HUMAN ACTIVITY RECOGNITION

Team 06: Shengbo Wang, Junmao Liao

Institute of Signal Processing and System Theory, University of Stuttgart, Germany

1. Input Pipeline

- HAPT Dataset
  - HAPT raw data with tri-axial acc and gyro data(50 Hz).
- Preprocessing and Augmentation
  - Remove unlabeled data from the dataset.
  - Z-score normalization for multi-channel data.
  - Sliding window for data augmentation: window size of 250 samples with 50% overlap.
  - Create TFRecords files for a Sequence-to-Sequence task.

2. Model and Training

We choose GRU- and LSTM-based models, sparse categorical cross-entropy as loss function and Adam optimizer to train network for 10000 steps, and save the checkpoints with the best validation accuracy.

- Hyperparameter Optimization for a GRU-based model
  - The model with relative more GRU layers and more units tends to have a better performance.
  - The best window size is still 250 with 50% overlap.

GRU layers	2	2	1	1	1
Dense layers	3	3	2	1	1
GRU units	512	128	32	256	256
Dense units	256	64	128	128	64
Window size	250	250	250	100	250
Shift size	125	125	75	50	125
Dropout rate	0.471	0.387	0.566	0.454	0.248
Val. accuracy	0.929	0.908	0.851	0.858	0.884

Table 1: *Hyperparameter Optimization*

- Try different kernel initializers and select the three best ones

He_normal	Glorot_uniform	Glorot_normal
0.930	0.929	0.941

Table 2: *Kernal initializers and test accuracy*

- Ensemble the models to reduce generalization error by averaging the predictions

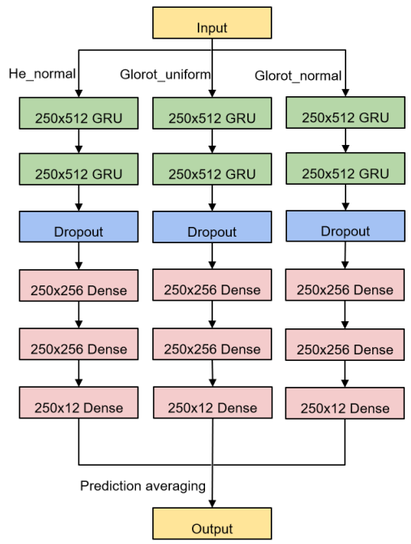

Figure 1: *Ensemble models*

3. Evaluation

We evaluate our ensemble model on test dataset by using:

- Visualization of the comparison between original label and predictions
  - High degree of correspondence proves the generalization ability of our model.

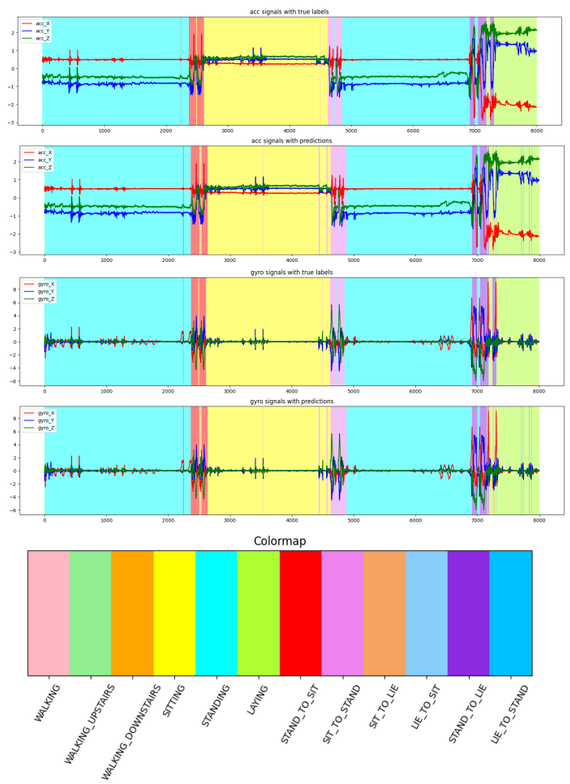

Figure 2: *Test labels and predictions*

- Confusion matrix
  - The performance on the first 6 classes(static activities) is satisfied, while the predictions for the last 6 classes(postural transitions) can still be improved.

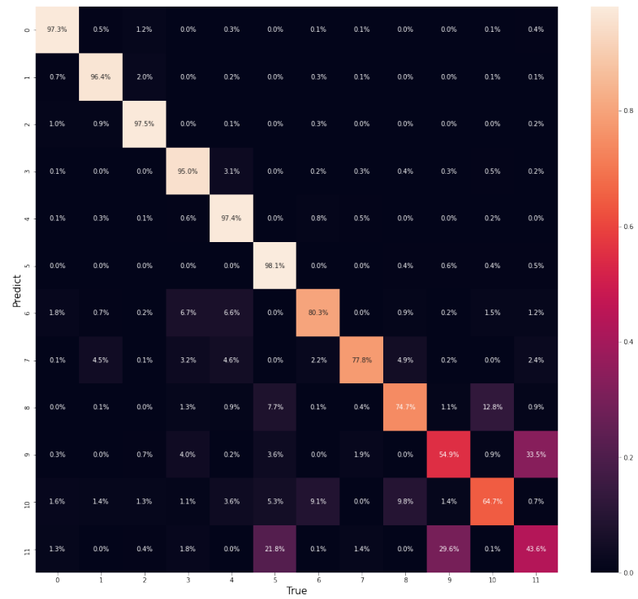

Figure 3: *Normalized Confusion Matrix*

- The best test accuracy of our model reached 94.48%

4. Results and Conclusion

Overall, we perform a GRU-based model with ensemble learning and achieve a better result than LSTM.

Architecture	LSTM	GRU
Test Accuracy	91.73%	94.48%

Table 3: *Results of ensemble learning*

Deep Learning Lab 2020, February 09, Stuttgart, Germany