IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 8, 2021, accepted August 23, 2021, date of publication August 24, 2021, date of current version September 3, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3107782

Coded-Aperture Computational Millimeter-Wave Image Classifier Using Convolutional Neural Network

RAHUL SHARMA[1], RAPHAEL HUSSUNG[2], ANDREAS KEIL[2], (Member, IEEE), FABIAN FRIEDERICH[2], THOMAS FROMENTEZE[3], MOHSEN KHALILY[4], (Senior Member, IEEE), BHABESH DEKA[5], (Senior Member, IEEE), VINCENT FUSCO[1], (Fellow, IEEE), AND OKAN YURDUSEVEN[1], (Senior Member, IEEE)

[1]Institute of Electronics, Communication and Information Technology, Queen's University Belfast, Belfast BT3 9DT, U.K.
[2]Department of Material Characterization and Testing, Fraunhofer Institute for Industrial Mathematics ITWM, 67663 Kaiserslautern, Germany
[3]XLIM, UMR 7252, University of Limoges, 87000 Limoges, France
[4]5G and 6G Innovation Centre (5GIC and 6GIC), Institute for Communication Systems (ICS), University of Surrey, Guildford GU2 7XH, U.K.
[5]Department of Electronics and Communication Engineering, Tezpur University, Tezpur 784028, India

Corresponding author: Rahul Sharma (rsharma05@qub.ac.uk)

This work was supported by the Leverhulme Trust through Research Leadership Award under Grant RL-2019-019.

ABSTRACT A millimeter-wave (mmW) classifier system applied to images synthesized from a coded-aperture based computational imaging (CI) radar is presented. A developed physical model of a CI system is used to generate the image dataset for the classification algorithm. A convolutional neural network (CNN) is integrated with the physical model and trained using the dataset comprising of synthesized mmW images obtained directly from the developed CI physical model. A *k*-fold cross validation technique is applied during the training process to validate the classification model. The coded-aperture CI concept enables image reconstruction from a significantly reduced number of back-scattered measurements by facilitating physical layer compression. This physical layer compression can substantially simplify the data acquisition layer of imaging radars, which is realized using only two channels in this article. The integration of the classification algorithm with the CI numerical model is particularly important in enabling the training step to be carried out using relevant system metrics and without the necessity for experimental data. Leveraging the CI numerical model generated data, training step for the classification algorithm is achieved in real-time while also confirming that the numerically trained CI classifier offers high accuracy with both simulated and experimental data. The classifier integrated physical model also enables performance analysis of the classification algorithm to be carried out as a function of key system metrics such as signal-to-noise (SNR) level, ensuring a complete understanding of the classification accuracy under different operating conditions. The trained CI system is tested with synthesized mmW images from the physical model and a classification accuracy of 89% is achieved. The proposed model is also verified using experimental data validating the fidelity of the developed CI integrated classifier system. A classification latency of 3.8 ms per frame is achieved, paving the way for real-time automated threat detection (ATD) for security-screening applications.

INDEX TERMS Millimeter-wave, imaging radars, computational imaging, neural networks, image classification, coded-aperture.

I. INTRODUCTION

Millimeter-wave (mmW) imaging has found applications in various areas ranging from remote sensing [1]–[3] to autonomous robotics [4]–[6] and medical applications such as breast cancer detection [7], [8]. Another such useful application is concealed weapon detection [9]–[12] because, unlike X-rays, radiation at mmW does not exhibit ionizing effects, hence posing no health hazards. As a result, mmW radar modalities have been widely used in places where

The associate editor coordinating the review of this manuscript and approving it for publication was Kathiravan Srinivasan.

security is of utmost importance. Various image processing techniques have been demonstrated on mmW images such as compressive sensing [13], [14], super-resolution [15], [16], denoising and image enhancement [17], [18]. Classification techniques on such images have been previously explored [19], [20], although the literature on these techniques is still scarce. The majority of classification research focuses on generating images with systems using traditional imaging techniques, relying on a raster scanning principle, i.e. carrying out a point-by-point scan of an imaged scene by means of a beam synthesis on the radar aperture. In this context, most commonly used techniques can be considered as variants of synthetic aperture radar (SAR) [11], [21], [22] and phased array based modalities [23], [24]. The SAR concept conventionally involves performing a mechanical scan on the aperture plane to directly measure the back-scattered radar data on a point-by-point basis. The phased array technique performs the raster scanning of the scene in an all-electronic manner by making use of a radiation pattern which can be steered electronically without physically moving the antennas. Although such techniques are known to produce excellent imaging results, they exhibit few shortcomings. The mechanical scanning on the aperture plane used in SAR concept poses significant challenges when it comes to real-time data acquisition, particularly in security-screening applications which require imaging of electrically large scenes. The all-electronic raster scanning in phased arrays makes use of an array of dedicated phase shifting circuits and power amplifiers to have full control of the individual array elements forming the aperture. This can significantly increase the complexity of the hardware layer and amount of power consumed by the aperture.

For a classification problem, the convolutional neural network (CNN) model requires a sufficient number of data samples to ensure that the system generalizes well and does not overfit to the training data. In the context of mmW security-screening and classification for threat detection, this requirement translates into the need for a sufficient number of mmW images of threat objects to successfully train the CNN model. Moreover, given the demand of the application to produce a fast and accurate classification system and keeping in mind the quality of mmW images, training the neural network model for mmW classification typically requires a large dataset. In light of these challenges, employing the traditional raster-scanning based methods to generate such large datasets and train the classification system can be counterproductive. This is because raster-scanning based imaging modalities conventionally suffer from slow data acquisition speeds and exhibit significant computational demands for image reconstruction due to the need to process large datasets caused by the dense sampling of the aperture layer, typically at the Nyquist limit. Moreover, generating the mmW images of real targets for training by means of experimental measurements can take a prohibitively long time. To relax the raster-scanning requirement, sparse computational imaging (CI) techniques [25]–[32] can be leveraged. In addition, the training of the CNN network can be done on synthesized mmW images.

CI differs from the conventional methods in that it makes use of compressive single-pixel apertures radiating spatio-temporally incoherent radiation patterns. Using these quasi-random bases, the entire scene information can be encoded and compressed onto a single channel at the antenna layer. This can substantially simplify the hardware architecture of mmW radars and reduce the number of data acquisition channels on the physical layer. This paper shows a CNN-enabled classification algorithm integrated into a coded-aperture CI model. To this end, a numerical CI forward model is used to generate mmW images from CAD models of various threat objects. The training of the model is achieved using the reconstructed mmW images generated by the developed near-field coded-aperture CI model.

The main novelty of the paper rests on three pillars:

- First, a numerical model of a single-frequency coded-aperture based compressive CI system is developed for the first time to facilitate the training of mmW CI based radar image classification. This is achieved using the synthetic mmW images reconstructed by the numerical model itself. As outlined in Section II, conventional classification research in mmW radar imaging has a significant limitation in that it mainly focuses on the learning algorithm – not the generation of the imaging data for training. This poses a lack of complete system analysis. The developed physical model in this paper represents an end-to-end system, taking into account all aspects of the radar imaging process, from data acquisition to image reconstruction and classification. The developed CI model not only can significantly simplify the hardware layer of conventional radar architectures due to the single-frequency physical layer compression (Section IV), it also paves the way for physical model development for imaging systems with deep learning to facilitate real-time classification (Sections V and VI). This is particularly important for security-screening applications, as real-time classification is the key element for automated threat detection (ATD).
- Second, a CNN classifier is developed and integrated in the signal processing layer of the CI system. To the best of our knowledge, this is the first time that the mmW CI technique leverages the classification concept and achieves it in real-time. CI facilitated by compressive sensing has recently received significant traction, including the author's pioneering works in this field [26], [29]–[35]. However, all of these works have relied on interpreting the reconstructed images, qualitatively, by means of a visual inspection. In security-screening applications, this is not a viable option for threat detection. Different from the previous mmW CI imaging works, we successfully integrate the developed CI physical model with a k-fold cross-validated classifier. The presented technique not only achieves high

accuracy classification from the CI radar reconstructed mmW images, it also achieves this in real-time, generating the first proven classification model for ATD in CI radars.
- Third, and finally, we provide a systematic, quantitative analysis of the robustness of CI based radars to system noise and its effect to the classification accuracy - both numerically and experimentally. In order for CI based radars to be considered a realistic technology for security-screening applications, they must be able to achieve classification under various signal-to-noise ratio (SNR) conditions. Currently, the robustness of CI based radars to system noise for classification is unknown. As well as answering this question (Section VI), we prove that the capability to generate our own mmW imaging dataset is vital to ensure optimum classification performance under various system noise levels. This is of significant importance in radar imaging because the developed physical model gives us the ability to alter the system noise level within the practical imaging environment. It also enables us to optimize the training data to match with the desired system parameters and improve the accuracy of the classification step (Section VI).

The main motivation of this work is to develop a coded-aperture based mmW imaging system that facilitates compressive sensing by leveraging physical layer compression and its integration with a CNN-based classifier as an enabling technology for ATD in real-time. Specifically, this work aims to be the first to achieve all of the following tasks *simultaneously*: (*i*) a coded-aperture based mmW CI radar is developed as an enabling physical model for classification, (*ii*) the mmW CI physical model is integrated with a CNN-classifier, (*iii*) the integrated model is verified both numerically and experimentally under various SNR conditions, and (*iv*) real-time ATD at mmW frequencies is achieved. The physical layer compression enables the reconstruction of mmW images from a reduced number of back-scattered measurements of the imaged scene - a concept that is achieved using single pixel compressive coded-apertures in this work. The integration of the CNN classifier into the developed CI model is vital to ensure that training of the classifier system can be achieved using the developed CI numerical model generated synthetic mmW images, eliminating the need for experimental data. In other words, the classification algorithm used by the system is trained using only the synthesized data from the developed numerical CI model alone and not with data acquired by any experimental measures. This approach is advantageous in security-screening applications for three reasons:

- First, getting access to real-life threat objects for data collection can be extremely difficult. The proposed approach makes the data generation for training easier in the sense that any CAD model of real threat objects can be used as a target and the developed CI model can be used to generate its image.
- Second, for an accurate classification of mmW images, wherein a large set of mmW images are required for training, the presented approach is feasible because, instead of using an experimental system for data collection, the developed CI model, which relies solely on fast data acquisition CI techniques instead of slow and expensive SAR or phased-array based conventional imaging techniques, can be used to generate as much data as required.
- Third, and finally, synthesizing the training data through the developed physical model enables us to consider the effect of system metrics, such as SNR level, on classification performance. This advantage makes it possible to systematically optimize the training data and improve the classification accuracy.

The outline of this paper is as follows: Section II provides a brief insight into mmW image classification to reflect on the current state-of-the-art in this field. Section III provides an overview of the system designed. In Section IV, we present the synthesized CI system. Section V gives the detailed information about the CNN model design whereas Section VI provides a discussion on the results obtained. Finally, Sections VII and VIII present the future work and the concluding remarks, respectively.

II. RELATED WORK

Object detection or classification of images using machine learning techniques has received substantial interest in the literature, particularly at optical frequencies [36]–[43]. However, mmW image classification, including data acquisition, to facilitate real-time ATD for security-screening is vastly understudied. In [44], classification was carried out based on a spectral analysis of mmW images using support vector machine (SVM). Despite encouraging results, it is important to emphasize that SVM is a more suitable technique for binary classification problems. As a result, using SVM for multi-class classification requires an additional pre-processing of dividing the problem into several binary classification problems. Also, when the size of the dataset is large, as is the case in this paper, research suggest that deep learning networks are known to perform better in image classification problems. [45], [46]. Works related to deep learning networks [47]–[50] have also been carried out in recent past. Although most of the work has shown that CNN is a very useful approach in mmW object detection, almost all of the work available in the literature focuses solely on the learning algorithm and not on the data collection for training. For instance, in [51], CNN was used for object detection in human-scaled targets. To achieve this, an experimental radar system consisting of an array of antennas and operating at 27 GHz with a bandwidth of 5 GHz was leveraged to collect approximately 3000 multi-angle mmW human images. Generating necessary training data via such a system can be extremely time consuming and not a feasible solution, particularly when the number of classes to detect is significant.

In [52] and [53], the use of CNN was shown for target detection in airborne SAR images. In these works, instead of fully connected layers at the end, sparsely connected layers were used. However, the classification was shown on an already available dataset that is a collection of thousands of SAR images. This approach is not ideal for classification studies because such datasets do not necessarily exhibit relevant system metrics in terms of operating frequency, bandwidth, aperture size, SNR, imaging distance and resolution. It is therefore vital to generate the training data using the actual physical model of a radar system under its relevant system parameters. An interesting work was presented in [54] where deep learning was facilitated as an enabling technique to solve the image reconstruction problem at terahertz (THz) frequencies. It should be mentioned that this is a fundamentally different concept from the classification problem studied in this paper. Moreover, because the work presented in [54] was not experimentally tested, it is not possible assess the accuracy and applicability of [54] for practical applications. For the security-screening application we consider, this poses a significant disadvantage.

From the available research on classification of mmW images, it can be concluded that most of the works were carried out on already available datasets consisting of SAR images or depend upon approaches that heavily rely on an experimental set-up to generate the images. Our system has an advantage over such approaches in the sense that we propose an end-to-end mmW CI model for ATD in real-time.

III. SYSTEM OVERVIEW

The developed system is a mmW CI system designed for classification of threat objects, which has the potential to be used in security-screening applications. It classifies an image to its appropriate class, namely Class '0', '1', '2' or '3' (details about the classes are given in Section V. A.). In order to do so, the system uses three modules: a simulator, a CNN layer and lastly a classifier. The overview of the whole system is given in Fig. 1.

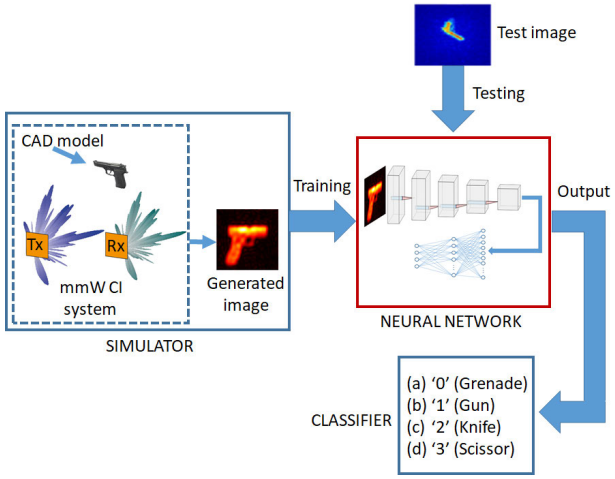


FIGURE 1. CI classifier system overview.

The simulator is used before the training phase whereas the CNN layer and the classifier are used during training and testing phases. Each module can be summarized as follows:

1) **Simulator** (*Section IV*): A physical model of a near-field coded-aperture CI system operating in a bi-static mode (one transmitter and one receiver aperture) and facilitated with a hardware layer of an FPGA architecture to reconstruct the mmW images in real-time.
2) **CNN Layer** (*Section V*): A CNN layer is designed having five convolutional layers, two fully connected layers and one output layer to extract the features from the input images and classify them accordingly to their respective classes. The CNN layer is trained with images generated by the simulator.
3) **Classifier** (*Section VI*): This is the testing phase where the trained CNN model is tested with a new set of mmW images numerically synthesized using the CI simulator and experimentally reconstructed mmW images of real targets. The classifier uses the trained CNN layer to decide on the class of each input test image. The predicted class for each image is compared with its true class to determine the accuracy of the model.

IV. CODED-APERTURE CI SYSTEM

CI techniques can be employed to synthesize quasi-random bases and relax the raster scanning requirements of SAR and phased array modalities. One such technique is the coded-aperture technique, which uses a set of masks with spatio-temporarily varying complex weights [55]–[58], through which the transmitted wavefronts probing the imaged scene and the received wavefronts collecting the back-scattered information are passed. This spatio-temporal mapping aperture concept can be thought of as a large aperture that can radiate complex, quasi-random radiation patterns. This enables the collection of the scene information in an *indirect* manner, by illuminating the scene with spatio-temporally incoherent field patterns and encoding the back-scattered information into a few set of measurements or modes. This indirect mapping of the scene information eliminates the need for a point-by-point raster scan, a significant advantage of CI based imaging modalities. Fig. 2 shows the setup for such a CI system. It consists of two dynamic apertures, operating in a bi-static mode (one as a transmitter, Tx, and other as a receiver, Rx). In this depiction, the radiated fields from the coded-apertures are shown for two different measurement mask configurations. In this paper, for a single imaging scenario, the synthesized CI system uses a total of 500 different mask configurations, suggesting that the back-scattered information from the imaging scene is encoded into 500 measurements. The size of the transmit and receive CI apertures is 0.5 m x 0.5 m each, hence the effective aperture size is of 1 m x 1 m. It should be noted here that synthesizing the same effective aperture using the conventional Nyquist SAR techniques would require 1600 sampling points across the aperture plane whereas the synthesized CI

system depicted in Fig. 2 uses only two channels to capture the scene information. The synthesized coded-aperture CI imaging scenario constitutes a near-field imaging problem where the imaging scene is placed at a distance of 0.5 m and has a cross-range size of 0.3 m x 0.3 m. Because the coded-aperture technique can synthesize spatio-temporally varying radiation patterns simply by means of modulating the aperture masks without the need for a frequency-sweep, we consider a single frequency operation of 12 GHz. However, it is worth mentioning that the operation of the CI system is not restricted to this particular frequency and can be extended to other frequency ranges as well. In the depiction of Fig. 2, the system is shown to be imaging a CAD model of a gun phantom as the threat object.

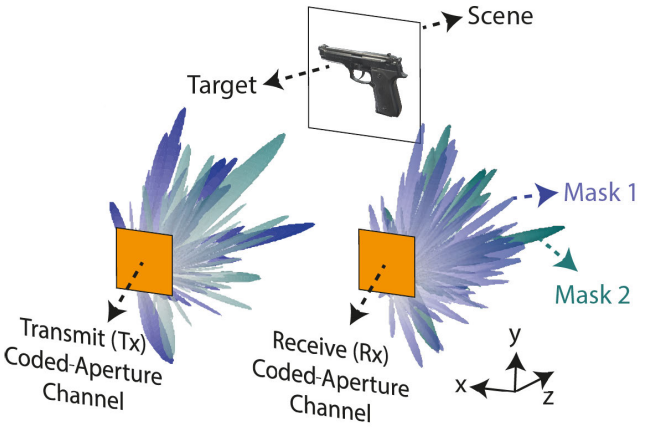


FIGURE 2. CI setup depicting a synthesized aperture consisting of transmit (Tx) and receive (Rx) coded-apertures operating in bi-static mode and imaging a gun target. Spatio-temporal variation in the radiated field patterns is shown for two masks as an example. Not drawn to scale.

From the radar imaging layout in Fig. 2, under the first Born approximation, the back-scattered radar measurements, given by $\mathbf{g}$, can be correlated to the imaged scene through the forward model, which is given as:

$$\mathbf{g}_{M\times 1} = \mathbf{H}_{M\times N}\mathbf{f}_{N\times 1} + \mathbf{n}_{M\times 1} \tag{1}$$

In Eq. (1), $\mathbf{f}$ signifies the reflectivity distribution of the discretized pixels in the imaged scene, $\mathbf{H}$ is the sensing matrix and $\mathbf{n}$ is the measurement noise, modelled as a Gaussian distribution with zero mean [33]. In this work, to ensure a realistic model, the back-scattered measurements, $\mathbf{g}$, exhibit a finite SNR of 20 dB [33], [34]. Subscripts M and N denote the number of measurement modes and the number of pixels used to define the scene, respectively. From the first Born approximation, the sensing matrix columns are formed by the dot product of the transmitter and receiver fields, $\mathbf{E}_{Tx}$ and $\mathbf{E}_{Rx}$ respectively, for every pixel of the scene as given in Eq. (2). Hence, the total number of columns is N whereas the number of rows is the number of measurement modes used, M. Therefore, the size of $\mathbf{H}$ is $M \times N$.

$$\mathbf{H} = \mathbf{E}_{Tx}\mathbf{E}_{Rx} \tag{2}$$

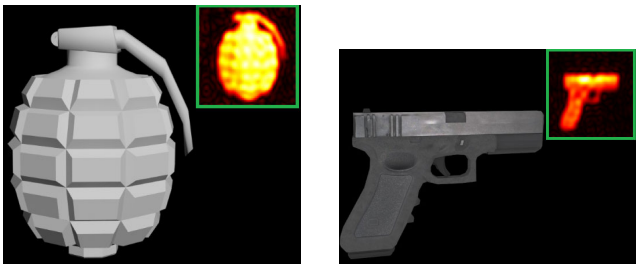


FIGURE 3. Images of two classes of threat objects (a) 2D CAD model of a grenade and CI reconstructed mmW image shown as inset (b) 2D CAD model of a gun phantom and CI reconstructed mmW image shown as inset.

From Eq. (1), it is evident that the retrieval of the scene information from the back-scattered measurements is an inverse problem and involves the adjoint operation of a large sensing matrix. In a typical CI setup, which uses multiple apertures and usually involves imaging a large scene, it has been shown in the past that processing of sensing matrices as large as 90 GB can be required [35]. Such a calculation tend to be intensive computationally, hence an FPGA architecture is employed in the reconstruction process to share the computational load with the CPU, as presented in [34]. This hardware architecture makes use of the FPGA logic blocks to carry out the calculation of the sensing matrix in quasi real-time. Such a calculation needs to be carried out only once as the sensing matrix is dependent only on the aperture and scene parameters, and exhibits no dependence on the imaged target. Once the sensing matrix is calculated, an estimate of the imaged scene ($\mathbf{f}_{est}$) can be retrieved by means of a single phase compensation, which is nothing but complex conjugate of the calculated sensing matrix ($\mathbf{H}^{\dagger}$), applied to the back-scattered measurements, $\mathbf{g}$, as given by:

$$\mathbf{f}_{est} = \mathbf{H}^{\dagger}\mathbf{g} \tag{3}$$

V. CNN CLASSIFIER

Once the sensing matrix is calculated, the CI model is ready to generate mmW images of any target within the defined scene to be imaged.

A. IMAGE DATASET

In this work, we make use of CAD models of several threat objects such as guns, grenades, knives and scissors to form the dataset for training. Fig. 3 shows CAD models of a gun and a grenade and their respective mmW images generated by the developed CI model outlined in Section III.

Since the dataset comprises of four types of threat objects, this classification problem will be a multi-class one, comprising of four classes:

- Class '0' or Grenade,
- Class '1' or Gun,
- Class '2' or Knife,
- Class '3' or Scissor.

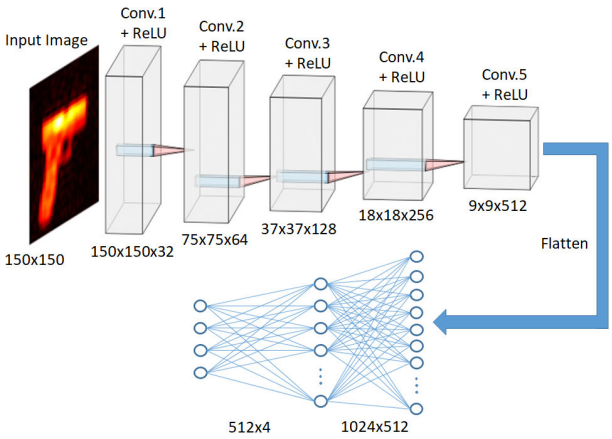


FIGURE 4. CNN architecture.

Initially, a total of 531 total images of four classes are generated. To further increase the size of the dataset for training the CNN model and achieve an accurate classifier, the concept of data augmentation is used. This process adds slightly modified copies of the existing data in which the CI reconstructed original image dataset are passed through a series of transformations, such as right or left shifting, rotation, cropping or flipping. Each transformation of an individual image creates a new version of it, and these newly created images are treated as separate images to increase the size of the dataset for training. In this case, each image is subjected to four transformations:

- Right shifting by 20%,
- Down shifting by 10%,
- Zooming by 30%,
- Horizontal flipping.

Using this method, four transformed copies of a single image are generated and this process is repeated for every image in the existing dataset, and used for training. Hence, the size of the dataset is increased from 531 to 2655 images.

B. CNN ARCHITECTURE

The developed CNN architecture is shown in Fig. 4.

As mentioned in Section V-A, the training dataset comprises of images reconstructed from the developed CI model. The properties of these images, such as colour space or pixel dimensions, are similar to the ones shown in Fig. 3 insets. Each input image for the network is of dimension 150×150. The CNN layer comprises of five convolution layers, layer 1 through layer 5. Each convolutional layer uses multiple filters to extract the useful features from the images. The size of each convolutional layer is specified in Fig. 4. For example, for the first convolution layer (layer 1), the size specified is $150 \times 150 \times 32$. It means that the width and height of this particular 2D convolutional window is 150 and 150, respectively, whereas the number 32 signifies the number of filters that the convolutional layer will learn.The first convolutional layers will filter the $150 \times 150 \times 3$ image with 32 kernels of size $3 \times 3 \times 3$. The second layer takes

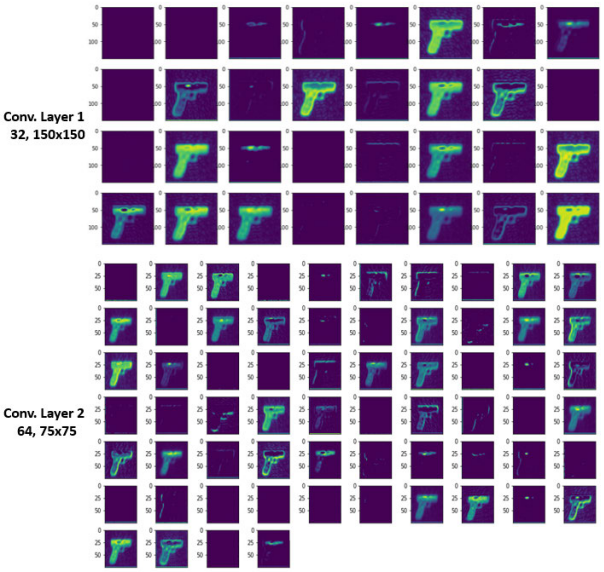


FIGURE 5. The convolutional filter output for the first two layers. Layer 1 has an output size of 32 filters each of 150 × 150 whereas layer 2 has an output size of 64 filters each of 75 × 75.

as input the output of first convolutional layers and filters it with 64 kernels of size $3 \times 3 \times 32$. This is followed for the rest of the convolutional layers. The filter outputs for the first two convolutional layers on the input 150×150 are given in Fig. 5. It should be mentioned here that whereas the input image example for the depictions in Fig. 4 and Fig. 5 is centralized, as indicated in Section V.A, the synthesized imaging dataset also includes non-centered mmW reconstructions. As a result, the presented technique does not require centralizing the objects to achieve classification. The important criterion is that the objects must be placed within the field-of-view (FoV) of the synthesized radar aperture. This is to ensure that the object can be seen by the radar aperture. However, this is not a fundamental limitation of the CI technique specifically. Rather, it is a physical phenomenon that applies to all imaging radars.

No spatial zero padding is used in the convolutional layer and the convolutional stride is fixed at 1 pixel. Each layer is equipped with an activation layer 'ReLU', which makes sure that the negative values are not activated onto the next layer. It is widely used in CNN and does not saturate whilst also not suffering from the Vanishing Gradient problem [59]–[61]. It is worth mentioning that deep neural networks with ReLU units train much faster than other activation units [62]. The selection is also justified with a study wherein training of the dataset has been carried out with two different activation units: ReLU and tanh for a fixed number of epochs and the training error for both the activation units have been recorded. As can be seen in Fig. 6, the network with ReLU units reaches the error rate of 11% faster, i.e. in less number of epochs as compared to the same network with tanh activation units. For this analysis, we choose the 11% error rate as baseline for comparison and note that a similar

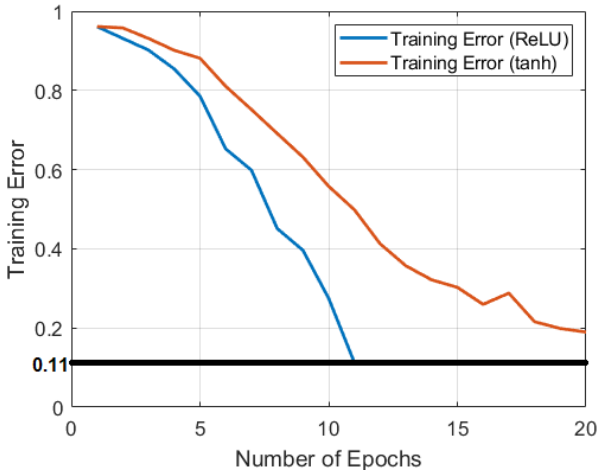


FIGURE 6. The neural network with ReLU units reaches 11% error rate much quicker than an equivalent network with tanh units.

outcome can be observed for other error rates without loss of generality.

Pooling is used to reduce the dimension of the output, hence decreasing the amount of computations and weights. The CNN architecture uses 2×2 MaxPooling which reduces the convolutional layer dimensions by half. As in Fig. 4, after applying MaxPooling to the output of the first convolution layer, the dimensions of the second layer are reduced to 75×75. After the creation of the convolutional base, in order to generate a prediction, the output of the convolutional base has to be flattened and helps to map the relationship between the input and the output. Two fully connected layers are used for this purpose. The first fully connected layer has 1024 neurons whereas the second one has 512 neurons. Each neuron in the first layer is connected to each neuron in the second layer.The last layer consists of four neurons as there are four classes to predict from. Since this particular problem is a case of multi-class classification, multinomial probability is used. In other words, a multinomial probability distribution for the four output nodes at the output layer needs to be calculated to predict class for a particular input image. This is carried out by the network by using the softmax activation function at the output layer. The probability distribution for the input gun image in Fig. 4 is shown in Fig. 7. As seen in the figure, the score at the output nodes are converted to probabilities by the formula given in the figure. It can be observed that the maximum probability is seen in the second node, which belongs to Class 1 or Gun. Hence, the model is correctly predicting that the input image is of a gun.

The 'categorical cross-entropy loss' is set as the loss function as it being a case of multi-class classification.

$$\text{Loss} = -\sum_{i=1}^{\substack{\text{total}\\ \text{classes}}} y_i \cdot \log \hat{p}_i \tag{4}$$

where $\hat{p}_i$ is the i-th probability value in the model output, y_i and 'total classes' is the number of classes in the model output (in this case, it is four).

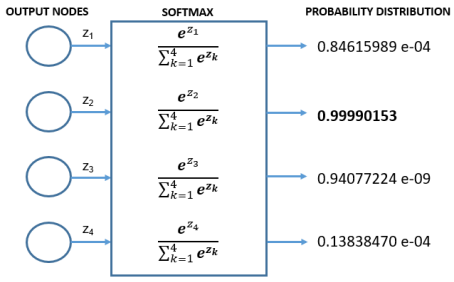


FIGURE 7. The probability distribution at the output layer is shown in detail for the input gun image in Fig. 4. Maximum probability is observed in the second node which belongs to Class 1, i.e. 'Gun'.

The model is optimized with the 'adam' optimizer [63]. As per studies in [64], for neural networks with ReLU activation functions, weights are initialized by zero-mean of Gaussian distribution with standard deviation of $\sqrt{\frac{2}{n}}$, where n is the number of inputs to each input node. A constant value of 0.1 is used as biases [65]. The learning rate is selected to be 0.001 initially. A common practice is to observe the validation accuracy during the training process and to scale down the learning rate by a factor of 0.1 or 0.5 if the validation accuracy stops improving [66]. A similar approach is followed wherein the learning rate is reduced by a factor of 0.1 after 50 epochs.

C. CNN TRAINING

Since the CNN uses synthesized mmW images from the CI physical model for training, an assurance on the accuracy of predictions is needed. In other words, we need to validate the model, i.e. check whether the model is underfitting/over-fitting/well generalised based upon its performances on unseen data. One such technique that can be used for this purpose is the *k*-folds cross-validation [67]–[70]. The *k*-folds cross-validation results in a less biased model compared to other validation methods such as the train-test split [71], [72] because, in this method, it is ensured that every observation from the original dataset has a chance of appearing in the training and validation set. This is extremely helpful in our particular case of limited dataset. For this work, the value of *k* is chosen to be 5 and the entire dataset is split into *k*=5 folds. The selection of *k*=5 ensures that there are sufficient samples in the training set to efficiently train the model as well as there are sufficient samples in the validation set to accurately evaluate the model. Choosing a higher value of *k* will result in a reduced validation set, hence affecting the evaluation process. On the other hand, choosing a lower value will result in a reduced training set, hence affecting the learning process of the model. The model is trained using the *k*-1 folds and the k^{th} fold is kept aside for validating the model and the accuracy is noted. This step is repeated until every *k* fold serves as validation set and accuracies are recorded. Each fold is run for 100 epochs. The average of all the accuracies serve as the performance metric of the model. The validation accuracies for each of the fold are recorded and tabulated in Table 1.

TABLE 1. Cross-validation accuracies of the model.

Iterations	Validation Accuracies
1	0.7851
2	0.7930
3	0.8818
4	0.8961
5	0.8184
Average	**0.8348**

The average accuracy of 0.8348 in Table 1 shows that the model has the ability to predict with the desired high precision in an actual testing environment on test images. It should be noted that the variation in the validation accuracies across different number of iterations in Table 1 is the consequence of how the dataset (which is random) is split in each iteration. In this context, every iteration will have a different training and validation set. The validation accuracies are a measure of what the model has learned from the training set at each iteration and the prediction by the model for the validation set chosen. Therefore, the accuracy varies for each iteration step and the variation pattern is random.

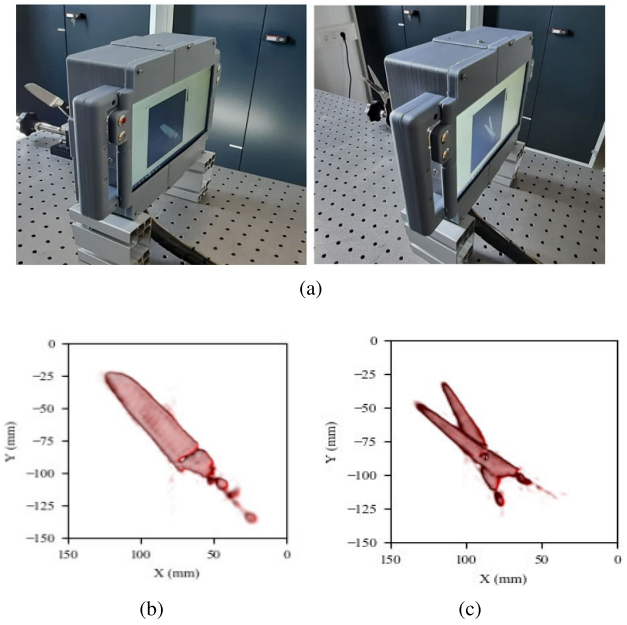


FIGURE 8. The experimental setup involving a handheld scanner imaging a (a) knife (left) and scissor (right) (b) reconstructed mmW image of the knife target (c) reconstructed mmW image of the scissor target.

VI. RESULTS AND DISCUSSION

A. CLASSIFICATION RESULTS

Following the training of the classifier-enabled CI model, two studies are conducted to test the classification performance: numerical and experimental.

For the numerical study, different set of CAD models of the classes used for training (grenades, guns, knives and scissors) are imported into the developed CI model to reconstruct the mmW images. Leveraging data augmentation on the generated mmW dataset, a sufficient number of data are obtained for numerically testing the classification accuracy. A total of 520 mmW images from all the four classes were generated for testing the model.

In addition to the numerical testing to characterize the classification accuracy, in order to demonstrate that the trained classifier model performs accurately with experimental mmW data, mmW images of real threat data are also considered and used as test samples. For this experimental demonstration, a millimeter-wave handheld imager [73] based on a multistatic sparse array [74], [75] is used to image a scissor and a knife, the setup of which is shown in Fig. 8(a).

These two images are also geometrically transformed leveraging the data augmentation process highlighted in Section IV. In addition to increasing the size of the dataset for experimental verification of the classification process, this step also enables us to test whether the model is able to accurately predict the class of an image if it is geometrically altered. Hence, after a series of transformations on the two images from Fig. 8(b) and 8(c), the test dataset is created with these transformed images.

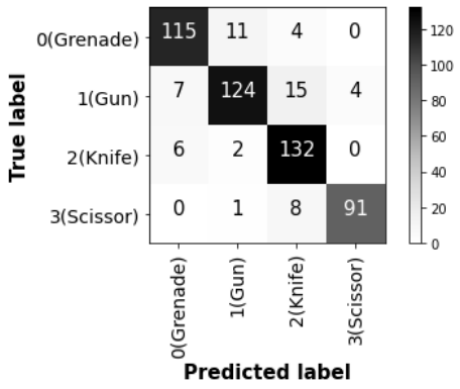
It takes the model approximately 2.05 seconds to predict the classes of all 520 simulated and 22 experimental images, resulting in a classification speed of 264 images per second. Based on the predictions by the system on the test images, performance metrics of the model is determined. Since this being a case of multi-class classification, confusion matrices are generated first for both the studies, numerical and experimental, to record all the predicted labels of the model against the true labels for all the test images. The two generated matrices are shown in Fig. 9.

Analyzing the numerical and experimental confusion matrices in Fig. 9, it is evident that a strong diagonal on both the matrices is obtained. This reflects that the trained CI model has achieved high accuracy on both these studies using synthetic and experimental data.

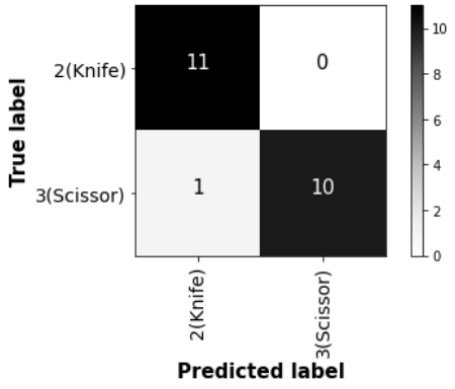
The system is designed so as to have the following abilities:

- The system should have a high accuracy of positive predictions, i.e. the ability of the model not to label an instance positive that is actually negative. In other words, the model should not have the tendency to give any false alarms. To measure this, we look into the 'precision' metric of the model based upon its performance on the test images.
- The system should also identify all the positive cases correctly, i.e. the ability of the model to find all the positive instances. In other words, the model should be able to correctly identify all the threat objects it is detecting. To measure this, we look into the 'recall' metric of the model based upon its performance on the test images.

To record the 'precision' and 'recall' metrics of the system, a classification report is generated. It gives the classification scores of the system for predictions on each of the classes and calculates an average of all the scores. Based on the above



(a)



(b)

FIGURE 9. Confusion matrices for (a) 520 synthesized test images (b) 22 experimental test images.

TABLE 2. Classification report for synthesized test images.

Class	Precision	Recall	f1-score	Support
Grenade	0.90	0.88	0.89	130
Gun	0.90	0.83	0.86	150
Knife	0.83	0.94	0.88	140
Scissor	0.96	0.91	0.93	100
Accuracy			**0.89**	**520**
Macro Avg	0.90	0.89	**0.89**	**520**
Weighted Avg	0.89	0.89	**0.89**	**520**

listed objectives, the system is expected to be a balanced one, i.e., a system that has a balanced precision and recall measures. The f1 score is helpful in this case, since it gives a harmonic mean of the precision and recall scores. The generated classification reports for both the studies are given in Table 2 and 3.

From the classification reports, it can be concluded that the CI-integrated classifier model has classified the test images consisting of both synthesized and experimental data with accuracies of 89% and 95%, respectively. It should be noted here that a major advantage of the developed classifier integrated CI model is that it can synthesize a large number of reconstructed mmW images, and unlike the numerical model, creating a large number of data samples, i.e. mmW

TABLE 3. Classification report for experimental test images.

Class	Precision	Recall	f1-score	Support
Knife	0.92	1.00	0.96	11
Scissor	1.00	0.91	0.95	11
Accuracy			**0.95**	**22**
Macro Avg	0.95	0.95	**0.95**	**22**
Weighted Avg	0.96	0.95	**0.95**	**22**

images, using the experimental scenario is not a feasible option. Therefore, whereas the synthetic mmW images reconstructed using the developed CI model constitutes the main approach to characterize the classification accuracy in this paper, the main reason behind the presented experimental study is to show that the classifier model performs well with experimental data.

At this point, a particular comment can be made regarding the selection of the CNN technique as the enabling method to be integrated with the CI physical model to achieve classification. In this context, other algorithms, such as K-Nearest Neighbors (kNN) [76], [77] and Random Forest [78], [79] techniques, can also be considered ideal for the multi-class classification problem. To this end, the mmW CI classification problem in this paper was also solved using these two algorithms, and the overall classification f1-score accuracies are reported to be 0.86 for the Random Forest algorithm and 0.83 for the kNN algorithm. It is evident that these values are lower than what we obtained using CNN. Moreover, a major reason for performing classification using CNN is because our problem is an image classification one and CNN specializes in processing data in a grid-like topology such as in images. Algorithms such as Random Forest only work on data that are in tabular form, whereas CNN can handle any type of data, regardless of being in a tabular form or as a collection of images. Also, such traditional machine learning algorithms are more suitable for smaller datasets. In our case, we have 2655 images with labels associated with them. Therefore, it is advisable to use neural networks for such a size of dataset. In addition, algorithms such as Random Forest are generally slow in generating predictions because they use multiple decision trees. When making a prediction, all the trees have to make a decision for the same given input and then perform voting. As a result, the more the number of trees used, the slower the prediction process becomes. Hence, real-time decision making is difficult to achieve with such algorithms.

B. PRESENCE OF NOISE AND CLASSIFICATION ACCURACY

As mentioned earlier, to model a realistic system, the reconstruction of the images were carried out by considering a SNR level of 20 dB. In this section, a new study is carried out to observe the effect of system SNR level, i.e., with an increase in noise level in the input images, on the classification accuracy. This an important study when it comes to the practical implications of the system, because the final

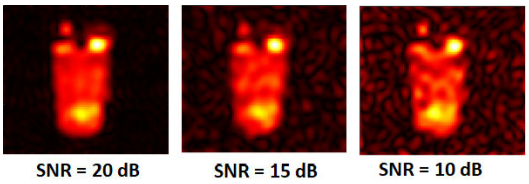
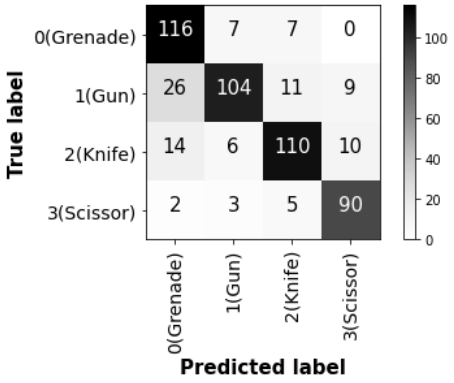
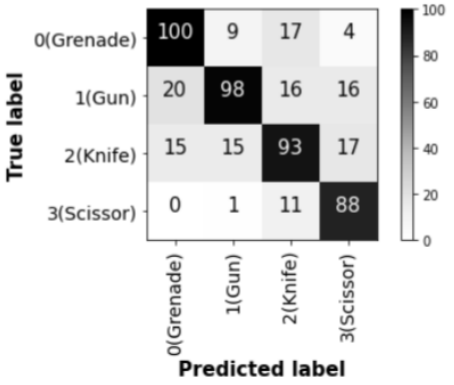


FIGURE 10. The effect of different SNR levels on a synthesized image. For this depiction, the reconstructed image of a grenade CAD model is used.



(a)



(b)

FIGURE 11. Confusion matrices for predictions on test dataset with (a) 15 dB and (b) 10 dB SNR level.

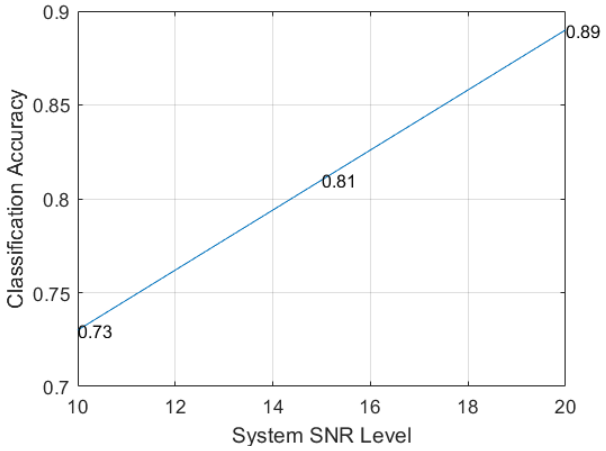


FIGURE 12. The SNR v/s accuracy curve for the classifier.

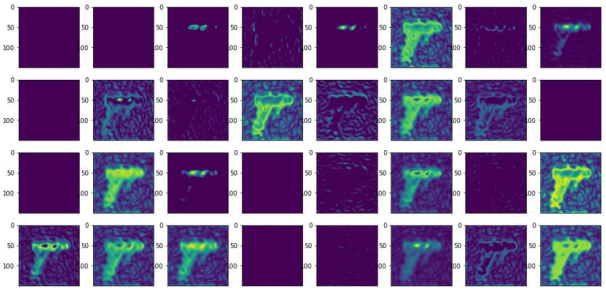
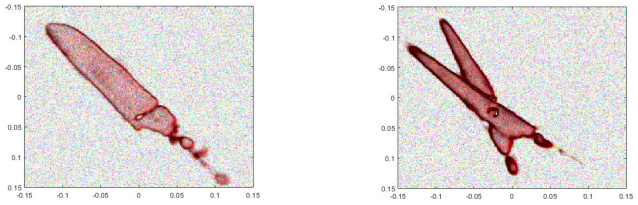


FIGURE 13. The output of the first convolutional layer for the same input image as in Fig. 5 but with 10 dB SNR level. The presence of noise in the outputs makes it difficult for the model to extract high frequency details such as shape of the image.



(a) (b)

FIGURE 14. Experimental images with speckle noise (a) knife target (b) scissor target.

goal of the system is to work with real mmW images and real world mmW images are susceptible to degradation such as motion-blur or speckle noise. Therefore apart from the previous study on 20 dB SNR level, the reconstructions were also carried out with two different SNR levels, 15 dB and 10 dB. The effect of different SNR levels on the input images can be seen in Fig. 10.

Two different test datasets were constructed out of the previous test dataset by considering the two SNR levels. The classification was carried out on both the datasets separately and the performances were recorded.

From the confusion matrices in Fig. 11, the weighted average f1-scores for all the predictions were compared. In case of images with 15 dB SNR level, an accuracy of 81% was observed, whereas for images with 10 dB SNR level, 73% was recorded. From these results, we can conclude that accuracy of image classification is highly affected by noise. The variation of the classification accuracy with the change in SNR level is shown in Fig. 12.

Analyzing the SNR v/s accuracy curve in Fig. 12, it can be concluded that the classification accuracy drops substantially as the system SNR is reduced. Underlying reason behind this behaviour can be seen in Fig. 13, where the output of the first layer is shown in case of an input image, similar to one in Fig. 5, but with 10 dB SNR level. By comparing the qualities of the filter outputs in these two figures, it is evident that the input image in Fig. 13 is distorted, posing a challenge for the model to extract meaningful features from it.

A similar study was conducted on the experimental images as well. These images were introduced to speckle noises as a function of varying the SNR level. The noise effect is shown in Fig. 14.

The same test dataset was used as in the case of the experimental study (Section VI.A.). But in this case, a noise element

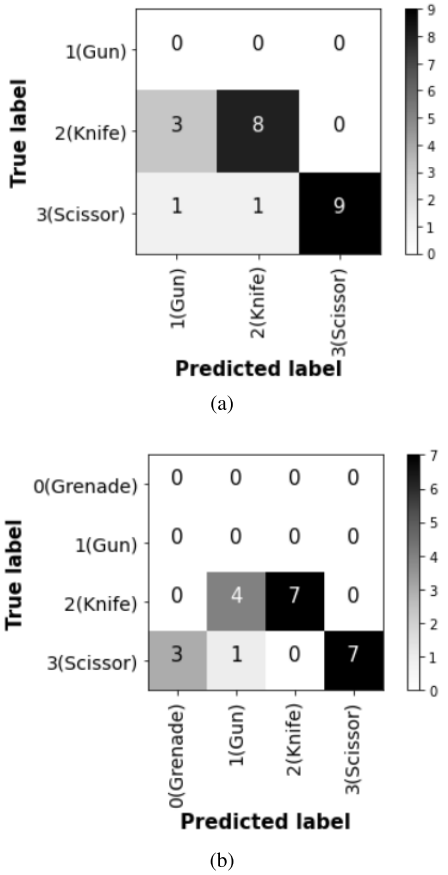


FIGURE 15. Confusion matrices for predictions on experimental test dataset with (a) 15 dB and (b) 10 dB SNR level.

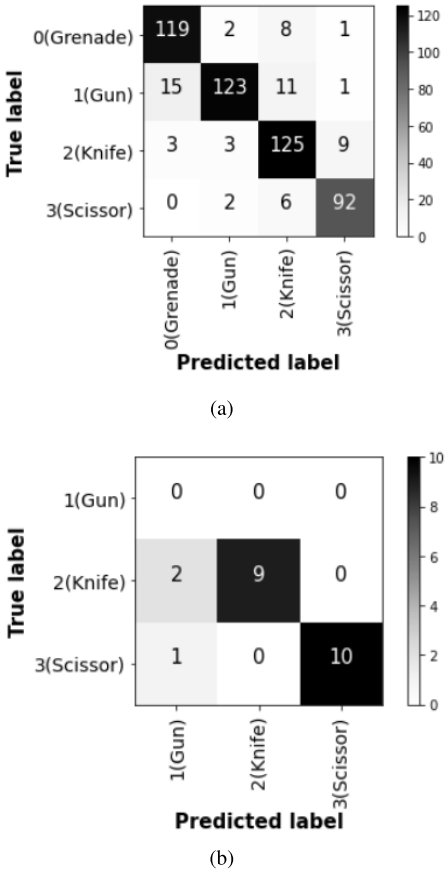


FIGURE 16. Confusion matrices for the predictions on (a) simulated, and (b) experimental noise images. The model was trained with increased noise presence.

was introduced to each image. The predictions were carried out by the model on those images and the confusion matrix is generated as shown in Fig. 15.

Comparing Fig. 9 (b) and Fig. 15, it is evident that the presence of noise has affected the model's ability to predict accurately the classes of input images on multiple occasions. The measured accuracies were recorded to be 85% and 78%, for the two SNR levels, 15 dB and 10 dB, respectively, which is a significant decrease as compared to the previous case. Same conclusion can be drawn that the presence of noise has significantly reduced the classification accuracy.

To find a solution for the decrease in accuracy in case of noisy test images, as a next step, the classification model was trained with images exhibiting a higher noise content. To facilitate this, the developed CI numerical model in Section III was used to generate the synthetic training data with varying SNR levels by adjusting the noise parameter in Eq. 1. The same training parameters such as number of layers in the CNN architecture, learning rate, optimizers, etc. were used as in the previous case. After training, the model was tested with the same noisy test dataset. The following predictions were recorded in both the noisy simulated as well as the experimental images, as shown in Fig. 16.

From the confusion matrices in Fig. 16, the classification accuracies on the simulated and experimental images were recorded to be 88% and 93%, respectively. This proves that the performance of the classifier on images with higher noise content can be improved if the system is trained with images having a similar noise content. This is also a testament to the importance of using the physical model which accurately represents the actual system parameters of a radar rather than using readily available datasets to generate the training data.

VII. FUTURE WORK

Although demonstrated for security-screening applications, particularly in the context of threat object classification, as future work, the proposed system can be adopted across a wide application spectrum. In the context of security-screening, the developed system can be readily integrated with the system developed in [80], wherein a W-band 'spot-light' imager is integrated with a K-band frequency-diverse imager. The K-band imager gives a low resolution image of the entire (FOV), revealing any presence of threat objects. If such an object is detected, its position is introduced to the W-band spotlight imager which extracts high resolution image of the present threat object. This system can be inte-grated with our developed classifier to classify accurately the threat object detected. As mentioned, our classification model predicts the classes of all the test images, 542 in total, in real-time, i.e. 2.05 seconds total (3.8 ms/image,

including the image reconstruction time per frame in this metric). This makes it possible for the model to be used in applications requiring real-time decision making. One such particular application example for this can be automotive radars for debris detection [81]. In [81], a frequency-diverse computational radar system has been developed for detecting debris on roads. This computational radar system can be integrated with the classification scheme developed in this paper. Our developed classification algorithm can be trained with radar images of objects that are found on roads such as other vehicles, traffic lights, road-signs, pedestrian, debris, etc. to achieve classification in the images generated by the automotive radar.

VIII. CONCLUSION

In this paper, we have presented an image classifier system by integrating a CNN layer with a physical model of a mmW CI system. The aperture of the CI system was synthesized using the coded-aperture technique, which illuminates the imaged scene with quasi-random spatio-temporally independent radiation patterns (or bases) and encodes the scene information into a set of compressed measurements. The system uses two channels (transmit and receive) to compress the back-scattered measurements and reconstruct the mmW image of the imaged target. The reconstructed images from the CI model were used to train the integrated CNN-based classifier algorithm. The algorithm was finally tested with both synthesized images generated using the developed numerical model as well as with mmW images of real objects reconstructed using a mmW handheld imager. The predictions record a high accuracy as shown in Tables 2 and 3. The obtained results validated that the integrated CNN algorithm, trained with synthesized images from the developed CI physical model, can accurately classify images of threat objects. This system has the potential to be used in security-screening applications, wherein different threat objects present in reconstructed mmW images can be detected and classified into their appropriate classes.

Since it is known that real-world images are susceptible to noise, the effect of presence of measurement noise in input images on the classification accuracy was also studied. For this analysis, no changes were made to the training dataset, but the test images were subjected to noise. It was observed that the accuracy decreases as the noise level increases. Hence, from these findings, it is recommended that in order to perform classification on noisy mmW images, the model must be trained with higher noise images. We justified this outcome by doing another study wherein the model was trained with images having a higher presence of noise. The performances of this model on the noisy test dataset clearly suggest that the accuracy improves when trained with images of higher noise content. It is worth mentioning that achieving training using the synthesized mmW images from the developed CI physical model made it possible to carry out this systematic analysis and understand the system classification performance as a function of varying SNR levels.

REFERENCES

[1] Q. Li, K. Chen, W. Guo, L. Lang, F. He, L. Chen, and Z. Xiong, "An aperture synthesis radiometer at millimeter wave band," in *Proc. Int. Conf. Microw. Millim. Wave Technol. (ICMMT)*, vol. 4, Apr. 2008, pp. 1699–1701.

[2] V. I. Piddyachiy, V. M. Shulga, V. V. Myshenko, A. M. Korolev, A. V. Myshenko, and A. V. Antyufeyev, "Ground-based 3 mm-wave radiometer for spectroscopic observations of atmospheric ozone and carbon monoxide," in *Proc. Int. KHARKOV Symp. Phys. Eng. Microw., Millim. (SUBMILLIMETER WAVES)*, Jun. 2010, pp. 1–3.

[3] S. H. Shehab, J. Feng, and N. Karmakar, "Trends on remote sensing technology: Receiver architectures and antenna systems," in *Proc. Int. Conf. Robot., Elect. Signal Process. Techn. (ICREST)*, Jan. 2019, pp. 227–232.

[4] I. Bilik, O. Longman, S. Villeval, and J. Tabrikian, "The rise of radar for autonomous vehicles: Signal processing solutions and future research directions," *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 20–31, Sep. 2019.

[5] J. B. Detlefsen, "Industrial applications of microwave imaging," in *Proc. 21st Eur. Microw. Conf.*, vol. 1, Oct. 1991, pp. 108–119.

[6] K. Sarabandi, M. Vahidpour, M. Moallem, and J. East, "Compact beam scanning 240 GHz radar for navigation and collision avoidance," in *Proc. 3rd Micro-Nanotechnol. Sensors, Syst., Appl.*, vol. 8031, May 2011, Art. no. 803113.

[7] M. T. Bevacqua, S. D. Meo, L. Crocco, T. Isernia, and M. Pasian, "Millimeter-waves breast cancer imaging via inverse scattering techniques," *IEEE J. Electromagn., RF Microw. Med. Biol.*, vol. 5, no. 3, pp. 246–253, Sep. 2021.

[8] S. D. Meo, G. Matrone, and M. Pasian, "Experimental validation on tissue-mimicking phantoms of millimeter-wave imaging for breast cancer detection," *Appl. Sci.*, vol. 11, no. 1, pp. 1–17, 2021.

[9] D. M. Sheen, D. L. McMakin, and T. E. Hall, "Detection of explosives by millimeter-wave imaging," in *Counterterrorist Detection Techniques of Explosives*, J. Yinon, Ed. Amsterdam, The Netherlands: Elsevier, 2007, ch. 9, pp. 237–277.

[10] S. Demirci, H. Cetinkaya, E. Yigit, C. Ozdemir, and A. A. Vertiy, "A study on millimeter-wave imaging of concealed objects: Application using back-projection algorithm," *Prog. Electromagn. Res.*, vol. 128, pp. 457–477, Jun. 2012.

[11] J. A. Martínez-Lorenzo, F. Quivira, and C. M. Rappaport, "SAR imaging of suicide bombers wearing concealed explosive threats," *Prog. Electromagn. Res.*, vol. 125, pp. 255–272, Feb. 2012.

[12] D. M. Sheen, D. L. McMakin, and T. E. Hall, "Three-dimensional millimeter-wave imaging for concealed weapon detection," *IEEE Trans. Microw. Theory Techn.*, vol. 49, no. 9, pp. 1581–1592, Sep. 2001.

[13] D. Ma, Q. Zhang, A. K. Rasahid, and X. Tang, "A low-cost two-dimensional near-field imaging system based on a Goubau line," in *Proc. Int. Conf. Microw. Millim. Wave Technol. (ICMMT)*, May 2019, pp. 1–3.

[14] W. Zhang, A. Molaei, J. Heredia-Juesas, L. Tirado, K. Graham, A. Bisulco, H. Gomez-Sousa, and J. A. Martinez-Lorenzo, "Experimental results of a 3-D millimeter-wave compressive reflector antenna imaging system," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 1, pp. 24–28, Jan. 2019.

[15] J. Lyu, D. Bi, X. Li, Y. Xie, and X. Xie, "Super-resolution image reconstruction of compressive 2D near-field millimetre-wave," *Electron. Lett.*, vol. 56, no. 19, pp. 978–980, Sep. 2020.

[16] T. Omori, Y. Isono, K. Kondo, Y. Akamine, and S. Kidera, "K-space decomposition based super-resolution three-dimensional imaging method for millimeter wave radar," in *Proc. IEEE Radar Conf. (RadarConf)*, Sep. 2020, pp. 1–6.

[17] K. V. Mishra and L. H. Nguyen, "Image enhancement with blind deconvolution in millimeter-wave 3-D FLoSAR," in *Proc. IEEE Radar Conf. (RadarConf)*, Sep. 2020, pp. 1–5.

[18] S. Nagayama, S. Muramatsu, H. Yamada, and Y. Sugiyama, "Millimeter wave radar image denoising with complex nonseparable oversampled lapped transform," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Dec. 2017, pp. 1824–1829.

[19] C. Wang, J. Shi, Z. Zhou, L. Li, Y. Zhou, and X. Yang, "Concealed object detection for millimeter-wave images with normalized accumulation map," *IEEE Sensors J.*, vol. 21, no. 5, pp. 6468–6475, Mar. 2021.

[20] K. Yu, X. Qi, T. Sato, S. H. Myint, Z. Wen, Y. Katsuyama, K. Tokuda, W. Kameyama, and T. Sato, "Design and performance evaluation of an AI-based W-band suspicious object detection system for moving persons in the IoT paradigm," *IEEE Access*, vol. 8, pp. 81378–81393, 2020.

[21] F. Qi, I. Ocket, D. Schreurs, and B. Nauwelaers, "A system-level simulator for indoor mmW SAR imaging and its applications," *Opt. Exp.*, vol. 20, no. 21, pp. 23811–23820, 2012.
[22] J. Laviada, A. Arboleya-Arboleya, Y. Alvarez-Lopez, C. Garcia-Gonzalez, and F. Las-Heras, "Phaseless synthetic aperture radar with efficient sampling for broadband near-field imaging: Theory and validation," *IEEE Trans. Antennas Propag.*, vol. 63, no. 2, pp. 573–584, Feb. 2015.
[23] G. L. Charvat, L. C. Kempel, E. J. Rothwell, C. M. Coleman, and E. L. Mokole, "An ultrawideband (UWB) switched-antenna-array radar imaging system," in *Proc. IEEE Int. Symp. Phased Array Syst. Technol.*, Oct. 2010, pp. 543–550.
[24] S. Withington, G. Saklatvala, and M. P. Hobson, "Partially coherent analysis of imaging and interferometric phased arrays: Noise, correlations, and fluctuations," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 23, no. 6, pp. 1340–1348, 2006.
[25] G. Barbastathis, "On the use of deep learning for computational imaging," in *Proc. 17th Opt. Trapping Opt. Micromanipulation*, vol. 11463, Sep. 2020, Art. no. 114631L.
[26] T. Fromenteze, E. L. Kpré, D. Carsenat, C. Decroze, and T. Sakamoto, "Single-shot compressive multiple-inputs multiple-outputs radar imaging using a two-port passive device," *IEEE Access*, vol. 4, pp. 1050–1060, 2016.
[27] J. Hunt, T. Driscoll, A. Mrozack, G. Lipworth, M. Reynolds, D. Brady, and D. Smith, "Metamaterial apertures for computational imaging," *Science*, vol. 339, no. 6117, pp. 310–313, Jan. 2013.
[28] A. Molaei, J. Heredia-Juesas, G. Ghazi, J. Vlahakis, and J. A. Martinez-Lorenzo, "Digitized metamaterial absorber-based compressive reflector antenna for high sensing capacity imaging," *IEEE Access*, vol. 7, pp. 1160–1173, 2019.
[29] O. Yurduseven, V. R. Gowda, J. N. Gollub, and D. R. Smith, "Printed aperiodic cavity for computational and microwave imaging," *IEEE Microw. Wireless Compon. Lett.*, vol. 26, no. 5, pp. 367–369, May 2016.
[30] J. N. Gollub, O. Yurduseven, K. P. Trofatter, D. Arnitz, M. F. Imani, T. Sleasman, M. Boyarsky, A. Rose, A. Pedross-Engel, H. Odabasi, and T. Zvolensky, "Large metasurface aperture for millimeter wave computational imaging at the human-scale," *Sci. Rep.*, vol. 7, pp. 1–9, Feb. 2017, Art. no. 42650.
[31] O. Yurduseven, J. N. Gollub, D. L. Marks, and D. R. Smith, "Frequency-diverse microwave imaging using planar mills-cross cavity apertures," *Opt. Exp.*, vol. 24, no. 8, pp. 8907–8925, 2016.
[32] O. Yurduseven, V. R. Gowda, J. N. Gollub, and D. R. Smith, "Multistatic microwave imaging with arrays of planar cavities," *IET Microw., Antennas Propag.*, vol. 10, no. 11, pp. 1174–1181, Aug. 2016.
[33] O. Yurduseven, M. Imani, H. Odabasi, J. Gollub, G. Lipworth, A. Rose, and D. Smith, "Resolution of the frequency diverse metamaterial aperture imager," *Prog. Electromagn. Res.*, vol. 150, pp. 97–107, Jan. 2015.
[34] R. Sharma, O. Yurduseven, B. Deka, and V. Fusco, "Hardware enabled acceleration of near-field coded aperture radar physical model for millimetre-wave computational imaging," *Prog. Electromagn. Res. B*, vol. 90, pp. 91–108, Jan. 2021.
[35] O. Yurduseven, J. N. Gollub, A. Rose, D. L. Marks, and D. R. Smith, "Design and simulation of a frequency-diverse aperture for imaging of human-scale targets," *IEEE Access*, vol. 4, pp. 5436–5451, 2016.
[36] R. Maree, P. Geurts, J. Piater, and L. Wehenkel, "Random subwindows for robust image classification," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Jun. 2005, pp. 34–40.
[37] S. Loussaief and A. Abdelkrim, "Machine learning framework for image classification," in *Proc. 7th Int. Conf. Sci. Electron., Technol. Inf. Telecommun. (SETIT)*, Dec. 2016, pp. 58–61.
[38] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Dec. 2001, p. 1.
[39] Z. Sun, G. Bebis, and R. Miller, "Object detection using feature subset selection," *Pattern Recognit.*, vol. 37, no. 11, pp. 2165–2176, Nov. 2004.
[40] C. Kwan, B. Chou, J. Yang, A. Rangamani, T. Tran, J. Zhang, and R. Etienne-Cummings, "Deep learning-based target tracking and classification for low quality videos using coded aperture cameras," *Sensors*, vol. 19, no. 17, p. 3702, Aug. 2019.
[41] C. Kwan, D. Gribben, B. Chou, B. Budavari, J. Larkin, A. Rangamani, T. Tran, J. Zhang, and R. Etienne-Cummings, "Real-time and deep learning based vehicle detection and classification using pixel-wise code exposure measurements," *Electronics*, vol. 9, no. 6, p. 1014, Jun. 2020.
[42] C. Kwan, D. Gribben, A. Rangamani, T. Tran, J. Zhang, and R. Etienne-Cummings, "Detection and confirmation of multiple human targets using pixel-wise code aperture measurements," *J. Imag.*, vol. 6, no. 6, p. 40, 2020.
[43] C. Kwan, B. Chou, J. Yang, A. Rangamani, T. Tran, J. Zhang, and R. Etienne-Cummings, "Target tracking and classification using compressive measurements of MWIR and LWIR coded aperture cameras," *J. Signal Inf. Process.*, vol. 10, no. 3, pp. 73–95, 2019.
[44] A. Shayei, M. Abbasi, A. Habiban, M. Shabany, and Z. Kavehvash, "A machine learning approach for material classification in MMW imaging systems based on frequency spectra," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2018, pp. 1–5.
[45] M. Miškuf and I. Zolotová, "Comparison between multi-class classifiers and deep learning with focus on industry 4.0," in *Proc. Cybern. Informat. (K I)*, Feb. 2016, pp. 1–5.
[46] S. Villon, M. Chaumont, G. Subsol, S. Villéger, T. Claverie, and D. Mouillot, "Coral reef fish detection and recognition in underwater videos by supervised machine learning: Comparison between deep learning and HOG+SVM methods," in *Proc. Int. Conf. Adv. Concepts Intell. Vis. Syst.* Cham, Switzerland: Springer, 2016, pp. 160–171.
[47] S. Lopez-Tapia, R. Molina, and N. P. de la Blanca, "Deep CNNs for object detection using passive millimeter sensors," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 9, pp. 2580–2589, Sep. 2019.
[48] J. Yuan and C. Guo, "A deep learning method for detection of dangerous equipment," in *Proc. 8th Int. Conf. Inf. Sci. Technol. (ICIST)*, Jun. 2018, pp. 159–164.
[49] J. Zhang, W. Xing, M. Xing, and G. Sun, "Terahertz image detection with the improved faster region-based convolutional neural network," *Sensors*, vol. 18, no. 7, p. 2327, Jul. 2018.
[50] X. Yang, T. Wu, L. Zhang, D. Yang, N. Wang, B. Song, and X. Gao, "CNN with spatio-temporal information for fast suspicious object detection and recognition in THz security images," *Signal Process.*, vol. 160, pp. 202–214, Jul. 2019.
[51] Z. Meng, M. Zhang, and H. Wang, "CNN with pose segmentation for suspicious object detection in MMW security images," *Sensors*, vol. 20, no. 17, p. 4974, Sep. 2020.
[52] S. Chen, H. Wang, F. Xu, and Y. Q. Jin, "Target classification using the deep convolutional networks for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4806–4817, Aug. 2016.
[53] M. Liu, L. Peng, X. Liu, L. Dong, M. Hui, and Y. Zhao, "SAR image classification based on CNN in real and simulation datasets," in *Proc. 9th Int. Conf. Graphic Image Process. (ICGIP)*, H. Yu and J. Dong, Eds., Apr. 2018, pp. 820–827.
[54] F. Gan, C. Luo, X. Liu, H. Wang, and L. Peng, "Fast terahertz coded-aperture imaging based on convolutional neural network," *Appl. Sci.*, vol. 10, no. 8, p. 2661, Apr. 2020.
[55] W. Chi and N. George, "Phase-coded aperture for optical imaging," *Opt. Commun.*, vol. 282, no. 11, pp. 2110–2117, Jun. 2009.
[56] M. L. Don, C. Fu, and G. R. Arce, "Compressive imaging via a rotating coded aperture," *Appl. Opt.*, vol. 56, no. 3, p. B142, 2017.
[57] T. Sleasman, M. F. Imani, J. N. Gollub, and D. R. Smith, "Dynamic metamaterial aperture for microwave imaging," *Appl. Phys. Lett.*, vol. 107, no. 20, Nov. 2015, Art. no. 204104.
[58] C. Watts, D. Shrekenhamer, J. Montoya, G. Lipworth, J. Hunt, T. Sleasman, S. Krishna, D. Smith, and W. Padilla, "Terahertz compressive imaging with metamaterial spatial light modulators," *Nature Photon.*, vol. 8, no. 8, pp. 605–609, 2014.
[59] N. Rusk, "Deep learning," *Nature Methods*, vol. 13, no. 1, p. 35, Dec. 2015.
[60] B. M. Wilamowski and J. Korniak, "Learning architectures with enhanced capabilities and easier training," in *Proc. IEEE 19th Int. Conf. Intell. Eng. Syst. (INES)*, Sep. 2015, pp. 21–29.
[61] W. Sun and F. Su, "Regularization of deep neural networks using a novel companion objective function," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 2865–2869.
[62] V. Nair and G. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. 27th Int. Conf. Mach. Learn. (ICML)*, 2010, pp. 807–814.
[63] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017, *arXiv:1412.6980*. [Online]. Available: https://arxiv.org/abs/1412.6980
[64] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1026–1034.

[65] M. Zeiler and R. Fergus, *Visualizing and Understanding Convolutional Networks* (Lecture Notes in Computer Science), vol. 8689. Basel, Switzerland: Springer, 2014.
[66] Y. Bengio, *Practical Recommendations for Gradient-Based Training Deep Architectures* (Lecture Notes in Computer Science), vol. 7700. 2012.
[67] C. Capraro, U. Majumder, J. Siddall, E. Davis, D. Brown, and C. Cicotta, "SAR object classification implementation for embedded platforms," in *Proc. SPIE Int. Soc. Opt. Eng.*, vol. 10987, 2019, Art. no. 109870F.
[68] E. P. Leite and C. R. de Souza Filho, "TEXTNN—A MATLAB program for textural classification using neural networks," *Comput. Geosci.*, vol. 35, no. 10, pp. 2084–2094, Oct. 2009.
[69] T. Oktay, S. Arik, I. Turkmen, M. Uzun, and H. Celik, "Neural network based redesign of morphing UAV for simultaneous improvement of roll stability and maximum lift/drag ratio," *Aircr. Eng. Aerosp. Technol.*, vol. 90, no. 8, pp. 1203–1212, Nov. 2018.
[70] M. Soylak, T. Oktay, and İ. Turkmen, "A simulation-based method using artificial neural networks for solving the inverse kinematic problem of articulated robots," *Proc. Inst. Mech. Eng. E, J. Process Mech. Eng.*, vol. 231, no. 3, pp. 470–479, Jun. 2017.
[71] D. Acharya, R. Jain, S. Panigrahi, R. Sahni, S. Jain, S. Deshmukh, and A. Bhardwaj, *Multi-Class Emotion Classification Using EEG Signals*, vol. 1367. Singapore: Springer, 2021.
[72] R. Hazra, M. Banerjee, and L. Badia, "Machine learning for breast cancer classification with ANN and decision tree," in *Proc. 11th IEEE Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON)*, Nov. 2020, pp. 522–527.
[73] R. Hussung, A. Keil, and F. Friederich, "Handheld millimeter wave imaging system based on a two-dimensional multistatic sparse array," in *Proc. 45th Int. Conf. Infr., Millim., THz Waves (IRMMW-THz)*, Nov. 2020, pp. 1–2.
[74] B. Baccouche, P. Agostini, S. Mohammadzadeh, M. Kahl, C. Weisenstein, J. Jonuscheit, A. Keil, T. Löffler, W. Sauer-Greff, R. Urbansky, P. H. Bolivar, and F. Friederich, "Three-dimensional terahertz imaging with sparse multistatic line arrays," *IEEE J. Sel. Topics Quantum Electron.*, vol. 23, no. 4, pp. 1–11, Jul. 2017.
[75] S. S. Ahmed, "Microwave imaging in security—Two decades of innovation," *IEEE J. Microw.*, vol. 1, no. 1, pp. 191–201, Jan. 2021.
[76] H. Ferdous, T. Siraj, S. J. Setu, M. M. Anwar, and M. A. Rahman, "Machine learning approach towards satellite image classification," in *Proc. Int. Conf. Trends Comput. Cognit. Eng.*, M. S. Kaiser, A. Bandyopadhyay, M. Mahmud, and K. Ray, Eds. Singapore: Springer, 2021, pp. 627–637.
[77] J. Jaruenpunyasak and R. Duangsoithong, "Empirical analysis of feature reduction in deep learning and conventional methods for foot image classification," *IEEE Access*, vol. 9, pp. 53133–53145, 2021.
[78] M. J. Sagayaraj, V. Jithesh, and D. Roshani, "Comparative study between deep learning techniques and random forest approach for HRRP based radar target classification," in *Proc. Int. Conf. Artif. Intell. Smart Syst. (ICAIS)*, Mar. 2021, pp. 385–388.
[79] Y. Mo, R. Zhong, and S. Cao, "Orbita hyperspectral satellite image for land cover classification using random forest classifier," *J. Appl. Remote Sens.*, vol. 15, no. 1, pp. 1–20, Mar. 2021.
[80] O. Yurduseven, D. L. Marks, T. Fromenteze, J. N. Gollub, and D. R. Smith, "Millimeter-wave spotlight imager using dynamic holographic metasurface antennas," *Opt. Exp.*, vol. 25, no. 15, pp. 18230–18249, Jul. 2017.
[81] O. Yurduseven, T. Fromenteze, C. Decroze, and V. F. Fusco, "Frequency-diverse computational automotive radar technique for debris detection," *IEEE Sensors J.*, vol. 20, no. 22, pp. 13167–13177, Nov. 2020.

RAHUL SHARMA received the Bachelor in Technology (B.Tech.) degree in electronics and communication engineering from Assam Don Bosco University, Guwahati, India, in 2012, and the Master in Technology (M.Tech.) degree in electronics design and technology from the Department of Electronics and Communication Engineering, Tezpur University, Tezpur, India, in 2015. He is currently pursuing the Ph.D. degree with the Institute of Electronics, Communication and Information Technology, Queen's University Belfast, U.K.

His current research interests include computational imaging, machine learning for image analysis and improvement, hardware architectures for image processing acceleration, and imaging radars in autonomous vehicles.

RAPHAEL HUSSUNG received the Bachelor of Science (B.Sc.) and Master of Science (M.Sc.) degrees in electrical engineering and computer engineering from TU Kaiserslautern, Kaiserslautern, Germany, in 2018 and 2020, respectively. During that time, he worked part-time with Pollux Electromechanical Systems GmbH, Kirchheimbolanden, Germany, a supplier of conditioning systems for the automotive industry, where he was tasked with hard- and software development. He joined Fraunhofer Institute for Industrial Mathematics, Kaiserslautern, as a Research Associate, in July 2020. His research interest includes millimeter wave and terahertz imaging with multistatic sensor arrays for non-destructive testing applications.

ANDREAS KEIL (Member, IEEE) was born in Kiel, Germany, in 1975. He graduated in physics from Christian Albrechts University, Kiel, in 2003. He received the Ph.D. degree in quantum information from the National University of Singapore, Singapore, in 2010.

From 2009 to 2013, he was involved in the research and development of terahertz imaging systems with SynView GmbH, Bad Homburg, Germany. Since 2013, he has been with Becker Photonik GmbH, Porta Westfalica, Germany, where he continues this line of work. Since 2018, he has been with Fraunhofer ITWM. His research interests include terahertz imaging, image processing algorithms, and high-performance computation on general purpose graphics processing units.

FABIAN FRIEDERICH graduated in optoelectronics from Aalen University Sciences, Germany, in 2007. He received the Ph.D. degree in physics from Goethe University Frankfurt am Main, Germany, for his work in the field of terahertz imaging with the Ultrafast Spectroscopy and Terahertz Physics Group, in 2012. During his studies, he was also associated with the Centre for Micro-Photonics, Swinburne University of Technology, Melbourne, VIC, Australia, and with Laser Zentrum Hannover, Germany. He joined the Institute of Technical Physics, German Aerospace Center (DLR), Stuttgart, Germany, in 2011, to establish laser-based concepts for monitoring space debris. He was granted a Fraunhofer Attract Funding to form a new research group in the field of millimetre-wave and terahertz measurement techniques at Fraunhofer Institute for Physical Measurement Techniques IPM, in 2013. Since 2017, his group has been with Fraunhofer Institute for Industrial Mathematics ITWM, Kaiserslautern, pursuing its millimetre-wave and terahertz activities with an even stronger focus on signal and image processing in the field of nondestructive testing. He is a member of the Program Committee of the Terahertz, RF, Millimetre, and Submillimetre-wave Technology and Applications Conference at SPIE Photonics West. He is an Active Member of the Microwave and Terahertz Methods Technical Committee of the German Society for Non-Destructive Testing and the VDI/VDE-GMA Technical Committee 8.17 Terahertz Systems of the Association of German Engineers (VDI). He is a Topical Editor of the *Sensors* journal (MDPI).

THOMAS FROMENTEZE received the Ph.D. degree from the University of Limoges, Limoges, France, in 2015. From 2015 to 2016, he was a Postdoctoral Researcher with Duke University, Durham, NC, USA. He is currently a Maître de Conférences (Assistant Professor) with the Xlim Research Institute, University of Limoges. He is also an Adjunct Assistant Professor with the Center for Metamaterials and Integrated Plasmonics, Duke University. His research interests include ultra-wideband microwave and millimetre-wave imaging, wave propagation in complex media, computational/compressive imaging, and the various associated inverse problems. He received the 11th EuRAD Young Engineer Prize during the European Microwave Week 2015.

MOHSEN KHALILY (Senior Member, IEEE) is currently a Lecturer in antenna and propagation with the Institute for Communication Systems (ICS), Home of the 5G and 6G Innovation Centres (5GIC and 6GIC), University of Surrey, U.K., where he was a Research Fellow on antennas and propagation from December 2015 to March 2019. Prior to joining the 5GIC, he was a Senior Lecturer with the Wireless Communication Centre (WCC), University Technology Malaysia (UTM). His research interests include surface electromagnetic, reconfigurable reflecting surface, 5G system, phased arrays, hybrid beamforming, and mm-wave and terahertz antennas and propagation. He has published almost 150 academic articles in international peer-reviewed journals and conference proceedings and has been the Principal Investigator on research grants totaling in excess of £1 M in these fields. He is also a fellow of the U.K. Higher Education Academy; a member of the IEEE Antennas and Propagation Society, the IEEE Communication Society, and the IEEE Microwave Theory and Techniques Society; and an Associate Editor of the IEEE Access. He was a Lead Guest Editor in several journals, including IEEE Antennas and Wireless Propagation Letters and IEEE Open Journal of Antennas and Propagation.

BHABESH DEKA (Senior Member, IEEE) received the Bachelor in Engineering (B.E.) degree in electronics and telecommunication engineering from Assam Engineering College, Guwahati, India, in 1999, the Master in Technology (M.Tech) degree from Tezpur University, Tezpur, India, in 2001, and the Ph.D. degree in electronics and electrical engineering from the Indian Institute of Technology, Guwahati, in 2011.

He is currently a Professor with the Department of Electronics and Communication Engineering, Tezpur University. He leads the Computer Vision and Image Processing (CVIP) Laboratory, Department of ECE, Tezpur University. He is also a Principal Investigator of two major research projects sponsored by All India Council for Technical Education (AICTE) and Indian Space Research Organisation (ISRO), Government of India. His research interests include image processing, particularly inverse ill-posed problems, computer vision, compressive sensing MRI, and machine learning for biomedical signal/image analysis. He is a fellow of the Institution of Electronics and Telecommunication Engineers (IETE) and a Life Member of the Indian Unit of Pattern Recognition and Artificial Intelligence (IUPRAI) affiliated to the International Association for Pattern Recognition (IAPR).

VINCENT FUSCO (Fellow, IEEE) received the bachelor's degree (Hons.) in electrical and electronic engineering, the Ph.D. degree in microwave electronics, and the D.Sc. degree from Queen's University Belfast (QUB), Belfast, U.K., in 1979, 1982, and 2000, respectively. He is currently the Director of Research with the ECIT Research Institute, QUB. His fundamental work on active antenna front-end techniques has provided generic advances in low-cost phased and self-tracking antenna array architectures. He has authored or coauthored more than 500 articles and two books. He holds a number of antenna-related patents. He is a fellow of the Institution of Engineering and Technology (IET) and the Royal Academy of Engineering. He was a recipient of the IET Senior Achievement Award and the Mountbatten Medal for seminal contributions in the field of microwave electronics and its impact on U.K. industry, in 2012, and the 2019 Royal Irish Academy Gold medal for Engineering Science.

OKAN YURDUSEVEN (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from Yildiz Technical University, Istanbul, Turkey, in 2009 and 2011, respectively, and the Ph.D. degree in electrical engineering from Northumbria University, Newcastle upon Tyne, U.K., in 2014. From 2014 to 2018, he was a Postdoctoral Research Associate with the Department of Electrical and Computer Engineering, Duke University, USA. From 2018 to 2019, he was a NASA Research Fellow with the Jet Propulsion Laboratory, California Institute of Technology, USA. He is currently a Senior Lecturer (Associate Professor) with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, U.K. He is also an Adjunct Professor with Duke University. He has authored more than 140 peer-reviewed technical journal and conference papers, and has been the Principal Investigator on research grants totaling in excess of £1.3m in these fields. His research interests include microwave and millimeter-wave imaging, multiple-input-multiple-output (MIMO) radars, wireless power transfer, antennas and propagation, and metamaterials. Since 2020, he has been serving as a Technical Program Committee Member in SPIE Defense and Commercial Sensing Conference. He is a member of the European Association on Antennas and Propagation (EurAAP). He was a recipient of the NASA Postdoctoral Program Fellowship administrated by Universities Space Research Association (USRA), in 2018. He received the Outstanding Postdoctoral Award from Duke University, in 2017. In 2019, in collaboration with the University of Limoges, France, he received the Alliance Hubert Curien Award funded by the British Council. In 2020, he was bestowed the Leverhulme Trust Research Leadership Award. He was a Guest Editor in several journals, including IEEE Antennas and Wireless Propagation Letters, IEEE Open Journal of Antennas and Propagation, and *Remote Sensing* (MDPI). He serves as an Associate Editor for the IEEE Antennas and Wireless Propagation Letters.