DIAT-RadHARNet: A Lightweight DCNN for Radar Based Classification of Human Suspicious Activities

Mainak Chakraborty, *Graduate Student Member, IEEE*, Harish C. Kumawat, *Graduate Student Member, IEEE*, Sunita Vikrant Dhavale, *Member, IEEE*, and Arockia Bazil Raj A., *Member, IEEE*

***Abstract*— Recognizing suspicious human activities is one of the critical requirements for national security considerations. Nowadays, designing the deep convolution neural network (DCNN) models suitable for micro-Doppler (m-D) signature-based human activity classification is rapidly growing. However, high computation cost and a huge number of parameters limit their direct/effective usability in field applications. This article introduces an m-D signatures' dataset "DIAT-μRadHAR" covering army crawling, boxing, jumping while holding a gun, army jogging, army marching, and stone-pelting/grenade-throwing, generated using an X-band continuous wave (CW) radar. This article also introduces a lightweight DCNN model, "DIAT-RadHARNet," designed for those human suspicious activity classification. To reduce the computation cost and to improve the generalization ability, DIAT-RadHARNet is designed with four design principles: depthwise separable convolutions, channel weighting (CHW) based on the importance, different size filters in the depthwise part, and operating different size kernels on the same input tensor. The network has 213 793 parameters with a total of 55 layers. Our extensive experimental analysis demonstrates that the DIAT-RadHARNet model efficiently classifies the activities with 99.22% accuracy, giving minimal false positive and false negative outcomes. The time complexity of the proposed DCNN model observed during the testing phase is 0.35 s. The same accuracy and time complexity are obtained even at adverse weather conditions, low-lighting environments, and long-range operations.**

***Index Terms*— Convolutional neural network, deep convolution neural network (DCNN)-based classification, human suspicious activity, micro-Doppler (m-Doppler) signatures, X-band continuous wave (CW) radar.**

I. Introduction

HUMAN activity recognition (HAR) has drawn significant attention in numerous fields such as security monitoring, counterterrorism, border surveillance, biomedical patient health monitoring, and early detection of public violent protests/attacks [1], [2]. The need of intelligent systems having the capability to detect and classify suspicious human activities becomes essential for triggering the necessary counter-action mechanisms, to control the situation, and/or for post situation/scenario analysis [3]. Nowadays, several HAR systems are available based on vision sensing [smart, closed-circuit television (CCTV), surveillance video, infrared, thermal, and acoustic sensors] and RF sensing (radars). More common vision-based approaches for HAR have relied on visual data, and hence have several limitations such as sensitivity to weather conditions, short-range operations [4], narrow field of view (for good accuracy), range focus/collimation tuning issues, accuracy issues of finderscope/red-dot-finder-based initial boresight alignment, and dependence on surrounding lighting conditions [5].

On the other hand, HAR using a radar system offers unique advantages, such as robustness for even adverse weather and/or low-lighting conditions, through-wall capability, foliage penetration/target-imaging, tunnel intruders' detection, and long-range operations [6], [7]. The more common human suspicious activities from the security point of view are person fight punching (boxing) during one-to-one attack, person intruding for preattack surveillance (army marching), person training (army jogging), person shooting (or escaping) with a rifle (jumping with holding a gun), stone/hand-grenade-throwing for damage/blasting (stone-pelting/grenade-throwing), and person hidden translation for attack execution or escape (army crawling) [8], [9]. Detecting/classifying these kinds of suspicious human activities correctly, better at the earliest, by processing their micro-Doppler (m-D) signatures is very much essential [10].

In the recent decades, researchers have used machine learning (ML) techniques that require separate time–frequency (T-F) map features' retrieve/classification algorithms [11], such as principal component analysis (PCA), multilayer perceptron, linear predictive coding (LPC), and support vector machine (SVM), for HAR [12], [13]. However, the efficiency of extracted features using such traditional techniques for classification is limited by prior information and intricacy of categorization problems [14], [15]. Also, the classification performance of this kind of approach degrades if the m-D signature similarity between classes is high and/or the number of classes is more [16]. Therefore, nowadays, deep learning (DL) gains the researchers' attention since it does concurrent automatic features' extraction, with improved self-actuation,

Manuscript received November 21, 2021; revised February 10, 2022; accepted February 14, 2022. Date of publication February 25, 2022; date of current version March 15, 2022. This work was supported by the Life Sciences Research Board (LSRB) in association with the Defence Institute of Psychological Research (DIPR) sanction letter no. LSRB/01/15001/M/LSRB-381/PEE&BS/2020, dated 15.03.2021. The Associate Editor coordinating the review process was Yuan Gao. *(Corresponding authors: Sunita Vikrant Dhavale; Arockia Bazil Raj A.)*

Mainak Chakraborty and Sunita Vikrant Dhavale are with the Department of Computer Science and Engineering, Defence Institute of Advanced Technology (DIAT), Pune 411025, India (e-mail: mainak.mail@gmail.com; sunitadhavale@gmail.com).

Harish C. Kumawat and Arockia Bazil Raj A. are with the Department of Electronics Engineering, Defence Institute of Advanced Technology (DIAT), Pune 411025, India (e-mail: harish.kumawat1@gmail.com; brazilraj.a@diat.ac.in).

Digital Object Identifier 10.1109/TIM.2022.3154832

self-regulation, and self-control due to its intrinsic computation and decision-making skills [17], [18], which motivated us to use the DL technique for the purpose of human suspicious activity classification. The main contributions reported in this article are as follows.

1) Indigenous development of an X-band continuous-wave (CW) 10-GHz radar, and its uses to prepare the m-D signatures' dataset, named as "DIAT-μRadHAR," corresponding to suspicious human activities: army marching, jumping with holding a gun, army jogging, army crawling, grenade-throwing/stone-pelting, and boxing.
2) Designing a lightweight deep convolutional neural network (DCNN), named as "DIAT-RadHARNet," using a diversified dataset ("DIAT-μRadHAR"), validating its classification/performance accuracy and comparing the computational complexities of the designed DCNN model with the state-of-the-art CNN models.
3) Following four design principles: depthwise separable convolutions, channel weighting (CHW) based on the importance, different size filters in the depthwise part, and operating different size kernels on the same input tensor, to reduce the computation cost and to improve the generalization ability of "DIAT-RadHARNet."
4) Computing the ablation study and classification report containing the statistics of layers, parameters, floating point operations (FLOPs), accuracy, precision, recall, F1-score, receiver operating characteristics (ROC), categorical cross-entropy loss, etc., to have thorough performance/classification accuracy analysis of the proposed DCNN model.
5) Analysis of test results of designed DCNN model's classification accuracy by conducting open-field experiments during different environmental conditions: normal/adverse weather conditions, low-lighting environments, and long-range operations.

The rest of this article is organized as follows: a brief review summary of related recent works on HAR is provided in Section II, Section III describes the developed RF sensor and dataset formation methodology, Section IV elucidate the developed "DIAT-RadHARNet" architecture and its implementation details, Section V presents the experimental results with performance analysis metrics, and finally Section VI gives the conclusion.

II. Related Recent Works

Recently, radar-based systems for recognizing different human activities [hand gesture [19], walking, running, jumping, crawling, cycling, sitting still/standing still [20], patient walk (with/without cane/walker), walking with a concealed rifle, walking while carrying bag/suitcase/laptop [21], etc.] are becoming very popular. Zenaldin and Narayanan [8] reported human activity classification based on a CW radar (6.5 GHz; target operation ranges are 0.5–9 m) using an SVM classifier for running, walking, and crawling classes. Chen and Ye [9] developed a 1-D CNN using a nondiversified dataset prepared using a commercial K-band CW radar ($\approx$24 GHz; target ranges are 0.5–5 m inside the laboratory) and used it for the classification of seven different types of human activities. Javier and Kim [13] have introduced a nondiversified m-D signatures' dataset, of size 1008, obtained using an S-band CW radar (2.4 GHz; target ranges are 2–8 m inside the laboratory). The LPC is used to extract the features, followed by an SVM-based categorization. Zhang and Cao [16] suggested a DCNN model trained using m-D samples, of size 1896, prepared using a w-band radar (77 GHz; target range is $\approx$17.7 m) for the recognition of walking, swinging hands, and falling human activities.

Gu *et al.*. [22] designed a 5.3-GHz one-transmitter and two-receiver-based CW radar system to classify hand gestures in the presence of human body movements. Li *et al*. [23] proposed a sign language/hand gesture classification system using an ultrawideband (UWB) radar module. Shrestha *et al.*. [18] prepared a dataset using a frequency-modulated indoor CW radar (5.8 GHz; with 15 human targets) and classified six different human activities using a long short-term memory (LSTM) network. The model was trained and validated with two different 35–45-s duration sequences, respectively. Alnujaim *et al.*. [15] presented a six-layer DCNN architecture designed using a 2-D convolutional and fully linked layers with fixed-size filters for HAR. This kind of design increases the computational cost and is incapable of capturing multilevel features. Furthermore, the developed model is trained using a nondiversified dataset of size 144, generated using a 2.4-GHz CW radar and 12 human targets. Lee *et al.*. [24] presented the human activity classification based on incremental learning of m-D signatures. The authors used an S-band CW radar (2.4 GHz; 12 human targets) and prepared a dataset (of size 1008) containing spectrogram images to classify seven different classes.

Thus, to the best of the authors' knowledge, numerous works are being conducted to classify general human activities, but only fewer works on suspicious human activities. Furthermore, from the literature, it is noted that the majority of the works have some flaws such as: 1) physical parameters (i.e., orientation) of the human target are not taken into account when acquiring the samples; 2) samples are collected in a laboratory environment and with short-range target operations; 3) smaller size datasets are used to train the models; 4) datasets are not available to the access of academicians/researchers; 5) multilevel features are not considered; 6) computationally costly; and 7) not suitable for on-device implementation. Sections III–V address these flaws to a greatest extent. In addition, we wish to find the answer for the following research questions through our extensive DCNN designs and validation experimental results/analysis.

1) Does the use of different size filters in depthwise separable convolutional layer increase the computational cost?
2) Does weighting the channels depending on the feature map's importance help improving the model's generalization capabilities?
3) Is multiscale filtering on the same input tensor useful in collecting information about human motion signals from a spectrogram image of suspected human activity?
4) Does combining CHW with multiscale feature maps increase network performance?

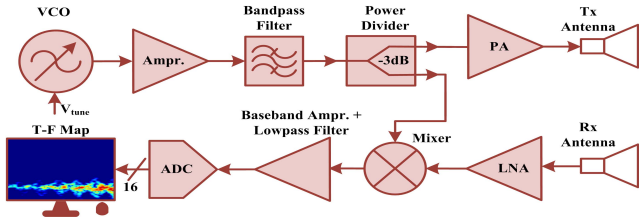


Fig. 1. Top-level schematic of the developed X-band CW radar.

Fig. 2. Photograph of an indigenously designed CW (X-band) radar in use during field tests.

III. Radar Experimental Setup and Dataset Formulation

This section describes the radar design and dataset preparation.

A. Radar Design and Specifications

An X-band CW radar, as shown in Fig. 1, operating at 10 GHz, is developed at our radar system design laboratory and is used to collect the experimental m-D signatures via different open-field experiments (refer Fig. 2). The developed radar consists of an X-band voltage controlled oscillator (VCO), source amplifier, bandpass filter, power divider, power amplifier (PA), and Tx horn antenna at the transmitter chain and Rx horn antenna, low-noise amplifier (LNA), and a mixer at the receiver chain. The main specifications of these subsystems are given in Table I. The radar signal acquisition module is designed with a baseband amplifier, low-pass filter (LPF), and analog-to-digital converter (ADC) [25]. This section receives the signal from the radar Rx chain and sends it, in digital form at a sampling rate of 10 ksps, to the m-D signature extraction algorithm developed in PC (MATLAB environment).

B. DIAT-μRadHAR Dataset Preparation

To have a diversified dataset, 30 different human targets, including males and females of different weights and heights, are directed to perform suspicious activities in the open experimentation field, at different ranges between 10 m and 0.5 km, in front of the radar. To generate the DIAT-μRadHAR dataset as a realistic one, all the suspicious activities are performed at seven different orientations (tilted positions): 0°, ±15°, ±30°, and ±45°, in radar run time. The radar raw data, i.e., the output of ADC (16 bits), are serialized by a MAX9258-based universal serial bus (USB) serializer and passed into the MATLAB environment. Radar time-domain signals, acquired during normal/adverse weather conditions, low-lighting environments, and long-range operations, are recorded for the radar run time of 3 s. Radar data acquired for 3 s are passed through a 600-Hz m-D software LPF, as shown in Fig. 3 which removes the high-frequency components and allows only the m-D spectrum for which our radar is designed, i.e., the maximum radial velocity is 9 m/s [26]. Since in coherent CW radar the background clutters bias the radar signal with a dc voltage [27], the clutter (stationary) effect in the input signal is filtered using a "dsp.DCBlocker" (first stage in clutter filter) MATLAB object. During the experiments, the radar cross section (RCS) (i.e., reflections to the X-band radar's CW signal) of grasses, tree leaves, wind fluctuations, or similar slow/random moving targets is very low; hence, they introduced very weak (very low power) and very low Doppler components (clutter power is concentrated about the zero frequency band). Thus, the Doppler effects of such clutters are removed by an appropriate design of a digital transversal filter (the second stage in clutter filter), as discussed in [27], giving a high degree of attenuation around 0 Hz (i.e., deep stopband around dc) and a high gain throughout the rest of the considered Doppler spectrum, to attain improved signal-to-clutter ratio (SCR). The clutter effect filtered radar signal is applied into short-time Fourier transform (STFT), as given in (1), for generating the T-F map

$$X[m,k] = \sum_{n=-\infty}^{+\infty} x[n]w[m-n]e^{-i2\pi nk/N} \quad (1)$$

TABLE I
Main Specifications of the Developed X-Band CW Radar

S.No.	RF Sub-system	Specification
1	VCO	RF Frequency: X-Band(10 GHz) V_{tune}: 10 V Power: 0 dBm
2	Amplifier	Bandwidth: 6 -18 GHz Gain: 24 dB
3	Bandpass Filter	Bandwidth: 9.75 - 11.25 GHz
4	Power Divider	Bandwidth: 1.8 - 12.6 GHz Isolation: 11.7-19.9 dB
5	Power Amplifier	Bandwidth: 8-12 GHz Gain: 50 dB
6	Mixer	Bandwidth: 3.7 -10.5 GHz
7	LNA	Bandwidth: 0.5 -12 GHz Gain: 17 dB
8	Tx/Rx Antenna	Bandwidth: 8 - 12 GHz Gain: 9 dBi Beamwidth: 15^o
9	Baseband Amplifier	Bandwidth: DC-3 kHz Gain: 50 dB
10	Lowpass filter	Bandwidth: DC-2 kHz
11	ADC	Sampling Frequency: 10 kHz

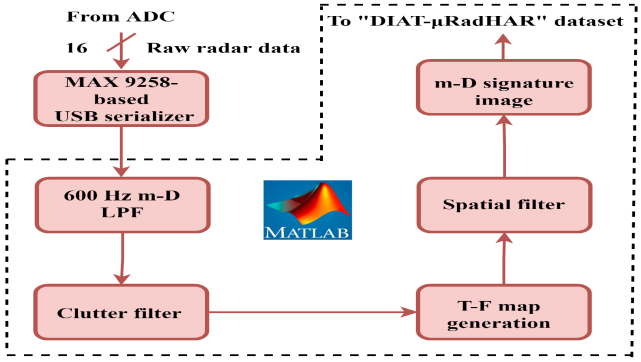


Fig. 3. Preprocessing pipeline on radar raw data to generate m-D signature images.

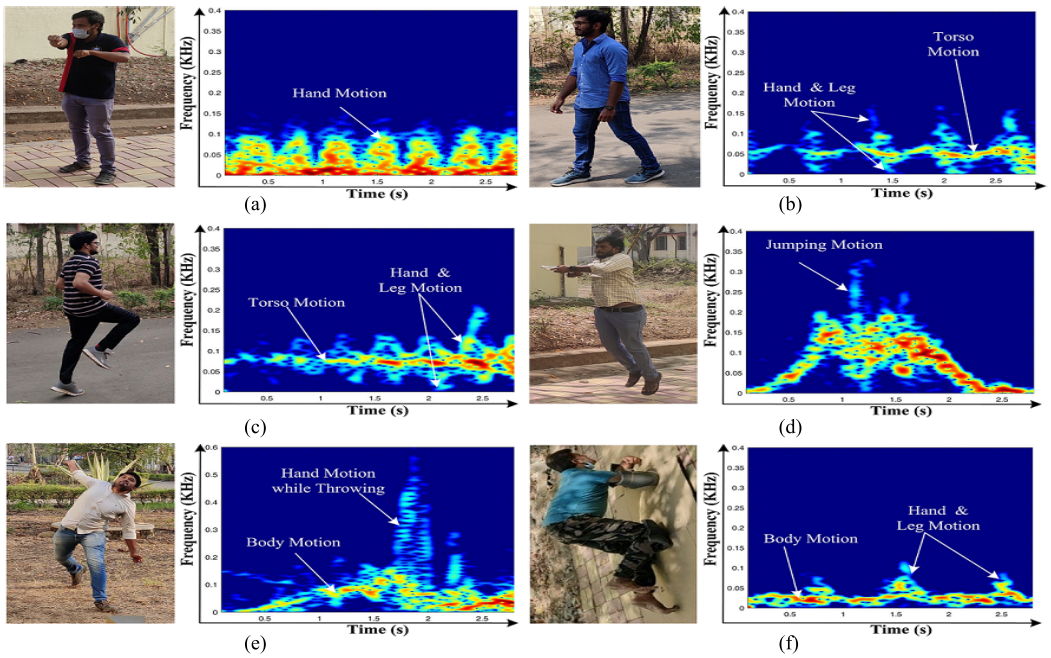


Fig. 4. m-D spectrogram samples of each suspicious human activity. (a) Boxing. (b) Army marching. (c) Army jogging. (d) Jumping with holding a gun. (e) Stone-pelting/grenade-throwing. (f) Army crawling.

TABLE II

HUMAN TARGETS' CLASSWISE PARAMETERS. THE TOTAL NUMBER OF SAMPLES IS 3780

Class Name	Human Target's Physical Parameters			Different Angle of Targets Orientation	Normal weather condition	Low-light environments	Adverse weather condition	Long-range operations	Total No. of Samples
	Height (cm)	Weight (kg)	Age (years)						
Boxing	165-190	51-90	24-35	0°, ± 15°, ± 30°, ± 45°	104	145	230	150	629
Army crawling	170-190	65-87	26-35	0°, ± 15°, ± 30°, ± 45°	150	150	149	152	601
Army jogging	160-189	58-85	24-40	0°, ± 15°, ± 30°, ± 45°	157	139	170	149	615
Army marching	147-183	55-93	24-52	0°, ± 15°, ± 30°, ± 45°	187	145	152	208	692
Stone pelting/ Grenades throwing	155- 190	51-90	25-35	0°, ± 15°, ± 30°, ± 45°	153	169	172	148	642
Jumping with holding a gun	170-185	65-90	25-38	0°, ± 15°, ± 30°, ± 45°	168	145	144	144	601

where $x[n]$ is the discrete signal, k is the index value for frequency, $\omega[n]$ is the window function, n is the sample number, and N is the number of samples in the analyzing window. In our work, the window length is 2048 (Hamming window), with 4096 points of discrete Fourier transform (DFT) and 2000 points of (97.65%) overlapping, and the signal length is 40 000. Noises, sourced from slight movements of the background clutter and radar subsystems, present in the T-F map are further removed using a spatial filter (the third stage, as shown in Fig. 3) designed as described in [28]. The noise canceled and background clutter effects' filtered spectrogram samples are stored in the respective class for designing our novel DCNN architecture. The classwise data collection details can be found in Table II. Fig. 4 shows the event photographs and their sample spectrogram images for each suspicious activity. The total number of spectrogram samples in our dataset is 3780. Out of the collected samples, 80% and 20% are used for training and validation purposes, respectively. The spectrogram images clearly exhibit how the features of different human activities, due to torso, hand, leg, and body movements, are correlated (i.e., interclass correlations) [29].

To facilitate the worldwide researchers working in this field, we decided to publish our complete dataset (as.mat file) and the proposed DCNN network (source code) at our institute website (https://www.diat.ac.in/view-profile/?id = 98) for free access for their academic/research purposes.

IV. PROPOSED NOVEL DCNN ARCHITECTURE

This section explains the design of a proposed novel DIAT-RadHARNet architecture along with its design considerations and its implementation details.

A. Proposed DCNN Architecture

The DIAT-RadHARNet architecture, as shown in Fig. 5, is designed from the inspiration of MobileNetV2 [30], InceptionV3 [31], SENet [32], and BlazeFace [33] architectures by considering the following five important design principles.

1) The depth of the feature map in a convolution operation denotes the number of convolutional kernels used to learn the various feature representations of the input tensor. Some of the feature maps are more significant than others. Thus, the simplest way to improve a DCNN's generalization capabilities is to weight the channels based on the importance of the feature map [32].

2) Depthwise separable convolutions reduce the parameters of the model and extra computing overhead substantially [30].
3) Most of the existing DCNN architectures used 3×3 filters in the depthwise separable convolutional layers. The computational cost on such convolution operation is dominated by pointwise convolution. For example; if the input tensor size is $D_s \times D_s \times M$ and the filter size is $D_k \times D_k$, then the total number of multiplication and addition operation for the depthwise part is ${D_s}^2 M {D_k}^2$, while the subsequent pointwise convolution on N channels require a total of ${D_s}^2 MN$ multiplication and addition operation, which is $(N/{D_k}^2)$ times than depthwise convolution. Thus, using different size $(1 \times 1,\ 3 \times 3,\ 5 \times 5)$ filters in the depthwise part does not increase extra computational overhead [33].
4) The occupied area of m-D signatures of suspicious human activities on the spectrogram images differs based on the type of activity, radar aspect angle, i.e., activities' orientation with respect to radar boresight, physical parameters, and the range at which activities take place, etc. Thus, finding the optimal filter size to learn those signatures is difficult. A larger filter produces a global distribution, while a smaller one produces a local distribution. On the other hand, increasing the depth is prone to overfitting. The solution to this problem is to operate different size kernels on the same input tensor [31].
5) The introduction of global average pooling (GAP) layers in any DCNN model, instated of the traditional fully connected layer, before the final classifier reduces the extra burden of trainable parameters [34].

The proposed DIAT-RadHARNet DCNN model comprises 55 layers as: 13 separable convolutional layers, one max-pooling layer, three multiply layers, six add layers, 13 batch normalization layers, six rectified linear unit (ReLU) activation layers, three dense layers with ReLU activation, three dense layers with sigmoid activation, one dropout layer, four GAP layer, one softmax, and one input layer. To reduce the computational overhead, we used separable convolutional layers throughout our network instead of 2-D DCNN. All the convolutional layers are activated by ReLU activation, and we use the same padding to maintain the same output feature maps' dimension as the input tensor. We included batch normalization layers in our model to improve the generalization error and accelerate the training process. The proposed DIAT-RadHARNet DCNN model can be expressed as follows:

$$\text{DIAT-RadHARNet(x)} = \beta(f^{\text{FC}}(f^{D}(f^{G}(f^{\text{MSFCA}_i}(f^{M}(f^{B}(\sigma \times (f^{C_i}(x * w_i + b_i))))))))) \quad (2)$$

where β, f^{FC}, f^{D}, f^{G}, f^{MSFCA_i}, f^{M}, f^{B}, σ, f^{C_i}, x, w_i, and b_i represent thesoftmax activation, classification layer, dropout layer, GAP layer, multiscale filtering and channel attention (MSFCA) layer, max-pooling layer, batch normalization layer, ReLU activation, first convolutional layer, input image, weight matrix, and bias, respectively. The proposed DCNN model takes the preprocessed spectrogram images as an input in the first phase. During the preprocessing phase, the images are first downsized to $256 \times 256 \times 3$, normalized using a 1/255 scaling ratio, and then augmented with zooming, brightening, and horizontal flipping [10], on-the-fly, using the "ImageDataGenerator" class of Keras library. We used separable convolutional layers to reduce the computational overhead and to make the model lightweight. On the other hand, introducing batch normalization layers after convolutional layers in our model helps reduce the generalization error and to speed up the training process.

The processed spectrogram samples are then passed through the first convolutional layer of the proposed DCNN model, where we apply 96 filters with stride two of size (3×3) to learn basic feature representations. The output of this layer of dimension $(128 \times 128 \times 96)$ is then given to the batch normalization layer, followed by the first max-pooling layer. The max-pooling layer with stride two reduces the input volume's spatial dimension to $(64 \times 64 \times 96)$ for subsequent layers. Then, we add three consecutive MSFCA blocks to its end as shown in Fig. 5. The MSFCA blocks are designed with the proposed design principles said in points 1–4 above. The MSFCA block takes the output of the previous max-pooling layer or activation layer. This block helps the model learn multilevel feature representations from the same input tensor and improve the channel interdependency. As a result, the overall performance of the network increased with minimal computational cost. Each MSFCA block consists of four parallel paths. Path one, two, and three are made up with one convolutional layer followed by one batch normalization layer which can be expressed by the following equations, respectively:

$$f^{P_1}(x) = f^{B}(\sigma(f^{C_i}(f^{x} * w_i + b_i))) \quad (3)$$

$$f^{P_2}(x) = f^{B}(\sigma(f^{C_i}(f^{x} * w_i + b_i))) \quad (4)$$

$$f^{P_3}(x) = f^{B}(\sigma(f^{C_i}(f^{x} * w_i + b_i))) \quad (5)$$

where f^{C_i}, σ, f^{x}, w_i, b_i, i, and f^{B} represent the convolutional layer, ReLU activation, output feature map of the previous max-pooling layer or activation layer, weights, bias, layer index, and batch normalization layer, respectively. To capture multilevel feature representation from the input tensor, we applied different size filters in paths one, two, and three, but the stride was one for all convolution operations as per the proposed design principles in points 2–4, above. The filter size for paths one, two, and three was 5×5, 3×3, and 1×1, respectively. The output tensors of the three paths are then elementwise added in the addition layer using

$$f^{A_1}(x) = f^{P_1}(x) \oplus f^{P_2}(x) \oplus f^{P_3}(x) \quad (6)$$

where $\oplus$ is the elementwise addition operation. The multilevel feature representation map $f^{A_1}(x)$, generated by the elementwise addition operation, is then passed through a batch normalization layer f^{B} followed by an ReLU activation layer, which is then elementwise added with the output of path four $f^{P_4}(x)$, as given in the following equation:

$$f^{A_2}(x) = \sigma(f^{B}(f^{A_1}(x))) \oplus f^{P_4}(x) \quad (7)$$

where σ is the ReLU activation function. Path four, as given in (8), helps model channel interdependences with only a small

increase in the computational cost, but including this path in the MSFCA block directly improves the overall network performance

$$f^{P_4}(x) = \sigma(f^{C_i}(f^{\text{Mul}}(\alpha(f^{\text{FC}}(\sigma(f^{\text{FC}}(f^{G}(f^{x})))))) * f^{x})) \tag{8}$$

where σ, f^{C_i}, f^{Mul}, α, f^{FC}, f^{G}, and f^{x} denote the ReLU activation, convolutional layer, multiplication layer, sigmoid activation, dense layer, GAP layer, and output of the previous max-pooling layer or activation layer, respectively. Path four takes an input feature map x of dimension $D_s \times D_s \times M$, reduces it to dimension $1 \times 1 \times M$ by GAP layer, and then passes this M-length tensor through a multilayer perceptron, which projected back to the same dimension $1 \times 1 \times M$ as the dimension of the input feature map, which is then elementwise multiplied with the input feature map x. The first dense layer in this path reduces the complexity of its output channel by a factor of 16. The advantage of including this path in the MSFCA block is shown above in the proposed design principle 1. The resultant tensor is then subsequently passed into a 1×1 convolutional layer to maintain the same dimension as $f^{A_1}(x)$. Finally, by combining (3) and (8), we can express the MSFCA block as follows:

$$f^{\text{MSFCA}_i} = \sigma(f^{A_2}(x)) \tag{9}$$

where f^{A_2} is the second addition layer of that block, and x is the output feature map of the previous layer, which can be represented by $f^{x} \in \mathbb{R}^{D}$. D is the output dimension of the previous layer. After that, we fed the output tensor of f^{MSFCA_3} block to a GAP layer (for the reason said in design principle 5). Then, the output feature map of the GAP layer is fed into the dropout layer to prevent overfitting. Finally, a dense layer of six neurons triggered by the softmax activation function is added to the proposed DCNN network to anticipate the final action label. The proposed network has 213 793 parameters. The developed new DCNN model is trained end to end using 3780 m-D images of six classes of suspicious human activities. Furthermore, the proposed model uses depthwise separable convolutional layers and a GAP layer to reduce computation costs (lightweight DCNN), which makes the proposed architecture more suitable for on-device [e.g., field-programmable gate array (FPGA)] implementation.

B. Implementation Details

The proposed novel DIAT-RadHARNet architectural model is implemented on a Windows 10 workstation with an Intel Xeon Gold 6140 2 Processor (2.30 GHz, 2.29 GHz), 256-GB memory, NVIDIA Quadro RTX 6000 (24.0 GB) Graphics card, CuDNN v7.6.0, and CUDA v10.0.130 Tool kit. Our proposed architecture is implemented using Keras 2.2.4, python 3.5.6, matplotlib 2.2.2, and TensorFlow-GPU v1.14.0. In our work, we used a trial-and-error method to tweak the model hyper-parameters. The DIAT-RadHARNet network is end-to-end trained on the captured m-D signature dataset, and the Adam optimizer is used to optimize the network. We used zooming by 0.1 scale, brightening between 0.9 and 1.1, and horizontal flipping data augmentation technique to increase the size of the training dataset. The learning rate of the model is initially set to 0.001 and reduced by 0.5 factor up to 0.00001 when validation accuracy saturated in subsequent two epochs. The batch size is set to be eight, and the maximum epoch is 300. The weights and bias in all the convolutional layers are initialized with glorot uniform and a constant value of 0.2, respectively. On the other hand, the normal weight initializer is used in the dense layers of the MSFCA block. The number of filters applied in each layer and its output dimension are shown in Fig. 5. We used the same padding in all convolution operations to maintain the same dimension as the input tensor. To make the network more generalized and to prevent overfitting, 40% dropout is applied before the softmax layer. We also monitor the validation loss; if no progress is seen in the next ten epochs, we use early stopping to avoid overfitting. The categorical cross-entropy loss function, as shown in (10), is used to measure the performance of the DIAT-RadHARNet network

$$L = -\sum_{k=1}^{c} t_k \log(p_k) \tag{10}$$

where p_k denotes the softmax probability for the kth class, c is the total number of class, and t_k denotes the truth label.

TABLE III
CLASSIFICATION REPORT OF THE PROPOSED DIAT-RADHARNET MODEL

Action categories	Precision	Recall	F1-Score
Boxing	1.00	0.98	0.99
Army crawling	0.95	1.00	0.98
Army jogging	1.00	1.00	1.00
Jumping with holding a gun	1.00	0.98	0.99
Stone pelting / Grenades throwing	1.00	0.98	0.99
Army marching	1.00	1.00	1.00
Accuracy			99.22
Macro avg	0.99	0.99	0.99
Weighted avg	0.99	0.99	0.99

V. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

To illustrate the efficacy of the proposed DCNN model, the ROC curves are created for each of the six distinct activity classes for a given set of predicted probabilities and ground-truth labels. We have used a technique known as "one vs all" to obtain the ROC curve. In this technique, we have used "roc_curve" metrics from "sklearn package" which takes: 1) true binary labels; 2) target scores; and 3) positive class, and returns: 1) false positive rate and 2) true positive rate. In this way, the ROC curve, for every class, is obtained against all other classes. The resulting ROC plots are depicted in Fig. 6; from which we can observe that the area under the curve (AUC) value is one for all the classes, which proves the model's classification excellency. As we know, a high AUC score indicates that the model is good at differentiating the six human suspicious activity classes.

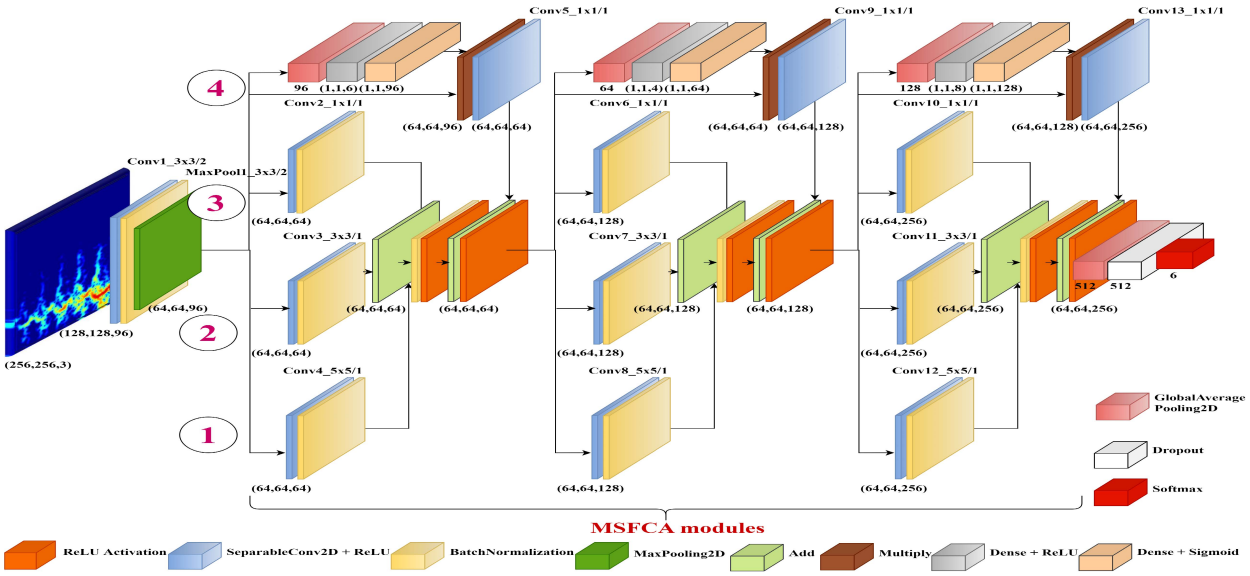


Fig. 5. Overall framework of the proposed DCNN architecture.

TABLE IV
COMPARISON OF THE DIAT-RADHARNET MODEL WITH OTHER POPULAR DCNN MODELS

DCNNs	Input Dimension	Layers	Size(Mb)	#Params(M)	FLOPs	Accuracy
VGG-16 [35]	$224 \times 224 \times 3$	23	528	138.3	15.5G	97%
VGG-19 [35]	$224 \times 224 \times 3$	26	549	143.6	19.6G	98%
1.0 MobileNet-224 [36]	$224 \times 224 \times 3$	92	16	4.2	0.57G	93%
MobileNetV2 [30]	$224 \times 224 \times 3$	156	14	3.5	0.30G	94.67%
DenseNet-201 [37]	$224 \times 224 \times 3$	709	80	20.2	4.32G	94.34%
Inception-v3 [31]	$299 \times 299 \times 3$	313	92	23.8	5.73G	84%
ResNet-50 [38]	$224 \times 224 \times 3$	177	98	25.6	3.88G	76%
BlazeFace [33]	$128 \times 128 \times 3$	73	3	0.6	0.30G	96.74%
DIAT-RadHARNet	$256 \times 256 \times 3$	55	2.75	0.21	0.83G	99.22%

A detailed/statistical performance study is conducted to evaluate the classwise performance and estimate the classification error of our DCNN model. The classwise F1-score, recall, and precision of our DCNN model are presented in the classification report given in Table III, which evidences for 99.22% of the overall classification accuracy of the proposed DCNN model. The loss and accuracy plot shown in Fig. 7 depicts the training process and the learning direction of our DCNN model. The training and validation curves in the accuracy plot converge, indicating that the network is not overfitted.

To perform a thorough test analysis on the classification accuracy of our proposed DCNN model, all the suspicious human activities shown in Fig. 4 are performed in front of the radar, with a new set of human targets, at different ranges between 10 m and 0.5 km during normal/adverse weather conditions, low-light environments, and long-range operations, and the respective (on-the-spot) test samples are collected. We used these on-the-spot test samples: normal (380 samples)/adverse (314 samples) weather conditions, low-light (211 samples) environments, and long-range (140 samples) target operations, to get thorough performance analysis. The classification statistics obtained during these open-field different environmental experiments are as follows.

1) Normal Weather Conditions: Fig. 8 gives the classification statistics of on-the-spot test results of the proposed DCNN model, obtained from the test experiments conducted during normal weather conditions. It can be observed from Fig. 8 that the model achieves 99.22% test classification accuracy, and the miss classification rate is only 0.78%. It also reveals that only three samples belonging to boxing, jumping with holding a gun, and stone-pelting are misclassified as army crawling.

2) Adverse Weather Conditions: The confusion matrix associated with the test experiments conducted during adverse weather conditions is shown in Fig. 9. The overall on-the-spot test accuracy of the model in adverse weather condition is 99.05%, and the misclassification rate is 0.95%. We can note that the model's false negative and positive scores for each class are very low.

3) Low-Light Environments: The performance statistics of the model in low-light conditions are given in the confusion matrix shown in Fig. 10. The model achieves 99.06% classification accuracy in these experimental conditions with only 0.94% of misclassification. It can also be observed that the model's false positive and negative results are minimal.

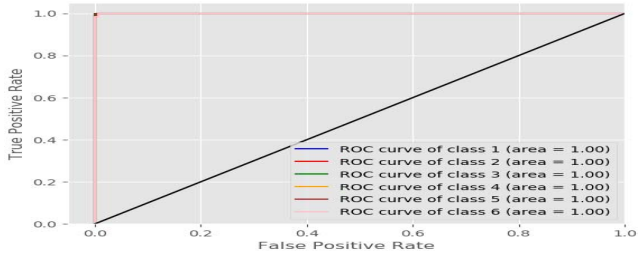
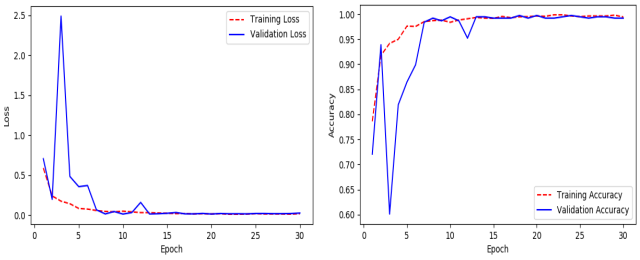


Fig. 6. ROC curve of the proposed DIAT-RadHARNet model.



(a) (b)

Fig. 7. Categorical cross-entropy loss. (a) Accuracy and (b) plots of the proposed DIAT-RadHARNet model.

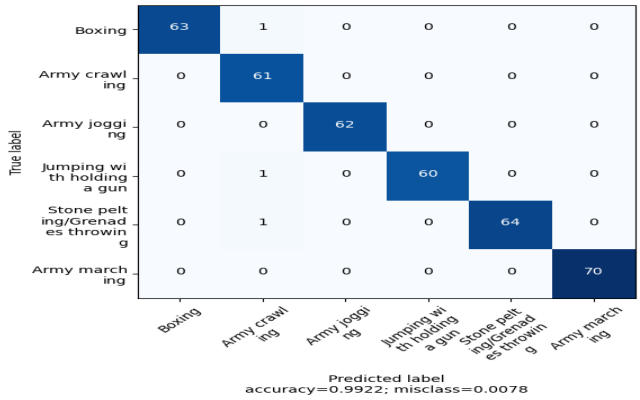


Fig. 8. DIAT-RadHARNet: confusion matrix—normal weather conditions.

4) Long-Range Target Operations: The proposed DCNN model achieves 99.29% of test classification accuracy in long-range target operations as reported in the confusion matrix shown in Fig. 11. The misclassification rate of our model in long-range operations is only 0.71%. We can note that only one sample belonging to the boxing category is misclassified as army crawling.

Thus, these open-field different environmental test experimental results evidence the good classwise classification performance/accuracy of our DCNN model on the newly, i.e., on-the-spot, collected test datasets/samples. It also reveals that the model's false negative and positive scores for each class are very low. Furthermore, the time complexity observed during the testing phase is 0.35 s, which is quite low, and it evidences the suitability of the designed DCNN model for on-the-fly classification applications.

Based on overall performance evaluations, the DIAT-RadHARNet model achieved a superior accuracy (99.22%) with 213 793 parameters and 55 layers. Table IV gives the comparison of DIAT-RadHARNet with that of other well-known models, trained, validated, and tested with our datasets.

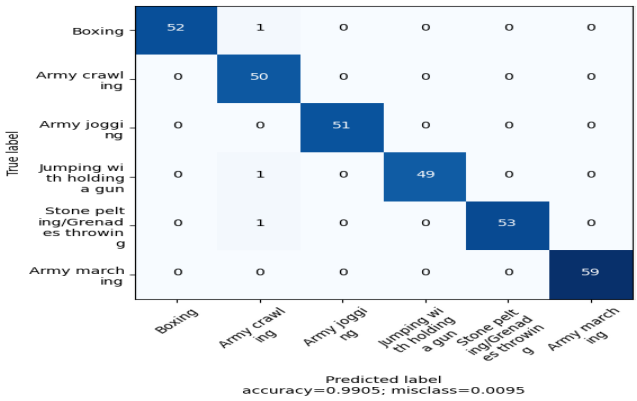


Fig. 9. DIAT-RadHARNet: confusion matrix—adverse weather conditions.

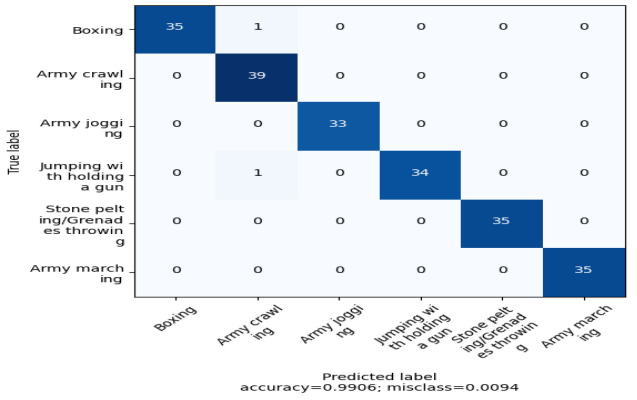


Fig. 10. DIAT-RadHARNet: confusion matrix—low-light conditions.

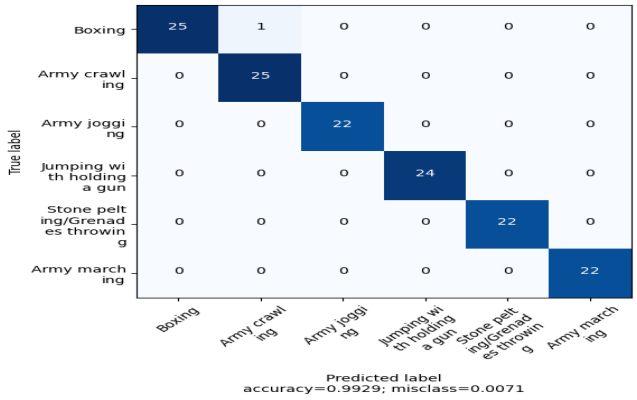


Fig. 11. DIAT-RadHARNet: confusion matrix—long-range operations.

It is noted from Table IV that the (layers), storing capacity (size), and computational complexity (parameters and FLOPs) are at the optimum tradeoff giving the classification accuracy at 99.22%. Even though the input dimension is high, i.e., 256 × 256 × 3, in our DCNN model, it gives a greater accuracy with a size of 2.75 Mb, parameters of 0.21M, and FLOPs of 0.83G. Although the layers of VGG-16 and VGG-19 are relatively low, all other design parameters are very high, and the classification accuracy is low, i.e., 97% and 98%, respectively. Similarly, FLOPs for 1.0 MobileNet-224, MobileNetV2, and BlazeFace are low as 0.57G, 0.30G, and 0.30G, respectively, but the respective classification accuracy is also low, i.e., 93%, 94.67%, and 96.74%, respectively.

TABLE V
ABLATION STUDY OF THE PROPOSED DIAT-RADHARNET MODEL. FLOPS DENOTE FLOATING POINT OPERATIONS

No of MSFCA blocks with/without channels weighting (CHW)	Layers	#Params(M)	FLOPs	Accuracy
1 MSFCA with CHW	23	0.03	0.12G	96.87%
2 MSFCA with CHW	39	0.07	0.27G	97.91%
3 MSFCA with CHW	55	0.21	0.83G	99.22%
4 MSFCA with CHW	71	0.49	1.96G	97.91%
5 MSFCA with CHW	87	1.05	4.16G	98.96%
6 MSFCA with CHW	103	2.16	8.55G	98.96%
3 MSFCA without CHW	31	0.10	0.62G	98.13%

Furthermore, through extensive designs and performance validation experiments, we found answers to all our research questions raised in Section II. The answers (findings) are as follows.

1) The optimum number of MSFCA blocks for our DCNN model is determined using an ablation study, as given in Table V. Therein, we can note that three consecutive MSFCA blocks are sufficient to have a test accuracy of 99.22%. Furthermore, it is observed that increasing the number of MSFCA blocks does not lead to increasing the accuracy, but it increases the computational complexity. Hence, we finalized only three MSFCA blocks in our proposed architecture. Adjusting the filter size in the depthwise separable convolutional layer resulted in no extra processing burden (as shown in Table V, the total FLOPs are only 0.62G when three MSFCA blocks are used without CHW).
2) Weighting the feature maps adds a minimal trainable parameter to the model, but the benefit is that it improves the model's generalization capabilities [as shown in Table V, the total FLOPs is only 0.62G when using three MSFCA blocks without CHW and 0.83G when using three MSFCA blocks with CHW, implying that the difference between these two cases is only 0.21G].
3) As shown in Table V, when three MSFCA blocks are used without CHW, the overall test accuracy of the model is 98.13%, indicating that the benefit of multilevel feature representation extraction plays a vital role in extracting valuable characteristics from spectrogram images.
4) Combination of feature map weighting based on priority and multiscale feature map extraction from m-D signatures of different human targets improves the network's generalization ability and overall accuracy (as shown in Table V, when three MSFCA blocks are used with CHW, the overall test accuracy of the model is 99.22% at the FLOPs of 0.83G).

VI. CONCLUSION

This article introduced a dataset "DIAT-μRadHAR" containing m-D signatures corresponding to six different human suspicious activities and designed a new lightweight DCNN architecture "DIAT-RadHARNet" for their detection and classification. The new DCNN model is trained/validated with the m-D signature samples recorded using an X-band (10 GHz) CW radar's open field, at different ranges (10 m–0.5 km), different weather/lighting conditions, diversified human targets, and different aspect angles, experiments. In this work, we increased the generalization capability of the designed DCNN model by weighting the feature maps and adding negligible trainable parameters. We also found that changing the filter size in the depthwise separable convolutional layer did not require additional computational overhead.

The designed DCNN model gives 99.22% classification accuracy with a very low false negative and positive outcome with 213 793 parameters and 55 layers. The time complexity observed during the testing phase of the designed DCNN model is 0.35 s which is quite low; hence, it is more suitable for on-device, i.e., FPGA implementation, which is one of our ongoing research works. However, the work reported in this article does not address the classification of suspicious activities of multiple subjects, inverse synthetic aperture radar (ISAR) imaging, classification of suspicious human activities based on ISAR signatures, cross-range/adaptive beam-forming effect analysis on m-D signatures of suspicious human activities, imaging the target's activities on a range-velocity map to demonstrate the separability of subject targets and slow-moving clutters, etc., which are our near-future research works.

ACKNOWLEDGMENT

The authors want to thank Dr. Suresh Kumar, DIPR, for his support and NVIDIA for the academic GPU research grant. The authors wish to thank all the students who supported them during data collection.

REFERENCES

[1] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1192–1209, Jul. 2013.
[2] D. F. Fioranelli, D. S. A. Shah, H. Li, A. Shrestha, D. S. Yang, and D. J. Le Kernec, "Radar sensing for healthcare," *Electron. Lett.*, vol. 55, no. 19, pp. 1022–1024, Sep. 2019.
[3] W. Huang, L. Zhang, W. Gao, F. Min, and J. He, "Shallow convolutional neural networks for human activity recognition using wearable sensors," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021.
[4] L. Ren, L. Kong, F. Foroughian, H. Wang, P. Theilmann, and A. E. Fathy, "Comparison study of noncontact vital signs detection using a Doppler stepped-frequency continuous-wave radar and camera-based imaging photoplethysmography," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 9, pp. 3519–3529, Sep. 2017.
[5] M. Chakraborty, A. Pramanick, and S. V. Dhavale, "Two-stream mid-level fusion network for human activity detection," in *Proc. Int. Conf. Innov. Comput. Commun.* Singapore: Springer, 2021, pp. 331–343.
[6] H. C. Kumawat and A. Bazil Raj, "Extraction of Doppler signature of micro-to-macro rotations/motions using continuous wave radar-assisted measurement system," *IET Sci., Meas. Technol.*, vol. 14, no. 7, pp. 772–785, Sep. 2020.
[7] R. M. Narayanan and M. Zenaldin, "Radar micro-Doppler signatures of various human activities," *IET Radar, Sonar Navigat.*, vol. 9, no. 9, pp. 1205–1215, 2015.
[8] M. Zenaldin and R. M. Narayanan, "Radar micro-Doppler based human activity classification for indoor and outdoor environments," *Proc. SPIE*, vol. 9829, May 2016, Art. no. 98291B.
[9] H. Chen and W. Ye, "Classification of human activity based on radar signal using 1-D convolutional neural network," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 7, pp. 1178–1182, Jul. 2020.
[10] H. C. Kumawat, M. Chakraborty, A. A. B. Raj, and S. V. Dhavale, "DIAT-μSAT: Small aerial targets' micro-Doppler signatures and their classification using CNN," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[11] W. Gao, L. Zhang, W. Huang, F. Min, J. He, and A. Song, "Deep neural networks for sensor-based human activity recognition using selective kernel convolution," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–13, 2021.
[12] W. Li, B. Xiong, and G. Kuang, "Target classification and recognition based on micro-Doppler radar signatures," in *Proc. Prog. Electromagn. Res. Symp.*, Nov. 2017, pp. 1679–1684.
[13] R. J. Javier and Y. Kim, "Application of linear predictive coding for human activity classification based on micro-Doppler signatures," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 10, pp. 1831–1834, Oct. 2014.
[14] Z. Chen, S. Xiang, J. Ding, and X. Li, "Smartphone sensor-based human activity recognition using feature fusion and maximum full *a posteriori*," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 3992–4001, Jul. 2020.
[15] I. Alnujaim, D. Oh, and Y. Kim, "Generative adversarial networks for classification of micro-Doppler signatures of human activity," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 3, pp. 396–400, Mar. 2020.
[16] R. Zhang and S. Cao, "Real-time human motion behavior detection via CNN using mmWave radar," *IEEE Sensors Lett.*, vol. 3, no. 2, pp. 1–4, Feb. 2019.
[17] S. K. Leem, F. Khan, and S. H. Cho, "Detecting mid-air gestures for digit writing with radio sensors and a CNN," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 4, pp. 1066–1081, Apr. 2020.
[18] A. Shrestha, H. Li, J. Le Kernec, and F. Fioranelli, "Continuous human activity classification from FMCW radar with Bi-LSTM networks," *IEEE Sensors J.*, vol. 20, no. 22, pp. 13607–13619, Nov. 2020.
[19] J. Höysniemi, P. Hämäläinen, L. Turkki, and T. Rouvi, "Children's intuitive gestures in vision-based action games," *Commun. ACM*, vol. 48, no. 1, pp. 44–50, Jan. 2005.
[20] A. F. Bobick and J. W. Davis, "The recognition of human movement using temporal templates," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 3, pp. 257–267, Mar. 2001.
[21] X. Geng, G. Li, Y. Ye, Y. Tu, and H. Dai, "Abnormal behavior detection for early warning of terrorist attack," in *Proc. Australas. Joint Conf. Artif. Intell.* Berlin, Germany: Springer, 2006, pp. 1002–1009.
[22] Z. Gu *et al.*, "Blind separation of Doppler human gesture signals based on continuous-wave radar sensors," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 7, pp. 2659–2661, Jul. 2019.
[23] B. Li, J. Yang, Y. Yang, C. Li, and Y. Zhang, "Sign language/gesture recognition based on cumulative distribution density features using UWB radar," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–13, 2021.
[24] D. Lee, H. Park, T. Moon, and Y. Kim, "Continual learning of micro-Doppler signature-based human activity classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
[25] A. A. B. Raj, *FPGA-Based Embedded System Developer's Guide*. Boca Raton, FL, USA: CRC Press, 2018.
[26] S. Kumar and A. A. B. Raj, "Design of X-band FMCW radar using digital Doppler processor," in *Proc. Int. Conf. Syst., Comput., Autom. Netw. (ICSCAN)*, Jul. 2021, pp. 1–5.
[27] G. L. Charvat, *Small and Short-Range Radar Systems*. Boca Raton, FL, USA: CRC Press, 2014.
[28] I. Pitas, *Digital Image Processing Algorithms and Applications*. Hoboken, NJ, USA: Wiley, 2000.
[29] A. S. Akella and A. A. B. Raj, "DCNN based activity classification of ornithopter using radar micro-Doppler images," in *Proc. 5th Int. Conf. Intell. Comput. Control Syst.*, May 2021, pp. 780–784.
[30] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.
[31] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.
[32] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.
[33] V. Bazarevsky, Y. Kartynnik, A. Vakunov, K. Raveendran, and M. Grundmann, "BlazeFace: Sub-millisecond neural face detection on mobile GPUs," 2019, *arXiv:1907.05047*.
[34] M. Chakraborty, S. V. Dhavale, and J. Ingole, "Corona-nidaan: Lightweight deep convolutional neural network for chest X-ray based COVID-19 infection detection," *Appl. Intell.*, vol. 51, pp. 3026–3043, Feb. 2021.
[35] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[36] A. G. Howard *et al.*, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.
[37] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.
[38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

Mainak Chakraborty (Graduate Student Member, IEEE) received the B.Tech. and M.Tech. degrees in information technology and computer science and engineering from the Maulana Abul Kalam Azad University of Technology, Kolkata, India, in 2008 and 2011, respectively. He is currently pursuing the Ph.D. degree with the Defence Institute of Advanced Technology, Pune, India.

He has been an Assistant Professor with the Department of Computer Science and Engineering, Dream Institute of Technology, Kolkata, since 2011. His research areas are deep learning, machine learning, human activity detection, and medical image analysis.

Mr. Chakraborty was a recipient of the AICTE-INAE Teachers Research Fellowship from the Indian National Academy of Engineering.

Harish C. Kumawat (Graduate Student Member, IEEE) received the B.Tech. degree in electronics and communication engineering from Rajasthan Technical University, Kota, India, and the M.Tech. degree in microwave and optical communication from Delhi Technological University, New Delhi, India.

He was an Assistant Professor with the Government Engineering College Bikaner, Bikaner, India, from August 2014 to January 2018. He is currently working in the field of radar signal processing with the Defence Institute of Advanced Technology, Pune, India.

Sunita Vikrant Dhavale (Member, IEEE) received the M.E. degree in computer science from Pune University, Pune, India, in 2009, and the Ph.D. degree in computer science from the Defence Institute of Advanced Technology (DIAT), Pune, in 2015.

She is currently an Assistant Professor with the Computer Science and Engineering Department, DIAT. Her research interests include computer vision, deep learning, and cybersecurity.

Dr. Dhavale is a member of many professional bodies, including IEEE, Association for the Advancement of Artificial Intelligence (AAAI), Association for Computing Machinery (ACM), Indian Society for Technical Education (ISTE), Institution of Electronics and Telecommunication Engineers (IETE), International Association of Engineers (IAENG), and Information Systems Audit and Control Association (ISACA). She was a recipient of the Atal Tinkering Lab Mentorship, the ISACA Academic Advocate Membership, the EC-Council's CEH-v8, the IETE M. N. Saha Memorial Award, and the Outstanding Woman Achiever Award from the Venus International Foundation.

Arockia Bazil Raj A. (Member, IEEE) received the B.E. degree in electronics and communication engineering from Bharathidasan University, Tiruchirappalli, India, and the M.E. degree in communication systems and the Ph.D. degree in information and communication technology from Anna University, Chennai, India.

He was an Assistant Professor with the Kings College of Engineering, Thanjavur, India, from 2002 to 2006. He is currently an Associate Professor with the Defence Institute of Advanced Technology (DIAT), Pune, India. Based on his teaching and research expertise, he wrote three books published by De Gruyter, GmbH, CRC Press, USA, and IET, U.K. His current research interests cover radar system design, radar signal processing, free-space optical communication (FSO), digital signal processing, and RF photonics.