TECHNISCHE UNIVERSITÄT DRESDEN

DRESDEN concept

Ljupka Titizova

Detection and Positioning of Barriers to Augment Indoor Maps

Master's Thesis Defense

Dresden // Friday, 18. September 2020

Motivation

- Outdoor positioning
- Indoor positioning
- Visually and mobility-impaired people (VMIP) in indoor environment
- Up-to-date maps
- Indoor map augmentation

Indoor Positioning Approaches

- Radio Frequency Indoor Positioning
  – Wi-Fi
  – Bluetooth Beacons

- Sensor-based IPS

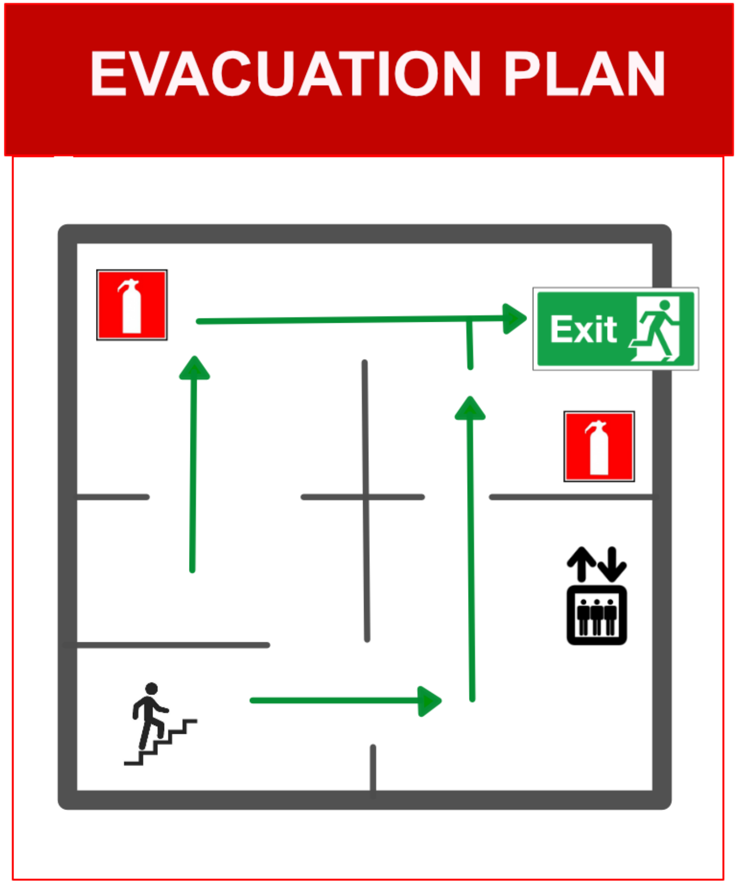
- Hybrid Technologies

- Computer Vision-based IPS

Components of the proposed Indoor Positioning System

1. Object detection
2. Indoor positioning based on evacuation plans
3. Object augmentation into indoor maps

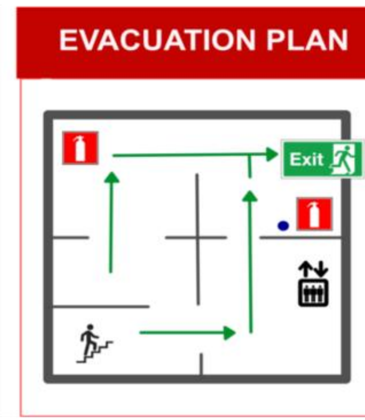
Evacuation plans for indoor positioning



Indoor Positioning Procedure

Object Detection

Indoor Positioning



Images Source: Indor Object Detection Dataset

Object Detection

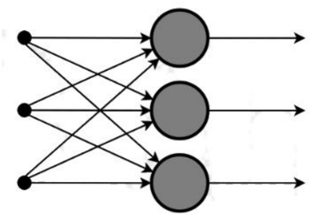
Objects to be detected

- CV approach

- *Features*: used for positioning
  - Evacuation plan objects and symbols

- *Barriers*: obstacles for the VMIP
  - Ground-level, body-level, head-level objects
  - Stairways, doorways, ramps, elevators

Object Detection

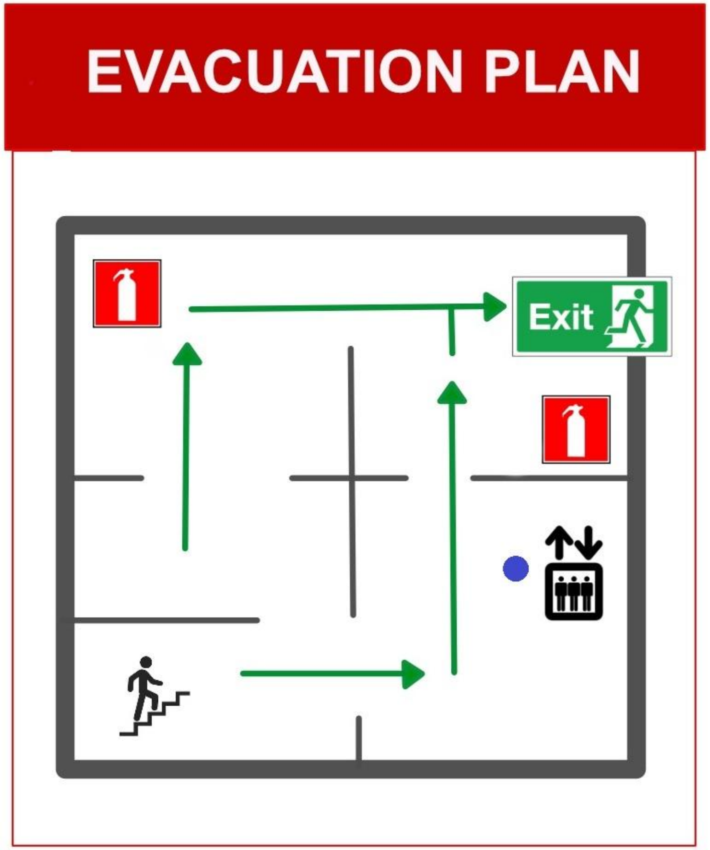


NN

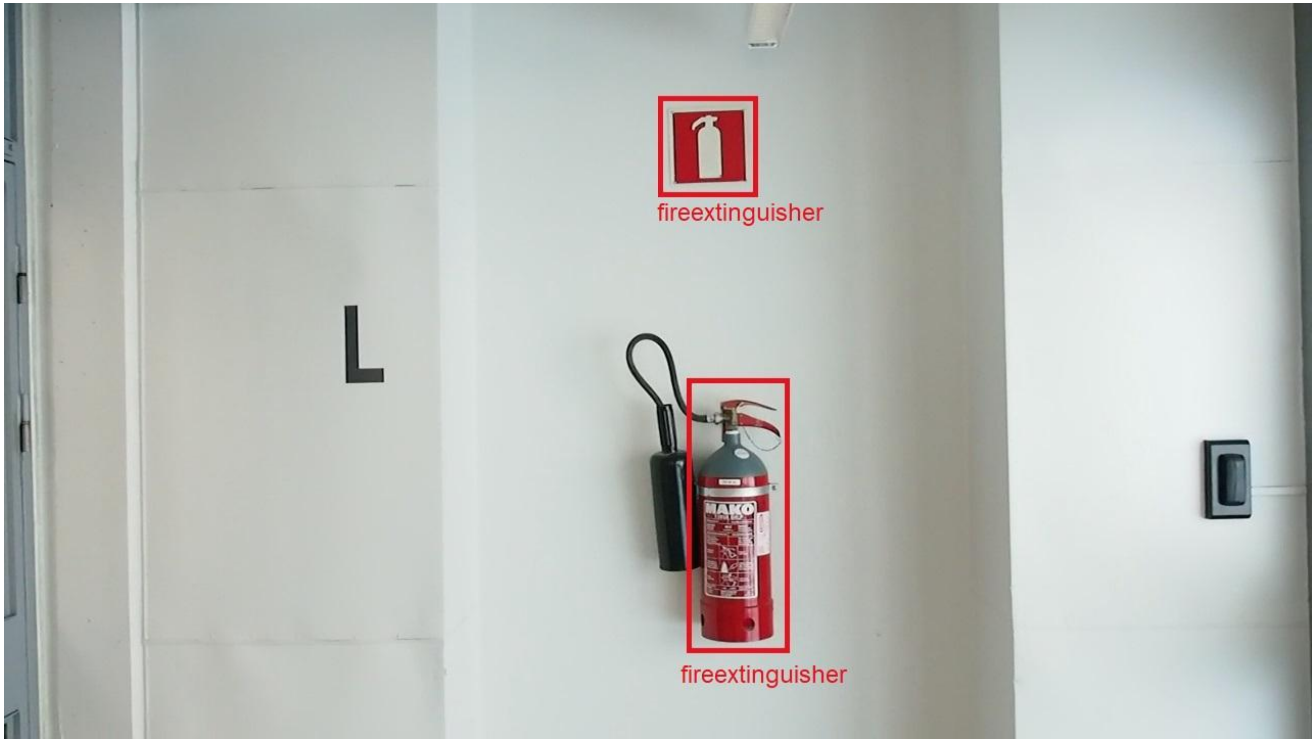
Detected Objects:
fire extinguisher

Indoor Positioning

Evacuation plan-based positioning

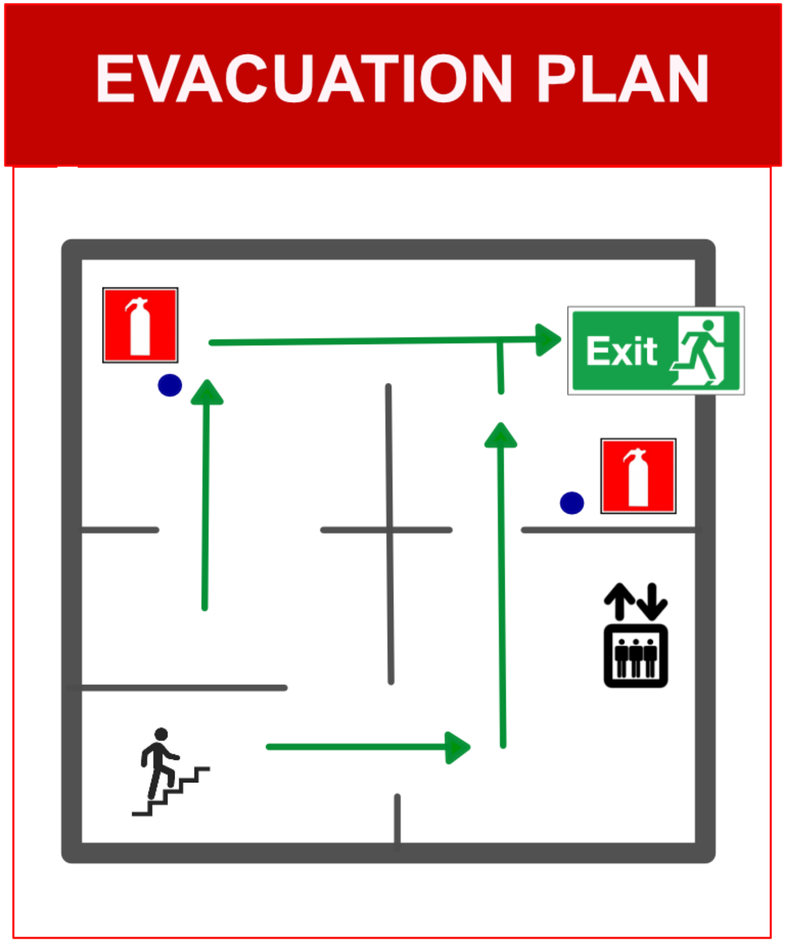


Use case 1



Ambiguous cases

- **2 fire extinguishers on 1. floor**
  - Position?
  - 50 % certainty

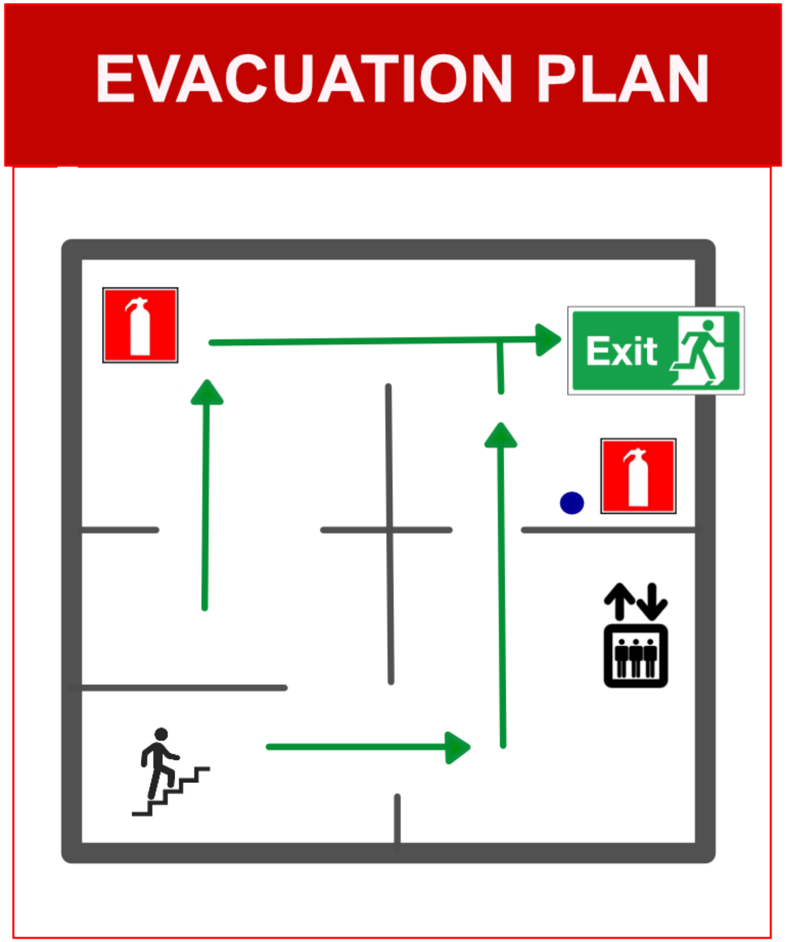


Use case 2

Ambiguous cases

- 2 Fire extinguishers on 1. floor
  – Position?
  – Exit sign next to one fire extinguisher

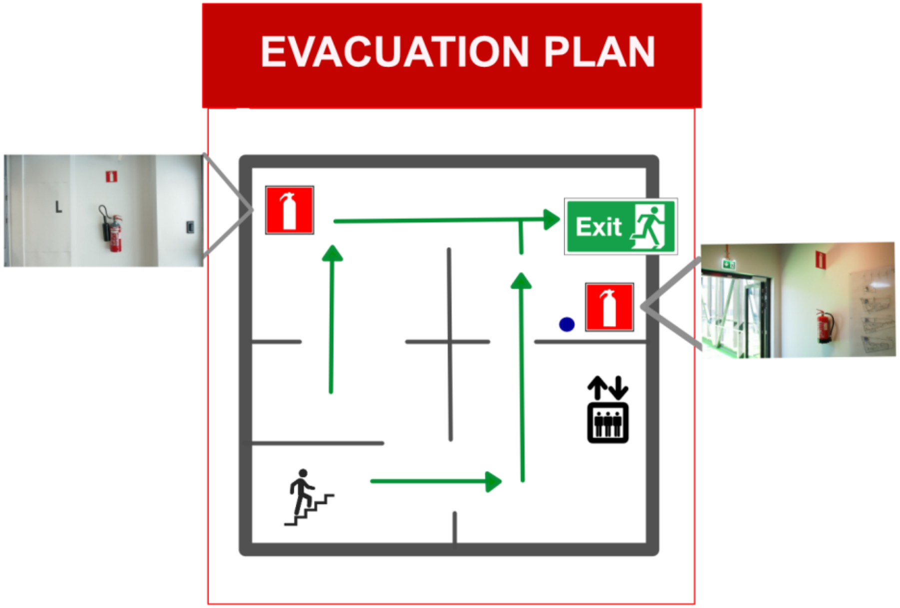
→ **position unique = 100% certainty**



What if no further detected objects?

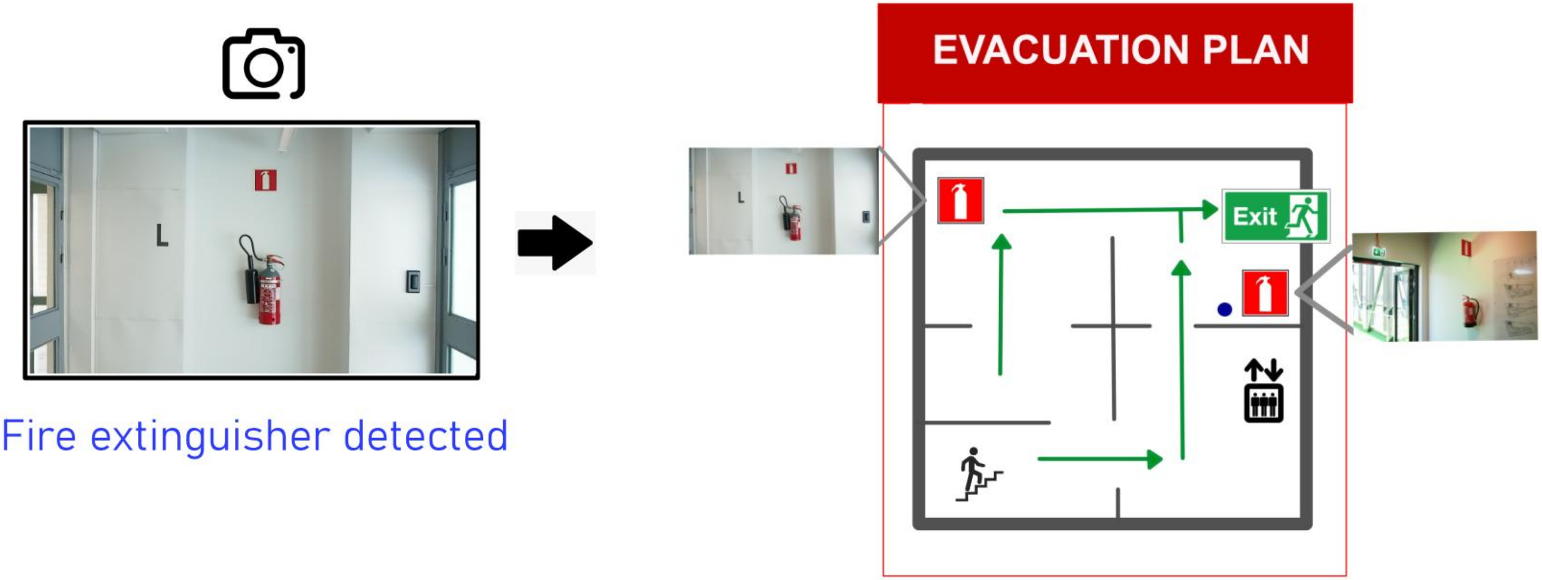
- **(image, position)** pairs

- Procedure:
  1. Capture image
  2. Select most similar image in the database



Similarity Learning

- Triplets (**query image, similar image, dissimilar image)**



Fire extinguisher detected

Augmenting an indoor map

- Same procedure as Indoor Positioning

- Ambiguous cases
  - Similarity Learning probably ineffective
  - Images of whole building costly
  - Wi-Fi Fingerprinting considerable

Implementation Details

Neural Network

- Object Detection NN
- Faster R-CNN ResNet-50 FPN
- Pretrained on COCO Dataset

COCO Classes

```
COCO_INSTANCE_CATEGORY_NAMES = [
    '__background__', 'person', 'bicycle', 'car',
'motorcycle', 'airplane', 'bus',
    'train', 'truck', 'boat', 'traffic light', 'fire
hydrant', 'N/A', 'stop sign',
    'parking meter', 'bench', 'bird', 'cat', 'dog', 'horse',
'sheep', 'cow',
    'elephant', 'bear', 'zebra', 'giraffe', 'N/A',
'backpack', 'umbrella', 'N/A', 'N/A',
    'handbag', 'tie', 'suitcase', 'frisbee', 'skis',
'snowboard', 'sports ball',
    'kite', 'baseball bat', 'baseball glove', 'skateboard',
'surfboard', 'tennis racket',
    'bottle', 'N/A', 'wine glass', 'cup', 'fork', 'knife',
'spoon', 'bowl',
    'banana', 'apple', 'sandwich', 'orange', 'broccoli',
'carrot', 'hot dog', 'pizza',
    'donut', 'cake', 'chair', 'couch', 'potted plant',
'bed', 'N/A', 'dining table',
    'N/A', 'N/A', 'toilet', 'N/A', 'tv', 'laptop', 'mouse',
'remote', 'keyboard', 'cell phone',
    'microwave', 'oven', 'toaster', 'sink', 'refrigerator',
'N/A', 'book',
    'clock', 'vase', 'scissors', 'teddy bear', 'hair drier',
'toothbrush'
]
```

Indoor Object Detection Dataset

- 7 classes
  - Fire extinguisher
  - Exit
  - Chair
  - Clock
  - Screen
  - Printer
  - Bin

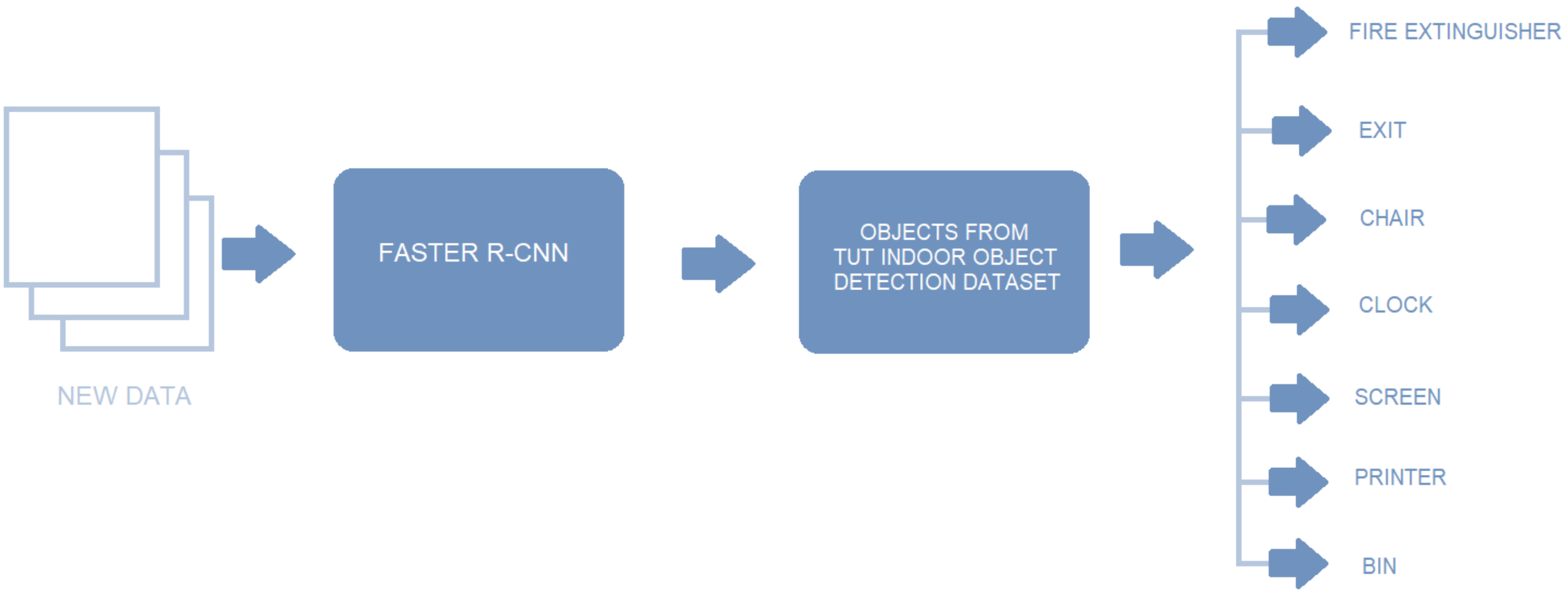
Source: https://arxiv.org/abs/1807.03142,
Accessed: 31.07.2020

TRANSFER OF LEARNING

The application of skills, knowledge, and/or attitudes that were learned in one situation to another **learning** situation (Perkins, 1992)

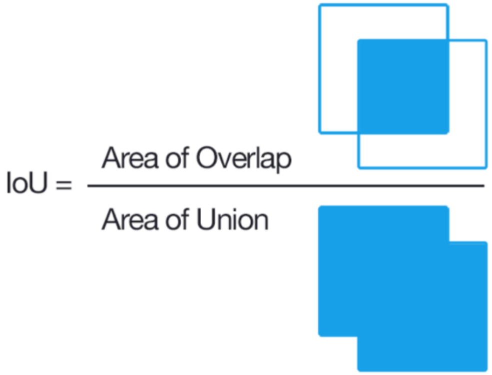
Source : https://serengil.wordpress.com/2017/12/10/transfer-learning-in-keras-using-inception-v3/, Accessed: 31.07.2020

Finetuning


NEW DATA
FASTER R-CNN
OBJECTS FROM
TUT INDOOR OBJECT
DETECTION DATASET
FIRE EXTINGUISHER
EXIT
CHAIR
CLOCK
SCREEN
PRINTER
BIN

Results

Category	AP
fire extinguisher	0.9978
chair	0.9996
exit	0.9836
clock	0.9995
trashbin	0.8485
screen	0.7405
printer	0.6911



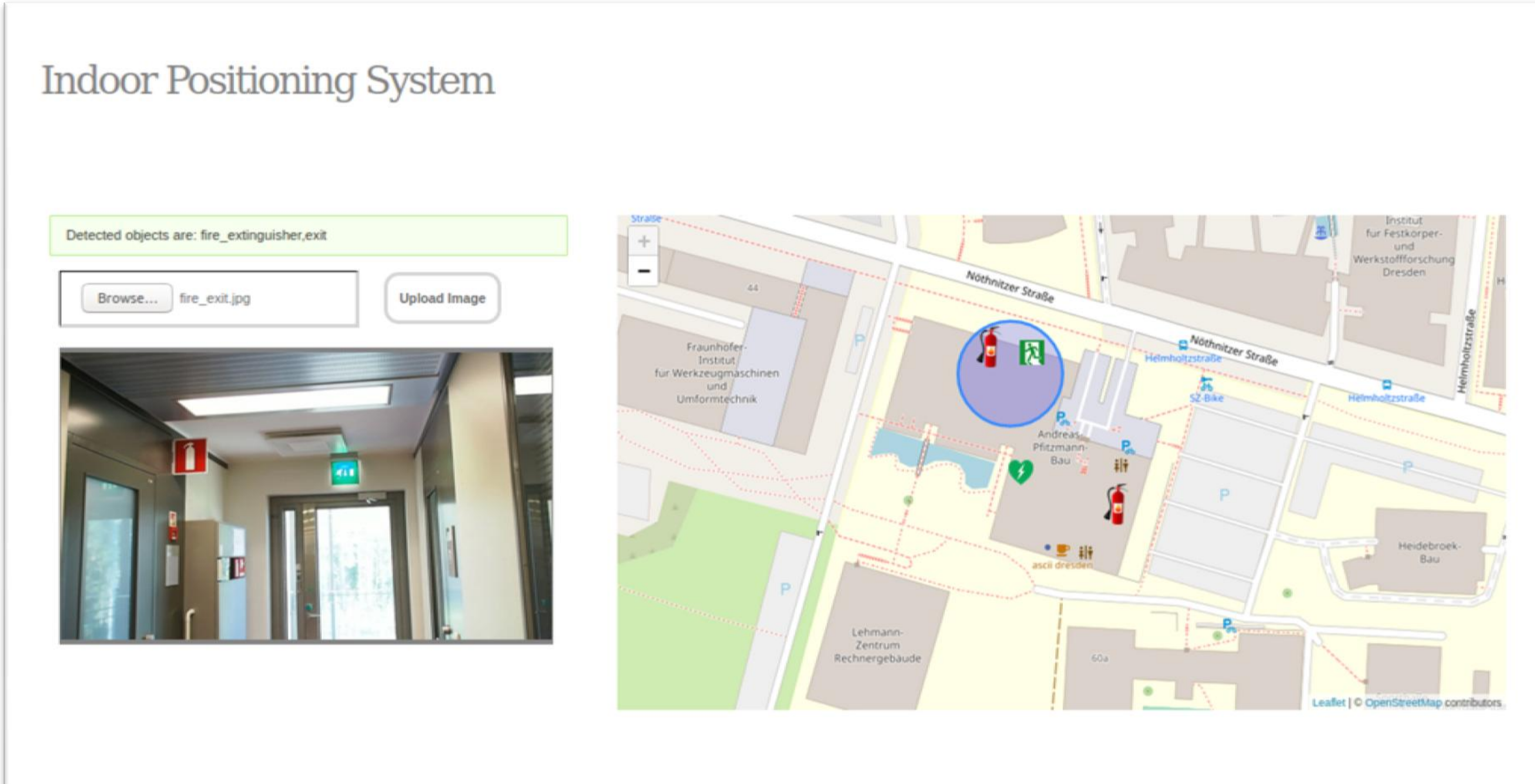
Metric	Result
AP @[IoU = 0.50]	0.9985
AP @[IoU = 0.75]	0.9682
AP @[IoU = 0.50:0.95]	0.7949

Images of the Indoor Object Detection Dataset

Images of APB at TU Dresden

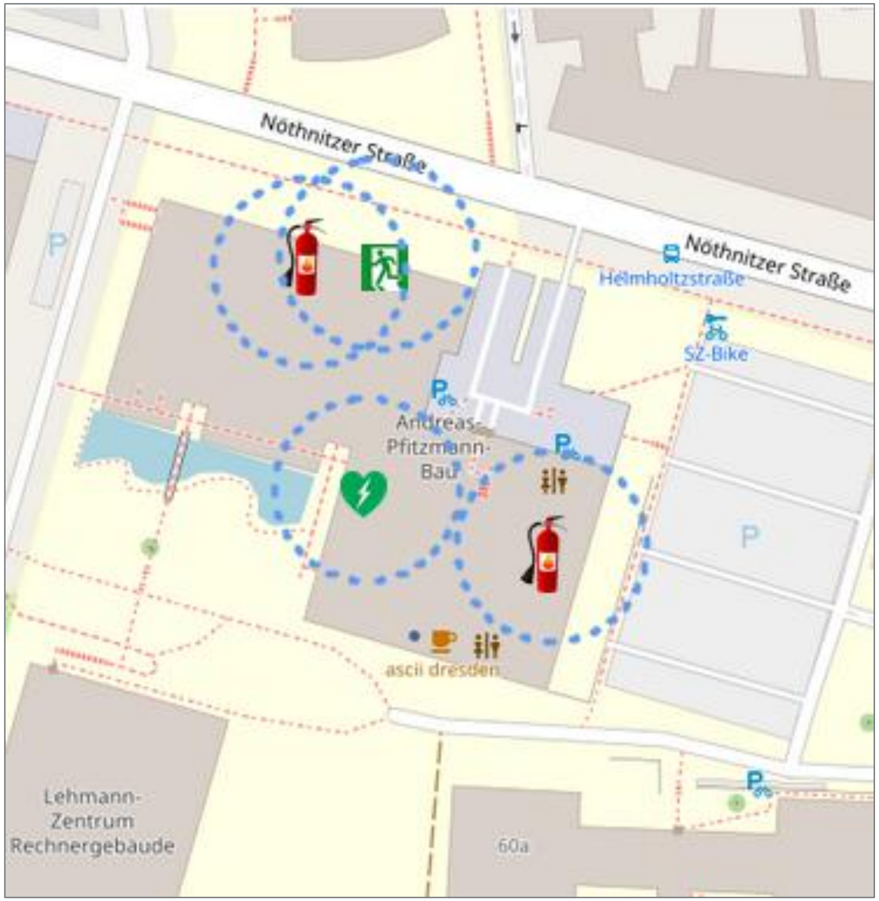
Images of APB at TU Dresden

Prototype

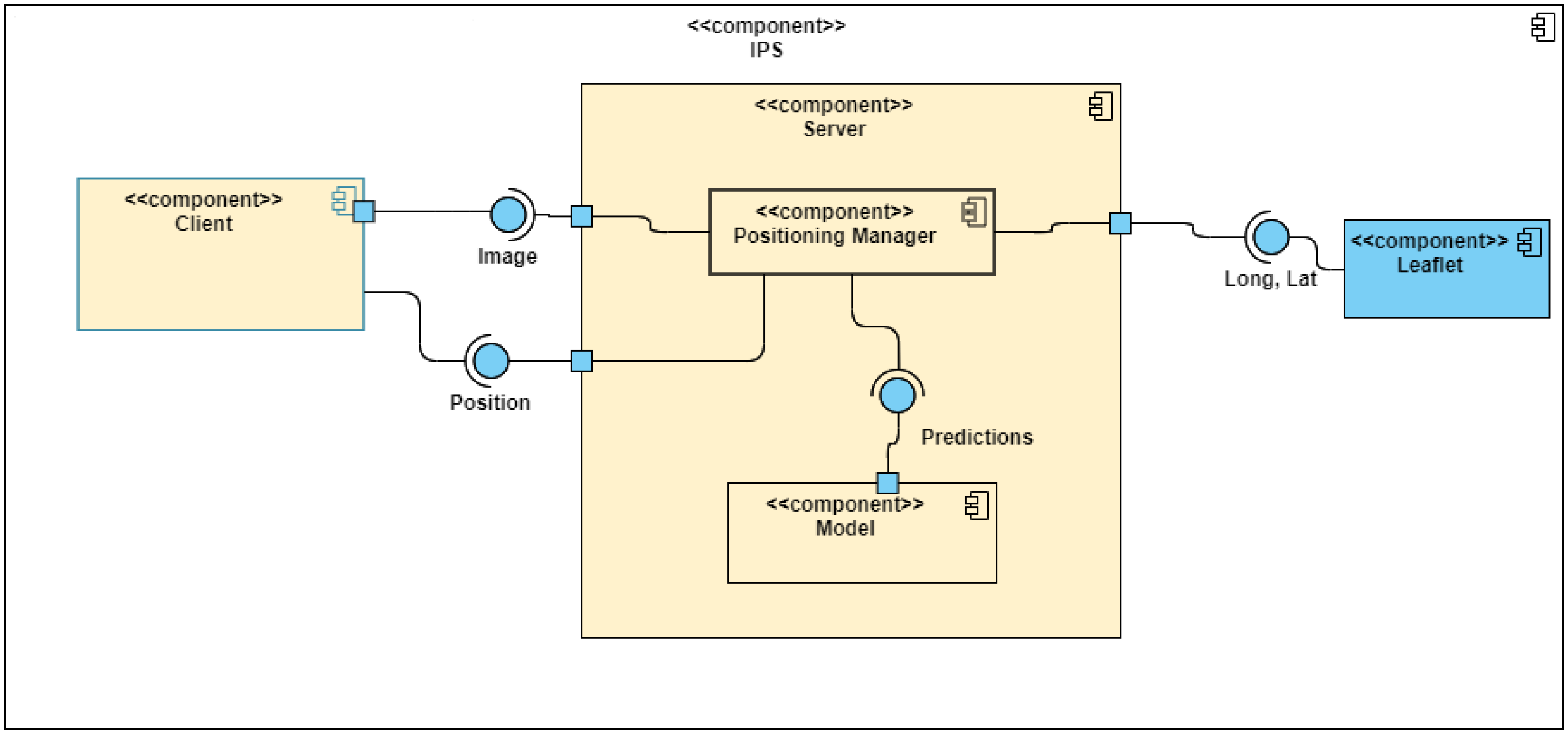

Indoor Positioning System
Detected objects are: fire_extinguisher,exit
Browse...
fire_exit.jpg
Upload Image
Nöthnitzer Straße
Fraunhofer-Institut für Werkzeugmaschinen und Umformtechnik
Institut für Festkörper- und Werkstoffforschung Dresden
Helmholtzstraße
SZ-Bike
Andreas-Pfitzmann-Bau
ascii dresden
Heidebroek-Bau
Lehmann-Zentrum Rechnergebäude
Leaflet | © OpenStreetMap contributors

Determining map position

- Type of object, coordinates
- Radius as a proximity metric



Prototype Architecture



Demo

Problems

- Red objects, rectangular areas, lighting conditions etc.

- FP in the image → false positioning

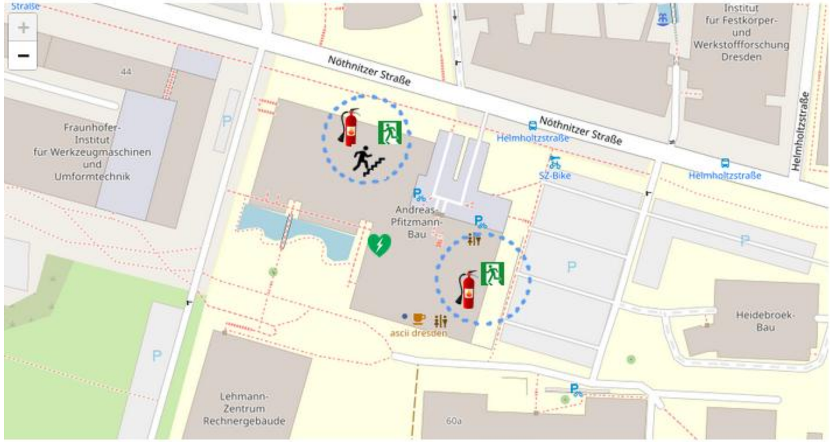
- Double detection of fire extinguishers

- Navigational Symbols

Future improvements

- Augmenting with building-specific features, room numbers etc.

- Enriched indoor maps
  - *Map-based positioning :* augmented objects as prospective features
  - Spatial relations between the objects, photogrammetry
  - Similarity Learning / Wi-Fi Positioning can be omitted



Summary

- Indoor Positioning System:
  - **Hybrid approach:** Object Detection NN + information from evacuation plans + Similarity Learning / Wi-Fi / Bluetooth

Object Detection

Indoor Positioning

Summary

- Object detection with NN
  - Works well on unseen data from APB of TU Dresden

- Frequent indoor map augmentation → Up-to-date indoor maps

- Information from evacuation plans for indoor positioning
  - Accurate results only when unique objects or object clusters detected
  - Augmented objects as prospective features

- Additional positioning methods required for ambiguous cases

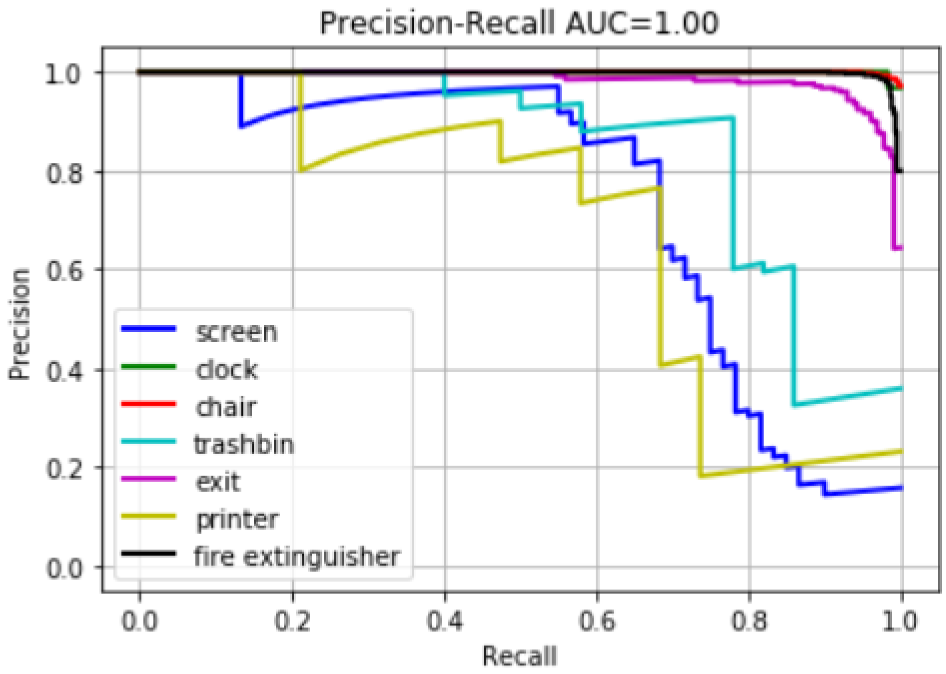
- Additional positioning methods can be omitted by enriching the indoor maps

Thank you for your attention!

☺

Questions?

Precision-Recall Curves



Average Precision

$$AP = \sum_n (R_n - R_{n-1}) P_n$$

R_n , P_n - the precision and recall at the n-th threshold.

AP = precision-recall curve as the weighted mean of precisions achieved at each threshold, with
the increase in recall from the previous threshold used as the weight

No confidence score filtering

Ground truth \ Predictions	no obj.	screen	clock	chair	trash bin	exit	printer	fire ext.
no object	N/A	314	0	350	87	122	59	151
screen	1	52	0	0	0	5	0	0
clock	1	0	59	0	0	0	0	0
chair	0	0	0	565	0	0	0	0
trash bin	3	1	0	0	41	0	3	0
exit	0	0	2	0	0	227	0	0
printer	0	4	0	0	1	0	14	0
fire extinguisher	2	0	0	0	0	0	0	595

Confidence score > 50 % filtering

Predictions

Ground truth	no obj.	screen	clock	chair	trash bin	exit	printer	fire ext.
no object	N/A	44	0	17	12	31	5	43
screen	8	45	0	0	0	5	0	0
clock	1	0	59	0	0	0	0	0
chair	1	0	0	564	0	0	0	0
trash bin	6	0	0	0	41	0	1	0
exit	3	0	0	0	0	226	0	0
printer	2	0	0	0	0	0	17	0
fire extinguisher	6	0	0	0	0	0	0	591

Confidence score > 90 % filtering

		Predictions							
		no obj.	screen	clock	chair	trash bin	exit	printer	fire ext.
Ground truth	object detection	N/A	0	0	8	2	16	1	14
	screen	28	26	0	0	0	4	0	0
	clock	1	0	59	0	0	0	0	0
	chair	3	0	0	562	0	0	0	0
	trash bin	7	0	0	0	41	0	0	0
	exit	7	0	0	0	0	222	0	0
	printer	6	0	0	0	0	0	13	0
	fire extinguisher	8	0	0	0	0	0	0	589

Prototype

PyTorch

node JS

python™

django

React

Outline

- Motivation
- Indoor Positioning Approaches
- Proposed Indoor Positioning System (IPS)
  - Object Detection
  - Indoor Positioning
- Implementation Details
  - Neural Network Implementation
  - Prototype
- Problems and Future Improvements
- Summary