ZIB Challenge

David Nolte

November 30, 2021

Docker containers

Two `docker` containers are created and connected via a `docker` network: one container holding the database (`MongoDB`, populated by the script `docker_db/populate_db.py`) and the other containing the `JupyterLab` environment for deep neural networks (with `tensorflow/keras`) and the notebooks for solving the given task (see folder `docker_jl`).

The repository's `README.rst` explains the procedure.

ECG Dataset

The dataset is a subset of the large ECG signal PTB-XL data set [1], considering only those ECG timeseries where the cardiac rhythm diagnosis was marked *sinus rhythm* (SR) or *atrial fibrillation* (AFIB). The dataset is partitioned into 10 folds, where the first 8 are recommended by the authors of [1] for the training of a neural network, and 9 and 10—the data with the highest confidence in the diagnosis—for testing and validation. All timeseries are normalized such that their be 0 and the standard deviation be 1.

Deep Neural Network

A deep neural network is trained to classify the ECG time series into 'SR' and 'AFIB'. A residual neural network (ResNet) is used, following [2, 3]. The implementation uses the `keras` library. See the code in the `jupyter` notebook `ecg_resnet.ipynb` and the references [2, 3] for details on the setup. Conveniently, this setup can be run on an old laptop without GPUs.

Results

The accuracy of the ResNet w.r.t. the training and the validation data over 75 training iterations is illustrated in Fig. 1.

The test dataset is classified with an accuracy of 0.9787. A useful metric for the quality of the trained network is the confusion matrix, showing the number of correctly and erroneously classified sets, see Fig. 2. 26 ECG signals labelled SR are wrongly classified as AFIB, and 13 AFIB data are wrongly predicted as SR.
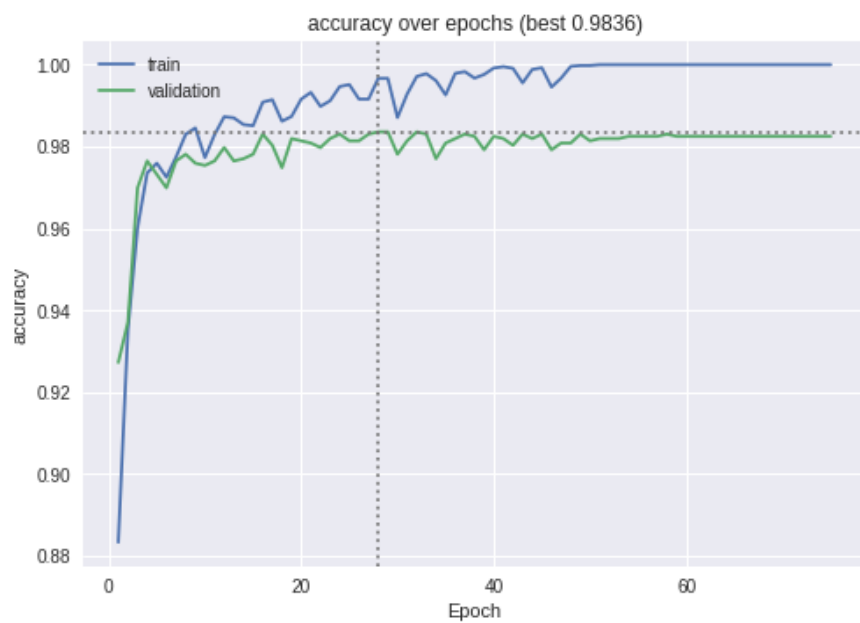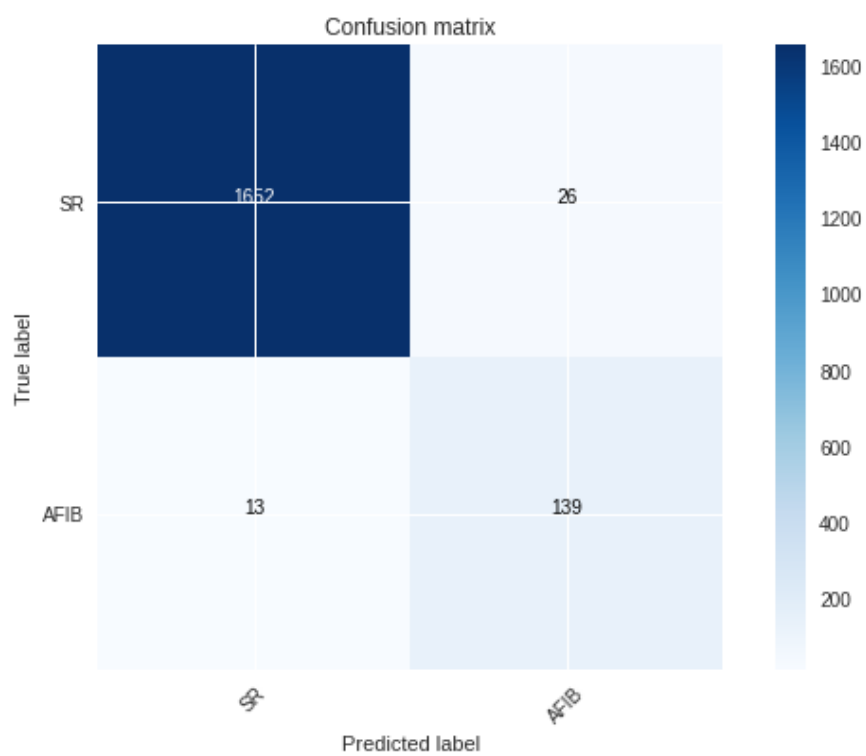
Figure 1: ResNet accuracy with training data

Figure 2: Confusion matrix for testing dataset

References

[1]   P. Wagner et al. "PTB-XL, a Large Publicly Available Electrocardiography Dataset." In: *Scientific Data* 7 (May 25, 2020), p. 154. ISSN: 2052-4463. DOI: 10.1038/s41597-020-0495-6. pmid: 32451379. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7248071/ (visited on 11/30/2021).

[2]   M. Kachuee, S. Fazeli, and M. Sarrafzadeh. "ECG Heartbeat Classification: A Deep Transferable Representation." In: *2018 IEEE International Conference on Healthcare Informatics (ICHI)*. 2018 IEEE International Conference on Healthcare Informatics (ICHI). June 2018, pp. 443–444. DOI: 10.1109/ICHI.2018.00092.

[3]   spdrnl. *Ecg*. Aug. 31, 2021. URL: https://github.com/spdrnl/ecg (visited on 11/30/2021).