Zero-Shot Buildings' Segmentation using SAM

Ali Ghandour
Ali Mayladan
Hasan Moughnieh

18 September 2023

GEOspatial Artificial Intelligence

GEOAI

- **Lead: Dr. Ali J. Ghandour**
- **Members: 16**
- **Webpage:** www.geogroup.ai
- **ranked 4th/730: "NASA Field Boundary Challenge"**

International Collaboration

IMT Nord Europe
École Mines-Télécom
IMT-Université de Lille

CERI-SN

SORBONNE UNIVERSITÉ

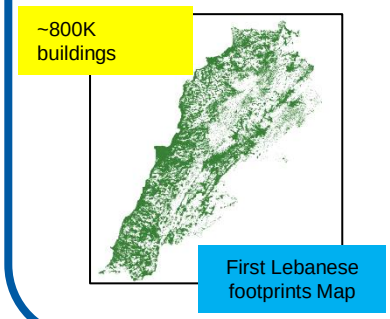
ISIR

TECHNISCHE UNIVERSITÄT MÜNCHEN

The League Of Arab States
Arab Organization For Agricultural Development

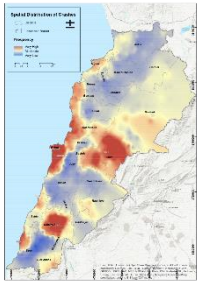
Ongoing Projects

GeoUrban-AI Tool

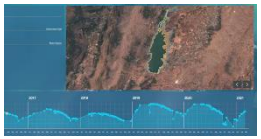


Solar rooftop potential map

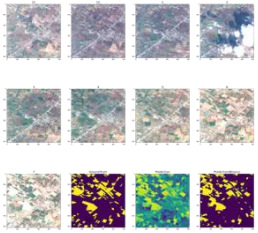
Road Crashes Observatory



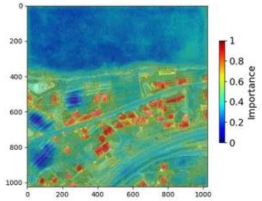
Water Body Monitoring



Winter Wheat Monitoring

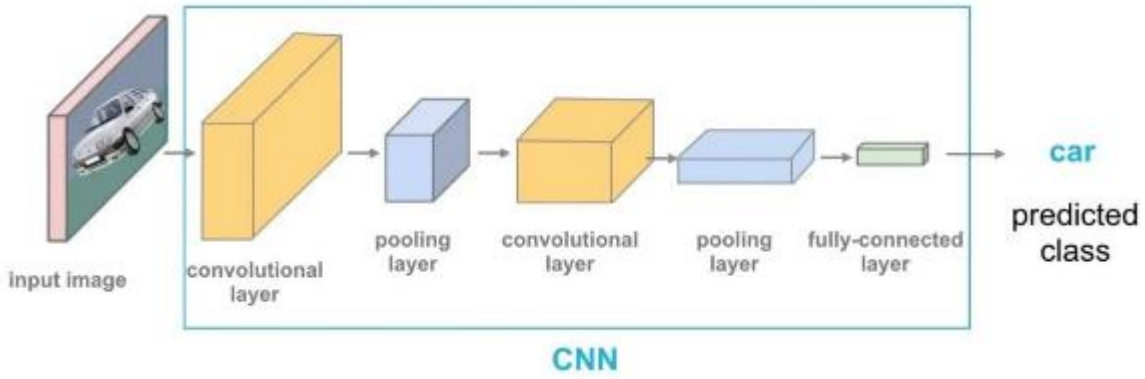


eXplainable AI (XAI)



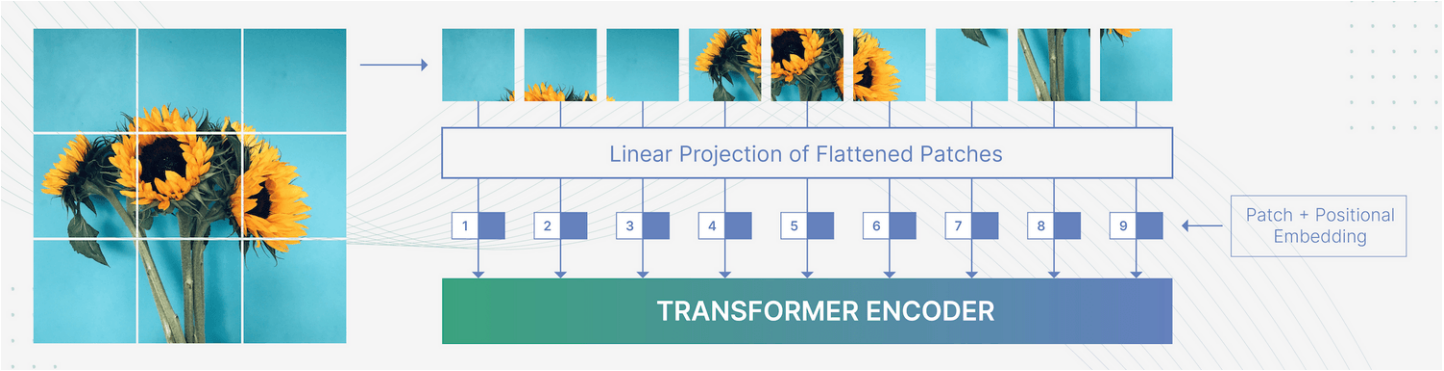
Deep Learning

- CNN:



Source: https://viso.ai/deep-learning/vision-transformer-vit/

- Transformers:



Source: https://www.picsellia.com/post/are-transformers-replacing-cnns-in-object-detection/

OOD Generalization

- Out-of-Distribution Generalization:
  - testing distribution is unknown and different from the training

source: https://www.gabriele-cavallaro.com/teaching/parallel-and-scalable-machine-learning-for-remote-sensing-big-data-ws-2020-2021

Foundation Models

- coined in July 2022:
  - https://arxiv.org/pdf/2108.07258.pdf
  - Large (deep neural network) model
  - trained on broad data (generally using self-supervision at scale)
  - usually multimodal
  - can be adapted to a wide range of downstream tasks

Large DL

- transformer model on a large scale.
- widely used in NLP: Large Language Models (LLMs)
- human brain:
  - ~86 billion neurons
  - 100 trillion synapses

Source: https://huggingface.co/blog/large-language-models/

SAM

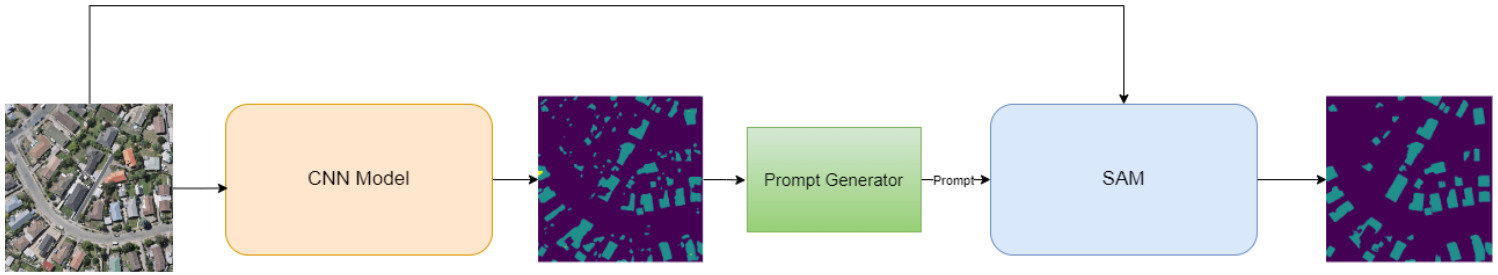
- foundation model Meta AI (April 2023):
  - https://arxiv.org/pdf/2304.02643.pdf

- SA-1B Dataset:
  - 11M diverse, high-res images (3300x4950 pixels)
  - 1.1B segmentation masks (99.1% auto-generated)
  - High-quality masks
  - Significantly larger than any segmentation datasets

Prompt Engineering SAM

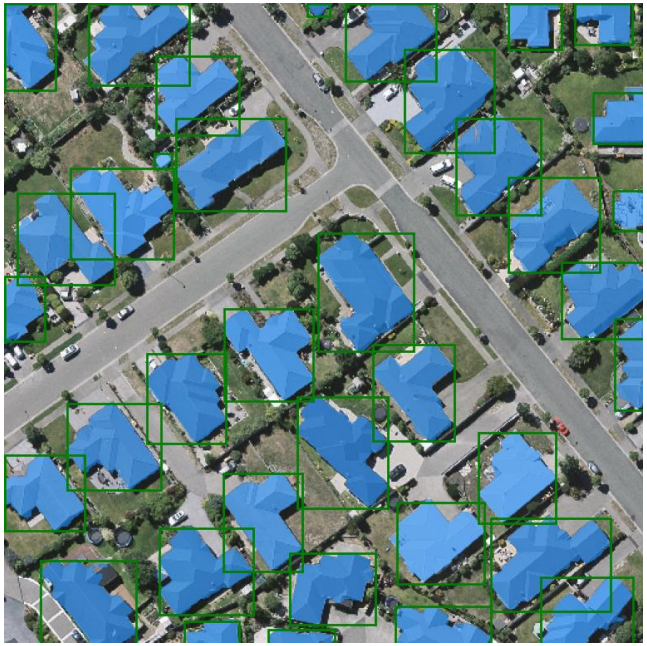
- PE SAM for satellite imagery:
  - CNN model as prompt generator

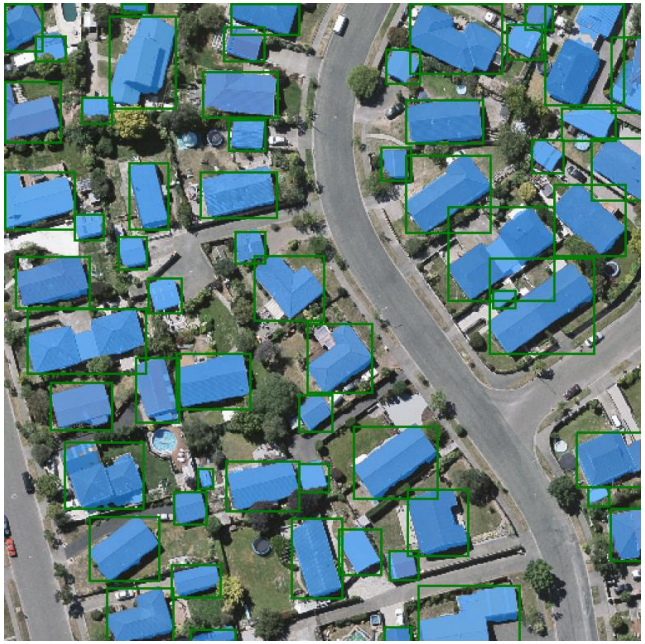
- Prompt types:
  - Single point
  - Bounding Box
  - Skeleton



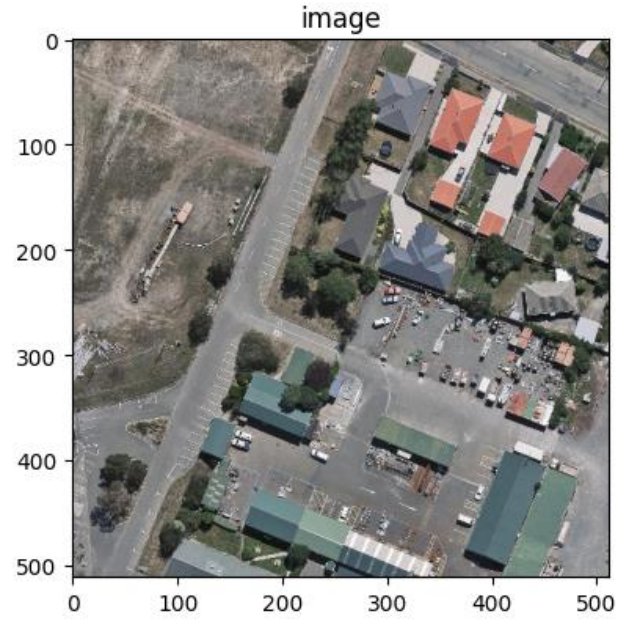
Single Prompt

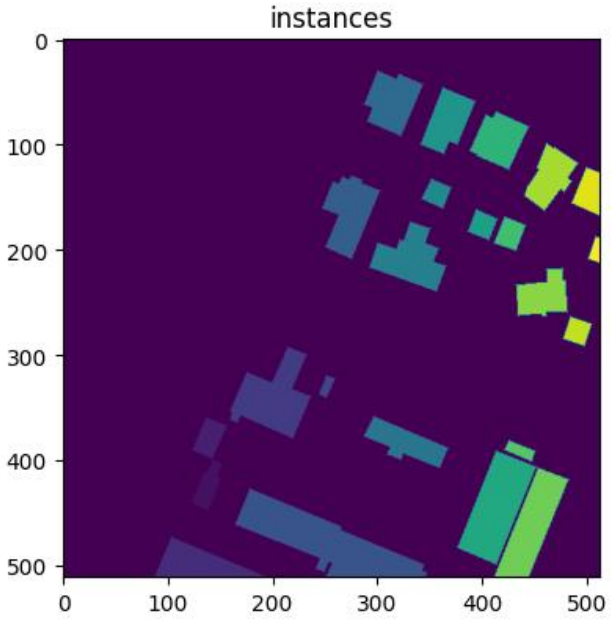
Bounding Box Prompt





Skeleton Prompt





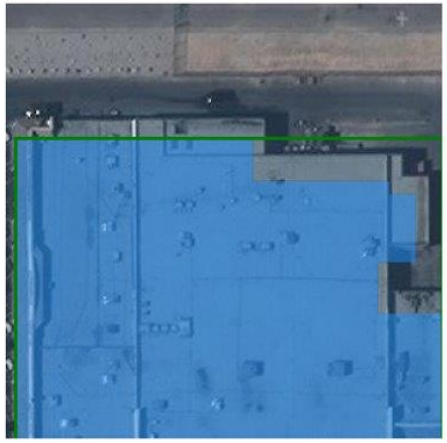
Crowd AI dataset

GroundTruth

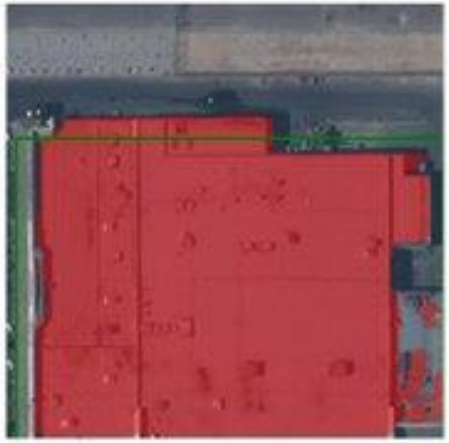
Prediction

Crowd AI dataset – *cont.*

GroundTruth



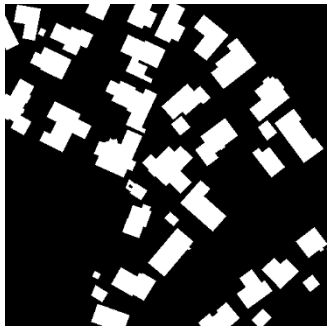
Prediction

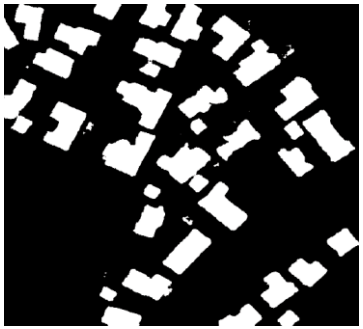


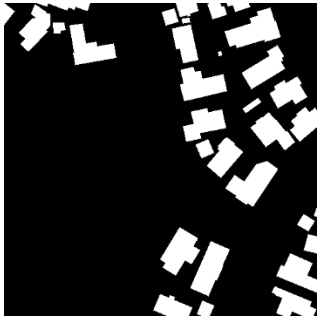
Ongoing Work

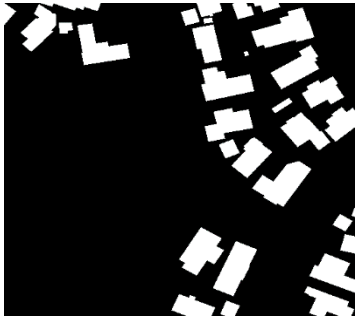
GT

predicted









TP-F1 score: 93%

Live Demo!

- Ali: graduate student at GEOAI group
- Hasan: research engineer at GEOAI group
  - Link

Thank you!

https://github.com/geoaigroup/buildingsSAM

GEOAI group

aghandour@cnrs.edu.lb

www.geogroup.ai