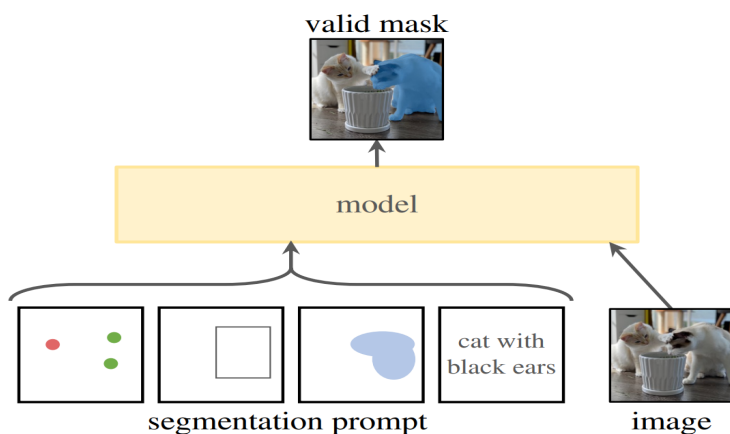


VISION AI USING VISUAL PROMPTING

Assignment

Visual prompting is a new approach to building computer vision models. Specifically, it has the potential to convert a unlabeled image dataset into a segmentation model with just a few user clicks or scribbles in a few images. Prompting is common for textual large language models, but this idea has now also been applied to images. The approach relies on pre-trained large vision models / foundation models that do not need further training or fine-tuning.

In this assignment we want to explore visual prompting approaches. We want to get a good overview of the landscape of methods and subsequently test some methods on specific datasets to assess their performance, strengths and weaknesses and potential improvements/extensions. From there, you are encouraged to proceed and develop visual prompting workflows that improve upon our baseline.



Internship overview

- Master
- Internship / Graduation
- Applied Mathematics, Data Science, Industrial Engineering
- Location: Eindhoven

Technologies

- Deep Learning
- Computer Vision
- Segmentation
- Foundation models
- Visual prompting
- Python



Activities

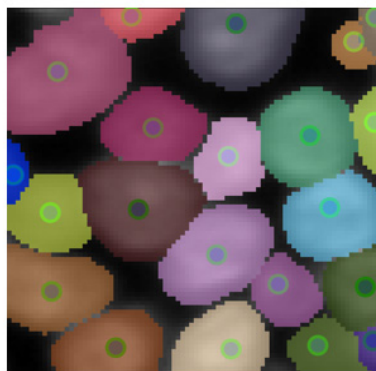
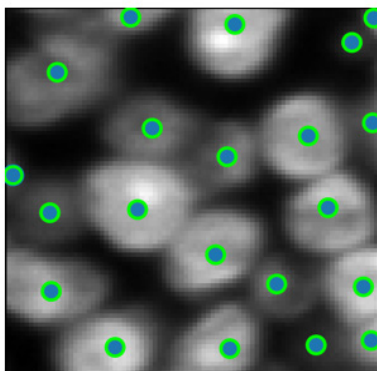
- Survey recent developments in visual prompting approaches with a focus on (instance) segmentation, but also including other tasks such as object detection.
- Experiment with some selected approaches and compare; identify potential improvements or extensions.
- As far as time allows: propose/develop/test potential improvements to the workflow.
- Draw up conclusions and recommendations.

Context

We specifically envision visual prompting workflows as useful with limited data, when new models are often needed or when labelling and training time is very constrained. But we want to understand in general the trade-off between visual prompting and traditional ML workflows, so we know better what to use when.

We have some specific applications of interest for visual prompting workflows. This concerns at least cell segmentation (and counting) and visual inspections. Though our focus is on (instance) segmentation, we are also interested in object detection, counting and potentially other tasks.

We also see methods that allow additional textual prompting. This is also interesting as this can further enhance performance. There is a direct connection with large vision/foundation models. They typically form an essential component of a visual prompting workflow. Investigation the impact of the actual foundation models used is also relevant.



Why choose Sioux?

- Working on innovative technology
- Challenging, dynamic and varied work
- A comfortable and personal work environment
- Plenty of opportunities for personal development
- Great career opportunities
- Contributing to a safe, healthy and sustainable society

Get in touch!

Would you like to know more about this student assignment?

Contact:

Mark van den Broek / Raffaele Imbriaco

werving_mathware@sioux.eu
+31 (0)40 - 2677 100