III-6

Application of Convolutional Deep Neural Network for Human Detection in Through the Wall Radar Signals*

Dalius NAVAKAUSKAS[1] Julius SKIRELIS[1] Eldar ŠABANOVIČ[1] Mantas KAZLAUSKAS[1]
Borisas LEVITAS[2] Irina NAIDIONOVA[2] Michail DROZDOV[2] Aleksandr PRISIAŽNYJ[2] Matvei KAZHAROV[2]

E-mail: dalius.navakauskas@vilniustech.lt, [1]Department of Electronic Systems, Vilnius Gediminas Technical University, [2]JSC "Geozondas"

Introduction

Ultra-wideband radars are capable of detecting people through walls or ruins in tactical and rescue operations. Frequency-modulated continuous-wave radars with at least one transmitting and two receiving antennas can provide information about the distance, angle and speed of objects. Typically, the signals are pre-processed using FFT to obtain 3D spectrograms in which the position of a human can be very difficult to discern because multiple moving human body parts can move in different directions and at different speeds. Radar signal spectrograms can be considered as images in which people's patterns can be detected using Convolutional Deep Neural Network (CNN).

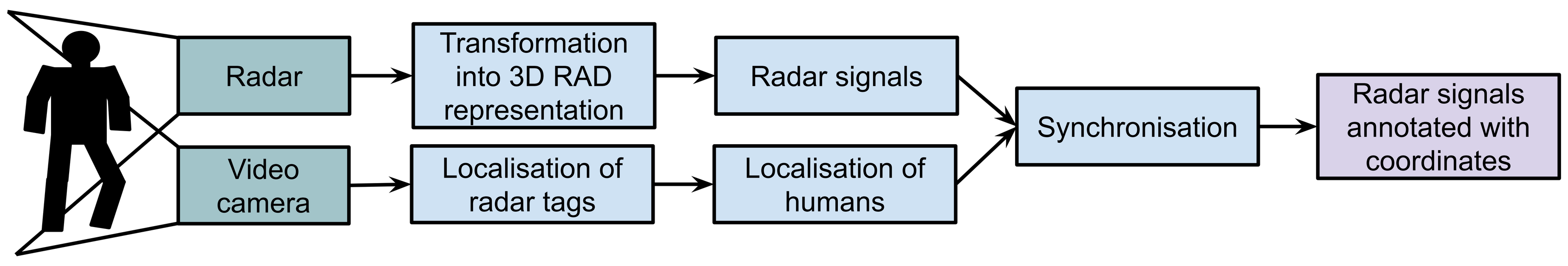
Aims

The goal of the study – to implement through-wall imaging radar signal processing using CNN.
The main objectives:

- form a radar signal dataset annotated with coordinates by using video and radar signal measurements of people in real-life situations that are performed in an experimental room;
- create CNN for human detection in frequency-modulated continuous-wave through-wall imaging radar signal processing;
- implement real-time CNN processing of radar signals on a gateway wirelessly connected to the radar.

Dataset Creation Process

To form the dataset, each frame of radar signal data is annotated based on the coordinates of the detected humans. In order to automate tagging, a process of data collection and processing has been developed, during which radar and video signals are recorded in parallel.



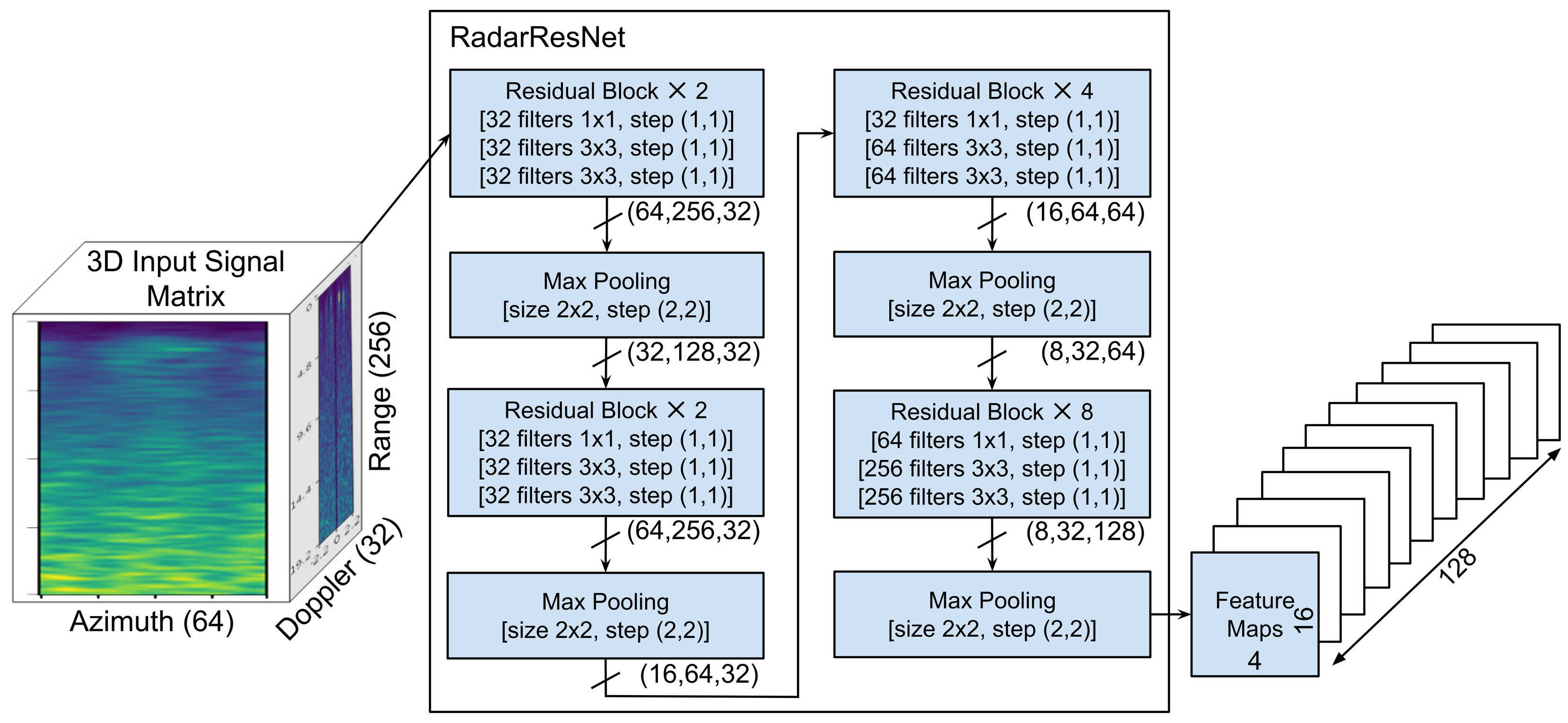
Positions of markers on people in space relative to the radar are determined and calculated. To determine the coordinates of a person, one must put a visual ArUco tag of size 20 cm^2 on the chest, back, and head, and a size of 10 cm^2 on the shoulders. For the synchronisation flag, waving was used with the metal sheet that is visible in the visual data and the radar signals.

CNN Implementation

The collected dataset was used with a modified version of the RADDet neural network based on ResNet backbone and YOLO detection heads for real-time processing of radar signals on an embedded computer – gateway, wirelessly connected to the radar.
Main changes to the RADDet:

- the Range-Azimuth-Doppler (RAD) 3D matrix dimensions are set to 256 for the range, 64 for azimuth, and 16 for doppler;
- set one output class – human;
- data loading was changed to support our specific data formats.

*Funded by a Agency for Science, Innovation and Technology grant No. TPP-04-032.

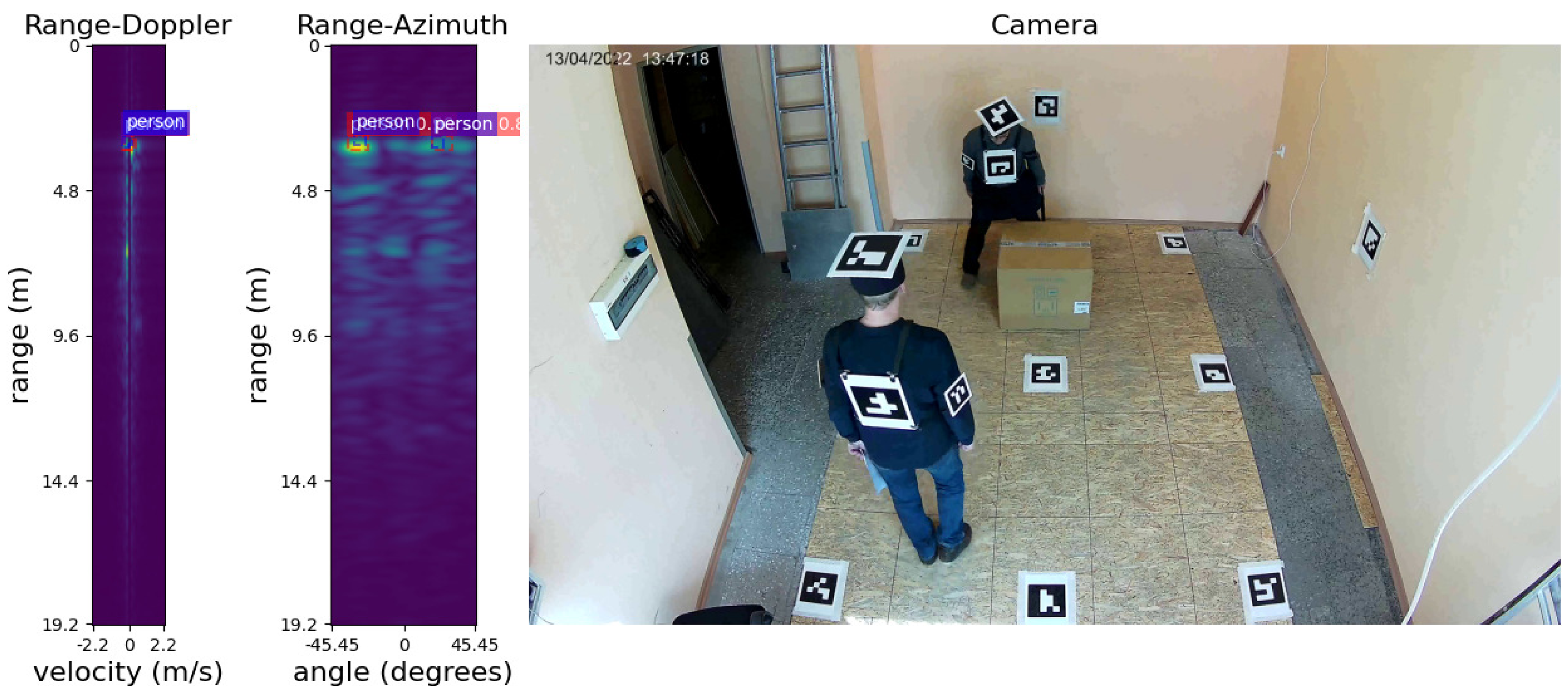


Modified RADDet CNN structure

Best obtained RADDet CNN training mean average precision was around 0.4.

Results

Final dataset consists of 291 synchronised records and over 1.5 million radar signals and RAD representation samples with annotations of human positions.



Modified RADDet CNN human detection with camera view of experimental room

Average duration of radar signal processing using different means

Process item	Radar	Radar & gateway	Two radars & gateway
Single chirp	94.7 ± 5.0 μs	50.70 ± 2.00 μs	55.10 ± 2.00 μs
RADDet CNN recall	99.2 ± 2.7 ms	14.20 ± 0.41 ms	21.50 ± 0.51 ms
Complete data frame	107.7 ± 2.0 ms	16.70 ± 0.43 ms	27.70 ± 0.55 ms

Processing two streams of data makes a more efficient use of gateway computational resources, as the data frame and CNN model processing time are less than doubled.

Conclusions

1. The modified RADDet CNN for human detection in through-wall imaging applications was implemented in a gateway ensuring 6–7 times faster detection compared to a single radar use.
2. An automated process of video and frequency-modulated continuous-wave radar signal acquisition was created based on human position detection from ArUco tags.
3. The unique dataset of Range-Azimuth-Doppler representation samples with human position annotations was created.

VILNIUS TECH
Department of Electronic Systems