

# Classification of Room Occupancy using a Convolutional Neural Network

Alessandro Amella

April 21, 2024

## 1 Introduction

The objective of this work is to develop a machine learning model capable of accurately classifying the number of persons in a room based on 60 GHz signal transmissions. By analyzing delay-Doppler snapshots that represent reflections from targets, the model aims to distinguish between four classes: machine only (zero persons), one person, two persons, and three persons in the room.

## 2 Problem Description

The dataset consists of delay-Doppler snapshots in PNG format, where each file represents a two-dimensional snapshot of reflections from targets (humans/machines) at varying distances and velocities. The snapshots capture the delay and Doppler frequency shifts generated by the targets. However, real-world records may be distorted by noise, reflections, and missed targets. The task is to classify the number of persons present in the room based on these snapshots (see Figure 1).

## 3 Methodology

The model development process involved data preprocessing, designing a CNN architecture, and hyperparameters optimization.

### 3.1 Data Preprocessing

The training data was loaded from PNG images and preprocessed by converting to RGB, cropping to remove white borders, resizing to a consistent dimension, and normalizing pixel values. The data was then split into training and validation sets. Image augmentation techniques were explored in an attempt to improve model performance. Various augmentation methods, including rotation, flipping, and scaling, were applied to the training dataset to increase its diversity and robustness. However, despite these efforts, the accuracy value did not improve. This suggests that the original dataset may already contain sufficient variability for the model to learn effectively.

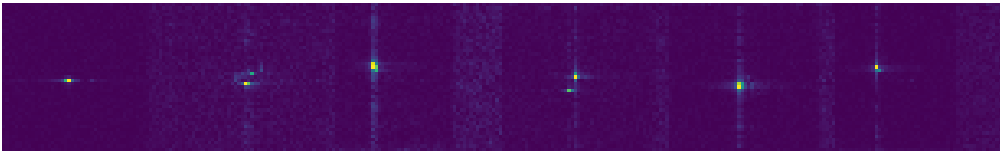


Figure 1: Examples of delay-Doppler snapshots for different room occupancy scenarios.

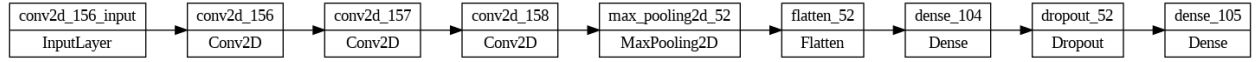


Figure 2: CNN architecture used for room occupancy classification.

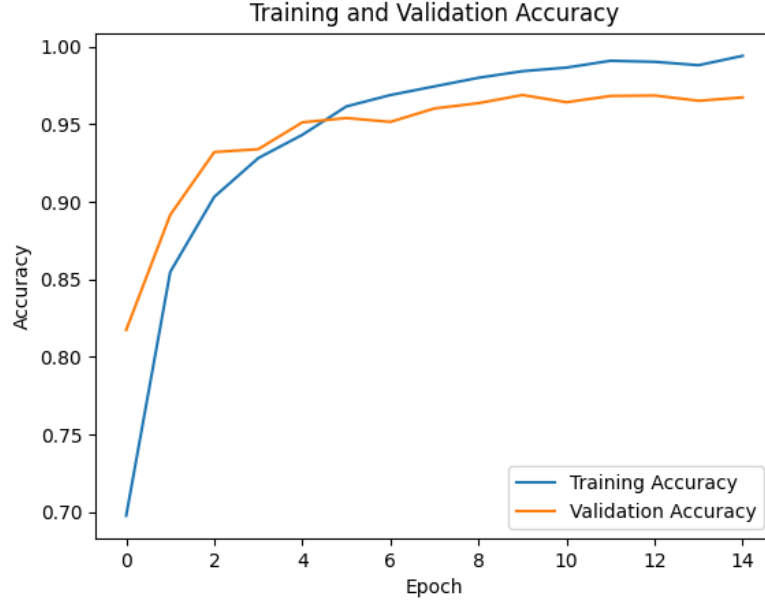


Figure 3: Training and validation accuracy curves during model training.

### 3.2 Model used

A CNN architecture was used for image classification. The model consists of multiple convolutional layers with ReLU activation, followed by max pooling, flattening, dense layers with ReLU activation, dropout regularization, and a final dense layer with softmax activation for multi-class classification (see Figure 2). This model proved to yield a high accuracy value.

### 3.3 Hyperparameter Optimization

Hyperparameter tuning was performed using Optuna to find the best combination of hyperparameters. The objective function defined the search space for the number of convolutional filters, dense units, dropout rate, and learning rate. Optuna optimized these hyperparameters based on the validation accuracy (see Figure 3).

## 4 Conclusion

The CNN model, optimized using Optuna, achieved a high validation accuracy of 96.94%. This work advances occupancy detection for applications in smart environments, security systems, and others. The source code is available on GitHub [here](#).