Deep Learning based Sensor Fusion for 6-DoF Pose Estimation

Michael Sorg

Final: Master Informatics

Supervisors: Adnane Jadid, Christian Eichhorn

Thesis is done at Daimler Protics in Ulm

Motivation

Motivation

Motivation

Problem Description: Issues

- Real-time rendering with 60 FPS
- Different sensor types (GPS, IMU, Odometry)
- Different sensor update rates (1 – 100 Hz)
- Low-cost sensors

- GPS outages
- IMU sensor drift

Existing Solutions / Related Work

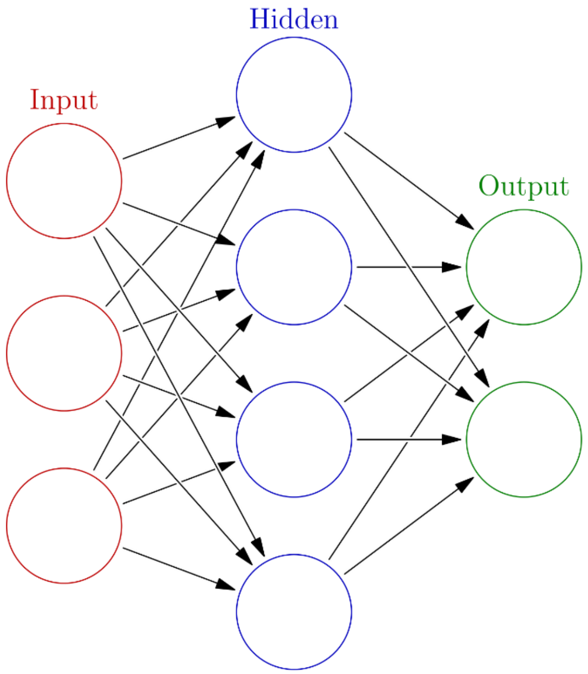
- Kalman filter is state of the art

- Kalman filter has several downsides:
  - Requires sensor error model (prior knowledge)
  - Noise is complicated, non-linear, correlated over time and can differ from sensor to sensor
  - Linearization of non-linear motion and sensor models leads to inaccuracies

Goal of this Thesis

- Evaluate if deep learning works for 6DoF pose estimation (feasibility study, proof of concept)

Approach



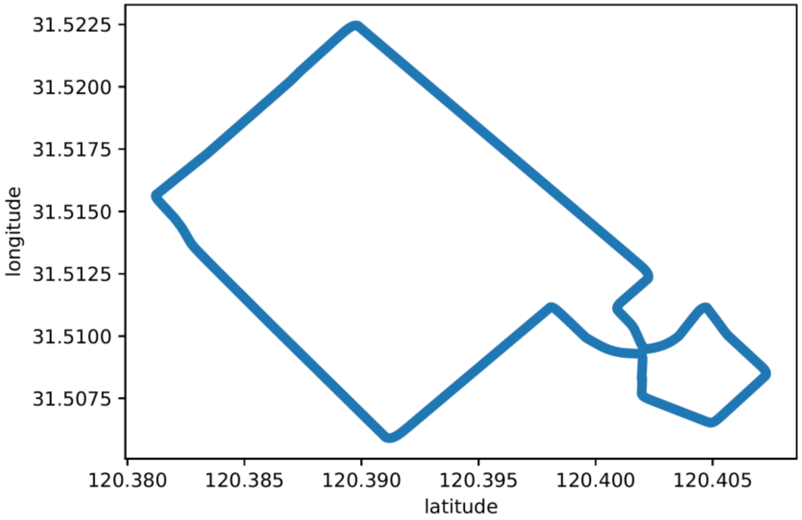
- 3 types of neural networks:
  - Dense (fully connected)
  - Recurrent neural network (RNN)
  - Convolutional neural network (CNN)

- Combinations possible

- Goal: Find a suitable model architecture

Approach

- Use Simulation data
  - less noise / controllable noise
  - more data
  - 100 Hz GPS measurements
  - no outages, jitter etc.



- Simulation data to simplify the problem and get first insights

Architecture search

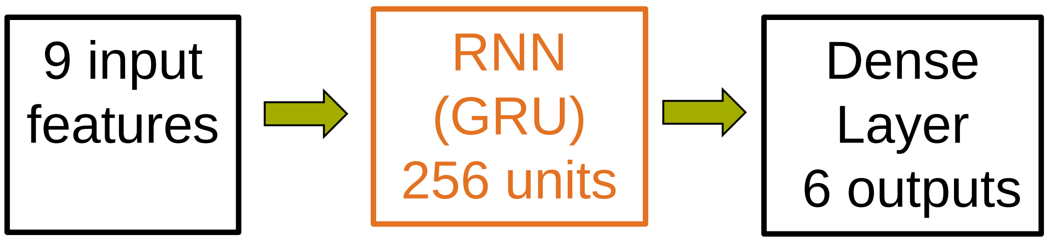
- Grid search for finding a good performing model

- RNN: LSTM, GRU
- CNN: Number of filters, filter size
- Batch normalization, Activation Layer, Dropout, regularization
- Loss function, Optimizer, batch size

Architecture search

- Best results with 5 – 10 seconds of the past as input

Best network architecture:



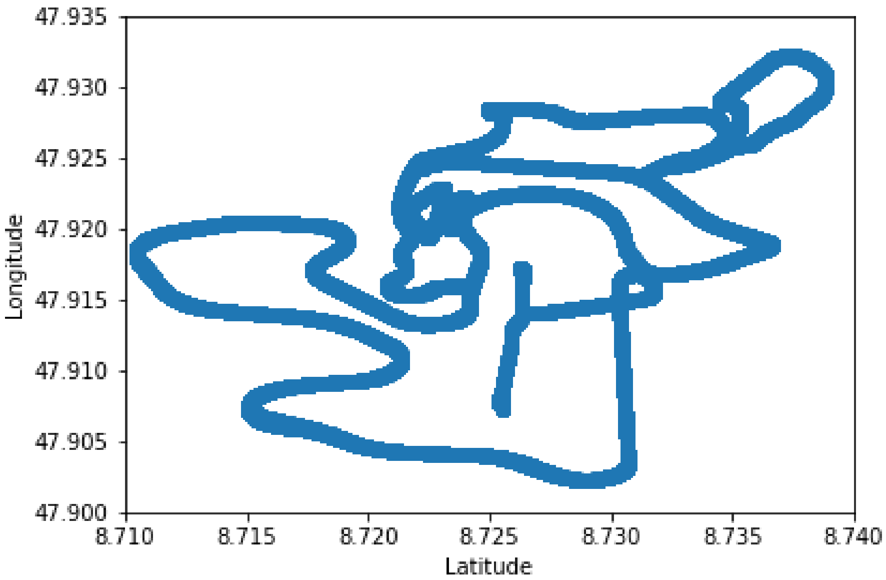
	Position MAE	Orientation MAE
RNN	3,5 m	1,5°

Increasing Dataset Size

	Position MAE	Orientation MAE
12 min.	15 m	120°
24 min.	9 m	16°
1 h	6 m	4,6°
2 h	4 m	1,8°
5 h	4 m	2,0°
10 h	2,5 m	0,9°
15 h	2 m	0,7°
20 h	0,95 m	0,3°

Training on Real Data

- 1,5 hours of data
- 100 Hz high quality ground truth data
- 1 Hz GPS
- Odometry data available (wheel ticks, wheel angle)



Training on real data

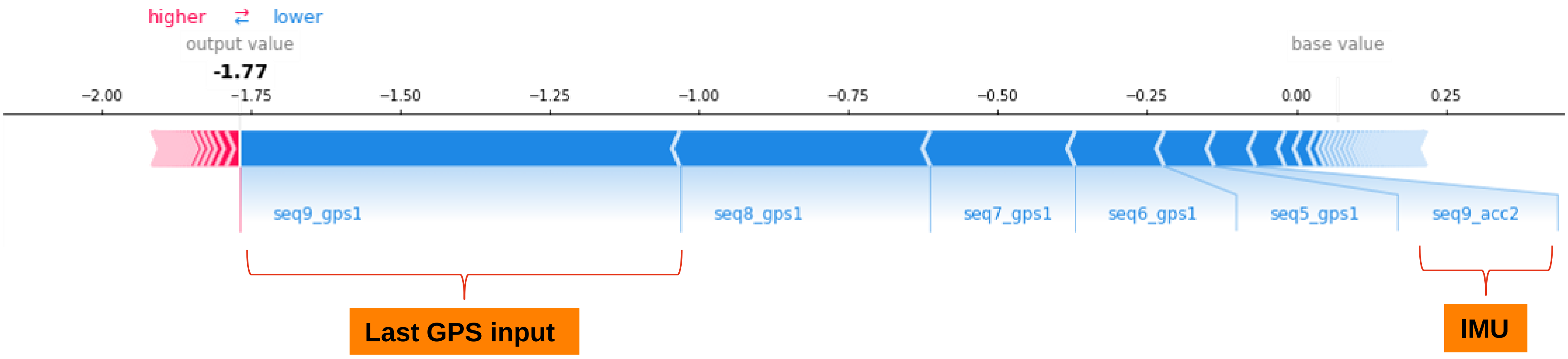
- Architecture search
- Results similar
- Best model was again GRU(256)

	Position	Pitch	Roll	Yaw
RNN	3,6 m	0,12°	0,22°	0,15°

Problems

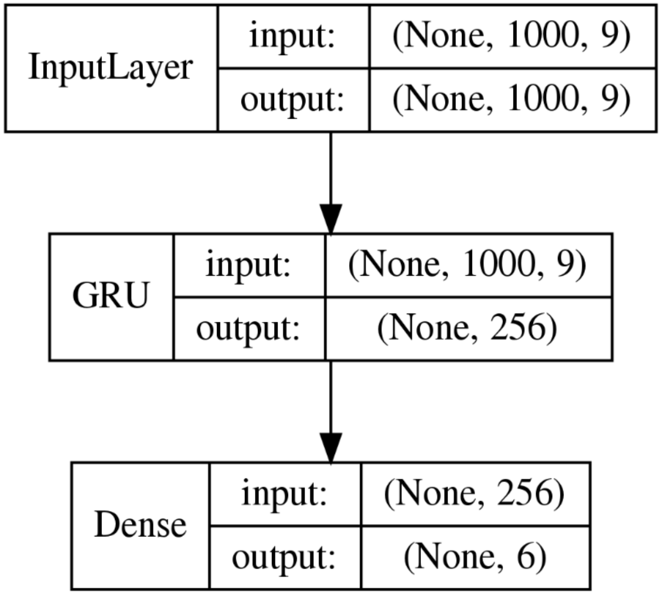
- *Explainable AI*: Explain model predictions

- Problem: Model only relies on GPS inputs for position prediction



Conclusion

- Predicting the orientation works even with a small dataset very well
- Accurately predicting the position needs a lot of data (more than 20h)
- Major Problem: Model only uses GPS inputs for position prediction



Outlook

- More data
- Use Kalman filter as preprocessing step
- *Deep Kalman Filter*
  - Sensor error is estimated by neural network

S. Hosseinyalamdary. “Deep Kalman filter: Simultaneous multi-sensor integration and modelling; A GNSS/IMU case study.” In: Sensors (Switzerland) 18.5 (2018)