Delft University of Technology

EE4675
Object Classification with Radar

Human Activity Classification using Radar

Group - 3

Authors:
Yanqi Hong 5884683 Y.Hong-8@student.tudelft.nl
Alexandros Theocharous 5930901 a.theocharous@student.tudelft.nl

June 11, 2024

1 Introduction

The aim of this project is the classification of human movement from raw radar data. The dataset that was utilized was the "Radar signatures of human activities" [1] dataset that includes six different movements that people performed in front of a radar sensor under steady conditions.

While in the literature it is common for such problems to be tackled with machine learning algorithms that utilize hand crafted features [2], such as SVMs or Random Forests, we decided as a group to implement a deep learning approach. CNNs are a usual way to extract information from images and even though radars provide lists of data, according to their measurements, those lists can be transformed into different plots. Potential plots are the Range-Time plot, the Range-Doppler plot, the Time-Doppler plot (Spectogram) and the phase of the signal. Publications exist, where CNN where applied onto the Spectogram and the Phase plots [3].

2 Approach

Our approach uses a combination of those two methods. While it applies a CNN on the images it uses a Constant False Alarm Rate (CFAR) detector to filter out the noise from the plots so only useful information passes to the trained model. Additionally since CNNs take RGB images as inputs, we exploited this feature by encoding the different plots into a different image channel, as it can be seen in Figure 1. This allows the model to train on the different plots created from a single measurement simultaneously and correlate them, without losing the unique spatial information in each one of them.
The images after being filtered, they are also normalized from 0 to 1 but also resized to 256 by 256 to be compatible with the CNNs input format.

Using this data processing method all the different combinations of the Range-Doppler, Doppler-Time, Range-Time plots where tested, both training on a custom CNN and transfer learning on the MobileNet V2 and ResNet V2. Our custom CNN consists of 6 Convolutional 2D layers, one Drop-out layer and two Dense ones for the final classification.

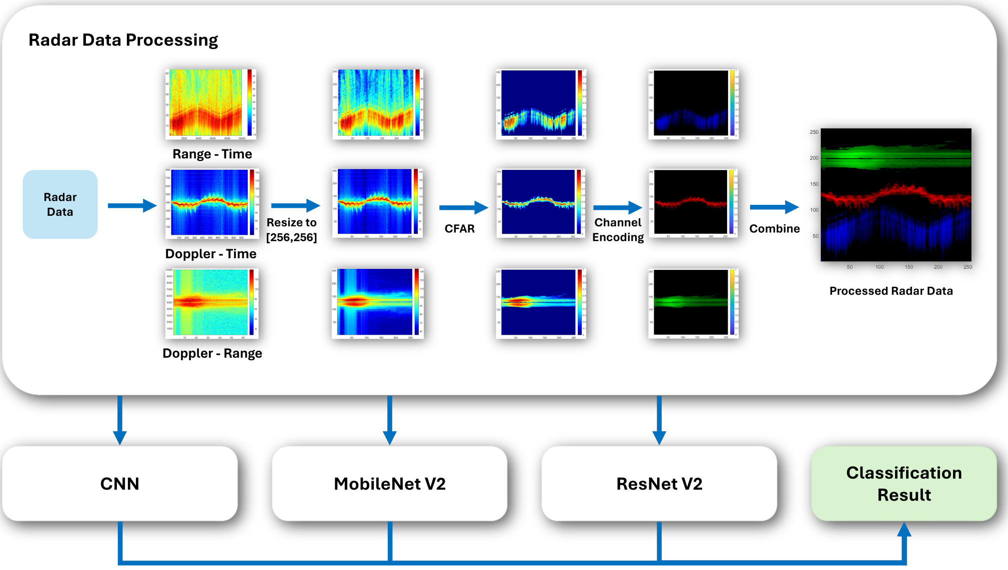

Figure 1: Workflow Overview

3 Results

After training the different models and evaluating on a unique test set that was not seen by the model during the training or the validation time, we recorded the accuracy scores and the most interesting observations are the following:

1. The custom CNN had he highest accuracy when all three domains were used: 95.01%.
2. The second highest performance was he spectogram alone, with: 94.53%. We believe that this is due to the simplicity of the dataset and the prepossessing method, as the only two movements that are being significantly mixed 25% are the drink and pick, which is also a problem mentioned in the papers that use more complex models than ours [4]. In more complex datasets we speculate that the difference between this and the model with three plots could be higher.
3. This inability to distinguish those two movements is due to the ambiguity of the task. Drink and pick can be interpreted in ways that make the final movements very similar if no specific instructions are given. If this mix of movements happens constantly by the participants the dataset can be deemed problematic.
4. The pre-trained Neural Networks: MobilenetV2 and ResnetV2 had 88.17% and 80.15% accuracy respectively. We believe that this is due to the high complexity of the model, or the fine tuning that was not done optimally by us.

4 Future Work

Due to time constrains we could not implement everything we wanted within the given deadline. Some things that we would like to additionally investigate in the future are:

1. Try the channel encoding method in a more complex dataset, or a benchmark dataset, to see how it generalizes and if the three plots will beat the accuracy of the spectogram alone by a meaningful margin.
2. Try to fine-tune the benchmark models in a similar way as they've done in literature and see if the channel encoding method improves their accuracy scores.
3. Find the heatmap of important features for our model, to see which part of the plots are actually important and build an intuition around meaningful plot characteristics.

References

[1] Fioranelli, F. , Shah, S. A. , Li, H., Shrestha, A., Yang, S. and Le Kernec, J. (2019) Radar signatures of human activities.

[2] Taylor, W., Dashtipour, K., Shah, S. A., Hussain, A., Abbasi, Q. H., & Imran, M. A. (2021). Radar sensing for activity classification in elderly people exploiting micro-doppler signatures using machine learning. Sensors, 21(11), 3881.

[3] Han, H., Kim, J., Park, J., Lee, Y., Jo, H., Park, Y., ... & Park, S. (2019, March). Object classification on raw radar data using convolutional neural networks. In 2019 IEEE Sensors Applications Symposium (SAS) (pp. 1-6). IEEE.

[4] Li, Z., Liu, Y., Liu, B., Le Kernec, J., & Yang, S. (2024). A holistic human activity recognition optimisation using AI techniques. IET Radar, Sonar & Navigation, 18(2), 256-265.