limited radar dataset [5–8]. In this way, transfer learning approaches can simultaneously mitigate the need for a more extensive database for training and reduce the time required to train the classifier. However, the performance using shallow networks, such as Alexnet, remained sub-optimal, possibly due to the low interpretability of the radar micro-Doppler signatures, especially at lower carrier frequencies.

To address this general problem, we propose a simple and yet effective solution, called attention mechanism [9–11]. The attention mechanism can automatically localize and highlight the salient regions of interest in the radar micro-Doppler signatures. In addition, it can improve model sensitivity and accuracy by suppressing feature activations in irrelevant regions. The attention modules are highly flexible and can be integrated

with any existing DCNN architecture without introducing significant computational overhead in model parameters. The attention-enhanced DCNN (AE-DCNN) can be trained similarly to any standard DCNN network.

Given an intermediate feature map, our proposed attention module jointly utilizes the global features computed at the network's last layer to highlight local salient regions of interest at intermediate layers. Since the attention module uses global features to refine intermediate layer features, we termed this as a global spatial attention module (GSAM). The attention-refined features from the intermediate layers are then aggregated with global features to yield the final predictions. In this study, we incorporate GSAM into a standard network – Alexnet, to demonstrate its effectiveness in automatically localizing the object of interest and improving the overall classification performance. We choose to evaluate our implementation on a publically available radar dataset in [12]. The dataset has been acquired using three synchronized RF sensors at three frequencies: 10 , 24, and 77 GHz. It comprises radar micro-Doppler signatures corresponding to eleven human activities of daily life. We test the performance of our attention-enhanced Alexnet (AE-Alexnet) using all three datasets, but one at a time. If the training is done on one sensor dataset, it is tested on the same sensor dataset. The results show that AE-Alexnet consistently improves prediction accuracy across different datasets while performing better than complex state-of-the-art DCNN models, such as Resnet and VGG. We further investigate a simple data augmentation scheme to increase the overall training support by adding additive white Gaussian noise (AWGN) and test the performance of both AE-Alexnet and Modified Alexnet with no attention.

Attention mechanisms have been commonly used in natural language processing tasks, such as image captioning [13] and machine translation [9–11, 14]. In computer vision, it has been applied to a variety of problems naming a few – Image classification, Image segmentation, action recognition, and Image captioning [15–21]. Attention models have also been exploited for medical report generation and medical image classification [22–24]. In the context of radar image analysis, attention models have been exploited for synthetic aperture radar image segmentation and classification problems [25, 26]. More recently, it has been used for the classification of the high-range resolution profile of radar targets [27]. However, this work uses a more complex cascaded network architecture for the desired task. Moreover, only a handful of works use attention mechanisms. The authors in ref. [28] used attention-augmented convolutional auto-encoder for human activity recognition using radar micro-Doppler signatures.

Important contributions of this study are as follows:

1. We use an attention mechanism to improve the classification performance of the existing DCNN architecture Alexnet without significantly increasing the computational overhead. The idea of attention mechanisms is to generate an attention map that assigns weights on the input data to highlight its salient features while suppressing the irrelevant regions, making the prediction more pronounced.
2. We exploit global features at the last layer to learn class-specific local features at intermediate layers obtaining more refined features. We further propose a feature aggregation strategy to improve the overall classification performance of the standard DCNN architecture.
3. We propose one of the first used cases of attention mechanism in a feed-forward convolution neural network (CNN) model applied to a radar micro-Doppler signature classification. The modified network is lightweight and end-to-end trainable.
4. We evaluate the performance of AE-Alexnet using a publicly available radar dataset comprising human micro-Doppler signatures corresponding to eleven different daily living activities at three different frequencies: 10, 24, and 77 GHz. The results demonstrate that the attention mechanism dramatically improves the overall performance across all three frequency datasets.
5. In addition, we demonstrate that the proposed attention mechanism provides refined feature maps that can be visualized using class activation mapping (CAM), helping better interpret predictions made by the network.
6. Furthermore, we use a simple data augmentation scheme to increase the overall training support. The results demonstrate that the augmentation scheme can further improve the performance of both AE-Alexnet and modified Alexnet with no attention.

Our paper is organized as follows. Section 2 introduces the proposed attention module along with the feature aggregation strategy to yield the final predictions. Section 3 describes the experimental dataset used for evaluating the performance of attention-enhanced Alexnet. This section also presents the qualitative and quantitative classification results of attention-enhanced Alexnet and benchmarks its performance against the standard Alexnet. Section 4 discusses the limitations and future possibility of using attention mechanisms for improved performance using other existing networks, such as VGG and Resnet, which have shown promising results on classifying radar micro-Doppler signatures. We finally conclude our study in Section 5.

2 | METHODOLOGY

This section introduces the proposed attention mechanism that can be incorporated into any existing CNN architecture to improve its performance; however, we chose a lightweight Alexnet to be our base architecture in this study.

2.1 | Alexnet

AlexNet contains eight layers; the first five are convolutional layers, the last three are fully connected layers,

followed by a classification layer with a SoftMax activation function [29]. Max-pooling layers follow the first two layers. It uses the rectified linear unit activation function for better training performance than other non-linear activation functions like tanh and sigmoid. The network has already been trained on optical imagery dataset comprising 1000 target classes. However, in recent years, multiple researchers have used Alexnet to classify radar targets' micro-Doppler signatures [1, 5]. It has been possible due to the introduction of the transfer learning approach, where the existing network weights trained on optical images or another domain dataset can be updated or fine-tuned to adapt to radar micro-Doppler signatures using a limited training sample support. However, the literature suggests that the classification performance of radar micro-Doppler signatures using 8-layered Alexnet remains sub-optimal compared to other deep networks, such as VGG16, GoogleNet, and ResNet50 [12].

2.2 | Attention-enhanced Alexnet

This work demonstrates that improved performance can be achieved by integrating attention modules in the standard Alexnet architecture. Furthermore, it does not require multiple additional layers or the training of multiple models. Instead, it progressively suppresses feature responses in irrelevant background regions without the requirement to crop a region of interest and enhances the response by putting more weight on the most crucial spatial structural information in the radar micro-Doppler signatures.

2.2.1 | Global spatial attention module

Figure 1 presents the proposed GSAM. Given the feature maps $F^l \in \mathbb{R}^{H\times W\times C}$ at a chosen intermediate layer $l \in 1, 2, ., L$, GSAM computes a two-dimensional spatial attention mask M_s, where the entries of $M_s \in [0, 1]$, in order to identify salient local information in the feature maps F^l and prune feature responses to suppress the information in the irrelevant regions. It does so by jointly utilizing the feature maps at the last convolutional layer (global features $\hat{G}$) and the feature maps F^l at any intermediate layer l. The deeper layers encode global information from a large spatial context to identify the location of the target objects in the images and model their relationship at a global scale. Therefore, these global features can provide flexibility regarding focussing on a regional basis and disambiguate irrelevant feature content present in intermediate layer features F_l. Here, H, W, and C are the feature maps' height, width, and the number of channels at any layer l, respectively.

In standard CNN architectures, the feature map is gradually down-sampled to capture sufficiently large receptive fields. Therefore, the resulting spatial resolution of each layer might be different. To generate the attention mask M_s, we can either up-sample the global feature maps $\hat{G} \in \mathbb{R}^{\hat{H}\times\hat{W}\times C}$ to match to the intermediate feature maps' F^l spatial resolution $(H \times W)$. Since the spatial resolution of the feature maps might differ from layer to layer, we resort to down-sampling each F^l to match the spatial resolution $(\hat{H} \times \hat{W})$ of global feature maps $\hat{G}$, where $\hat{H}$, $\hat{W}$, and C are the height, width, and number of channels of $\hat{G}$. We used the bicubic interpolation method to down-sample the features by computing the weighted average

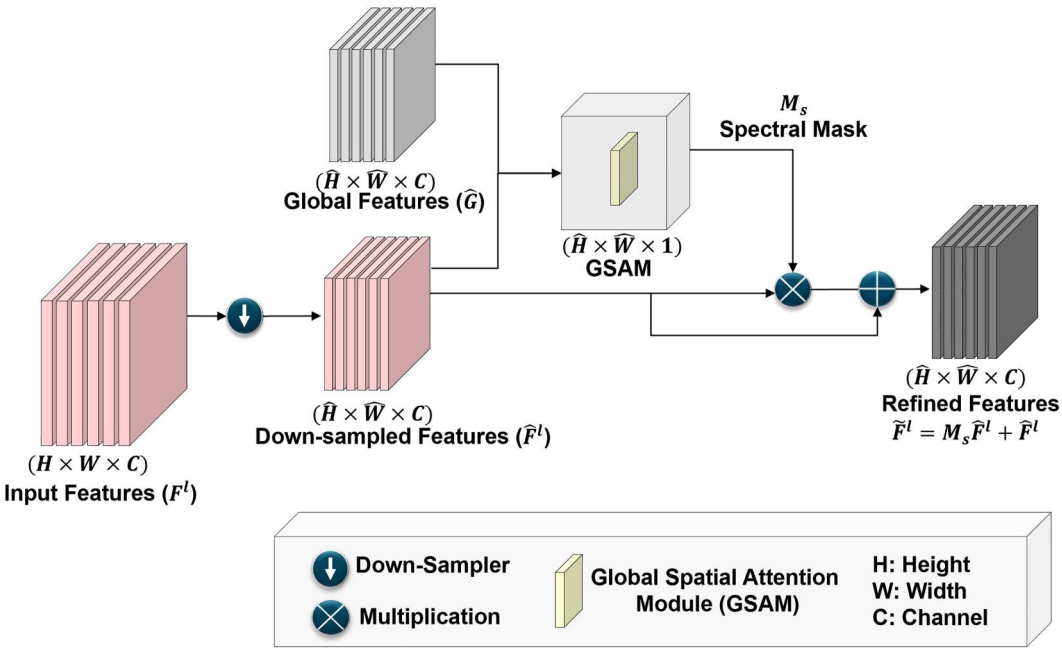

FIGURE 1 Schematic of the proposed attention module. The intermediate feature maps F^l are adaptively refined and scaled with spatial attention mask M_s computed through the proposed attention module at different convolutional blocks of the Alexnet. Spatial regions are selected by jointly analysing the global features $\hat{G}$ and the intermediate layer features F^l. The idea is to attend the features on a regional basis that are most relevant for the human activity recognition task. The spatial grid re-sampling of the input feature maps F^l is performed to obtain feature maps $\hat{F}^l$ of the size equivalent to global features $\hat{G}$.

of pixels in the nearest 4-by-4 neighborhood, which is a type of avg pooling operation. The output of GSAM is $\tilde{F}^l = M_s \hat{F}^l$, where each feature map $\hat{F}^l \in \mathbb{R}^{\hat{\mathbb{H}} \times \hat{\mathbb{W}} \times \mathbb{C}}$ is scaled by the 2D spatial attention map $M_s \in \mathbb{R}^{\hat{\mathbb{H}} \times \hat{\mathbb{W}}}$.

The detailed global attention process is depicted in Figure 2. To compute the spatial attention, we first apply two pooling operations – average and max-pooling along the channel axis of both – the intermediate feature maps $\hat{F}^l$ and the global feature maps $\hat{G}$. The operations result in the generation of four efficient feature maps $\hat{F}^l_{Avg}$, $\hat{G}_{Avg}$, $\hat{F}^l_{Max}$, and $\hat{G}_{Max}$ each of size $\hat{H} \times \hat{W}$. The average-pooled features $\hat{F}^l_{Avg}$ and $\hat{G}_{Avg}$ are added together and passed through the non-linear activation function σ_1 to focus on the informative regions in $\hat{F}^l$ relative to global information. The same process is repeated for the max-pooled features. The average-pooled and max-pooled feature descriptors are finally added together and passed through a convolution layer to generate a global spatial attention map M_s (F^l) encoding the regions to emphasize or suppress. In short, the global spatial attention map can be formulated as

$$M_s\left(F^l\right) = \sigma_2\left(f^{1\times1}\left(\sigma_1\left(\hat{F}^l_{Avg} + \hat{G}_{Avg}\right) + \sigma_1\left(\hat{F}^l_{Max} + \hat{G}_{Max}\right)\right)\right) \quad (1)$$

where $F^l_{Avg} = AvgPool(F^l)$, $\hat{G}_{Avg} = AvgPool(\hat{G})$, $F^l_{Max} = MaxPool(F^l)$, $\hat{G}_{Max} = MaxPool(\hat{G})$, and σ_2 is the normalization function, which can be sigmoid or softmax operation to restrict $M_s \in [0, 1]$. $f^{1\times1}$ represents a convolution operation with the filter of size 1×1. However, we used element-wise sigmoid operation to normalize the spatial mask.

The final refined features $\tilde{F}^l$ are computed by the element-wise multiplication $\otimes$, of the attention mask M_s with the down-sampled intermediate-layer feature maps $\hat{F}^l$ as shown below

$$\tilde{F}^l = M_s \otimes \hat{F}^l \quad (2)$$

During multiplication, the spatial attention map M_s of size $\hat{H} \times \hat{W}$ are copied along the channel dimension of the $\hat{F}^l \in \mathbb{R}^{\hat{\mathbb{H}} \times \hat{\mathbb{W}} \times \mathbb{C}}$, resulting in the overall size of $\hat{H} \times \hat{W} \times C$.

2.2.2 | GSAM-enhanced Alexnet for classification

Figure 3 presents the attention-gated classification model of Alexnet. The proposed attention units are incorporated into the 2nd, 3rd, and 4th layers of the Alexnet to exploit local information present in these intermediate layers. We find that the attention maps are less effective if applied to the first-layer feature maps as the first layer represents very low level features that are not discriminative enough to require attention.

We use activation maps of the 5th layer as our global features. Generally, the global feature maps must encode global spatial contextual information; it is usually obtained from the layer just before the final softmax layer. However, in Alexnet, the layer before the softmax layer is the fully connected layer. In the context of radar micro-Doppler signatures, since most signatures of interest are highly localized, flattening may have the disadvantage of losing important spatial contextual information. Therefore, we consider the 5th activation maps as our global features (right before any flattening is done).

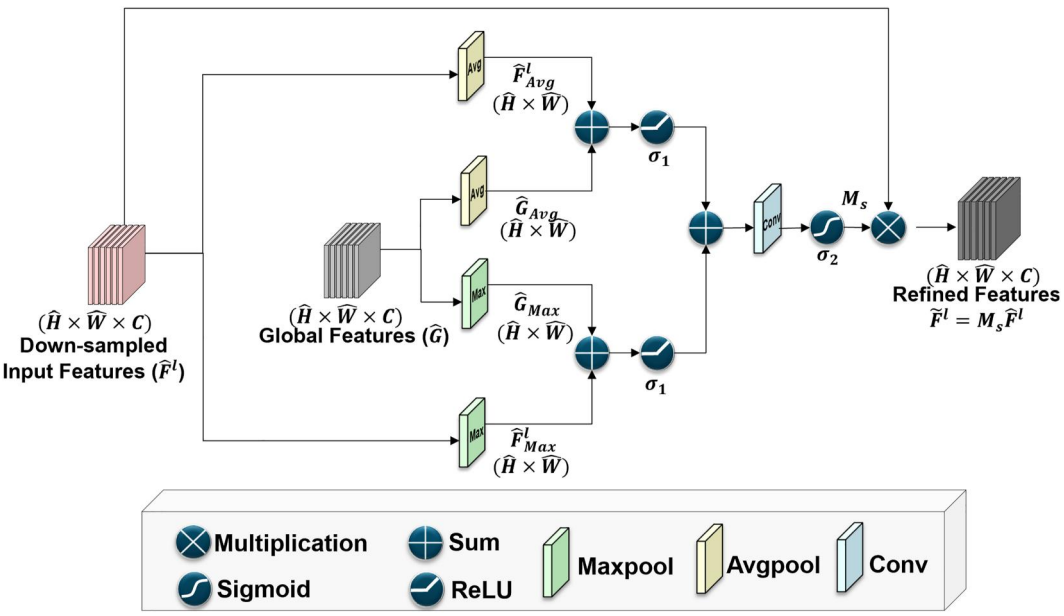

FIGURE 2 The architecture of global spatial attention module (GSAM). As illustrated, GSAM jointly utilizes intermediate feature maps and the global features maps to compute the sum of the respective max-pooled and average-pooled features along the channel axis. The resulting 2D spatial maps are added and forwarded to a convolution layer. The range of the 2D-spatial attention mask M_s is restricted between [0, 1] through an element-wise sigmoid operation.

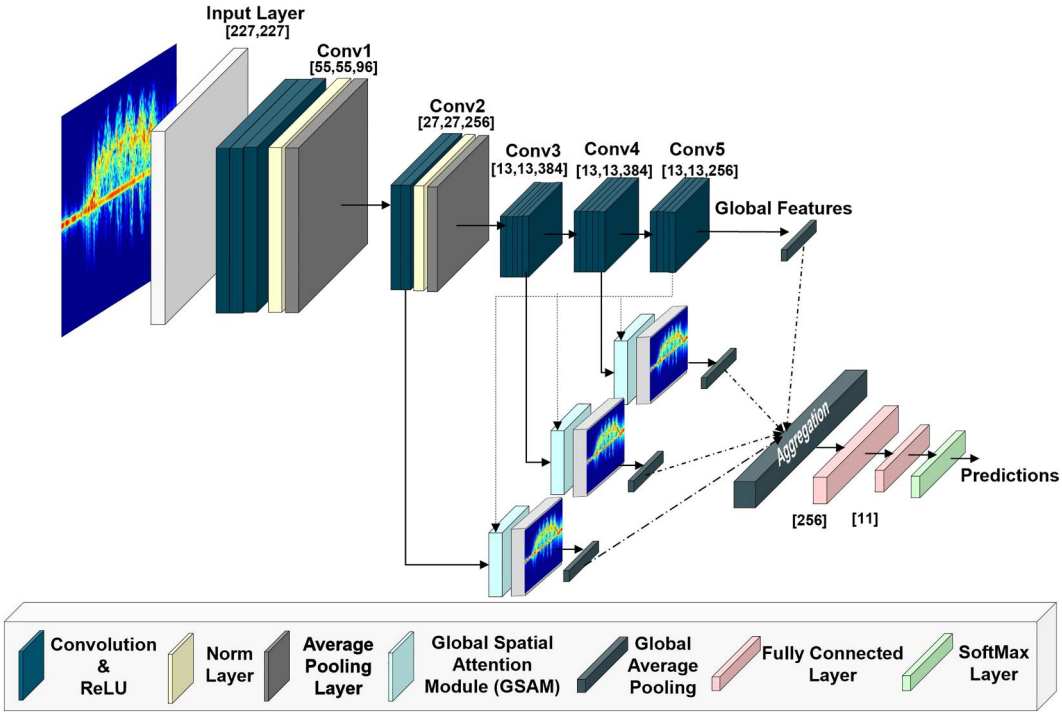

FIGURE 3 The architecture of the proposed AE-Alexnet for human activity recognition

The local feature maps at 2nd, 3rd, and 4th layers are passed through the GSAM along with the global feature maps to obtain the attention-refined feature maps. Then, we aggregate these attention-refined features and the global features together to yield the final predictions. In order to do so, we first compute the global average pooling along the spatial axis, resulting in a vector of length equal to the number of channels in refined feature maps. In addition, we also perform the global average pooling on the global feature maps. Subsequently, the average pooled features are concatenated and passed through two fully connected layers. Finally, a softmax operation is applied to the resulting flattened vector, and the entry with maximum activation is selected as the prediction.

3 | EXPERIMENTAL DATASET DESCRIPTION AND RESULTS

3.1 | Evaluation datasets

This section presents the radar micro-Doppler datasets used in classification experiments. We test the performance of the AE-Alexnet on the publically available radar dataset acquired from three synchronized RF sensors at the following three frequencies: 10 , 24, and 77 GHz [12]. Figure 4 presents the experimental setup used for data acquisition where all the sensors were placed side-by-side at the height of 1-m from the ground with the test subject moving between 0.5 and 3 m in front of the sensors. For the detailed explanation of the data acquisition protocol, we refer the readers to [12]. The dataset consists of micro-Doppler signatures of six participants of different heights,

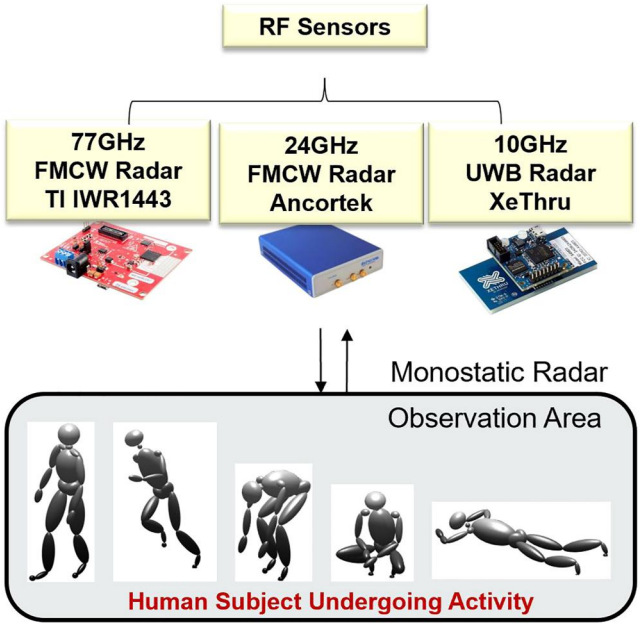

FIGURE 4 Experimental setup comprises three RF sensors configured to operate as monostatic radars. The sensors include Texas Instruments IWR1443 77 GHz FMCW radar, Ancortek's 24 GHz FMCW radar, and XeThru's X4 10 GHz UWB impulse radar.

gender, and ages groups performing eleven different activities of daily living as listed in Table 1. These activities are mainly inspired by intelligent home applications, where monitoring of daily living can help support non-intrusive health monitoring

and enabling healthy living [4, 30, 31]. Each participant repeated these activities ten times, resulting in 60 radar signatures per class per sensor. Note that all the experiments are performed in line-of-sight conditions.

The micro-Doppler signatures from all the three RF sensors for all the eleven activities are shown in Figure 5. The figures demonstrate that the micro-Doppler spectrograms corresponding to high-carrier frequency data show a finer resolution than lower carrier frequency data.

3.2 | Network training parameter settings

This subsection presents the chosen network training parameter settings for the desired classification task. We empirically found the following parameter settings to be the most effective optimization with adaptive moment estimation (ADAM) with an initial learning rate of α of 0.0001, gradient decay factor of $\beta_1 = 0.9$, and squared gradient decay factor of $\beta_2 = 0.99$. The learning rate is updated for every 100 epochs with a batch size fixed to 10. All the attention modules at intermediate layers are randomly initialized. Note that the global attention module does not require multiple separate training stages. It can be trained similar to any standard DCNN, thus simplifying the whole training process.

Since the number of cores, memory size, and speed efficiencies of graphical processing unit (GPU) cards increase with each new generation, these cards are now flexible enough to perform numerical computing tasks like training a neural network as they can process multiple computations simultaneously. Therefore, we perform training of our AE-Alexnet on Matlab 2020b, where all the variables are stored as 64-bit floats, with the following GPU configuration, such as GeForce GTX 1650 Ti, Compute Capability of '7.5' with a multi-processor count of 16.

TABLE 1 Human activity dataset description

Activity number	Activity ID	Activity
1	WLKT	Walking towards the radar
2	WLKA	Walking away from the radar
3	PICK	Pick up an object from the ground
4	BEND	Bending
5	SIT	Sitting on a chair
6	KNEEL	Kneeling
7	CRWL	Crawling towards the radar
8	WTOES	Walking on both toes
9	LIMP	Limping with right leg stiff
10	SHSTEP	Walking with short steps
11	SCSSR	Scissor gait

3.3 | AE-Alexenet classification results and comparison to state-of-the-art Alexnet framework

This section presents 5-fold classification results of radar micro-Doppler signatures captured from three RF sensors. The network is trained and tested on the data captured from the same sensor data. If the network is trained on data from one sensor, it is tested on the same sensor data. The proposed AE-Alexnet model is benchmarked against the standard Alexnet.

We perform 5-fold cross-validation on our dataset, where the entire dataset is split into fivefolds with each fold used as a testing set at some point. The first fold is used to test the model in the first iteration, and the rest are used to train the model. The second fold is used as the testing set in the second iteration, while the rest serves as the training set. This process is repeated until each fold of the 5-folds has been used as the testing set.

The 5-fold classification results corresponding to three sensor datasets at 10, 24, and 77 GHz are presented in Tables 2–4, respectively. The performance difference over the standard Alexnet is presented in the brackets. The highlighted values represent an improvement of over 1% compared to the standard Alexnet. We used the following metrics for the class-wise classification performance evaluation: accuracy, F1-Score, precision, and recall. As we can observe, AE-Alexnet improves the results at all metric levels, and the performance is consistent for all three frequency datasets. It achieves higher

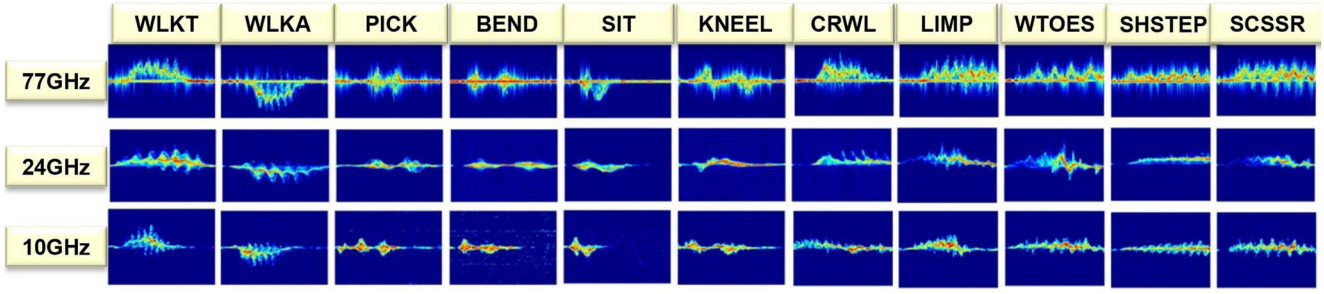

FIGURE 5 Micro-Doppler signatures for each human activity class acquired by three RF sensors operating at three different frequencies: 77, 24, and 10 GHz

TABLE 2 5-fold class-wise classification performance for AE-Alexnet using the micro-Doppler radar signature dataset acquired at 10 GHz. The improvement over the standard Alexnet is presented in the brackets. Values colored red highlight the improvement of more than 1%

Network	Accuracy	F1-score	Precision	Recall
WLKT	0.921 (−2)	0.926 (0.8)	0.931 **(4)**	0.921 (−2)
WLKA	1 **(1.1)**	1 **(1.2)**	1 **(1.4)**	1 **(1)**
PICK	0.964 **(9.4)**	0.963 **(10)**	0.963 **(10.5)**	0.964 **(9.4)**
BEND	0.985 **(10.5)**	0.977 **(11.4)**	0.970 **(12.1)**	0.985 **(10.5)**
SIT	1 **(7.7)**	1 **(5.5)**	1 **(3.2)**	1 **(7.7)**
KNEEL	0.978 **(6.5)**	0.988 **(9.1)**	1 **(11.6)**	0.978 **(6.5)**
CRWL	1 **(5)**	1 **(3)**	1 **(1)**	1 **(5)**
WTOES	0.85 **(7)**	0.838 **(1.7)**	0.827 (−3)	0.85 **(7)**
LIMP	0.859 (0.2)	0.892 (0)	0.929 (−1.8)	0.859 (0.2)
SHSTEP	0.925 **(3.4)**	0.935 **(6.8)**	0.947 **(10.1)**	0.924 **(3.4)**
SCSSR	0.847 (−7)	0.809 (−7)	0.775 (−7)	0.846 (−7)

TABLE 3 5-fold class-wise classification performance for AE-Alexnet using the micro-Doppler radar signature dataset acquired at 24 GHz. The improvement over the standard Alexnet is presented in the brackets. Values colored red highlight the improvement of more than 1%

Activity	Accuracy	F1-score	Precision	Recall
WLKT	0.968 **(2.2)**	0.981 **(6.3)**	0.994 **(10.4)**	0.968 **(2.2)**
WLKA	0.995 (0.6)	0.992 (0.4)	0.989 (0.4)	0.995 (0.5)
PICK	0.922 **(5.2)**	0.927 **(6.3)**	0.931 **(7.4)**	0.922 **(5.2)**
BEND	0.948 **(6.7)**	0.938 **(7.5)**	0.929 **(8)**	0.948 **(6.8)**
SIT	0.964 **(4)**	0.967 **(2.2)**	0.969 (0.2)	0.964 **(4.1)**
KNEEL	0.959 **(4.6)**	0.957 **(5.9)**	0.955 **(7.1)**	0.959 **(4.6)**
CRWL	0.983 **(3.3)**	0.984 **(1.5)**	0.986 (−0.2)	0.983 **(3.3)**
WTOES	0.959 **(17.9)**	0.944 **(12.3)**	0.930 **(6.2)**	0.959 **(17.9)**
LIMP	0.954 **(9.7)**	0.946 **(4.6)**	0.939 (−0.8)	0.954 **(9.7)**
SHSTEP	0.932 **(4.1)**	0.938 **(7.1)**	0.946 **(10)**	0.931 **(4.1)**
SCSSR	0.947 **(2.8)**	0.954 **(7.3)**	0.962 **(11.4)**	0.947 **(2.9)**

precision and reduces the false-positive rate, likely because the attention mechanism suppresses irrelevant background in the radar signatures and forces the network to predict based on class-specific features. Moreover, we see that the precision improves by more than 5% in four target classes for the 10 GHz dataset -PICK, BEND, KNEEL, and SHSTEP. The same performance improvement for the 24 GHz dataset is achieved in 7 target classes, which are significant enough to demonstrate the effectiveness of the attention mechanism. We can also observe similar performance improvement for the 77 GHz radar dataset. As expected, the improvement is greater for 24 and 77 GHz data since the higher frequency data has a higher resolution and more discriminative signatures, making it easy for the network to discern between activities.

TABLE 4 5-fold class-wise classification performance for AE-Alexnet using the micro-Doppler radar signature dataset acquired at 77 GHz. The improvement over the standard Alexnet is presented in the brackets. Values colored red highlight the improvement of more than 1%

Network	Accuracy	F1-score	Precision	Recall
WLKT	1 **(3.5)**	0.951 **(1.8)**	0.906 (0.3)	1 **(3.5)**
WLKA	1 (0)	1 (0.8)	1 **(1.7)**	1 (0)
PICK	0.964 **(5.2)**	0.932 **(2.8)**	0.903 (0.5)	0.964 **(5.3)**
BEND	0.915 (0)	0.937 **(2.3)**	0.962 **(4.8)**	0.915 (0)
SIT	0.982 **(1.8)**	0.982 **(1.8)**	0.982 **(1.8)**	0.982 **(1.8)**
KNEEL	0.983 **(1.5)**	0.991 **(1.6)**	1 **(1.7)**	0.983 **(1.5)**
CRWL	1 **(4.9)**	1 **(2.5)**	1 (0)	1 **(4.9)**
WTOES	0.845 **(5)**	0.876 **(8.8)**	0.910 **(12.8)**	0.845 **(5.1)**
LIMP	0.862 (−5)	0.908 **(2.5)**	0.960 **(10.6)**	0.862 (−5)
SHSTEP	0.983 **(4.9)**	0.966 **(3.3)**	0.950 **(1.7)**	0.983 **(4.9)**
SCSSR	0.898 **(22.1)**	0.881 **(16.1)**	0.865 **(9.6)**	0.898 **(22.1))**

To further rigorously evaluate our attention module, we perform additional classification experiments for 2-Fold, 3-Fold, and 4-Fold partitions and compare its performance with the 5-Fold dataset. We follow the same protocol specified itn the previous section and benchmark its performance with the standard Alexnet. Table 5 summarizes our experimental results. The improvement over the standard Alexnet is presented in the brackets and highlighted in red. We observe that the AE-Alexnet outperforms the baseline, and a consistent improvement in performance can be observed across all the folds and all frequency sensor data. It implies that GSAM boosts the accuracy of baselines significantly for 5-Folds and favorably improves the performance of more challenging 3-Fold and 2-Fold scenarios where limited data is used for training (indicating low-training sample support). The results demonstrate that our proposed approach is powerful, showing the efficacy of a new attention mechanism that generates richer spatial feature descriptors with a quite small overhead in terms of parameters and computation. It motivates us to apply our proposed module GSAM to lightweight networks like Alexnet and demonstrate its great potential for applications on low-end devices.

The above study indicates that the refined features learned through the attention mechanism are very discriminating and focus heavily on class-specific features resulting in improved performance without increasing the overall complexity of the standard architecture. Furthermore, the proposed attention module is highly flexible and can be integrated into existing baseline architectures, like VGG-16, VGG-19, Resnet-18, and Resnet-50, to improve their performances further.

We perform an additional interesting experiment further to support the usefulness of the proposed attention mechanism and investigate the performance of the modified Alexnet structure presented in Figure 6. The network architectures presented in Figure 6 differ from AE-Alexnet presented in Figure 3 in terms of the attention module. The modified

architecture in Figure 6 gives the prediction by fitting a fully connected layer on the concatenated feature vector from the same set of layers. However, this time, the features have not been refined through any attention mechanism. The resulting multi-Fold classification accuracies are presented in Table 6. The results demonstrate that the classification performance of the standard networks can be improved by simply concatenating the features from multiple intermediate layers. However, this is true only for the higher-resolution radar signatures dataset. For datasets acquired at lower frequencies like 10 GHz, the performance remained suboptimal. It indicates that simple concatenation of intermediate-layer features for low-resolution radar signatures cannot discriminate between different target classes since most share feature space, especially at lower carrier frequencies.

3.4 | Network visualizations with class activation mapping

We use CAM for the qualitative analysis and determine which part of the input signatures is responsible for network predictions. We compare the CAM visualization results of attention-refined features with standard Alexnet features. The resulting visualization maps are presented in Fig.7–9 corresponding to target class WLKT at three different frequencies,

TABLE 5 AE-Alexnet classification results. The improvement over the standard Alexnet is presented in the brackets and highlighted in red

Training	Testing	Baseline accuracy Alexnet 5-fold accuracy	Modified Alexnet with attention 5-fold accuracy	Modified Alexnet with attention 4-fold accuracy	Modified Alexnet with attention 3-fold accuracy	Modified Alexnet with attention 2-fold accuracy
77 GHz	77 GHz	90.83	94.84 **(4.01)**	94.03 **(3.2)**	94.23 **(3.4)**	92.22 **(1.39)**
24 GHz	24 GHz	90.15	95.74 **(4.91)**	94.85 **(4.02)**	94.89 **(4.06)**	92.83 **(2)**
10 GHz	10 GHz	90.93	93.89 **(3.06)**	92.65 **(1.82)**	91.75 **(0.92)**	90.71 **(−0.12)**

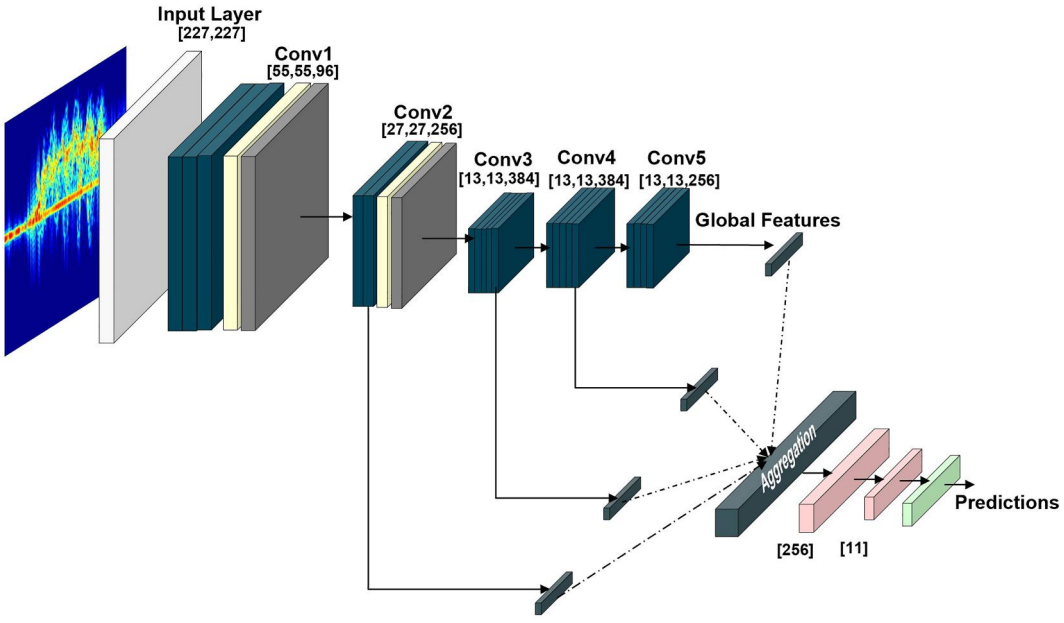

FIGURE 6 The architecture of the modified Alexnet for human activity recognition. The feature vectors of each layer are directly concatenated and passed through fully connected layers to yield final predictions without using the attention mechanism.

TABLE 6 Modified Alexnet (w/o attention) classification results. The improvement over the standard Alexnet is presented in the brackets. Values colored red highlight the improvement over the standard Alexnet

Training	Testing	Baseline accuracy Alexnet 5-fold accuracy	Modified Alexnet w/o attention 5-fold accuracy	Modified Alexnet w/o attention 4-fold accuracy	Modified Alexnet w/o attention 3-fold accuracy	Modified Alexnet w/o attention 2-fold accuracy
77 GHz	77 GHz	90.83	91.83 **(1)**	91.60 **(0.77)**	91.07 **(0.24)**	90.94 **(0.1)**
24 GHz	24 GHz	90.15	94.61 **(3.78)**	93.82 **(2.99)**	93.89 **(3.06)**	91.68 **(0.85)**
10 GHz	10 GHz	90.93	90.77 **(−0.06)**	89.93 **(−0.9)**	87.23 **(−3.6)**	86.3 **(−4.53)**

10, 24, and 77 GHz, respectively. The first column presents the ground truth signatures, the second presents the CAM visualization of raw features, and finally, the third column presents the CAM visualization of attention-refined features obtained at the following intermediate convolutional layers: 2nd, 3rd, and 4th.

In Figures 7–9, we can see that the CAM masks of the AE-Alexnet cover the class-specific activity regions better than unrefined features. The network bases its classification on the entire signature, but the most decisive input comes from the red areas. It shows that GSAM can effectively calculate the importance of the spatial locations in convolutional layers and detects the region that highly agrees with the object of interest. Although the attention map outlines the discriminant region, it does not coincide with the entire activity region.

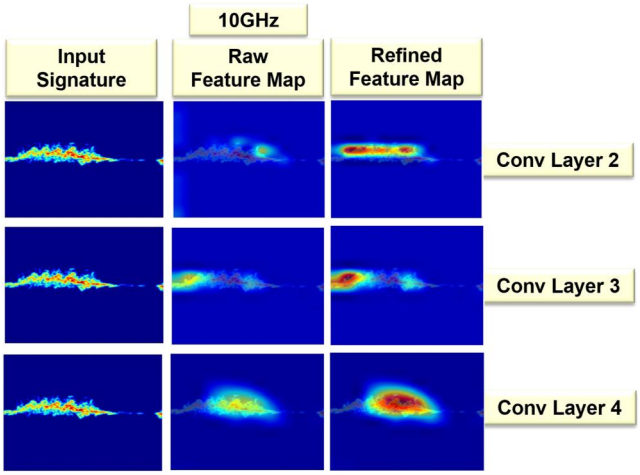

FIGURE 7 Class activation mapping (CAM) visualization results from AE-Alexnet for radar signature classification at 10 GHz. Red regions contribute the most. The detected region highly agrees with the object of interest.

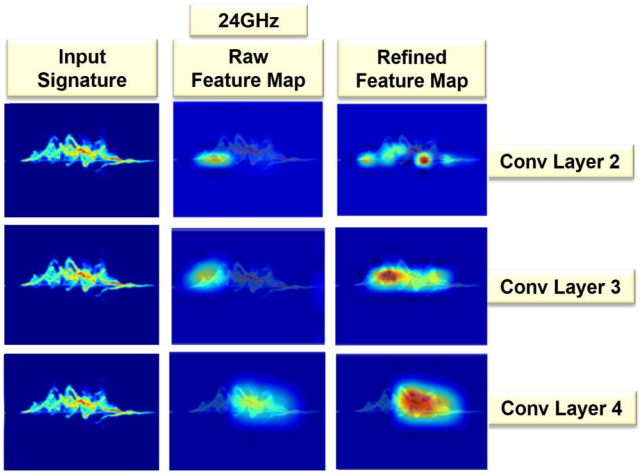

FIGURE 8 Class activation mapping (CAM) visualization results from AE-Alexnet for radar signature classification at 24 GHz. Red regions contribute the most. The detected region highly agrees with the object of interest.

3.5 | Performance bench-marking

In this section, we benchmark the performance of our AE-Alexnet with two baseline attention mechanisms popularly known as the SAM and convolutional block attention module (CBAM) [19]. Like the GSAM proposed in this work, SAM and CBAM can be integrated into any neural network to enhance classification performance. However, GSAM differs from SAM and CBAM in terms of the feature refinement method. GSAM refines the spatial features based on the global layer features, while SAM and CBAM refine the intermediate features locally without considering global layer features. Spatial attention module utilizes the interspatial relationship of features to generate spatial attention and focusses on where is an informative part. On the other hand, CBAM exploits both spatial and channel-wise attention. A channel attention map is generated by exploiting the interchannel relationship of features.

In order to make a fair comparison, we introduce SAM and CBAM at every convolutional block of the Alexnet architecture. Table 7 presents the resulting classification performance for data corresponding to three frequencies. As we can see, SAM and CBAM-enhanced Alexnet perform better than the standard Alexnet due to the feature refinement at each intermediate layer. However, the performance improvement is less significant than the proposed GSAM-enhanced Alexnet. This could be possibly because GSAM uses the global layer to enhance the features at the intermediate layer whereas both the SAM and CBAM use online local feature refinement, which could not possibly highlight the desired regions.

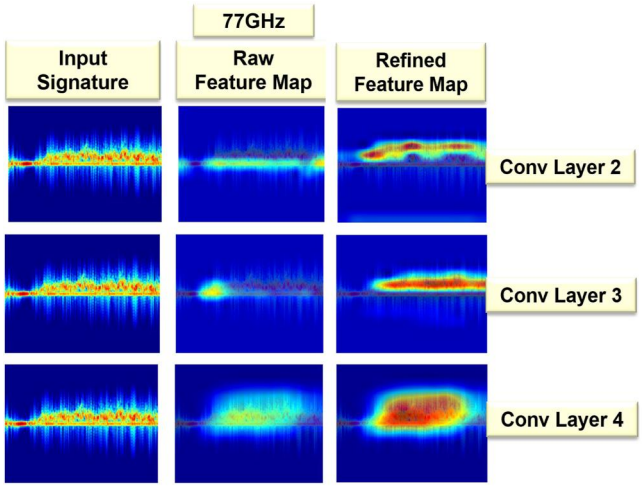

FIGURE 9 Class activation mapping (CAM) visualization results from AE-Alexnet for radar signatures classification at 77 GHz. Red regions contribute the most. The detected region highly agrees with the object of interest.

3.6 | Classification performance with data augmentation

The literature suggests that the addition of slightly modified copies of already existing training data through operations, such as adding noise, image flipping, and time-shifting, can improve the classification performance of any neural network [32, 33]. It does so by reducing the over-fitting during the training of a neural network. Figure 10 presents one such data augmentation scheme, where we artificially add AWGN with zero mean and variance of $\sigma = 0.05$ to our measurement data to increase the overall training support. We test the performance of AE-Alexnet with varying augmentation ratios (AR). Augmentation represents the ratio of noisy measurements to the size of original measurements. If AR = 1, it represents that the augmentation scheme adds noise to the one copy of the original dataset, and the resulting dataset size is twice that of the original dataset size. Similarly, if the AR-5, the noise is added to 5 copies of the original dataset, and the resulting size becomes 6 times.

Table 8 presents the augmentation results as a function of AR for the dataset with AWGN noise and data with no augmentation. The values in brackets are colored red to highlight the improvements over the standard Alexnet. We observe an increase in average classification accuracy by another 1.6% when the original dataset is augmented with a noisy copy of itself. The results suggest that a data augmentation scheme can be used to further improve activity recognition without the burden of collecting significant quantities of experimental training data.

Table 9 presents the class-wise augmentation results for AR = 5 and **shows improvement over AE-Alexnet in the brackets**. Values in brackets are colored red to highlight the improvement of more than 0.01 (that is 1%). The classification metrics clearly show that more than five activities are now more discernible and have improved classification metrics compared to the no augmentation case. This improvement could be attributed to the increased measurement training support size, giving the neural network enough data to be trained well and extract more features from this diverse dataset.

We also investigate the data augmentation performance for modified Alexnet with no attention and report our augmentation scheme results in Table 10 for AR = 1, 3, and 5. The improvement over the standard Alexnet is highlighted in red. We note that the average classification accuracy increases with overall training support. The performance improvement is the highest for AR = 5, suggesting that a simple data augmentation scheme, such as adding AWGN noise, can improve any neural network's classification performance.

TABLE 7 5-fold classification performance comparison with baseline methods: CBAM, Spatial attention module (SAM)

Training	Testing	Baseline accuracy Alexnet 5-fold accuracy	Proposed AE-Alexnet 5-fold accuracy	Modified Alexnet w/o attention 5-fold accuracy	Baseline SAM 5-fold accuracy	Baseline CBAM 5-fold accuracy
77 GHz	77 GHz	90.83	94.84 **(4.01)**	93.89 (3.06)	91.72 (0.89)	91.51 (0.68)
24 GHz	24 GHz	90.15	95.74 **(4.91)**	94.61 (3.78)	91.85 (1.73)	90.88 (0.73)
10 GHz	10 GHz	90.93	93.89 **(3.06)**	90.77 (−0.06)	91.43 (0.5)	91.25 (0.32)

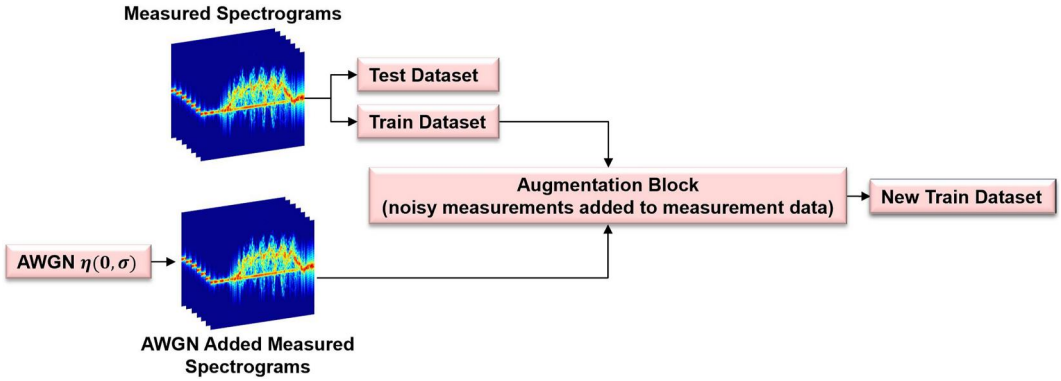

FIGURE 10 Augmentation: The ratio of additive white Gaussian noise (AWGN) added measurement data augmented with original training measurement data is varied to study the impact of data augmentation on the classification performance.

TABLE 8 Data augmentation classification performance for AE-Alexnet using the dataset acquired at 77 GHz. Values colored red highlight the improvements over the standard Alexnet

Training	Testing	Baseline accuracy Alexnet 5-fold accuracy	Modified Alexnet with attention 5-fold accuracy	AR = 1 modified Alexnet with attention 5-fold accuracy	AR = 3 modified Alexnet with attention 5-fold accuracy	AR = 5 modified Alexnet with attention 5-fold accuracy
77 GHz	77 GHz	90.83	94.84 **(4.01)**	95.31 **(4.48)**	95.52 **(4.69)**	96.44 **(5.61)**

4 | DISCUSSION AND FUTURE DIRECTIONS

This work proposes an attention mechanism that uses global features to effectively exploit the regions of interest and suppress the background noise and the region of no-interest in the intermediate layer features. In particular, we investigate several aspects, including-spatial attention mechanism and feature aggregation strategy with and without attention mechanism and a simple data augmentation scheme. However, many open-research problems still need to be addressed.

1. The attention refined feature vectors are aggregated with the global feature vector to yield the final predictions. However, to do so, a global average pooling operation is performed along the spatial axis of the AE feature maps, resulting in a vector of length equal to the number of channels at each intermediate layer. Future work will investigate the aggregation of average pooled features along the channel axis, resulting in a feature matrix of size equal to the spatial dimension of the input maps. This 2D matrix can be flattened to generate a feature vector and aggregate these vectors for investigating better network predictions.
2. In this work, we consider Alexnet to be our base architecture for investigating the performance of the attention mechanism. However, the literature suggests that the VGG-16 network seems to perform well over radar micro-Doppler signatures [12]. Therefore, we believe introducing an attention mechanism into VGG-16 or more complex networks like ResNet could significantly improve the classification performance. Moreover, we hope to investigate the most practical combination of intermediate layers to give the best performance in the case of deeper networks, such as VGG-16, VGG-19, and Resnet, is another open-research problem.
3. The dataset used to investigate the network performance comprises some poor quality radar signatures that could have significantly influenced the network's performance. Our previous work presented an open-source simulation tool for the passive WiFi radar scenario, called SimHumalator [34] where we leveraged human micro-Doppler data generated using SimHumalator to augment our measured data [3]. The classification results presented in [4, 34, 35] suggest that the data augmentation or replacement of poor quality data with synthetic signatures could further greatly enhance the performance. Therefore, our future investigations will drop/replace the poor quality signatures and use data augmentation schemes instead to increase the overall training support.
4. In this work, we test a simple data augmentation scheme and find that adding AWGN noise data to the training support further improves the performance of AE-Alexnet by more than 1.5% for the 77 GHz dataset. In our previous works [3, 4, 35], we investigated the gains from data augmentation using a training dataset generated using our simulator, SimHumalator, with two different types of noise added to the SimHumalator data. These were based on generative adversarial networks and style transfer networks to mimic more complex real-world scenarios and were derived directly from the measurement data. The results highlight that the classification performance can be significantly improved for cases in which only limited experimental data is available for training. However, the detailed experiments combining the two studies-attention mechanism and different data augmentation schemes are still under investigation and will form part of future research.

TABLE 9 Class-wise data augmentation classification performance for AE-Alexnet using the dataset acquired at 77 GHz. The improvement over AE-Alexnet is presented in the brackets. Values colored red highlight the improvement of more than 1%

Network	Accuracy	F1-score	Precision	Recall
WLKT	0.965 (−3)	0.965 **(1.4)**	0.966 **(6)**	0.965 (−3)
WLKA	1 (0)	1 (0)	1 (0)	1 (0)
PICK	0.965 (0.1)	0.974 **(4.1)**	0.983 **(7.9)**	0.965 (0.1)
BEND	0.95 **(3.4)**	0.949 **(1.1)**	0.949 (−1)	0.95 **(3.4)**
SIT	1 **(1.8)**	0.991 (0.9)	0.983 (0)	1 **(1.8)**
KNEEL	0.984 (0.1)	0.983 (−0)	0.983 (−1)	0.984 (0.1)
CRWL	1 (0)	1 (0)	1 (0)	1 (0)
WTOES	0.931 **(8.6)**	0.907 **(3)**	0.883 (−2)	0.931 **(8.6)**
LIMP	0.896 **(3.4)**	0.937 **(2.8)**	0.981 **(2.1)**	0.896 **(3.4)**
SHSTEP	1 **(1.7)**	0.983 **(1.7)**	0.967 **(1.7)**	1 **(1.7)**
SCSSR	0.915 **(1.7)**	0.915 **(3.3)**	0.915 **(5)**	0.915 **(1.7)**

5 | CONCLUSION

This work presents a SAM that can be easily integrated into the existing classification architecture. We apply the proposed attention module to a lightweight network – Alexnet and demonstrate its great potential for applications on low-end devices. The module jointly exploits the global information to highlight the local regions of interest in the intermediate

TABLE 10 5-fold class-wise data augmentation classification performance for Modified-Alexnet (without attention) using the dataset acquired at 77 GHz. Values colored red highlight the improvements over the standard Alexnet

Training	Testing	Baseline accuracy Alexnet 5-fold accuracy	Modified Alexnet w/o attention 5-fold accuracy	AR = 1 modified Alexnet w/o attention 5-fold accuracy	AR = 3 modified Alexnet w/o attention 5-fold accuracy	AR = 5 modified Alexnet w/o attention 5-fold accuracy
77 GHz	77 GHz	90.83	91.83 **(1)**	93.86 **(3.03)**	94.08 **(3.25)**	94.55 **(3.72)**

network layers, thus significantly improving the classification performance compared to the standard network. The publically available dataset from three different RF sensors is used for investigating the classification performance. The experimental results demonstrate that the proposed attention module can benefit radar micro-Doppler signature classification. It generates richer spatial feature descriptors with a pretty small overhead in terms of parameters and computation, which significantly improve the overall performance, especially the precision.

The CAM visualization results show that the detected region that contributes the most to the predicted class highly agrees with the regions of interest. The proposed attention module could support explainable deep learning, a vital research area for automatic radar signature classification.

Furthermore, this study presents a simple data augmentation scheme where AWGN is added to the original dataset to increase the overall training support. The results showed that the AE-Alexnet, when trained on augmented data, outperformed the AE-Alexnet trained on original training data by 1.5%. Similar performance improvement is observed for modified Alexnet with no attention.

AUTHOR CONTRIBUTIONS

Shelly Vishwakarma contributed towards conceptualization, data curation, formal analysis, investigation, developing methodology, validation, visualization, and writing of the original draft. Wenda Li contributed towards data curation, investigation and in the writing – review and editing. Chong Tang contributed towards developing the algorithms and its validation, visualization, and writing – review and editing. Karl Woodbridge provided supervision on this work alongside reviewing the draft. Ravi Raj Adve supervised the work along with data investigation and validation. He also reviewed the draft of the paper. Kevin Chetty contributed towards acquisition of funding for the project and its administration. He also provided supervision on this work alongside reviewing the draft.

ACKNOWLEDGEMENTS

This work was funded under the OPERA Project, the UK Engineering and Physical Sciences Research Council (EPSRC), Grant EP/R018677/1.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

Shelly Vishwakarma https://orcid.org/0000-0003-1035-3259
Wenda Li https://orcid.org/0000-0001-6617-9136

REFERENCES

1. Park, J., et al.: Micro-Doppler based classification of human aquatic activities via transfer learning of convolutional neural networks. Sensors 16(12), 1990 (2016). https://doi.org/10.3390/s16121990
2. Amin, M.: Radar for Indoor Monitoring: Detection, Classification, and Assessment. CRC Press (2017)
3. Tang, C., et al.: Augmenting experimental data with simulations to improve activity classification in healthcare monitoring. In: In: Accepted for IEEE Radar Conference (RadarConf21). IEEE (2021)
4. Vishwakarma, S., et al.: Gan based noise generation to aid activity recognition when augmenting measured wifi radar data with simulations. In: IEEE International Conference on Communications (2021)
5. Seyfioğlu, M.S., Gürbüz, S.Z.: Deep neural network initialization methods for micro-Doppler classification with low training sample support. Geosci. Rem. Sens. Lett. IEEE 14(12), 2462–2466 (2017). https://doi.org/10.1109/lgrs.2017.2771405
6. Seyfioglu, M.S., et al.: Diversified radar micro-Doppler simulations as training data for deep residual neural networks. In: 2018 IEEE Radar Conference (radarConf18), pp. 0612–0617. IEEE (2018)
7. Chen, Z., et al.: Personnel recognition and gait classification based on multistatic micro-Doppler signatures using deep convolutional neural networks. Geosci. Rem. Sens. Lett. IEEE 15(5), 669–673 (2018). https://doi.org/10.1109/lgrs.2018.2806940
8. Gurbuz, S.Z., Amin, M.G.: Radar-based human-motion recognition with deep learning: promising applications for indoor monitoring. IEEE Signal Process. Mag. 36(4), 16–28 (2019). https://doi.org/10.1109/msp.2018.2890128
9. Vaswani, A., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems, 5998–6008 (2017)
10. Bahdanau, D., Cho, K., Bengio, Y.: Neural Machine Translation by Jointly Learning to Align and Translate (2014). arXiv preprint arXiv:14090473
11. Luong, M.T., Pham, H., Manning, C.D.: Effective Approaches to Attention-Based Neural Machine Translation (2015). arXiv preprint arXiv:150804025
12. Gurbuz, S.Z., et al.: Cross-frequency training with adversarial learning for radar micro-Doppler signature classification (rising researcher). In: Radar Sensor Technology XXIV, vol. 11408. International Society for Optics and Photonics 114080A, (2020)
13. Anderson, P., et al.: Bottom-up and top-down attention for image captioning and visual question answering. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 6077–6086 (2018)
14. Shen, T., et al.: Disan: directional self-attention network for RNN/CNN-free language understanding. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32 (2018)
15. Ren, M., Zemel, R.: End-to-end instance segmentation and counting with recurrent attention. arxiv 2016', arXiv preprint arXiv:16050 9410, **2**
16. Zhao, B., et al.: A survey on deep learning-based fine-grained object classification and semantic segmentation. Int. J. Autom. Comput. 14(2), 119–135 (2017). https://doi.org/10.1007/s11633-017-1053-3
17. Wang, F., et al.: Residual attention network for image classification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3156–3164 (2017)
18. Lu, J., et al.: Knowing when to look: adaptive attention via a visual sentinel for image captioning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 375–383 (2017)
19. Woo, S., et al.: CBAM: convolutional block attention module. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 3–19 (2018)
20. Jetley, S., et al.: Learn to Pay Attention (2018). arXiv preprint arXiv: 180402391
21. Liu, J., et al.: Global context-aware attention LSTM networks for 3d action recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1647–1656 (2017)
22. Zhang, Z., et al.: Tandemnet: distilling knowledge from medical images using diagnostic reports as optional semantic references. In: International

Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 320–328. Springer (2017)
23. Guan, Q., et al.: Diagnose like a Radiologist: Attention Guided Convolutional Neural Network for Thorax Disease Classification (2018). arXiv preprint arXiv:180109927
24. Schlemper, J., et al.: Attention gated networks: learning to leverage salient regions in medical images. Med. Image Anal. 53, 197–207 (2019). https://doi.org/10.1016/j.media.2019.01.012
25. Gao, F., et al.: A semi-supervised synthetic aperture radar (SAR) image recognition algorithm based on an attention mechanism and bias-variance decomposition. IEEE Access 7, 108617–108632 (2019). https://doi.org/10.1109/access.2019.2933459
26. Shi, B., et al.: Synthetic aperture radar SAR image target recognition algorithm based on attention mechanism. IEEE Access 9, 140512–140524 (2021). https://doi.org/10.1109/access.2021.3118034
27. Pan, M., et al.: Radar HRRP target recognition model based on a stacked CNN-BI-RNN with attention mechanism. IEEE Trans. Geosci. Rem. Sens. 60, 1–14 (2021). https://doi.org/10.1109/tgrs.2021.3055061
28. Campbell, C., Ahmad, F.: Attention-augmented convolutional autoencoder for radar-based human activity recognition. In: 2020 IEEE International Radar Conference (RADAR), pp. 990–995. IEEE (2020)
29. Alom, M.Z., et al.: The History Began from Alexnet: A Comprehensive Survey on Deep Learning Approaches (2018). arXiv preprint arXiv: 180301164
30. Li, W., Tan, B., Piechocki, R.J.: Wifi-based passive sensing system for human presence and activity event classification. IET Wirel. Sens. Syst. 8(6), 276–283 (2018). https://doi.org/10.1049/iet-wss.2018.5113
31. Shah, S.A., Fioranelli, F.: RF sensing technologies for assisted daily living in healthcare: a comprehensive review. IEEE Aero. Electron. Syst. Mag. 34(11), 26–44 (2019). https://doi.org/10.1109/maes.2019.2933971
32. Al Hadhrami, E., et al.: Transfer learning with convolutional neural networks for moving target classification with micro-Doppler radar spectrograms. In: 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD), pp. 148–154. IEEE (2018)
33. Gurbuz, S.Z., Amin, M.G.: Radar-based human-motion recognition with deep learning: promising applications for indoor monitoring. In: Signal Processing Magazine. IEEE (2019)
34. Vishwakarma, S., et al.: Simhumalator: an open source end-to-end radar simulator for human activity recognition. IEEE Aero. Electron. Syst. Mag. 37(3), 6–22 (2021). https://doi.org/10.1109/maes.2021.3138948
35. Vishwakarma, S., et al.: Neural Style Transfer Enhanced Training Support for Human Activity Recognition (2021). arXiv preprint arXiv: 210712821

How to cite this article: Vishwakarma, S., et al.: Attention enhanced Alexnet for improved radar micro-Doppler signature classification. IET Radar Sonar Navig. 1–13 (2022). https://doi.org/10.1049/rsn2.12369