2025

Classification 5G base stations

MPA-MLF - FINAL PROJECT

JAKUB PACHEL, TOMÁŠ PAULYSKO

1. Introduction

Cellular networks are a common and important technology in today's world. For example, according to the Czech Statistical Office, 99% of people over the age of 16 in the Czech Republic use a mobile phone [1]. To make sure the network works for so many users, populated areas need to be covered by base stations called eNodeB or gNodeB.

But not all base stations are real or safe—some can be set up by attackers who take advantage of weak spots in mobile network security. With the help of special hardware and software, these fake stations can listen to private communication, send harmful messages, or follow the location of users. One well-known device used for this is called a False Base Station (FBS). It pretends to be a real base station and tries to make mobile phones connect to it.

In this project, our team worked on classifying the type of signal that was received. The goal was to decide whether the signal came from a real base station owned by T-Mobile in a nearby building (class 0), or from a fake station placed inside the building by an attacker. In this case, there are two fake stations: class 1 means the first location, and class 2 means the second location.

2. Data

We use a simple pipeline to prepare our data. First, each frame is randomly shifted by up to 2 % in both directions, scaled by 0.98–1.02, and has Gaussian noise added (σ=0.01). This augmentation boosts diversity and helps generalization. Next, all samples are normalized. We then split the data 80/20 into training and validation sets with stratified sampling for reproducibility. Finally, we add a channel dimension to each frame and compute class weights from the training labels so that rarer classes are weighted more strongly during training. An example of such a frame can be seen in *Figure 1*, which shows a random sample from the dataset.

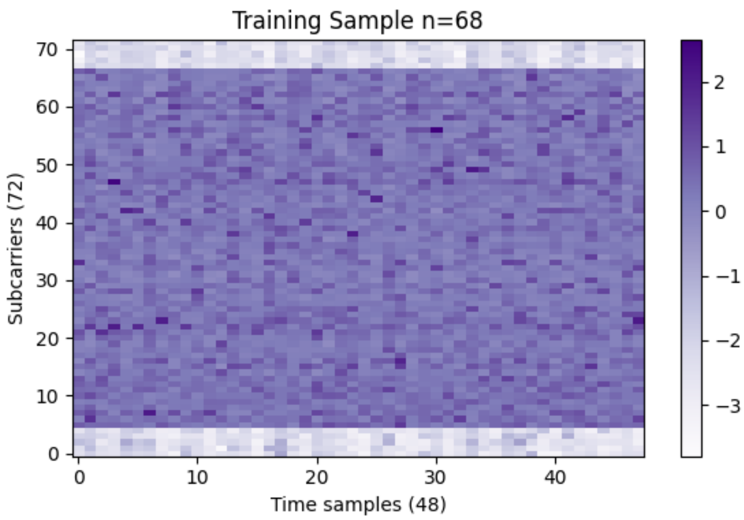

Figure 1 Data example

3. Model

We used a convolutional neural network made of three convolutional stages, each with 32, 64, and 128 filters of size 3×3 and L2 regularization (0,0001). After each convolution we applied 2×2 max pooling and batch normalization. A global average pooling layer followed these stages to reduce each feature map to one value. We then added a fully connected layer with 64 units and ended with a softmax layer matching the number of classes. The model was trained end-to-end with the Adam optimizer (learning rate set as in our config), using sparse categorical cross-entropy loss and accuracy as our metric. The layer diagram is in *Figure 2*. For training, we set 120 epochs and a batch size of 8.

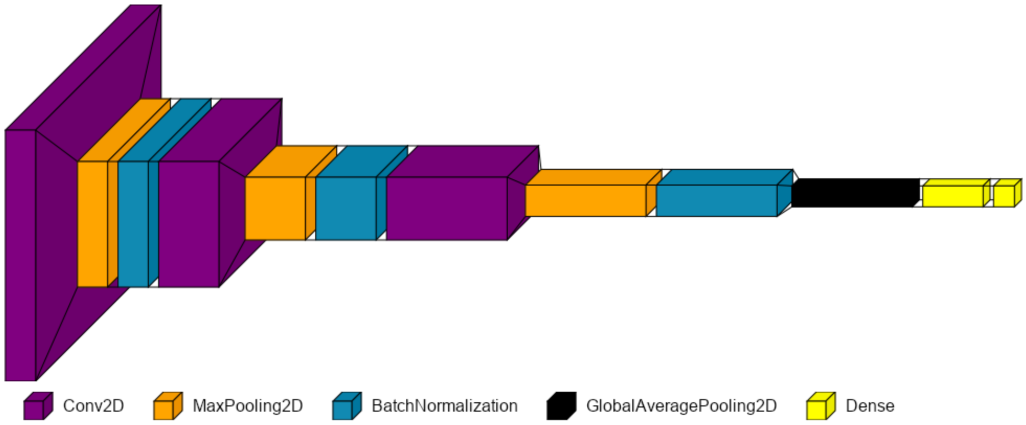

Figure 2 CNN structure diagram

4. Model evaluation

The graphical representation of the model's accuracy and loss throughout the training process is provided in *Figure 3*. These graphs illustrate how the model's performance evolved with each epoch. As training progressed, the model gradually improved, ultimately reaching a final accuracy of 100%. Simultaneously, the loss decreased to a final value of 1.87%.

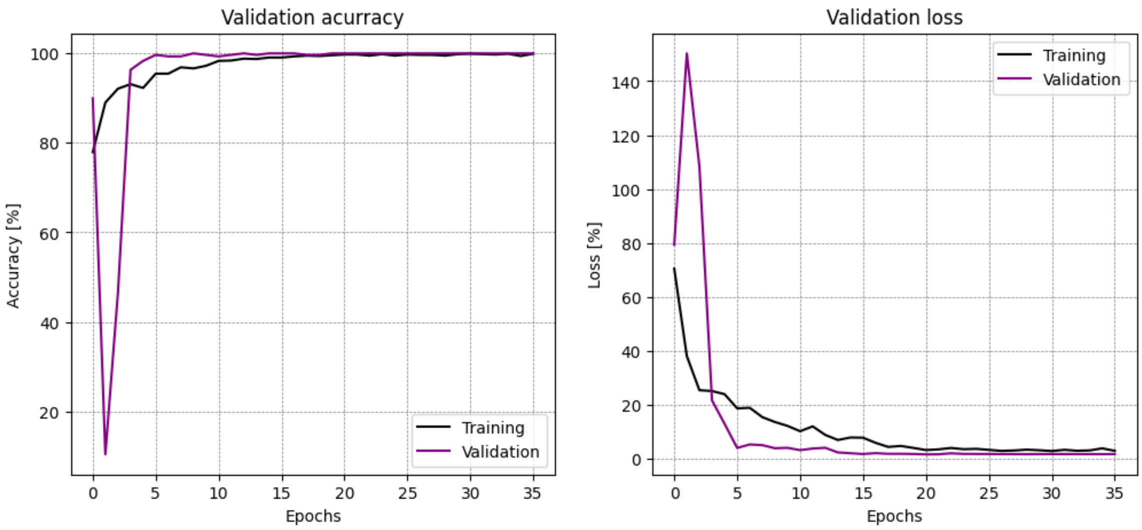

Figure 3 Accuracy and loss graphs

Figure 4 is the confusion matrix for the validation set. It shows 243 correct predictions for class 0 and 28 correct predictions each for classes 1 and 2, with zero entries off the main diagonal. In other words, the model made no errors on the validation data. When we submitted our results to Kaggle, we achieved a score of 0,99166 %. Our Jupyter Notebook code is uploaded to GitHub.

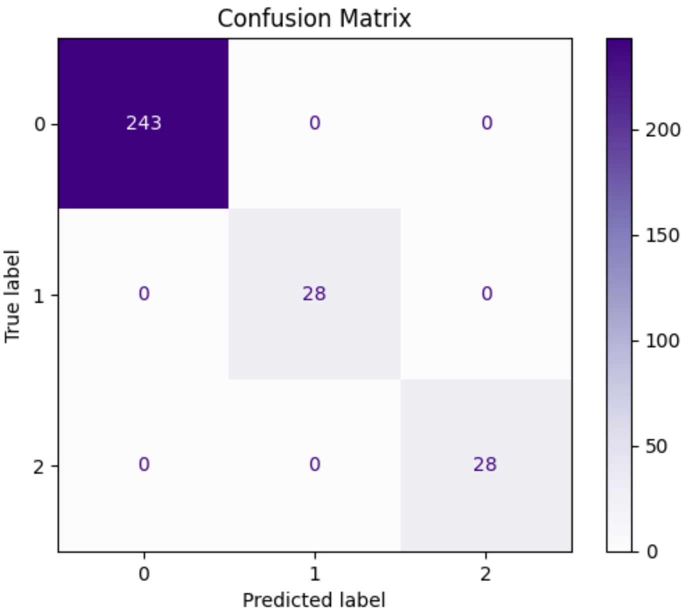

Figure 4 Confusion matrix

5. Conclusion

Our simple CNN scored 99.166% on the Kaggle public leaderboard, putting us among the top teams. This means it only misclassified one base station in the test set. The result shows that using three convolutional layers with L2 regularization and batch normalization works very well. Even with a fairly small model and standard training steps, we reached almost perfect accuracy.

6. References

[1] V Česku používá chytré telefony již 82 % osob. Online. In: Https://csu.gov.cz/. 09. 11. 2023. Available from: https://csu.gov.cz/produkty/v-cesku-pouziva-chytre-telefony-jiz-82-osob. [cit. 2025-05-03].