

Human motion classification by micro-doppler radar using intelligent algorithms

Andres Felipe Arias Ballen, Edith Paola Estupiñán Cuesta, Juan Carlos Martínez Quintero

Telecommunications Engineering Program, Universidad Militar Nueva Granada, Bogota, Colombia

Article Info

Article history:

Received Jun 12, 2024

Revised Sep 11, 2024

Accepted Oct 1, 2024

Keywords:

Convolutional neural networks

Micro-doppler radar

Motion detection

VGG-16

VGG-19

ABSTRACT

This article introduces a technique for detecting four human movements using micro-doppler radar and intelligent algorithms. Micro-doppler radar exhibits the capability to detect and measure object movements with intricate detail, even capturing complex or non-rigid motions, while accurately identifying direction, velocity, and motion patterns. The application of intelligent algorithms enhances detection efficiency and reduces false alarms by discerning subtle movement patterns, thereby facilitating more accurate detection and a deeper understanding of observed object dynamics. A continuous wave radar setup was implemented utilizing a spectrum analyzer and radio frequency (RF) generator capturing signals in a spectrogram centered at 2,395 MHz. Six models were assessed for image classification: VGG-16, VGG-19, MobileNet, MobileNet V2, Xception, and Inception V3. A dataset comprising 500 images depicting four movements—running, walking, arm raising, and jumping—was curated. Our findings reveal that the most optimal architecture in terms of training time, accuracy, and loss is VGG-16, achieving an accuracy of 96%. Furthermore, precision values of 96%, 100%, and 98% were obtained for the movements of walking, running, and arm raising, respectively. Notably, VGG-16 exhibited a training loss of 4.191E-04, attributed to the utilization of the Adam optimizer with a learning rate of 0.001 over 15 epochs and a batch size of 32.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Edith Paola Estupiñán Cuesta

Telecommunications Engineering Program, Universidad Militar Nueva Granada

Sede Calle 100, Bogotá, Colombia

Email: edith.estupinan@unimilitar.edu.co

1. INTRODUCTION

The detection of human motion has become increasingly relevant over time, especially in surveillance mechanisms where the identification of suspicious activities, intrusions, or abnormal behaviors can prevent criminal acts such as theft and provide a rapid response. In the medical field, monitoring a patient's breathing and detecting abnormal body movements can facilitate immediate attention from medical personal. In the realm of traffic control, detecting vehicular movement and classifying the number of individuals can regulate traffic flow and enhance efficiency. Micro-doppler radar, in addition to its applications in detection and security, unfolds its potential in human-computer interaction. In environments like virtual reality and motion control systems, this radar captures users' gestures and subtle movements, translating them into digital commands through intelligent algorithms. This article introduces a technique for detecting four distinct human movements using micro-doppler radar and intelligent algorithms.

Micro-doppler radar, in addition to its applications in detection and security, reveals its potential in human-computer interaction. In environments such as virtual reality and motion control systems, this radar captures users' gestures and subtle movements, translating them into digital commands through intelligent

algorithms. This enhances the user experience and enables natural and immersive gesture-based interfaces. This improves the user experience and enables natural and immersive gesture-based interfaces. Studies have addressed this topic, proposing various scenarios and techniques for detecting human movement [1]–[8]. Several research endeavors stand out in this field. In study [3], the power of millimeter wave (mmWave) radar for precise motion data capture is combined with deep image processing using convolutional neural networks (CNNs). The proposed approach involves mmWave data collection, preprocessing, and training of a CNN designed to recognize specific behavior patterns. Evaluated results demonstrate the viability of this methodology for real-time detection and classification, with implications for surveillance, security, and related fields in human behavior monitoring. In study [6], an ultra-wideband radar technique is employed to detect human movement through walls, incorporating a convolutional neural network (CNN) for image processing and object classification. Spectrograms taken through walls achieve high accuracy when a subject walks behind them. In study [7], an algorithm is developed for the recognition of American sign language (ASL) signals, aiding individuals with speech difficulties, achieving a recognition rate of 72.5% for 20 signals. In study [8], the detection of drones using software defined radio (sdr) radar and radiofrequency signals is presented, highlighting the diverse applications of micro-doppler radar in motion detection. Tivive *et al.* [9] proposes a methodology for capturing subtle features of human gait through micro-doppler detection, which are small fluctuations in radar return frequency caused by body part movements during walking. Using signal processing, the authors extract relevant information from doppler spectrograms and develop a classification method that identifies distinctive walking patterns among different individuals. This approach demonstrates the feasibility of using radar micro-doppler information to differentiate and classify gait patterns, with potential applications in biometrics and health monitoring. These research efforts underscore the use of micro-doppler radar as a promising technique for motion detection, capable of detecting unique features of a moving object, such as characteristic human movement patterns like walking and running. Unlike other sensors, such as video cameras, micro-doppler radar is minimally affected by environmental conditions such as lighting, rain, or fog, providing reliability in image processing [10]. In study [11], the training of an intelligent algorithm for the classification of each movement is presented. From the project, it can be concluded that the proposed model achieves an accuracy of 92.65% for human motion detection using micro-doppler radar and intelligent algorithms featuring long short-term memory (LSTM) with 36 cells and 82.33% for deep convolutional neural networks (DCNN), surpassing previously studied existing methods.

Previous research, such as references [3], [9], [10] focus on specific applications or employ particular approaches for motion detection (*e.g.*, CNNs in [3] and specific gait classification techniques in [9]), this article stands out by exhaustively evaluating and comparing six different deep learning models (visual geometry group-16 (VGG-16), VGG-19, MobileNet, MobileNet V2, Xception, and Inception V3). This provides a broader perspective on which models are most effective for detecting human movements using micro-doppler radar. This research distinguishes itself by focusing on the capability of micro-doppler radar to detect not only common movements such as walking and running but also more subtle and complex movements like jumping or arm-raising. This contrasts with works like [9], which focus on gait classification, and significantly expands the scope of the study. This research employs six deep learning models for human motion detection using micro-doppler radar. The models utilized include VGG-16, VGG-19, MobileNet, MobileNet V2, Xception, and Inception V3. Key metrics such as accuracy, training time, loss, and the confusion matrix are assessed. A dataset comprising 500 samples named WARJ MAXWELL of the implemented scenario was created, with defined motion categories: walking, running, jumping, and arm raising.

Despite advances in human motion detection, accurately distinguishing subtle and complex movements, especially in challenging environments, remains a significant challenge. Existing approaches are often affected by environmental conditions or exhibit a high rate of false positives, limiting their effectiveness in critical applications such as surveillance and medicine. Micro-doppler radar exhibits the exceptional ability to capture and measure precise details of human movements, even those with complex or non-rigid patterns. Its integration with intelligent algorithms enhances detection efficiency and reduces false positives. For instance, the radar's capability to distinguish between the subtle movements of walking and running. This multidisciplinary approach enables more precise detection and a better understanding of human movement patterns, with significant implications in applications such as surveillance and medicine. The comparison of key metrics such as precision, recall, F1 score, and training and validation loss, a thorough assessment of each model's performance in spectrogram classification is provided. Particularly noteworthy is the superior performance of the VGG-16 model, positioning it as a standout tool for accurate spectrogram classification in this context.

Furthermore, the article distinguishes its value by recognizing the feasibility of radar techniques for detecting slow-speed movements, especially when combined with intelligent algorithms and advanced

technologies. This exploration of potential radar applications adds a significant dimension to the study's conclusions, broadening its scope beyond the evaluation of deep learning models. In summary, the combination of a thorough evaluation of deep learning models with innovative research on radar applications in motion detection establishes this article as a valuable and distinctive contribution to the field. Finally, the article is organized into sections. Section 1 provides a review of the state of the art. Section 2 introduces the proposed setup, equipment used, the methodology applied, and signal processing. Section 3 presents the results of training intelligent algorithms using a dataset consisting of 125 images for each motion category, followed by the conclusions drawn.

2. METHOD

This research employed a phased methodology, as illustrated in Figure 1. Phase A identify the primary hardware and software requirements to define the scenario and conduct motion detection along with its specific characteristics, including distance, movements, and hardware configuration parameters. Phase B define signal acquisition and processing for dataset creation, utilizing Anritsu's master tools software for acquiring spectrograms and measurement data from the spectrum analyzer. Finally, in Phase C is designed and trained the 6 selected classification models using metrics such as accuracy, training time, loss, recall, and F1 score. Table 1 show the equipment used in this research and the proposed scenario.

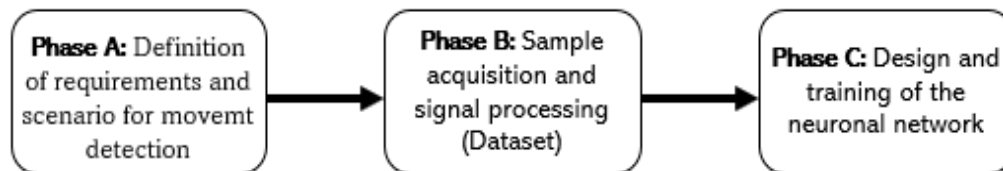


Figure 1. Research methodology workflow

The selected metrics allow for the measurement of accuracy in classification models, with a focus on minimizing false positives-cases where the model incorrectly predicts a positive outcome when it is actually negative. The accuracy metric is defined as a parameter that evaluates the model's ability to correctly classify samples into the desired categories. It is calculated by dividing the number of correct predictions by the total number of predictions and is typically expressed as a value between 0 and 1, where 1 signifies perfect accuracy and 0 indicates no accuracy. Training time refers to the period required for a machine learning model to process and analyze a training dataset in order to learn patterns and relationships. The loss parameter serves as an indicator of the model's learning level during training, with the goal of minimizing it by adjusting model parameters [12]. The recall metric evaluates the model's ability to correctly identify all positive examples in the dataset, focusing on minimizing false negatives. The F1 score combines precision and recall metrics into a single value to assess the performance of a classification model, especially in binary classification problems.

Table 1. Materials

Equipment	Characteristics/installation requirements	Objective
Spectrum analyzer Anritsu S332E	Frequency: 100 kHz to 4 GHz, Average noise level: 152 dBm to 10 Hz RBW, Phase noise: 100 dBc/Hz max, Connections: Ethernet, universal serial bus (USB) cable, memory USB, RS-232 [13]	Measure the power distribution of a signal as a function of frequency and time.
RF generator R&S SMB 100 A	Phase noise SSB: -108 dBc (tip.) at 10 GHz and compensation of 20 kHz, Broadband noise: -138 dBc at 10 GHz and 30 MHz compensation, Max output power +27 dBm [14]	Generate the continuous RF wave signal that functions as the radar signal. It is synchronized with the receiver (Spectrum analyzer).
Can antennas	Frequency range antenna 1: 2.2-2.7 GHz; Frequency range antenna 2: 2.16-2.5 GHz. Approximate bandwidth: Antenna 1: 500 MHz; Antenna 2: 340 MHz	Self-implemented antennas tuned to the radar's operating frequency.
Coaxial cable	Insulation resistance: 5.000 MΩ min, Impedance 50 Ω, VSWR: 1,3 max, Range frequency: 0-4 GHz [15]	Cable to synchronize the clock of the RF generator and the spectrum analyzer.
SMA to SMA cables	Frequency range: max 18 GHz. Impedance: 50 Ω [16]	Transmit high-frequency signals with low signal loss between equipment and antennas.
Software master tools Anritsu	Software proprietary of Anritsu [17]	Acquisition, handling, storage, and interpretation of the data obtained during the tests.

2.1. Phase A: scenario definition

The selected movements for the detect and evaluate of neural network models were: running, jumping, arm raising, and walking. These common movements that can enable non-verbal communication and are used in everyday life. Humans run to catch a bus, jump to avoid obstacles in their path, raise their arms for active breaks or stretching after sitting for an extended period, and walk to move from one place to another. These movements can also characterize specific activities, facilitating the identification and understanding of behaviors. Once the movements were defined, the number of repetitions for data collection and subsequent dataset creation was specified. Table 2 describes the movements and the repetition frequencies at which measurements were taken.

Table 2. Movements with their frequencies and distances

Movements	# of repetitions	Distance (m)	Number of tests
Walk	1	9.4	3
Jump	6	-	3
Run	1	9.4	3
Raise Arms	6	-	3

The scenario was implemented using the equipment and materials listed in Table 1, and its diagram is depicted in Figure 2. The continuous-wave radar was implemented by employing a radio frequency generator as the transmitting equipment configured to transmit a sine wave at a specific frequency and power level. The receiver is a spectrum analyzer with spectrogram functionality and ethernet communication with a personal computer (PC). Synchronization between the transmitter and receiver is achieved using a coaxial cable connected to the Bayonet Neill-Concelman (BNC) connectors designated for this purpose on each equipment. The custom-made controller area network (CAN-type) antennas were tuned using a vector network analyzer. The master tools software on the PC was used for spectrogram acquisition.

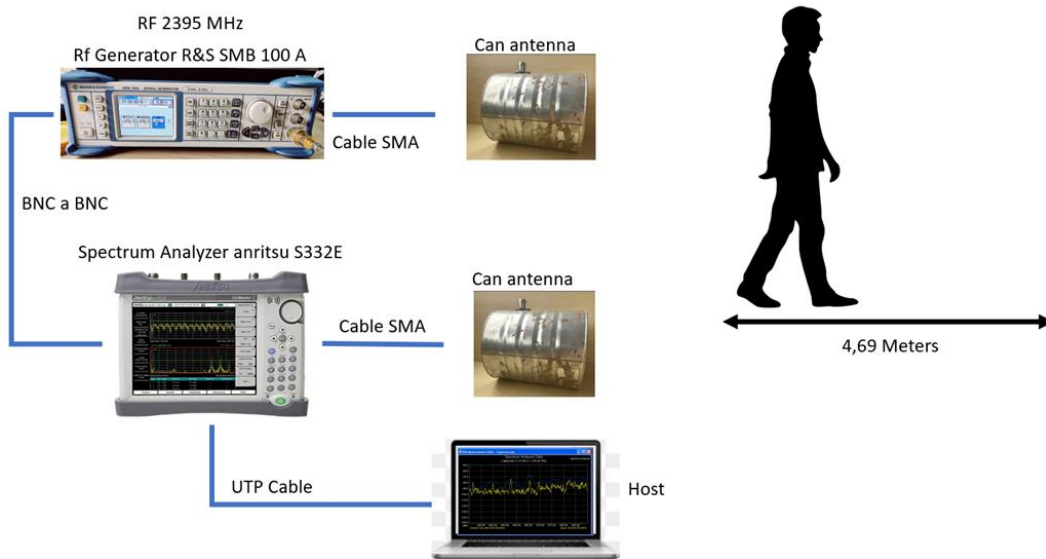


Figure 2. Proposed scenario for data collection

2.2. Phase B: sample acquisition and signal processing

In this phase, sample acquisition and signal processing were conducted for analysis. A continuous-wave signal frequency of 2,395 MHz was selected with a transmission power of -10 dBm. Spectrograms were configured to be received with a reference level of -4 dBm, a span of 201 Hz, a resolution bandwidth (RBW) of 10 Hz, and an acquisition time of 45 seconds. The received power is approximately -49 dBm. It is important to consider the proximity between the transmitting and receiving antennas and the absence of isolation between them, apart from the distance. Figure 3 allows us to observe one of the obtained spectrograms. This spectrogram is a visual representation that illustrates the energy variation of different frequencies in a radio signal over time [18].

The use of colors represents this frequency of energy or intensity at a given moment. The darker colors typically indicate lower power, while warmer colors such as green or yellow indicate higher power. The frequency shifts to the right of the central frequency in the spectrogram frequency increase represents the movement of a person approaching the radar. The greater the shift, the higher the person's velocity. The frequency shifts to the left frequency decrease represents a person moving away from the radar. These shifts are caused by the doppler effect, which occurs when there is relative motion between the signal source, in this case, the person, and the signal receiver that is recording the movement [19]. The Doppler effect manifests as a change in signal frequencies as a person moves. When a person is closer to the antennas, the signal will have more power and tend towards warmer colors, as observed in the previous figure. Conversely, if the person moves away from the antennas, the signal will exhibit cooler colors.

Figure 4 depicts the implemented scenario, where the antennas are spaced 30 cm apart. With this configuration, we conducted the acquisition of 500 sample images capturing various movements. The spectrogram images provided by the software were resized to 700x274 pixels with a depth of 8 bits, resulting in a file size of 11 kB.

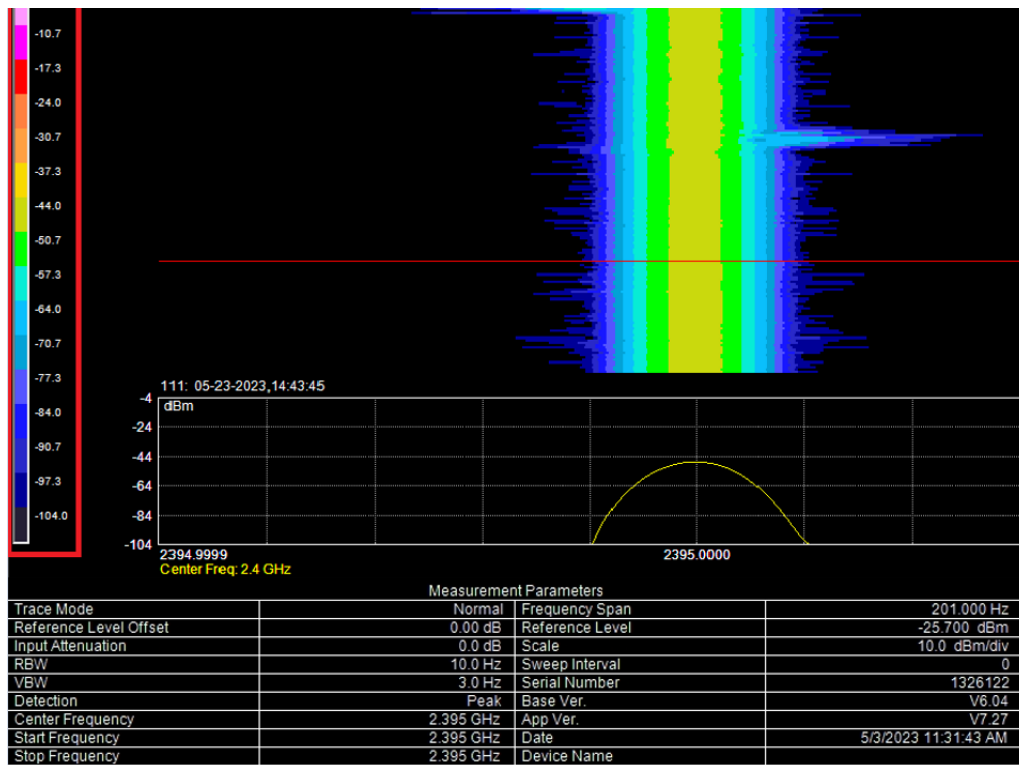


Figure 3. Spectrogram characteristics

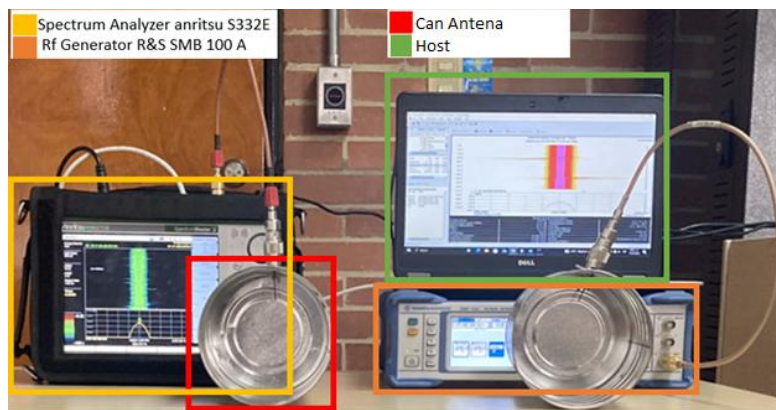


Figure 4. Proposed scenario with 30 cm spacing between antennas

To mitigate overfitting and ensure robust model evaluation, the dataset was divided into 80% of images for training and 20% for validation. This approach reserves a portion of the data for validation, enabling assessment of the model's performance with data not utilized during training. By dividing the data into training and validation sets, this strategy ensures improved generalization of the model and better adaptation to new data, thus enhancing its ability to address diverse scenarios in the future. The dataset of 500 images was labeled as WARJ MAXWELL sample collection took place in the laboratories of Universidad Militar Nueva Granada, involving a total of 60 individuals aged between 18 and 50 years.

Figure 5 shows the scenario with the equipment used and the way samples are captured for two of the four selected movements. Figure 5(a) shows the equipment used in the implemented scenario. Figure 5(b) demonstrates how samples are acquired while the target person raises their arms, and Figure 5(c) shows the acquisition process when the person is walking. The person must stand directly in front of the prototype to perform the movement because the antennas are directional and have a narrow beamwidth. Figure 6 shows the four spectrograms for each of the four movements to be classified: running Figure 6(a), walking Figure 6(b), jumping Figure 6(c), and raising arms Figure 6(d). The central part of the spectrogram blue color corresponds to the carrier signal, which is visualized with higher power. The lateral components yellow and red colors correspond to the doppler shift in frequencies due to the movement.

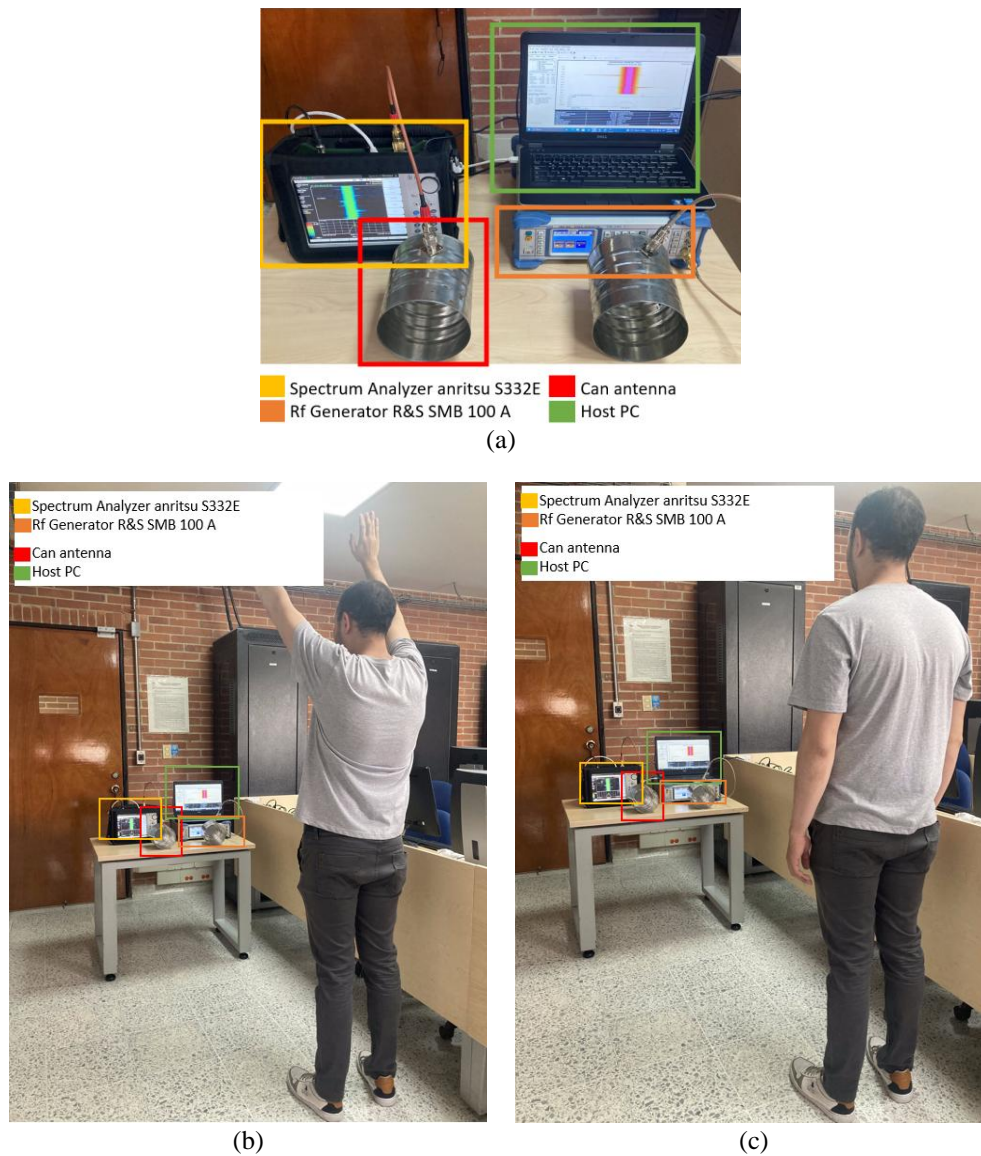


Figure 5. Proposed scenario with equipment used for detecting arm-raising and walking movements (a) scenario proposed with used equipment, (b) spectrogram acquisition of raised arms, and (c) spectrogram acquisition while walking

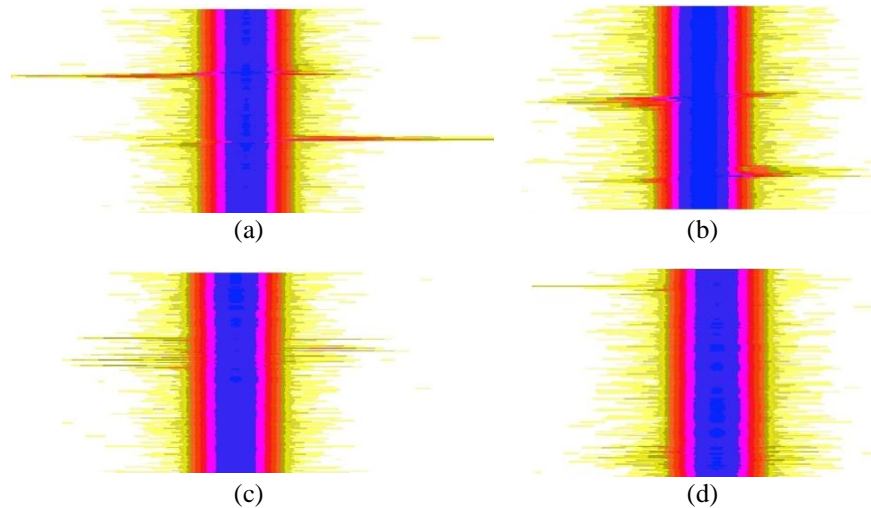


Figure 6. Spectrogram (a) running, (b) walking, (c) jumping, and (d) raising arms

2.3. Phase C: network architecture selection and training

For the classification of movements, 6 neural network models were selected: VGG-16, VGG-19, MobileNet, MobileNet V2, Xception, and Inception V3. Some of the characteristics of these models are shown in Table 3. In this case, “Size (MB)” represents the size in megabytes (MB) of the model after being trained, “parameters” indicates the number of parameters or coefficients the model has learned during the training process, and “depth” represents the depth or number of layers the model has. The more layers a model has, the deeper it is.

Table 3. Technical specifications of the used models

Model	Size (MB)	Parameters	Description	Depth
VGG-16	528	138.4 M	The VGG-16 architecture consists of 16 layers, including 13 convolutional layers with 3×3 filters and zero padding, followed by 2×2 Max pooling layers to reduce dimensionality. Afterward, there are 3 fully connected layers responsible for classification tasks. This structure enables VGG-16 to progressively learn from simple features like edges to more complex features in images, making it an effective deep network for computer vision [20], [21].	16
VGG-19	549	143.7 M	VGG-19 is an extended version of the VGG-16 architecture, characterized by its depth and uniformity in the arrangement of 19 layers, including both convolutional and fully connected layers. Like VGG-16, it employs small filters and Max-pooling layers to extract features from images. Used in tasks such as classification, object detection, and more, VGG-19 has been significant in computer vision, demonstrating solid performance, although accuracy may vary depending on the dataset and specific task [22], [23].	19
MobileNet	16	4.3 M	The exact number of layers in a MobileNet can vary depending on the specific version and modifications made. However, in general, MobileNet architectures typically consist of multiple convolutional and pooling layers, as well as fully connected layers at the end for classification or a specific task. For example, MobileNetV1 consists of approximately 55 layers. These networks are designed for applications on mobile devices and embedded systems, striking a balance between complexity and computational efficiency [24].	55
MobileNet V2	14	3.5 M	It utilizes building blocks called “inverted residuals” that optimize feature representation in deep networks while having a reduced number of parameters. This allows for a balance between accuracy and speed. In this project, it can be used in tasks such as object detection, image classification, and other vision-related tasks, making it suitable for classifying movements [25].	105
Xception	88	22.9 M	It utilizes building blocks called “inverted residuals” that optimize feature representation in deep networks while having a reduced number of parameters. This allows for a balance between accuracy and speed. In this project, it can be used in tasks such as object detection, image classification, and other vision-related tasks, making it suitable for classifying movements [25].	81
Inception V3	92	23.9 M	It stands out for the implementation of factorized convolutions, in which standard 3×3 convolutions are separated into two smaller convolutions (1×3 and 3×1), with a total of 189 hidden layers. This allows for capturing spatial patterns and reducing computational load. Additionally, Inception-v3 employs smaller-sized filters to improve efficiency and reduce the number of parameters. Its distinctive feature is the “Inception” modules, which perform parallel convolutions with different filter sizes and then concatenate their results, enabling the network to capture information at multiple scales [23].	189

The experimentation with the selected models was conducted using TensorFlow Keras, alongside the requisite Python libraries and modules for image processing and deep learning model construction. To optimize the processing speed of training and validation data, cache and prefetch methods were employed. In order to prevent overfitting of the model to the original data, a data augmentation model was implemented to generate new images from the training set through random transformations. A flexible function was developed to accept various input parameters, defining a convolutional neural network model and training it on both training and validation datasets. These input parameters encompass critical configurations pivotal in constructing and training a deep neural network.

The choice of model architecture, be it VGG16, VGG19, or MobileNet, determines the organization of layers within the network. The decision of whether the layers should be trainable conditions their adaptability to new data or retention of static properties. During this process, the optimizer such as Adam or stochastic gradient descent (SGD) facilitates parameter adjustments, utilizing the learning rate to regulate the magnitude of these adjustments in each iteration. In addition to the last convolutional block, the inclusion of extra layers and techniques like dropout influence the depth and robustness of the network. Fully-connected layers refine finer details, and the number of layers and neurons can vary accordingly. Preserving the model and its weights retains the knowledge acquired during the process, while the number of epochs specifies the frequency at which the training data will be traversed. Proper calibration of these parameters is crucial for achieving optimal performance in the desired task by the neural network. The input parameters encompass the model architecture, and the function returns a history object that records training and validation accuracy and loss throughout the training process. Figure 7 elucidates this description, while Table 4 delineates the selection of hyperparameters utilized for all models.

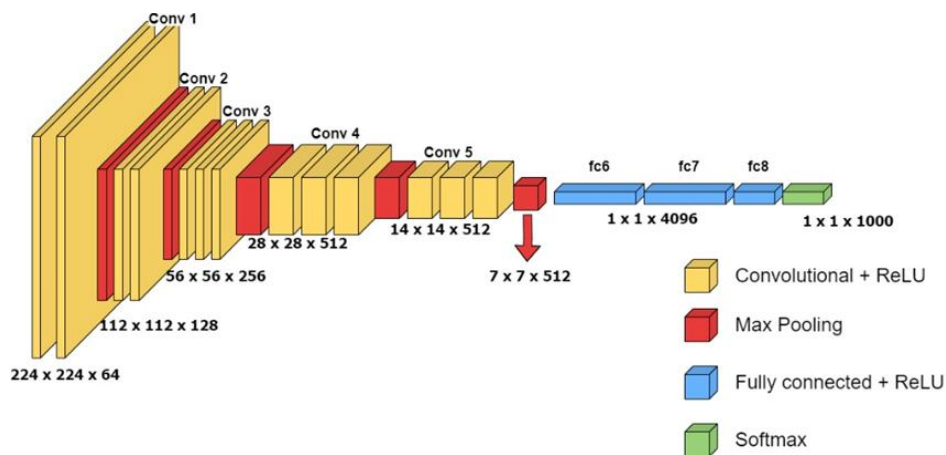


Figure 7. VGG-16 architecture

Table 4. Hyperparameters used

Metric	Value
Learning rate	0.01
Drop rate	0.01
Epochs	10
Max pooling	2

3. RESULTS AND DISCUSSION

Table 5 summarizes the results obtained from training the models. The VGG-16 model with the Adam optimizer achieves a training accuracy of around 100% and a validation accuracy of approximately 96%. Additionally, both training and validation losses are lower compared to other models. This suggests that the VGG-16 model is capable of successfully classifying images into the two target classes. VGG-16 is a robust and widely used model, especially in image classification tasks. It is known for its deep and uniform architecture, making it effective at extracting features from images of different scales and complexities. However, due to its depth, it has a relatively large number of parameters, which can result in a larger model size and require more computational resources for training and execution. Thanks to the -10 dB configuration, the dataset effectively captures human motion without reflections or interference from the environment. Regarding the time required to train the model, there is a notable similarity in the results. These

times were obtained using a graphics processing unit (GPU) runtime, with a total time of 2 minutes. The average training time for the 6 models was approximately 1.16 minutes. All models are capable of correctly classifying the 4 types of movements with high accuracy in a short training time, as shown in Table 5. Table 6 resume accuracy, recall, and F1 score metrics results. The best train accuracy results were obtained by the VGG-16, VGG-19, MobileNet, and MobileNetV2 models. However, according to validation accuracy, the model that best generalizes the validation data is VGG-19, considering that the other validation accuracy results are above 0.900. Despite VGG-19 showing the lowest train loss, it is evident that the model is not overfitting because it generalizes the validation data correctly with a validation accuracy of 0.970 and a train accuracy of 1.

Table 5. Results obtained; learning rate: 0.01; dense layers: 1024; number of epochs: 10

Model	Optimizer	Val accuracy	Train accuracy	Train loss	Val loss	Time (M)
VGG-16	SGD	0.849	0.943	0.232	0.421	2
	Adam	0.961	1.000	4.191E-04	0.430	2
VGG-19	Adam	0.970	1.000	1.608E-04	0.099	3
	SGD	0.207	0.7037	14.606	112.455	3
MobileNet	Adam	0.934	0.997	0.022	0.251	0
	SGD	0.910	1000	0.015	0.261	0
MobileNet V2	Adam	0.930	0.990	0.029	0.363	0
	SGD	0.920	1.000	0.021	0.256	0
Xception	Adam	0.860	0.820	0.383	0.703	1
	SGD	0.840	0.972	0.184	0.382	1
Inception V3	Adam	0.640	0.815	0.516	1.152	1
	SGD	0.880	0.905	0.331	0.347	1

Table 6. Results of evaluated metrics for the 6 models used

Model	Movements			
	Run	Jump	Raise arms	walk
VGG-16	Precision=1	Precision=0.92	Precision=0.98	Precision=0.96
	Recall=0.98	Recall=0.98	Recall=0.94	Recall=1.00
	F1 Score=0.98	F1 Score=0.96	F1 Score=0.96	F1 Score=0.98
VGG-19	Precision=0.92	Precision=0.94	Precision=0.96	Precision=0.94
	Recall=0.94	Recall=1	Recall=0.94	Recall=0.96
	F1 Score=0.95	F1 Score=0.96	F1 Score=0.94	F1 Score=0.98
MobileNet	Precision=0.95	Precision=0.94	Precision=0.96	Precision=0.94
	Recall=0.94	Recall=0.92	Recall=0.94	Recall=0.92
	F1 Score=0.94	F1 Score=0.94	F1 Score=0.98	F1 Score=0.92
MobileNet V2	Precision=0.96	Precision=0.86	Precision=0.92	Precision=0.92
	Recall=0.94	Recall=0.90	Recall=0.94	Recall=0.92
	F1 Score=0.94	F1 Score=0.90	F1 Score=0.91	F1 Score=0.96
Xception	Precision=0.90	Precision=0.86	Precision=0.90	Precision=0.90
	Recall=0.92	Recall=0.88	Recall=0.84	Recall=0.94
	F1 Score=0.90	F1 Score=0.92	F1 Score=0.88	F1 Score=0.96
Inception	Precision=0.90	Precision=0.86	Precision=0.90	Precision=0.90
	Recall=0.94	Recall=0.86	Recall=0.84	Recall=0.86
	F1 Score=0.86	F1 Score=0.84	F1 Score=0.92	F1 Score=0.84

The model faced challenges in accurately classifying the “jumping” movement, exhibiting the lowest precision among all categorized movements. Conversely, the movement classified with the highest precision by the models is “running.” This distinction can be attributed to the spectrogram's typically more pronounced presence, facilitating its differentiation from other movements. Figure 8 showcases the performance results of the VGG-16 model through a confusion matrix. With an accuracy exceeding 92% in recognizing arm movements, the model demonstrates commendable overall architectural proficiency. Nonetheless, it is noteworthy that “arm movement” occasionally incurs confusion with “jumping,” given their shared characteristics. Despite this, “jumping” maintains a precision of 100%, underscoring the algorithm's adeptness in image classification and the VGG-16 architecture's successful performance.

In Figure 9, the prediction of an image classification model with input dimensions of 150×150 pixels and 3 color channels are presented. The main objective is to recognize human actions, and specifically, in this instance, the action of raise arms. The model's prediction yielded a highly accurate result, where the class Raise Arms has a probability of approximately 68%, confirming the robustness and effectiveness of the model in identifying this specific action with certainty. This precision supports the suitability of the proposed approach, highlighting its applicability in environments where accurate identification of human actions is essential, such as in security monitoring systems or medical applications for the analysis of specific movements.

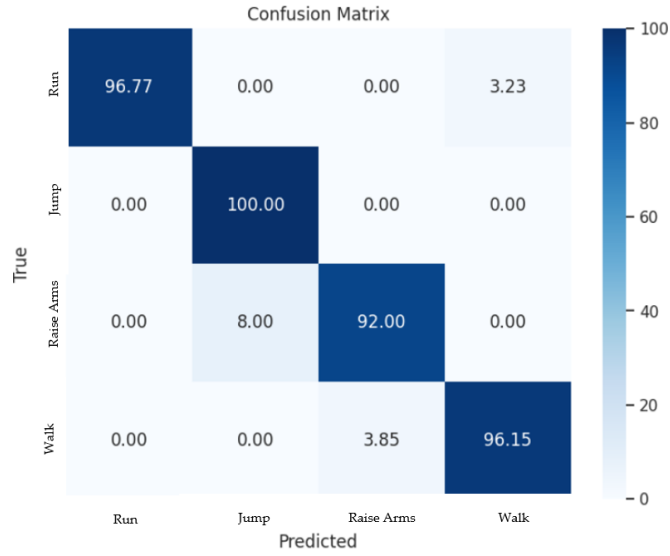


Figure 8. Confusion matrix in percentage

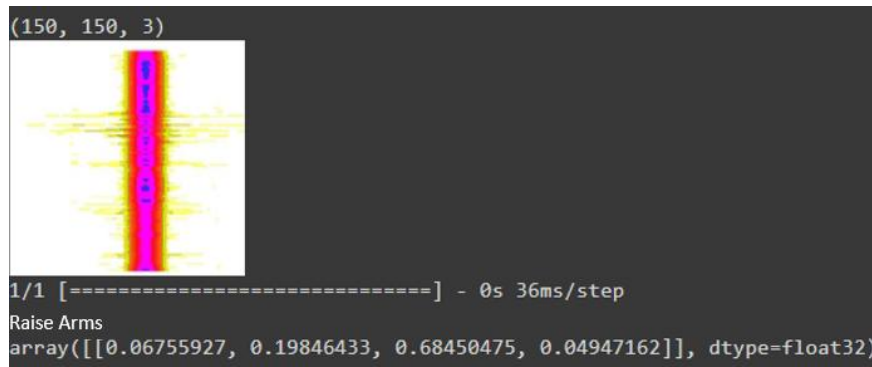


Figure 9. Prediction

4. CONCLUSION

The VGG-16 model achieves outstanding precision results for movement identification, with values as follows: walking 96%, running 100%, and arm raising 98%. For the “jumping” movement, both VGG-19 and MobileNet surpassed the VGG-16 model, achieving a precision of 94%. However, Xception and Inception models delivered the least favorable precision values for identifying the “jumping” movement, with both models scoring 86%. In terms of recall, VGG-16 stands out with an average value of 0.975. Notably, the VGG-19 model achieved a perfect recall score of 1 for the “jumping” movement, the highest among all models evaluated.

Regarding the F1 score metric, an average value of 0.97 was obtained for identifying the four movements. The Inception model obtained the lowest score in this metric, with a value of 0.865. In terms of training loss, VGG-16 achieved the lowest value at 4.191E-04. Conversely, for validation loss, the MobileNet model yielded the most favorable result, with a score of 0.251. The Inception model recorded the least favorable results for both training and validation loss, with a value of 1.152. Overall, the VGG-16 model emerges as one of the most effective tools for spectrogram classification. Furthermore, our implemented scenario underscores the viability of radar techniques for detecting slow-speed movements of people or objects when combined with intelligent algorithms and technologies.

ACKNOWLEDGEMENTS




This work was developed within the GISSIC research group, as a product derived from the Maxwell research seedbed, endorsed by the vice-rectorate for research at the Universidad Militar Nueva Granada.

REFERENCES

- [1] M. E. Paz, G. Friedrich, and C. L. Galasso, "Signal processing displayed on a spectrogram," *Elektron*, vol. 4, no. 1, pp. 35–39, 2020, doi: 10.37537/rev.elektron.4.1.90.2020.
- [2] W. Li, B. Tan, and R. J. Piechocki, "Non-contact breathing detection using passive radar," in *2016 IEEE International Conference on Communications*, 2016, pp. 1–6, doi: 10.1109/ICC.2016.7511389.
- [3] R. Zhang and S. Cao, "Real-time human motion behavior detection via CNN using mmWave radar," *IEEE Sensors Letters*, vol. 3, no. 2, pp. 1–4, Feb. 2019, doi: 10.1109/LSENS.2018.2889060.
- [4] E. Marchetti *et al.*, "Radar reflectivity and motion characteristics of pedestrians at 300 GHz," in *European Microwave Week 2017: "A Prime Year for a Prime Event", EuMW 2017 - Conference Proceedings; 14th European Microwave Conference, EURAD 2017*, 2017, pp. 57–60, doi: 10.23919/EURAD.2017.8249146.
- [5] Y. Wang and A. E. Fathy, "Range-time-frequency representation of a pulse Doppler radar imaging system for indoor localization and classification," in *2013 IEEE Topical Conference on Wireless Sensors and Sensor Networks (WiSNet)*, 2013, pp. 34–36, doi: 10.1109/WiSNet.2013.6488625.
- [6] N. Nourshamsi, S. Vakalis, and J. A. Nanzer, "Joint detection of human and object motion using harmonic micro-Doppler radar and harmonic tags," *IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 6, pp. 930–934, Jun. 2020, doi: 10.1109/LAWP.2020.2983012.
- [7] S. Z. Gurbuz *et al.*, "American sign language recognition using RF sensing," *IEEE Sensors Journal*, vol. 21, no. 3, pp. 3763–3775, Feb. 2021, doi: 10.1109/JSEN.2020.3022376.
- [8] H. Fu, S. Abeywickrama, L. Zhang, and C. Yuen, "Low-complexity portable passive drone surveillance via SDR-based signal processing," *IEEE Communications Magazine*, vol. 56, no. 4, pp. 112–118, Apr. 2018, doi: 10.1109/MCOM.2018.1700424.
- [9] F. H. C. Tivive, A. Bouzerdoum, and M. G. Amin, "A human gait classification method based on radar doppler spectrograms," *Eurasip Journal on Advances in Signal Processing*, vol. 2010, pp. 1–12, 2010, doi: 10.1155/2010/389716.
- [10] A. Hanif, M. Muaz, A. Hasan, and M. Adeel, "Micro-doppler based target recognition with radars: a review," *IEEE Sensors Journal*, vol. 22, no. 4, pp. 2948–2961, Feb. 2022, doi: 10.1109/JSEN.2022.3141213.
- [11] M. Wang, Y. D. Zhang, and G. Cui, "Human motion recognition exploiting radar with stacked recurrent neural network," *Digital Signal Processing: A Review Journal*, vol. 87, pp. 125–131, 2019, doi: 10.1016/j.dsp.2019.01.013.
- [12] J. Terven, D. M. Cordova-Esparza, A. Ramirez-Pedraza, E. A. Chavez-Urbiola, and J. A. Romero-Gonzalez, "Loss functions and metrics in deep learning," *arXiv:2307.02694*, Jul. 2023.
- [13] Anritsu, "Site master cable and antenna analyzer with spectrum analyzer S332E," *Anritsu.com*. <https://www.anritsu.com/en-us/test-measurement/products/s332e> (accessed Oct. 23, 2023).
- [14] S. A. Rohde and S. España, "R&S@SMB100A microwave signal generator," *Rohde and Schwarz España, S.A.* https://www.rohde-schwarz.com/es/productos/test-y-medida/generadores-de-senales-analogicas/rs-smb100a-microwave-signal-generator_63493-9379.html (accessed Oct. 26, 2023).
- [15] B. A. Forouzan, "Data communications and networking," 4th ed., New York, NY, USA: McGraw-Hill Higher Education, 2007, pp. 195–197.
- [16] M. Steer, "Microwave and RF design : a system approach," in *Microwave and RF design*, Scitech Publishing, 2010, pp. 189–190.
- [17] Anritsu, "Master software tools MST," *Anritsu.com*. <https://www.anritsu.com/En-US/Test-Measurement/Products/Mst> (accessed Oct. 26, 2023).
- [18] E. Kurtoglu, A. C. Gurbuz, E. Malaia, D. Griffin, C. Crawford, and S. Z. Gurbuz, "RF micro-Doppler classification with multiple spectrograms from angular subspace projections," in *Proceedings of the IEEE Radar Conference*, 2022, pp. 1–6, doi: 10.1109/RadarConf2248738.2022.9763904.
- [19] J. C. M. Quintero, E. P. E. Cuesta, and J. S. G. Ramírez, "Reconfigurable software-defined radar testbed with built-in validation," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 2, pp. 825–837, 2022, doi: 10.11591/eei.v11i2.3568.
- [20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *3rd International Conference on Learning Representations (ICLR 2015)*, 2015, pp. 1–14.
- [21] S. Mascarenhas and M. Agarwal, "A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for image classification," in *Proceedings of IEEE International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications, CENTCON 2021*, 2021, pp. 96–99, doi: 10.1109/CENTCON52345.2021.9687944.
- [22] A. Kumar and S. Sachar, "Deep learning techniques in leaf image segmentation and leaf species classification: a survey," *Wireless Personal Communications*, vol. 133, no. 4, pp. 2379–2410, 2023, doi: 10.1007/s11277-024-10873-2.
- [23] P. K. Rai *et al.*, "Localization and activity classification of unmanned aerial vehicle using mmWave FMCW radars," *IEEE Sensors Journal*, vol. 21, no. 14, pp. 16043–16053, Jul. 2021, doi: 10.1109/JSEN.2021.3075909.
- [24] A. G. Howard and Others, "MobileNets: efficient convolutional neural networks for mobile vision applications," *arXiv:1704.04861*, 2017.
- [25] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: inverted residuals and linear bottlenecks," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510–4520, doi: 10.1109/CVPR.2018.00474.

BIOGRAPHIES OF AUTHORS






Juan Carlos Martinez Quintero    received the M.Sc. degree in autonomous systems of production from in Universidad Tecnologica de Pereria in 2013. He is currently professor at Militar Nueva Granada University. His research interests include mobile networks, software-defined radio (SDR), communication systems and digital signal processing. He can be contacted at email: juan.martinezq@unimilitar.edu.co.



Edith Paola Estupiñán Cuesta    received the M.Sc. degree in electronic engineering from Pontifical Xavierian University in 2013. She is currently working at Militar Nueva Granada University, Colombia. Her research interests include mobile network, traffic analysis and data and management network. She can be contacted at email: edith.estupinan@unimilitar.edu.co.



Andres Felipe Arias Ballen    is a Telecommunications Engineer from Universidad Militar Nueva Granada. During his major, he joined the GISSIC research group (Maxwell hotbed of research), which led him to learn about some research methods. He can be contacted at email: est.andres.arias1@unimilitar.edu.co.