FACULTY OF SCIENCE
PAVOL JOZEF ŠAFÁRIK UNIVERSITY
IN KOŠICE

ics.upjs.sk/indora

PAVOL JOZEF ŠAFÁRIK UNIVERSITY
IN KOŠICE

Building Parts Classification using Neural Network

Miroslav Opiela, Viktória Mária Štedlová, Šimon Horvát, Ľubomír Antoni, Lucia Hajduková

miroslav.opiela@upjs.sk

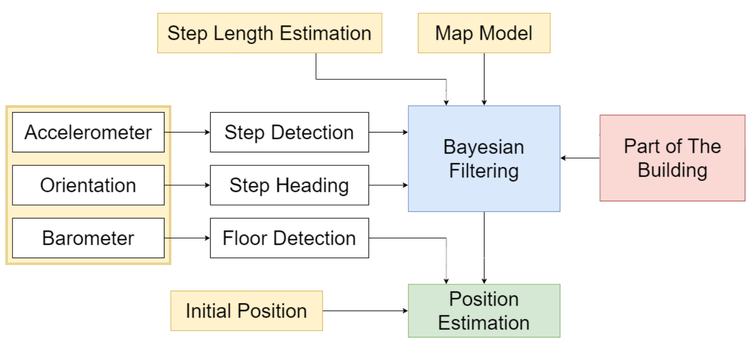
Institute of Computer Science, P. J. Šafárik University, Faculty of Science, Košice, Slovakia

Positioning System

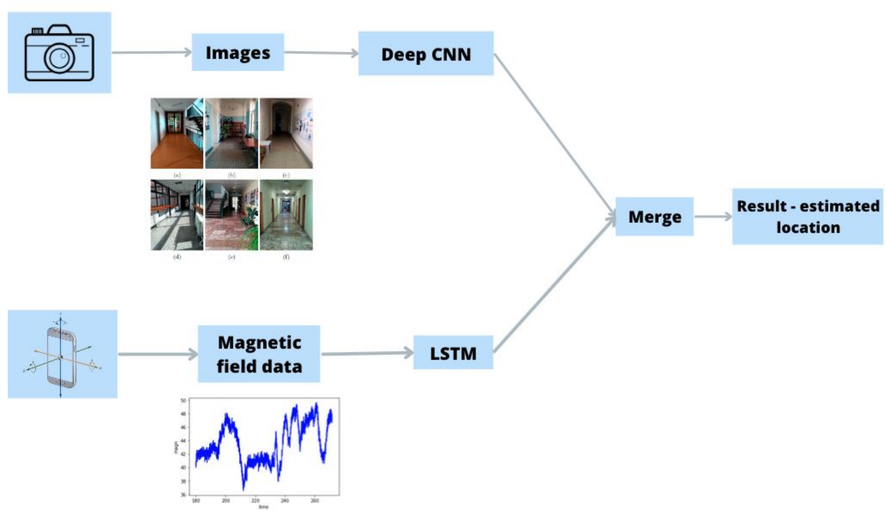
Positioning using **smartphones**. Accuracy depends on the building.
Turns, narrow corridors **improve the accuracy**.

Error increases especially on straight paths
with **incorrect step detection and step length estimation**.

Possible solution: **another source of information.**



Solution



Task

Classification of selected building parts using neural network trained on the particular building.

Merge

Scheduled as future work.

1. Output dense layers from **Deep CNN** and **LSTM** are removed.
2. **Output vectors are combined** and transformed into a unified vector representation of inputs.
3. A **supplementary layer** is added to perform the **classification**.
4. The classification is triggered either by a magnetic field measurement or new camera image.

Magnetometer

LSTM *(Long short-term memory)*

- 4 LSTM layers (60 units each) + dense layer
- Calibrated magnetic field values retrieved with **5Hz frequency**.
- Transformation of measurements at time t:

Rotatioon matrix provided by Android.

$$\mathbf{m}_t^n = \mathbf{R}_t^{nb}\mathbf{m}_t^b$$

$$\mathbf{m}_t^b = (m_{x,t}^b, m_{y,t}^b, m_{z,t}^b) \in \mathbb{R}^{3\times 1}$$

Measurement in device coordinate system

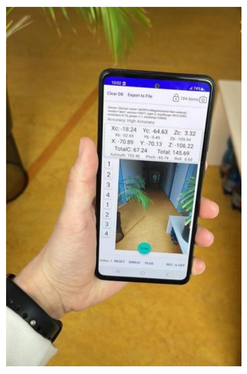
$$\mathbf{m}_t^n = (0, m_{h,t}^n, m_{v,t}^n) \in \mathbb{R}^{3\times 1}$$

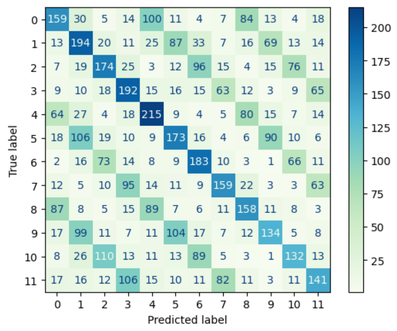
Transformed value in world coordinate system.

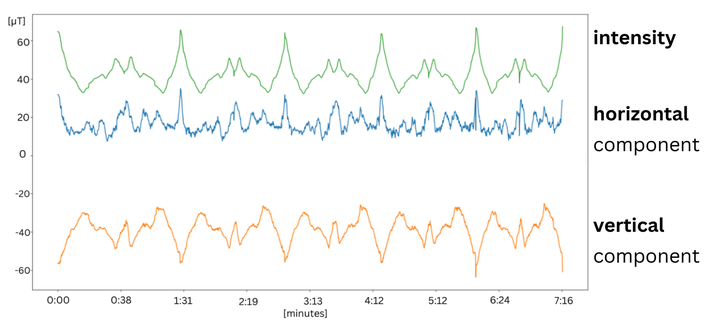
- **Feature vector** (horizontal, vertical component + magnetic-field intensity)

$$\mathbf{m} = (m_h, m_v, \sqrt{m_h^2 + m_v^2})$$

- **Input for LSTM:** 10 values (window of 2 seconds)







5 trajectories both directions on the same corridor

Evaluation

- **Difficult scenario for positioning:** Visually similar parts of the building on 3 floors
- Low overall accuracy (35%) - **misclassification** in floor affiliation, mostly **correct part of the corridor**
- **Addressing weakness of positioning system** on straight trajectory**.** Floors are detected separately.
- Classification on individual floors. Accuracy:
  - 1st floor: **89%**
  - 2nd floor: **94%**
  - 3rd floor: **74%**

34m corridor, 4 classes (two directions, 17m length)

Camera

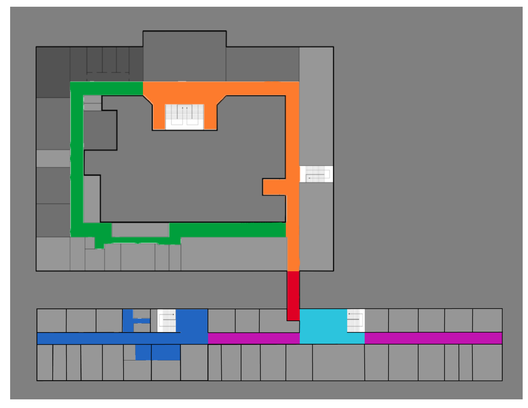
CNN *(Convolutional neural network)*

3 models for evaluation:

1. **CNN without pretraining** (4 pairs of convolutional and pooling layers + flatten layer + classification layer) **accuracy 93%, F1-score 0.85**
2. **VGG16 model without pretrained weights** (small convolution kernels used for reducing computation demands) **accuracy 96%, F1-score 0.86**
3. **Pretrained VGG16 model** (frozen pretrained layers +flatten layer + 2 fully connected dense layers + classification layer) **accuracy 98%, F1-score 0.94**

Evaluation

6 visually distinguishable parts selected manually



Dataset

30 videos. 4000 images used. Images distribution is not uniform. Various time of day, light conditions, weather and season included.
Data augmentation for extending the dataset (blurring, cropping, rotating, scaling, translating, shearing, contrast and brightness changing)