

Deep Learning Based Human Detection Using UWB Radar Signals

Shreyas D¹, Roopitha G Nayak², Umamahesh P³

^{1,2} Department of Information Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, KA

³Department of Electronics and Communication Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, KA

Abstract. Ultra-Wideband (UWB) radar technology has become a promising alternative for detecting humans in settings where conventional camera systems are either ineffective or unsuitable. This study introduces a method that uses deep learning to identify human presence through UWB radar signal data. Radar signals are transformed into visual formats and examined with Convolutional Neural Networks (CNNs) to determine if a person is present or not. Two datasets consisting of radar signal images gathered under different conditions, including varying angles, distances, and human postures, are utilized for the experiments. The system uses preprocessing methods such as adjusting image size, standardizing data, and converting labels into numerical form prior to training the model. Several CNN models with different numbers of convolutional layers, pooling techniques, and dropout settings are tested to identify the best-performing model for detecting humans using radar data. Performance is measured through evaluation metrics like accuracy, precision, recall, F1-score, and confusion matrices. The experimental results show that the suggested deep learning method effectively differentiates between human and non-human radar signals with high precision. The research emphasizes the promise of integrating UWB radar sensing with deep learning methods for privacy-conscious human detection in areas like security surveillance, healthcare, smart homes, and search-and-rescue missions.

Index Terms: UWB Radar, Human Detection, Deep Learning, Convolutional Neural Networks, Radar Signal Processing, Machine Learning, Indoor Sensing, Tensorflow.

I. INTRODUCTION

Radar-based sensing is widely used for detecting objects and motion by transmitting electromagnetic waves and analyzing their reflections. Unlike vision-

based systems, radar can operate effectively in low-light and obstructed environments, making it suitable for indoor and safety-critical applications.

Ultra-Wideband (UWB) radar extends these capabilities by transmitting short pulses over a wide frequency spectrum, enabling high-resolution detection. It can identify very small movements, such as human breathing or slight body motion, and can penetrate materials like walls. Additionally, since it does not capture visual information, it offers a privacy-preserving alternative to camera-based systems.

Detecting human presence using UWB radar involves analyzing complex reflected signals that contain both useful information and environmental noise. Traditional signal processing techniques often struggle to extract meaningful patterns from such data.

Deep learning provides an effective solution to this problem by automatically learning features from raw data. In particular, Convolutional Neural Networks (CNNs) are well-suited for analyzing radar signals when they are represented as images. These models can capture spatial patterns and distinguish between human and non-human signals without manual feature design.

In this work, TensorFlow is used to implement and train CNN-based models for radar signal classification, enabling efficient experimentation and performance optimization.

II. EXISTING AND PROPOSED SYSTEM

Conventional human detection systems primarily rely on cameras, infrared sensors, or motion detectors. While effective in controlled conditions, these systems

face limitations such as sensitivity to lighting conditions, inability to detect through obstacles, and concerns related to user privacy.

Radar-based approaches provide an alternative; however, many existing methods depend on handcrafted features or traditional machine learning models. These approaches often lack robustness and struggle to generalize across different environments due to the complexity of radar signals.

To address these challenges, this work proposes a deep learning-based framework for human detection using UWB radar data. The radar signals are converted into image-like representations, which are then used as input to CNN models.

Multiple CNN architectures are explored by varying parameters such as the number of convolutional layers, pooling techniques, and the use of dropout for regularization. Additionally, cross-validation strategies are employed to ensure reliable performance evaluation.

This approach enables automatic feature extraction from radar data and improves classification accuracy compared to traditional methods.

III. PROBLEM STATEMENT

Human detection systems based on cameras and conventional sensors are limited by environmental conditions and privacy concerns. Although UWB radar provides a promising alternative, accurately interpreting its complex signals remains a challenge.

This work aims to develop a CNN-based model that can effectively learn patterns from UWB radar data and accurately classify human and non-human scenarios, while improving robustness and generalization.

IV. RELATED WORK

Deep learning has significantly improved the ability to analyze complex patterns in both image and signal data. Early breakthroughs such as AlexNet by Krizhevsky et al. [1] demonstrated the effectiveness of deep Convolutional Neural Networks (CNNs) for large-scale image classification tasks. This was followed by architectures like ResNet [2], which introduced residual connections to enable deeper networks, and VGGNet [3], which showed that increasing network depth can improve performance.

Later, Inception models [4] focused on improving computational efficiency while maintaining accuracy. These advancements laid the foundation for applying deep learning to more complex domains, including radar signal processing.

CNNs have since been widely adopted for analyzing structured data such as images and time-series signals. In the context of radar applications, Le Minh et al. [6] applied CNNs to UWB impulse radar data for indoor people counting, demonstrating promising accuracy in controlled environments. Zhong et al. [7] extended this work by combining CNNs with LSTM networks to capture both spatial and temporal features for real-time motion recognition. Similarly, Bartolini et al. [8] explored machine learning techniques for human detection in non-line-of-sight (NLOS) scenarios, highlighting the challenges of generalization in complex environments.

Before the rise of deep learning, traditional computer vision techniques such as Histogram of Oriented Gradients (HOG) [9] and the Viola-Jones method [10] were widely used for human detection. While these methods were effective for image-based detection, they relied heavily on handcrafted features and struggled to perform well in noisy or cluttered environments, especially when applied to radar data.

With the advancement of UWB radar technology, researchers have increasingly focused on combining radar sensing with deep learning models. UWB radar provides high-resolution signal data that can be transformed into image-like representations, making it suitable for CNN-based analysis. Studies have shown that CNN-based approaches can achieve higher accuracy and robustness compared to traditional methods, particularly in indoor and obstructed environments.

However, existing approaches still face certain limitations. Many models are trained on limited datasets collected in controlled conditions, which affects their ability to generalize to real-world scenarios. Additionally, some methods rely on fixed architectures without exploring variations such as different pooling strategies, dropout regularization, or validation techniques.

Motivated by these gaps, this work explores multiple CNN architectures and training strategies to improve

the robustness and generalization of human detection using UWB radar data.

V. REQUIREMENTS SPECIFICATION AND METHODOLOGY

The proposed system is designed to detect human presence using Ultra-Wideband (UWB) radar signals by leveraging deep learning techniques. The system operates on radar data that is first converted into signal plots and categorized into two classes: human and non-human. These images serve as the primary input for the learning model. Before feeding the data into the network, preprocessing steps such as resizing and normalization are applied to ensure uniformity and stability during training. The system is capable of supporting multiple Convolutional Neural Network (CNN) architectures, allowing experimentation with different configurations such as varying the number of layers, pooling strategies, and the use of dropout for regularization.

To ensure reliable model evaluation, the dataset is divided into training, validation, and testing sets, with a portion reserved for final testing on unseen data. In addition, cross-validation techniques are used in certain configurations to improve robustness and reduce bias. The trained models are evaluated using standard performance metrics including accuracy, precision, recall, and F1-score, along with confusion matrices to better understand classification performance. The system also supports saving the best-performing model and provides visual outputs for easier interpretation of results.

From a design perspective, the system follows a structured pipeline that begins with radar data acquisition and progresses through preprocessing, model training, and evaluation. The dataset consists of radar reflections captured in both human-present and non-human scenarios, including variations in distance, angle, and posture to ensure diversity. These signals are transformed into image-like representations that capture spatial and temporal variations in the data.

The CNN models used in this work are designed to automatically extract features from these images. Each model typically consists of convolutional layers for feature extraction, followed by pooling layers to reduce dimensionality, and fully connected layers for classification. Dropout is optionally applied to

improve generalization and prevent overfitting. Different model variants are explored to analyze how architectural changes affect performance.

Training is carried out using supervised learning with a fixed number of epochs and batch size. In some cases, k-fold cross-validation is used to evaluate model consistency across multiple data splits. A separate test dataset is used at the final stage to assess how well the model generalizes to unseen data. The overall approach ensures that the system remains flexible, efficient, and capable of handling variations in radar signal patterns.

The implementation is carried out using Python and widely adopted libraries such as TensorFlow and Keras, along with supporting tools for data processing and visualization. The system is designed to run on standard hardware, with optional GPU support to accelerate training.

Overall, the methodology combines radar sensing and deep learning to build a practical and privacy-preserving human detection system, while maintaining flexibility for further improvements and real-world deployment.

VI. SYSTEM IMPLEMENTATION

The system is implemented using a structured pipeline that processes Ultra-Wideband (UWB) radar signal data and classifies it into human and non-human categories using deep learning models. The dataset consists of radar signal plots stored as images, organized into two classes representing human presence and non-human scenarios. These images are collected under different conditions such as variations in distance, angle, and subject posture, ensuring diversity in the dataset.

Each image is loaded using Python-based libraries such as TensorFlow, Keras, or OpenCV, and is resized to a fixed dimension of 128×128 pixels to maintain consistency across the dataset. The pixel values are normalized to a range between 0 and 1, which helps improve numerical stability during training. Labels are assigned automatically based on the directory structure, with one class representing human presence and the other representing non-human cases.

To ensure proper model evaluation, the dataset is divided into training, validation, and testing sets. In some configurations, a standard split of 80% for training, 10% for validation, and 10% for testing is used, while in others, k-fold cross-validation is applied to improve reliability and reduce bias. These approaches allow the system to evaluate performance under different training conditions and ensure better generalization.

The core of the implementation is based on Convolutional Neural Networks (CNNs), which are designed to learn patterns from the radar signal images. The models consist of convolutional layers that extract features, followed by pooling layers that reduce dimensionality and highlight important information. Fully connected layers are then used for classification, with a sigmoid activation function producing the final output for binary classification.

To study the impact of architectural design choices, multiple CNN configurations are implemented by varying the number of layers, pooling strategies, and the use of dropout for regularization. Some models use shallow architectures for faster computation, while others use deeper networks to capture more complex patterns. Dropout layers are selectively included to reduce overfitting and improve generalization.

Training is performed using supervised learning with a fixed number of epochs and batch size. The Adam optimizer is used for efficient gradient-based optimization, along with binary cross-entropy as the loss function. During training, validation performance is monitored to ensure that the model does not overfit and maintains consistent behaviour across different data splits.

Model performance is evaluated using standard classification metrics such as accuracy, precision, recall, and F1-score. In addition, confusion matrices are used to analyze the distribution of correct and incorrect predictions. These evaluations help in understanding how well the model distinguishes between human and non-human signals.

After training, the best-performing model is selected based on validation performance and tested on a separate dataset that was not used during training. This provides a realistic measure of how well the

model can generalize to unseen data. The final results are compared across different model configurations to identify the most effective architecture.

The implementation is carried out using Python along with widely used libraries such as TensorFlow, Keras, NumPy, and Scikit-learn. The system is designed to run on standard hardware, with optional GPU support to accelerate training and improve efficiency.

Overall, the implementation focuses on building a flexible and reliable system that combines UWB radar sensing with deep learning techniques. By exploring multiple model configurations and evaluation strategies, the system achieves improved performance while remaining adaptable for future enhancements and real-world applications.

VII. COMPARATIVE ANALYSIS OF MODELS

A total of eleven CNN-based models (M1–M11) were implemented to evaluate the effectiveness of different architectural and training choices for human detection using UWB radar signals. These models vary in terms of network depth, pooling strategy, use of dropout, dataset selection, and validation techniques.

The initial set of models (M1–M3) used shallow architectures with two convolutional layers and no dropout. Among these, M1 achieved the best performance with an average accuracy of approximately 85.99%, mainly due to the use of extensive k-fold validation along with a separate test set. In contrast, M2 and M3 showed similar accuracy levels around 76%, indicating that shallow models without regularization have limited ability to generalize.

Models M4 and M5 introduced deeper architectures with three convolutional layers and dropout regularization. M4, which used MaxPooling, performed better than M5, which used AveragePooling, suggesting that MaxPooling is more effective in preserving important features from radar signals.

Models M6 and M7 used a fixed train-validation-test split and achieved perfect accuracy on both datasets. These models highlight the advantage of combining

deeper architectures with dropout and structured data splitting.

Further comparisons were made using models M8 to M11. M8 and M9, which do not use dropout, achieved high accuracy but are likely to overfit, especially on smaller datasets. When dropout was introduced in M10 and M11, the models maintained high accuracy while improving stability and generalization.

Overall, the comparative study shows that deeper CNN architectures with dropout and MaxPooling consistently outperform simpler configurations.

VIII. PERFORMANCE EVALUATION

The performance of all models was evaluated using standard classification metrics including accuracy, precision, recall, and F1-score, along with confusion matrix analysis.

Dropout regularization was found to significantly improve model generalization. Models that included dropout (M4, M6, M7, M10, M11) showed more stable training behavior and better performance on unseen data compared to models without dropout.

The depth of the CNN also played an important role. Models with three convolutional layers were able to capture more complex patterns from radar signal images, resulting in improved classification performance.

Pooling strategy had a noticeable impact as well. MaxPooling consistently outperformed AveragePooling, as it preserved stronger feature representations that are critical for distinguishing between human and non-human signals.

Dataset size and variability further influenced performance. Models trained on Dataset 2 performed more reliably due to the presence of more diverse samples, whereas Dataset 1 models were slightly more prone to overfitting.

In terms of evaluation strategy, k-fold cross-validation provided robust internal validation, while fixed train-test splits offered a more realistic measure of performance in practical scenarios.

IX. RESULTS

The final results indicate that most models achieved high classification accuracy, with several

configurations reaching near-perfect performance. Models M6, M7, M8, M9, and M10 achieved accuracy close to or equal to 100%, along with very high precision, recall, and F1-scores.

Models M1, M2, and M3 achieved moderate performance, with accuracies ranging between approximately 76% and 86%, serving as baseline references for comparison.

Among all configurations, models M6 and M7 provided the most balanced performance, combining high accuracy with better generalization due to their use of deeper architectures, dropout regularization, and proper dataset splitting.

These results demonstrate that combining UWB radar data with deep learning techniques, particularly CNNs with appropriate architectural design and training strategies, can achieve highly accurate and reliable human detection.

X. CONCLUSION

This work presented a deep learning-based approach for human detection using Ultra-Wideband (UWB) radar signal data. The main goal was to develop a reliable and privacy-preserving alternative to conventional vision-based systems by using radar signals instead of visual inputs.

The study involved preparing a labeled dataset of radar signal plots and evaluating multiple Convolutional Neural Network (CNN) architectures under different configurations. A total of eleven models were implemented, varying in terms of network depth, dropout usage, pooling strategy, and evaluation method. This allowed a detailed comparison of how different design choices influence performance.

The experimental results show that deeper CNN architectures are more effective in capturing meaningful patterns from radar data. The inclusion of dropout further improves generalization by reducing overfitting, particularly when working with smaller datasets. It was also observed that MaxPooling consistently performs better than AveragePooling for this task, as it preserves stronger feature representations. Additionally, models trained on the larger and more diverse dataset demonstrated more stable and reliable performance.

Overall, the system achieved high accuracy in distinguishing between human and non-human radar signals, demonstrating the potential of combining UWB radar with deep learning techniques. The results confirm that radar-based human detection can serve as a practical and privacy-friendly solution in scenarios where traditional camera-based systems are limited or unsuitable.

This work provides a strong foundation for future improvements, including real-time implementation, larger datasets, and more advanced model architectures for enhanced performance in real-world applications.

XI. FUTURE SCOPE

Although the proposed system demonstrates strong performance in detecting human presence using UWB radar data, there are several directions in which this work can be extended.

One important improvement would be integrating the model with real-time radar data instead of relying on pre-generated signal plots. This would allow the system to perform live detection, making it more suitable for practical applications such as surveillance and healthcare monitoring.

The current approach is limited to binary classification. In future work, the system can be extended to support multiclass detection, such as identifying different human postures or distinguishing between multiple types of objects. This would significantly increase its usefulness in real-world scenarios.

Another potential enhancement is incorporating temporal information. At present, the model processes individual signal images, but radar data naturally contains time-based patterns. Using models that capture temporal dependencies, such as LSTM or 3D CNNs, could improve detection accuracy, especially for moving subjects.

From a deployment perspective, optimizing the model for edge devices would be a valuable step. Techniques such as model compression and quantization can enable the system to run efficiently on devices like Raspberry Pi or embedded platforms, making it more practical for portable and low-power applications.

Future research can also focus on improving robustness and security. While radar-based systems are inherently more privacy-friendly than cameras, it is still important to ensure that the model is resistant to noise, interference, or potential spoofing scenarios.

In addition, transfer learning using pretrained models could be explored to improve performance, especially when working with limited data. Training on larger and more diverse datasets will further enhance generalization and reliability across different environments.

Finally, improving model interpretability is another important direction. Techniques such as heatmaps or attention-based visualization can help explain how the model makes decisions, which is particularly useful for building trust in critical applications.

Overall, these improvements can make the system more accurate, efficient, and adaptable, enabling its use in a wider range of applications such as healthcare, smart environments, and safety systems.

REFERENCES

- [1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.
- [3] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [4] C. Szegedy et al., "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.
- [5] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.
- [6] L. M. Le Minh, H. D. Ngo, and A. Subramanian, "Deep learning for indoor people counting using ultra-wideband impulse radar," *IEEE Sensors Letters*, vol. 5, no. 4, pp. 1–4, Apr. 2021.
- [7] X. Zhong, H. Zeng, and Y. Zhang, "Real-time human motion recognition with CNN-LSTM using

- UWB radar,” *Sensors*, vol. 22, no. 7, pp. 1–15, 2022.
- [8] N. Bartolini, M. Maresca, A. Miceli, and G. Vessella, “Generalization of ML models on NLOS human detection with UWB radars,” *Sensors*, vol. 22, no. 12, pp. 1–17, 2022.
- [9] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005, pp. 886–893.
- [10] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001, vol. 1, pp. 511–518.