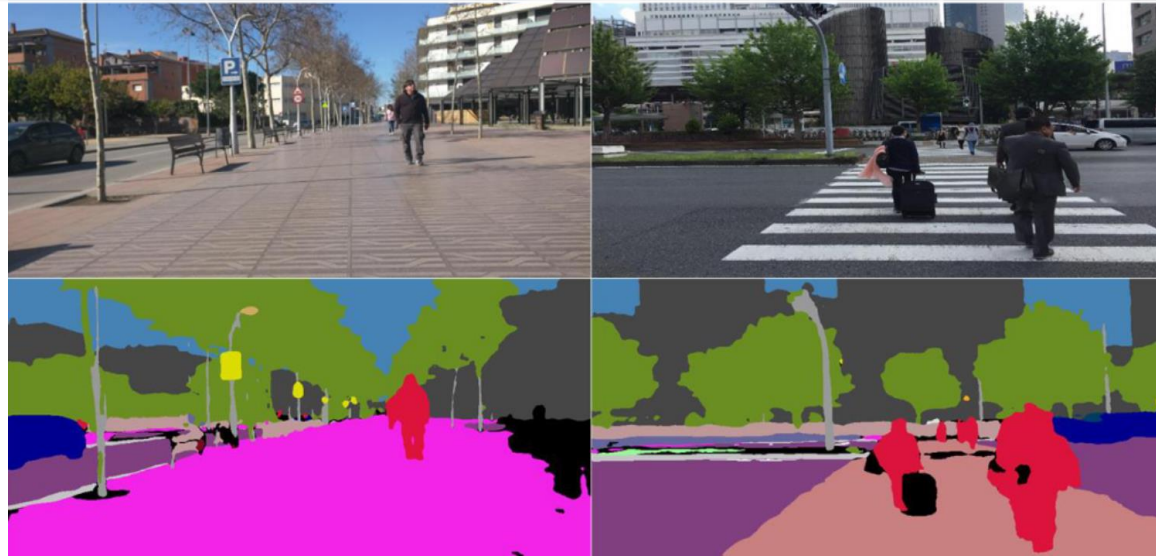


# Practical Course: Computer Vision for Human-Computer Interaction

SS 21

Constantin Seibold

COMPUTER VISION FOR HUMAN COMPUTER INTERACTION LAB  
INSTITUTE FOR ANTHROPOMATICS AND ROBOTICS



# What will you learn?

- Apply algorithms from lectures and papers
- Hands on experience
- Get comfortable with machine learning tools
- Learn about common problems and applications in machine learning & vision
- Find solutions to difficult problems

# Example Of Last Semester

## Human Drone Interaction

# General Information

## Weekly meeting

- Compulsory attendance
- Talk about intermediate results & problems
- Ask for help and guidance
- Weekly goal: stay „on track“

## 3 Students per Team

- Use version control (e.g. git)
- Internal git repos provided via the SCC's GitLab (<https://git.scc.kit.edu/>)
- Divide work into separate tasks and distribute within group

# At the end of the Practical Course...

- Final presentation of each group
  - 15 Minute talk
  - Each member talks about their contribution
  - The presentation should be about:
    - Goals and usefulness of your chosen topic
    - Your proposed approach
    - Results
- Written report describing the topic/approach/results
  - 4-pages in standard paper format
    - Abstract/Introduction/Method/Results/Conclusion
    - References can be fit on an extra page
  - Written in a conference paper template
    - <http://cvpr2020.thecvf.com/sites/default/files/2019-09/cvpr2020AuthorKit.zip>
- Final Code Submission
  - Working implementations of Algorithms
  - A Readme-file describing how the code can be used to reproduce the results
    - If the team agrees -> make code publicly available to the community

# Topics

- Walkable Path Discovery Utilizing Drones
- Image Analysis of Structured Visual Content
- Zero-Shot Action Recognition
- Anomaly Detection in Construction Sites

# Walkable Path Discovery Utilizing Drones

- Dynamically integrate drone usage into daily life
  - Assist for visually impaired
  - Drone-assisted filming
- Criteria:
  - Person following
  - Obstacle avoidance
  - Traffic orientation



Avila, Mauro, Markus Funk, and Niels Henze. "Dronenavigator: Using drones for navigating visually impaired persons." *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*. 2015.

# Walkable Path Discovery Utilizing Drones

## ■ Potential datasets

### ■ Datasets including sidewalks

- Cityscapes
- Mapillary
- Vistas

- <https://github.com/qweawq/Blind-road-and-crosswalk-dataset>



## ■ You will use the Parrot ANAFI or Dji Tello

### ■ Camera, WiFi-Connector, Remote Controller

### ■ Light-weight

### ■ Development using **dronekit-python** (<https://github.com/dronekit/dronekit-python>)

### ■ Working within drone regulations

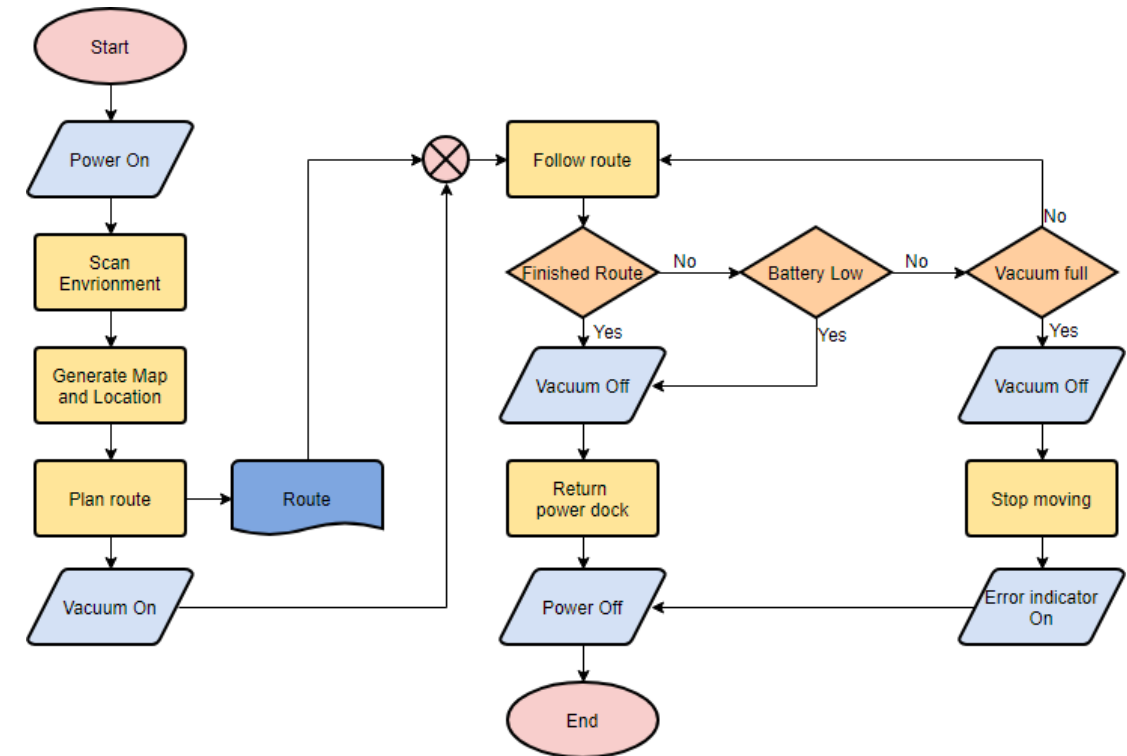
- [https://www.bmvi.de/SharedDocs/DE/Publikationen/LF/flyer-die-neue-drohnen-verordnung.pdf?\\_\\_blob=publicationFile](https://www.bmvi.de/SharedDocs/DE/Publikationen/LF/flyer-die-neue-drohnen-verordnung.pdf?__blob=publicationFile)





# Image Analysis of Structured Visual Content

- Extracting information from structured visual content is a non
- Employ deep learning methods for automated extraction of information
- Topics:
  - Document Segmentation
  - Graph Neural networks



# Image Analysis of Structured Visual Content

- Dataset:
  - Internal manually annotated dataset utilizing DISKNET
  - Containing images with paired graph representations
  - 579 papers, 4705 constructs and 7653 relations



The Digital Scientific Knowledge Network (DISKNET) Project provides a web-based platform that enables extraction, exploration, and aggregation of knowledge from scientific publications.

Aggregating the knowledge of 579 papers, 4705 constructs, and 7653 relations...

# Zero-Shot Action Recognition

- Given a video classify the occurring action despite the action not appearing in the training data
- Link visual and semantic representations
  - Word embeddings
  - Textual descriptions
  - Attributes



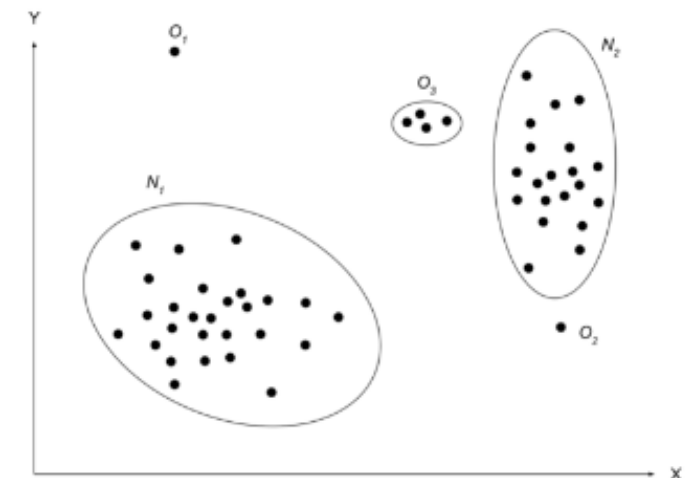
# Zero-Shot Action Recognition

- Usable Datasets:
  - HMDB51
  - UCF101
  - ActivityNet
  - Kinetics 400/700



# Anomaly Detection in Construction Sites

- Given an image of the construction site scene
  - Detect technical problems in parts of the image (scratches, cracks, flakes)
- Problem formulated as anomaly detection
  - Identify which parts of the image appear as outlier/faulty compared to the underlying distribution of unimpaired image regions



# Anomaly Detection in Construction Sites

- Usable Datasets:
  - ADE20K (S) (<https://groups.csail.mit.edu/vision/datasets/ADE20K/>)
  - COCO-Stuff (S,D) (<https://github.com/nightrome/cocostuff>)
  - Internal construction site dataset
  
- Evaluation using real life construction site images

# Topic Selection

- Find a team of three people
- Each team sends us
  - A ranking of their preferred three topics
  - until 18<sup>th</sup> 23:59 of April
  - per Email at [constantin.seibold@kit.edu](mailto:constantin.seibold@kit.edu)
- Scenarios
  - Re-Implement not publicly available model from paper
  - Change publicly available model by trying out parameters/losses
  - Fuse two different architectures in some way
  - Use existing model for a novel task



# Create a Plan

- Check related work
  - What has been done in this topic specifically?
  - What has been done for the overarching task?
    - What pretrained models do exist?
    - What datasets do exist?
- How can you use related work for your task?
- Who focuses on what?
- Create a rough schedule for how you approach your task



# Implementation

- Choose Framework
  - **TensorFlow**, see tutorial <https://www.tensorflow.org/tutorials>
  - **PyTorch**, see tutorial <https://pytorch.org/tutorials/>
  - Torch, Theano, Caffe
  - ...
- Each team can use an 8GB GPU on our servers
- Plot learning curves/results and show at weekly meeting
- Split work equally between team members

# Evaluation

- Split Dataset into distinct training, validation and test set
- Use training and validation to tune your model
- The test set is used at the very end
  
- Check out metrics in related work
  - Classification: Accuracy
  - Detection: mAP
  - Segmentation: mIOU
  
- If your project requires a user study, check in related work how similar studies are performed

# Presentation

- 15min Presentation per Team
- Explain your topic, approach and results
  - If applicable, a demo/video would be appreciated
- Allocate enough time for each team member to talk about their contributions
- What if your solution does not work?
  - No problem at all.
  - A presentation showing your development and why it didn't work is perfectly acceptable as well

# Organization

- Select teams of 3 students each
- Meeting schedule
  
- Week 0 [12.04.21] : Introduction
- Week 1: Present ideas on how to approach the problem
- Week 2: Read related work and consider how to incorporate these into your task
- Week 3: Implementation
- ...
- Week 13 [19.07.21]: Presentations
  
- Weekly meeting for discussion and status updates with corresponding supervisor
  - Set a consistent date for weekly meetings
  
- Register Projektpraktikum with KIT's Studienbüro...(Modulhandbuch Nummer #7500279, Teilleistung #T-INFO-110325)
- For these slides, other information, announcements and updates → check website [coursemember/321meins]