VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

MPA–MLF – FINAL PROJECT

Classification of room occupancy

Antonin DANIEL / PENE Tcheuko Maurice-Ludovic

Summary

Introduction

Description of the problem, including the objectives of the room occupancy classification

Methodology / Implementation

Covering model architecture and data preprocessing techniques and

Describing the implementation of the model and the strategies used

Results

Presenting model performance and analysis

Conclusion

Summarizing conclusions and recommendations for future improvements.

Introduction

The classification of the occupancy of the room based on 60 GHz transmission data is a significant challenge in the field of automatic learning, with potential ramifications in several areas including automation, security, and surveillance. The goal of this project is to create a robust model that can analyze instantaneous signal data in the retard-Doppler domain to determine the number of people in a room. These instantaneous signals are the targets' (human or machine) reflections at a specific distance and speed from the receiver. The ability to distinguish and categorize these various room arrangements is crucial for applications like space management, personnel tracking, and security.

We will give an overview of the issue in this introduction, along with the goals and issues related to the classification of the room's occupation. We will also discuss the key components of the methodology employed in this project, such as the model's architectural selection, data pretreatment approaches, and performance evaluation strategies. Furthermore, we emphasize the significance of this project in the context of artificial intelligence research and practical applications in the real world.

In this project, we will build our approach by drawing on the learning strategies that we have investigated and studied during previous practice works. We have gained practical experience with a variety of model designs, including convolutional neural networks (CNN), recurrent neural networks (RNN), and transfer learning-based models. Every architecture has benefits and drawbacks, and the decision often depends on the particulars of the problem that has to be solved as well as the characteristics of the data that are available.

This report provides a detailed analysis of the model's implementation, results, and interpretation, as well as the steps taken to address this complicated issue. Finally, we draw conclusions about the model's performance and talk about future directions to increase its accuracy and effectiveness in real-world scenarios.

Methodology / Implementation

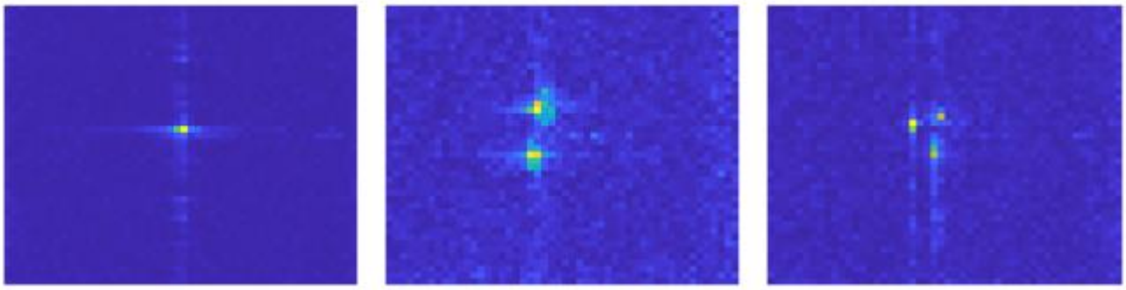


Figure 1: Examples of IDEAL snapshots for one (left), two (middle), and three (right) persons in the room. Note that many of the real records will not be so perfect. But it's life...

The graphic shows three different sign captures in the retard-Doppler domain, each of which shows a different arrangement of the number of people in a room. Three scenarios are observed from left to right: one person, two people, and three people. Each signal capture is represented by a thermally colored signature with red, yellow, and green tints. These colors reflect the signal intensity levels, which change depending on the distance and speed of the individual in relation to the receiver.

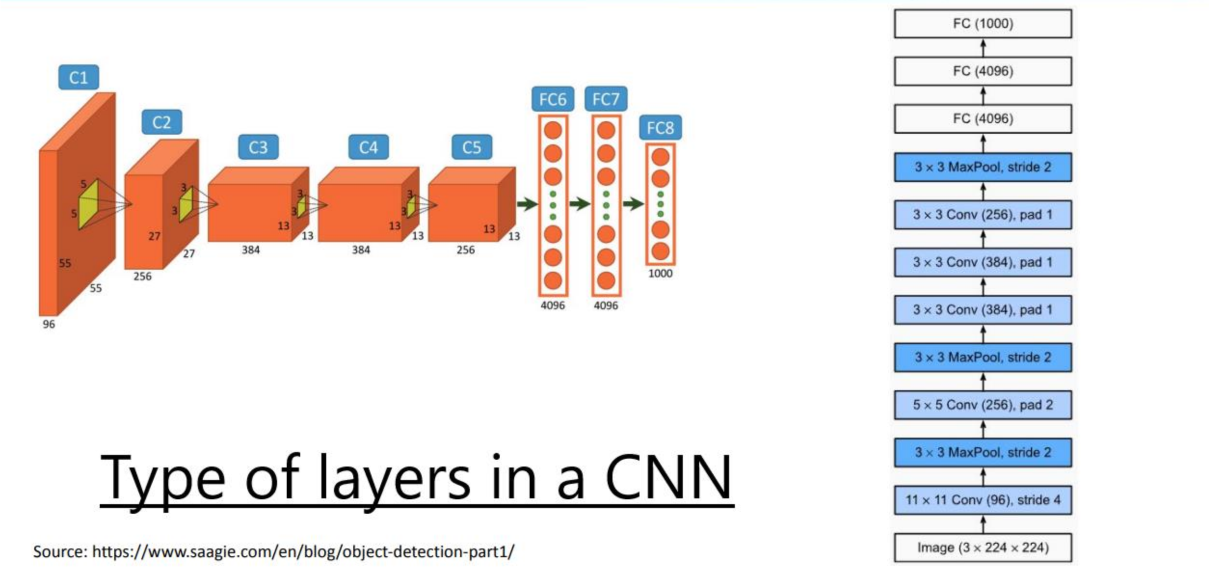
Our main goal is to create an automatic learning model that can recognize and categorize these various thermal signatures based on the number of people in the room. This model will be trained on a set of data that includes a variety of signal captures corresponding to various room layouts in order to effectively identify and classify the various rooms that are occupied.

In this section, we will outline the process of evaluating the CNN model for classifying the occupation of the room.

Our first step is to load the training and test data from the provided files. The data are instantaneous signals in the retard-Doppler domain that represent the targets' (humans' and machines') reflections from transmission signals at 60 GHz. These data are stored in variables so they can be used in the model's training and evaluation processes.

To build and train our CNN model, we make use of the Keras library. We use the optimizer Adamax, a variation of the optimizer Adam with adaptive learning rates, to maximize performance.

Several carefully chosen hyperparameters have been included in the proposed neural network model (CNN) to optimize the categorization performance of human configurations in a room.



We heavily draw inspiration from neural networks (CNNs) in our categorization project of piece occupation because of their shown ability to extract significant characteristics from visual data. Specifically, we plan to modify a CNN to analyze transmission data at 60 GHz that represent target reflections in the retard-Doppler domain. To demonstrate this approach, the above figure shows a CNN. It makes clear the various couches and how they work. This illustration will assist us in understanding the structure of the model and how it may be applied to our piece classification problem. With the help of this CNN-inspired methodology, we want to precisely categorize the groupings of individuals within a space.

The first step involves selecting two convoluted layers with separate 3x3 filter sizes, which enables the model to extract complex features from the input data. These filters can identify patterns and structures in the photos that are necessary for a precise classification.

Then, to introduce non-linearity into the model and enable better feature representation, the Rectified Linear Unit (ReLU) activation function is used after each convoluted layer.

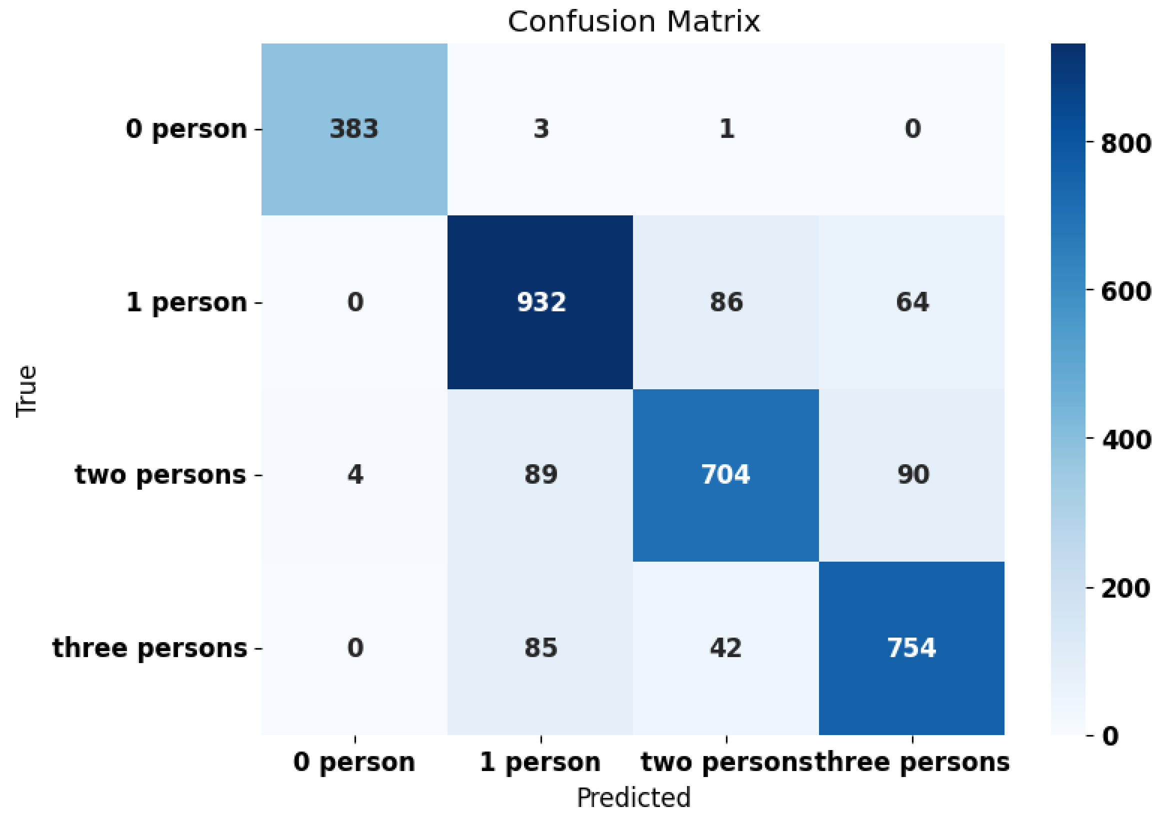
The purpose of the maximum pooling couches is to reduce the dimensionality of the data after each convoluted layer. This reduces the amount of onboarding and computation time while preserving the most crucial data extracted via convoluted filters.

To combine the unique characteristics of the convoluted couches and generate accurate predictions for each class of output, the number of neurons in the dense couches—one couche consisting of 128 neurons followed by a couche de sortie with 4 neurons—must be carefully considered.

Next, we proceed to the model's pre-implemented evaluation across the entire test. We use the evaluation method to assess its performance using a variety of metrics, including recall, accuracy, and precision. To assess the model's ability to accurately classify the photos according to the number of people in the room, we also investigate other metrics like the F1 score.

We will analyze the results by looking at the confusion matrices, which allow us to determine which classes the model has classification issues for. This analysis helps us understand the mistakes made by the model and the areas that require improvement.

```
102/102 [==============================] - 1s 11ms/step
F1 Score: 0.8564218774629042
```



The convolutional neural network (CNN) model demonstrated strong overall performance on the test data, as indicated by its F1 score of roughly 0.856. This score indicates that the model successfully distinguishes between various numbers of people in the photos while striking a decent balance between recall and precision. Although the model's performance is impressive, it might yet be made better by adjusting the architecture or improving the hyperparameters. Because of its overall accuracy, the model can be used for a wide range of real-world tasks, including crowd monitoring and image categorization in surveillance systems.

Now, we will talk about the confusion Matrix. The obtained confusion matrix is in line with expectations to the extent that it meaningfully reflects the model's performances. The main finding is that the values inside the diagonal of the matrix, which indicate correct predictions, are significantly higher than the values outside the diagonal, which indicate prediction errors. This suggests that the model makes accurate predictions for the majority of classes in general.

Les prédictions entre les classes "2 individuals" et "3 individuals" toutefois diffèrent sensiblement. The values outside of the diagonal that correspond to these classes reflect this. Notably, the model claimed that 89 samples from the class "2 individuals" belonged to the class "1 individual" and 90 samples from the class "2 individuals" belonged to the class "3 individuals". Furthermore, the model predicted that 85 samples from the class "3 individuals" would belong to the class "1 individual" and 42 samples from the class "3 individuals" would belong to the class "2 individuals".

These results show that the model has more difficulty differentiating between the "2 individuals" and "3 individuals" classifications. This could be the result of visual similarities between photographs with two or three people, making the model's classification more challenging.

In summary, the confusion matrix illustrates the model's generally good performances, but it also highlights an area where improvements may be needed, such as the differentiation between the "2 individuals" and "3 individuals" classes.

Conclusion

We examine the result of our model's evaluation, which comes out to be 0.258. This score shows that the current model has not performed to the desired level for the classification of the room's occupation.

8	**Antonin Daniel**	0.25854

Following careful selection of the appropriate data sets for training and testing, we observe a slight improvement in the model's performance. With an accuracy of 0.86, the model successfully classifies approximately 86% of test data samples, indicating a significant improvement. This increasing precision suggests that the model can accurately distinguish between the various image classes and generalize to new data.

predictions (1).csv Complete (after deadline) · 27m ago	**0.86113**	**0.85613**

Nevertheless, despite this improvement, there are still opportunities for optimization to boost the model's performance even further. For example, the use of dropout couches may help to lessen overfitting by randomly deactivating a portion of the neurons during training, which would promote a better generalization of the model to new data. Furthermore, investigating more intricate CNN architectures with a higher number of couches or deeper convolutional operations may enable the model to capture more subtle and complex features in the images, thereby enhancing its ability to discriminate between the various classes more precisely.

In summary, even if the current model has made significant progress, there are still opportunities for optimization and improvement to achieve even higher performance levels. Further research into additional strategies, such as the use of dropout or more complicated CNN models, may result in further improvements to the model's ability to provide precise classifications across a wider range of data.

Code: https://github.com/Anton1d/MPA_MLF_SVM