Performance Analysis of Classification Algorithms for Millimeter-Wave Imaging

Performance Analysis of Classification Algorithms for Millimeter-wave Imaging

Rahul Sharma[1], Raphael Hussung[2], Andreas Keil[2], Fabian Friederich[2],
Thomas Fromenteze[3], Mohsen Khalily[4], Bhabesh Deka[5], Vincent Fusco[1], Okan Yurduseven[1]

[1]ECIT, Queen's University Belfast, Belfast BT3 9DT, UK, email: rsharma05@qub.ac.uk
[2]Department of Material Characterization and Testing,
Fraunhofer Institute for Industrial Mathematics ITWM, Kaiserslautern 67663, Germany
[3]University of Limoges, XLIM, UMR 7252, F- 87000, Limoges, France
[4]Institute for Communication Systems, University of Surrey, Guildford GU2 7XH, UK
[5]Department of Electronics and Communication Engineering, Tezpur University, Tezpur, Indi, 784028.

***Abstract*—A detailed analysis of the performance of traditional machine learning and deep learning techniques applied on a representative classification problem of millimeter-wave (mmW) images is presented in this paper. The algorithms chosen for this analysis are the k-Nearest Neighbors (KNN), Random Forest (RF) and Convolutional Neural Network (CNN). All algorithms presented here are modeled using 'keras' library inside TensorFlow and 'scikit-learn' module. The dataset for training and testing are generated via a developed near-field coded aperture computational imaging (CI) physical model. The use of a physical model of an imaging system that implements CI techniques instead of an experimental set-up makes the whole dataset generation process facile and less time consuming. The training data, in case of the RF and KNN algorithms, are presented in tabular form whereas for the CNN technique, the synthesized images from the physical model itself are used for training. The models are tested with both synthesized as well as experimental data, generated from the physical model and a mmW handheld imager, respectively. Upon testing, it is observed that the KNN and RF algorithms are able to classify the test samples with accuracies of 82% and 87%, respectively, whereas an accuracy of 90% is observed in case of the CNN classifier. Also, an inference speed test is conducted on all the three algorithms. It was observed that CNN is the fastest to predict classes for all of the test samples with a frame rate of 3.8 ms/sample whereas RF is the slowest, with a frame rate of 65.9 ms/sample. These findings establish the fact that when it comes to image classification, CNN based classifiers perform better than any traditional machine learning algorithms with more accurate and faster predictions, paving the way for various real-time applications such as automatic threat detection.**

***Index Terms*—Convolutional Neural Networks, K-nearest neighbours, random forest, classifiers, Image classification, computational imaging, millimeter-waves.**

I. Introduction

Millimeter-waves (mmW) for security screening applications are widely used nowadays. Because of their non-ionizing affects, mmW do not pose any health hazards, unlike X-rays. As mmW can penetrate clothing materials, any concealed threat objects carried on personnel can be detected. This study is focused on such a system wherein it has the ability to detect and prompt the user about the type of threat object it is detecting in real-time, hence facilitating real-time automated threat detection (ATD).

To build such a system, the development of the classification or the object detection algorithm is of utmost important. There are previous studies in the literature on the use of different machine learning algorithms and deep neural networks to achieve image classification. In [1], target classification of mmW radar data for autonomous driving is shown using neural networks such as Artificial Neural Network (ANN) and Convolutional Neural Network (CNN) with good classification accuracies. In [2], a binary classification problem is presented wherein a wavelet-based Support Vector Machine (SVM) is used to detect deformations in ceramic tiles inside concealed packaging from images of a mmW radar system. Recently, in [3], an anomaly detection model on terahertz frequency modulated continuous wave (FMCW) images based on a multivariate Gaussian distribution has been presented. Studies have also been conducted on the use of a region based CNN (RCNN) for concealed weapon detection in the human body, as shown in [4]. These studies are sufficient to establish the fact that machine learning algorithms and deep learning networks can successfully be employed for object detection applications. It is worth mentioning that all these mmW radar image classification studies have been shown on readily available datasets or on data which were collected through experimental means. When it comes to mmW radar imaging, there are traditional imaging techniques such as SAR-based as well as phased array based methods [5]–[8]. These methods, although producing very good quality reconstructions, rely on a raster scanning of the imaging scene, resulting in a slow data acquisition process. Also, they use very complex hardware settings and are power consuming. Such shortcomings can prove to be a hindrance while generating the data for training. This is because, given the quality of mmW images, a large dataset is often required to design an accurate image classification algorithm. Relying on such traditional techniques for data generation can be a tedious and time consuming process. To overcome this, a physical model of a computational imaging (CI) system exploiting the coded aperture method [9]–[11] has been employed for data generation. This method is advantageous on two fronts:

- Firstly, it does not rely on any experimental means, it uses a numerical model of the imaging system wherein, two dynamic apertures, one transmitter and one receiver, are synthesized operating in bi-static mode. These apertures contain spatio-temporarily varying complex weights through which transmitted wavefronts from a single transmitter channel probing the scene as well as back-scattered information from the scene are passed into a single receiver channel. This phenomenon emulates a large aperture that radiates complex quasi-random radiation patterns. The entire scene information is collected in an indirect manner by encoding the back-scattered information onto a few measurement modes. This deviates from the traditional method of point-by-point scanning of the imaged scene, hence making the hardware layer much simpler by using only a single channel each for transmitter and receiver. This physical model was used in our previous works [12], [13].
- Secondly, because the classification algorithm is trained with reconstruction data generated from the physical model, physical targets are not necessarily needed for reconstructions. Instead, CAD models of threat objects are used. This is beneficial because this particular classification problem uses reconstruction data of threat objects. Using CAD models of threat objects for imaging is a much more convenient than using real physical threat objects. Hence, the data collection process becomes easier when a physical model is used instead of an experimental system.

II. Methodology

A. Classification Problem

A brief description of the problem at hand is given here. The problem here is a multi-class classification one, wherein four classes of threat objects are listed: Grenade, Gun, Knife and Scissor. These classes are labeled as '0', '1', '2' and '3', respectively. Three classifiers based on KNN and RF algorithms as well as a CNN architecture have been built. Based on the training of each of the classifiers, the developed classification framework predicts a class for a given input data. The nature of the input data is explained in details in the next section. All the three models are trained using the synthesized data from the physical model, as mentioned in Section I and tested with both synthesized as well as experimental data. There are a total of 2655 training and 542 testing samples.

B. Learning Models

In this section, details about the two machine learning algorithms as well as the CNN architecture used to solve the given classification problem are presented. The learning algorithms selected for this study are KNN and RF Algorithms. The performance of these two algorithms are compared with a straightforward CNN model. The details about these techniques are presented as follows:

1) k-Nearest Neighbour: KNN algorithm assumes that similar things exist in close proximity [14] . In other words, this algorithm calculates distances between two data points. In machine learning terms, this algorithm relies on the distance between feature vectors [15]. The most common distance metric and the one used for this particular problem is Euclidean distance, given by the formula:

$$d(\mathbf{a},\mathbf{b}) = \sqrt{\sum_{i=1}^{N}(b_i - a_i)^2} \tag{1}$$

where **a** and **b** are two datasets consisting of N samples. The other most important parameter to choose is the value of *K*. Choosing a lower value of *K* will result in less stable predictions. Increasing the value of *K* will make the predictions more stable because of majority voting/averaging hence the accuracy will improve. However, if the value of *K* is increased further, at some point there will be an increase in number of errors in the predictions, establishing the fact that higher value of *K* does not necessarily mean higher accuracy. The optimal value of *K*, on most of the problems, depends upon the nature of samples in the training data. To this end, we carried out a study to measure the relationship between mean error rate and value of *K*. The curve is plotted in Fig. 1. We see that the least error is obtained in the interval 20 to 40. Based on this finding, we choose *K* to be 23.

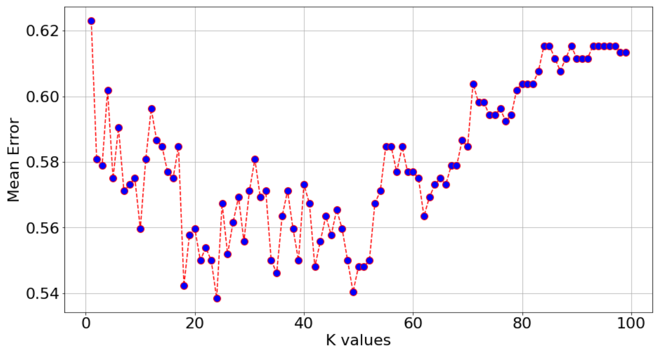

Fig. 1: Variation of the mean error with k values.

2) Random Forest: RF algorithm builds an ensemble of decision trees. In other words, it builds a multiple of decision trees and merges them to get a more accurate prediction [16]. Creating multiple decision trees can cause overfitting. RF algorithm overcomes that by creating random subsets of features, building smaller trees using those subsets and combining them later. However, there is a downside to this process in that it makes the computation process slower. Therefore, the most important parameter to select is the number of estimators, i.e. the number of trees in the random forest classification. Choosing the optimal value of trees varies from dataset to dataset. The more number of trees selected, the more accurate will be the prediction. However, at some point, the cost of cost of collecting a larger sample will be higher than the benefit in accuracy obtained from such larger sample. To this end, we conducted a study to learn the relationship between the mean error rate and number of estimators, which has been shown in Fig. 2. We observe that the least error is obtained at 60. We select the estimator value to be 60.

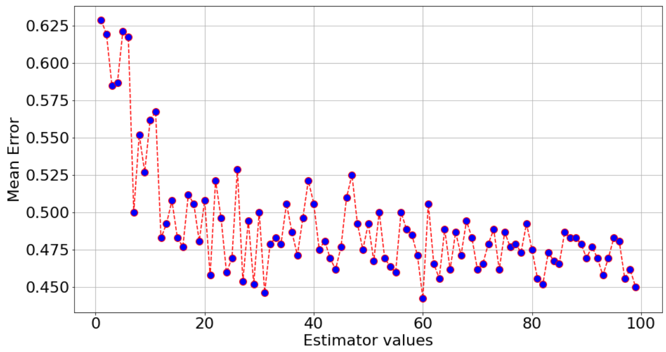

Fig. 2: Variation of mean error with the number of estimators used.

3) CNN: The performances of the two learning algorithms on our dataset are compared with a CNN model, as shown in Fig. 3, based on [12]. It consists of five convolutional layers, followed by two dense networks. The number of filters used in each layers are 32, 64, 128, 256 and 512, respectively. Each convolutional layers are followed by a Max pooling layer and a 'relu' activation layer. After the convolutional layers, the output is flattened and two dense layers of 1024 and 512 neurons are used. In the output layer, the 'softmax' activation function is used to carry out the prediction. During the training, categorical cross-entropy is chosen to be the loss function and the model was optimized with 'adam' optimizer [17].

C. Training and Testing Data Pre-processing

Machine learning algorithms and neural networks differ in their training data format. A brief description on the nature of the data for each of the technique is presented here.

1) KNN and RF Algorithms: Any image obtained from the CI physical model is first resized to a fixed set of 150 $\times$ 150 pixels and then the raw pixel intensities are extracted from the

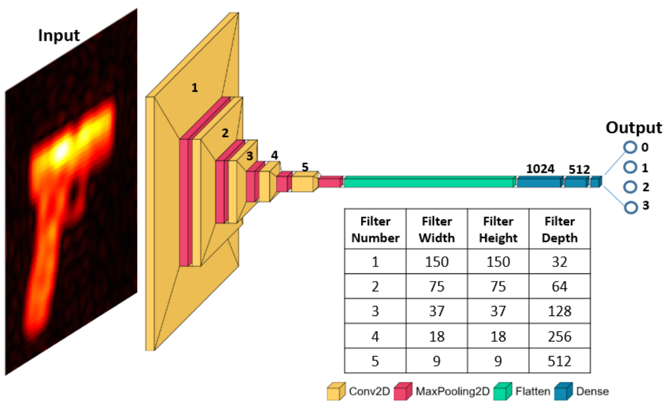

Fig. 3: Architecture of the CNN model used.

Fig. 4: Figure shows how input data is processed before training in kNN and RF algorithms. The pixel intensities are extracted, listed under 'feature vector' and the appropriate label class is associated against it.

image. The intensities are then flattened into a single list of numbers. This will be the feature vector which will contain the RGB pixel intensities of the image. The list of a single feature vector contains 150 $\times$ 150 $\times$ 3 = 67500 numbers. After extracting the pixel intensities, every feature is normalized (zero mean and unit variance) using 'StandardScaler' from sklearn. This process is repeated for each 2655 images in the dataset. The class labels (0, 1, 2 and 3) are initialized against each of these pixel intensities list. The whole process is laid out in Fig. 4.

2) CNN Algorithm: Since it is known that CNN can handle any type of data, regardless of being in image form or in tabular form, there is no need to do any kind of data pre-processing before training. The images from the CI physical model, of 150 $\times$ 150 pixels, are directly introduced to the model for training (Fig. 3). The training is carried out using k-fold cross-validation [18], [19]. The number of folds (k) was chosen to be 5, ensuring that there are sufficient samples in the training set for efficient learning as well as there are sufficient samples in the validation set to accurately evaluate the model while training.

D. Results

The classification performance of the three algorithms are presented here. The training process was completed on CUDA platform with a NVIDIA GRID M60-8Q GPU having 8GB memory size. All testing was carried out on Intel hexa-core i7 CPU with 16 GB RAM and Windows 10 operating system. A total of 542 samples were selected for testing, out of which, 521 samples were generated from the developed physical model whereas the rest 21 images were experimentally generated with the help of a handheld mmW imager [20], based on a multi-static sparse array [21], [22]. Classification reports are generated which lists the precision, recall as well as the f1-score of the predictions. We also do a comparison on the inference speed of each of the algorithm to establish which algorithm is best suited for real-time threat detection. The classification reports are listed below:

In Tables I, II and III, we take into account the f1-score to compare the three algorithms' performances. This is because we are looking for a balanced precision and recall score, i.e. the model is desired to have a very high accuracy of positive predictions (precision) as well as the model should find all the positive instances (recall). The f1-score gives the harmonic

Table I: Classification report for KNN algorithm

Class	Precision	Recall	f1-score	Support
Grenade	0.81	0.86	0.83	130
Gun	0.85	0.79	0.81	150
Knife	0.77	0.81	0.78	151
Scissor	0.89	0.87	0.87	111
Accuracy			**0.82**	**542**

Table II: Classification report for RF Algorithm

Class	Precision	Recall	f1-score	Support
Grenade	0.86	0.88	0.87	130
Gun	0.91	0.83	0.87	150
Knife	0.81	0.87	0.84	151
Scissor	0.89	0.90	0.90	111
Accuracy			**0.87**	**542**

Table III: Classification report for CNN

Class	Precision	Recall	f1-score	Support
Grenade	0.90	0.88	0.89	130
Gun	0.90	0.83	0.86	150
Knife	0.87	0.97	0.92	151
Scissor	0.98	0.91	0.94	111
Accuracy			**0.90**	**542**

mean between precision and recall ($\frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$). The fact that we consider the harmonic mean and not simple average is because the former punishes any extreme precision/recall scores. A classifier with a precision score of 1.0 and a recall score of 0.0 will have a harmonic mean of 0 but a simple average of 0.5. This shows that the former method will give a more balanced score of precision and recall. Hence, it is the ideal score to compare the accuracies of the three algorithm. It is observed that the accuracy is higher when CNN model is used as the classification model on our test dataset. It is known that CNN specializes in processing data in grid-like topology such as in image.

Since we are investigating different classification algorithms for real-time threat detection, a comparison was also drawn on the inference speed of the three algorithms. A bar chart is shown (Fig. 5) detailing out the testing speed of each algorithm on the 542 test samples.

From Fig. 5, it is evident that the CNN model takes less time in predicting classes for the whole test dataset, with a frame rate of 3.8 ms/sample, whereas, in case of KNN and RF algorithms, prediction rate of 41.7 ms/sample and 65.9 ms/sample, respectively were observed. This clearly implies that the CNN classifier model is best suited for real-time threat detection.The majority of the time in the two machine learning algorithms is taken up by the input pre-processing step of extracting pixel intensities of the input image, whereas the CNN model does not require any additional input processing step. The maximum time is taken by the RF algorithm. The reason is that the RF algorithm uses multiple trees (in this case

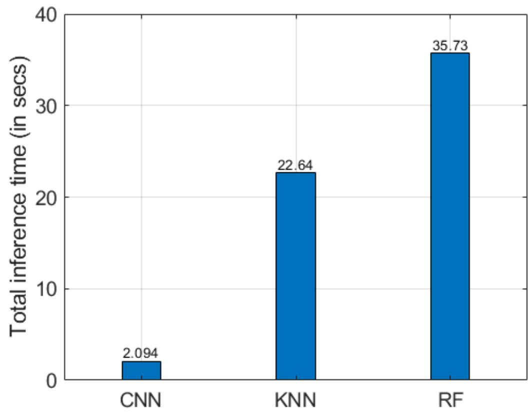

Fig. 5: A bar chart comparing the inference speed of the three algorithm used.

60) to reach a prediction. When making a prediction, all the trees have to makes a decision and then perform voting. Hence, a higher number of trees will result in slower predictions, making real-time decision making difficult. However, it should be noted that various parallel processing techniques can be accommodated in the training and the testing phase of RF algorithm to speed-up the whole process.

III. Conclusion

In this study, two machine learning algorithms: KNN and RF algorithms, and the CNN deep learning technique are compared to investigate their performance in a multi-class classification problem for real-time security screening applications. A quantitative analysis of the f1-score revealed that the CNN technique provides a better accuracy than the KNN and RF techniques for the threat object classification problem. Also, it was observed that the rate at which the CNN model makes prediction is also faster than the other two, at 3.8 ms/image. All these findings establish the fact that the CNN technique is better suited for accurate real-time image classifier in mmW CI classifier.

Acknowledgment

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

References

[1] X. Cai, M. Giallorenzo, and K. Sarabandi, “Machine learning-based target classification for mmw radar in autonomous driving,” *IEEE Transactions on Intelligent Vehicles*, pp. 1–1, 2021.

[2] S. Agarwal and B. Kumar, “Svm based concealed target quality monitoring system using millimeter wave radar,” in *2016 11th International Conference on Industrial and Information Systems (ICIIS)*, pp. 915–919, 2016.

[3] M. Bauer, R. Hussung, C. Matheis, H. Reichert, P. Weichenberger, J. Beck, U. Matuschczyk, J. Jonuscheit, and F. Friederich, “Fast fmcw terahertz imaging for in-process defect detection in press sleeves for the paper industry and image evaluation with a machine learning approach,” *Sensors*, vol. 21, no. 19, 2021.

[4] Y. Jiang, J. Cui, Z. Chen, and X. Wen, “Concealed threat detection based on multi-view millimeter wave imaging for human body,” in *2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP)*, pp. 1–4, 2019.

[5] F. Qi, I. Ocket, D. Schreurs, and B. Nauwelaers, "A system-level simulator for indoor mmw sar imaging and its applications," *Opt. Express*, vol. 20, pp. 23811–23820, Oct 2012.

[6] 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 Transactions on Antennas and Propagation*, vol. 63, no. 2, pp. 573–584, 2015.

[7] 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 *2010 IEEE International Symposium on Phased Array Systems and Technology*, pp. 543–550, 2010.

[8] S. Withington, G. Saklatvala, and M. P. Hobson, "Partially coherent analysis of imaging and interferometric phased arrays: noise, correlations, and fluctuations," *J. Opt. Soc. Am. A*, vol. 23, pp. 1340–1348, Jun 2006.

[9] M. L. Don, C. Fu, and G. R. Arce, "Compressive imaging via a rotating coded aperture," *Appl. Opt.*, vol. 56, pp. B142–B153, Jan 2017.

[10] T. Sleasman, M. F. Imani, J. N. Gollub, and D. R. Smith, "Dynamic metamaterial aperture for microwave imaging," *Applied Physics Letters*, vol. 107, p. 204104, 11 2015.

[11] C. Watts, D. Shrekenhamer, J. Montoya, G. Lipworth, J. Hunt, T. Sleasman, D. Smith, and W. Padilla, "Terahertz compressive imaging with metamaterial spatial light modulators," *Nature Photonics*, vol. 8, pp. 605–609, 08 2014.

[12] R. Sharma, R. Hussung, A. Keil, F. Friederich, T. Fromenteze, M. Khalily, B. Deka, V. Fusco, and O. Yurduseven, "Coded-aperture computational millimeter-wave image classifier using convolutional neural network," *IEEE Access*, vol. 9, pp. 119830–119844, 2021.

[13] 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," *Progress in Electromagnetic Research B*, vol. 90, pp. 91–108, Jan. 2021.

[14] N. Ali, D. Neagu, and P. Trundle, "Evaluation of k-nearest neighbour classifier performance for heterogeneous data sets," *SN Applied Sciences*, vol. 1, 11 2019.

[15] Y. Li and B. Cheng, "An improved k-nearest neighbor algorithm and its application to high resolution remote sensing image classification," in *2009 17th International Conference on Geoinformatics*, pp. 1–4, 2009.

[16] Y. Wang, Y. Huang, Z. Chen, S. Fan, Z. Liu, and H. Xu, "Complicated interference identification via machine learning methods," in *2021 IEEE 4th International Conference on Electronic Information and Communication Technology (ICEICT)*, pp. 400–405, 2021.

[17] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017.

[18] C. Capraro, U. Majumder, J. Siddall, E. Davis, D. Brown, and C. Cicotta, "SAR object classification implementation for embedded platforms," in *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 10987, 2019.

[19] E. Leite and C. de Souza Filho, "TEXTNN-A MATLAB program for textural classification using neural networks," *Computers and Geosciences*, vol. 35, no. 10, pp. 2084–2094, 2009.

[20] R. Hussung, A. Keil, and F. Friederich, "Handheld millimeter wave imaging system based on a two-dimensional multistatic sparse array," in *2020 45th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)*, pp. 1–2, 2020.

[21] B. Baccouche, P. Agostini, S. Mohammadzadeh, M. Kahl, C. Weisenstein, J. Jonuscheit, A. Keil, T. Löffler, W. Sauer-Greff, R. Urbansky, P. H. Bolívar, and F. Friederich, "Three-dimensional terahertz imaging with sparse multistatic line arrays," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 23, no. 4, pp. 1–11, 2017.

[22] S. S. Ahmed, "Microwave imaging in security — two decades of innovation," *IEEE Journal of Microwaves*, vol. 1, no. 1, pp. 191–201, 2021.