Karlsruher Institut für Technologie

cv:hci @KIT Computer Vision for Human-Computer Interaction

Master's Thesis on Assessing and Improving in-Context Learners on Out-of-Domain Datasets

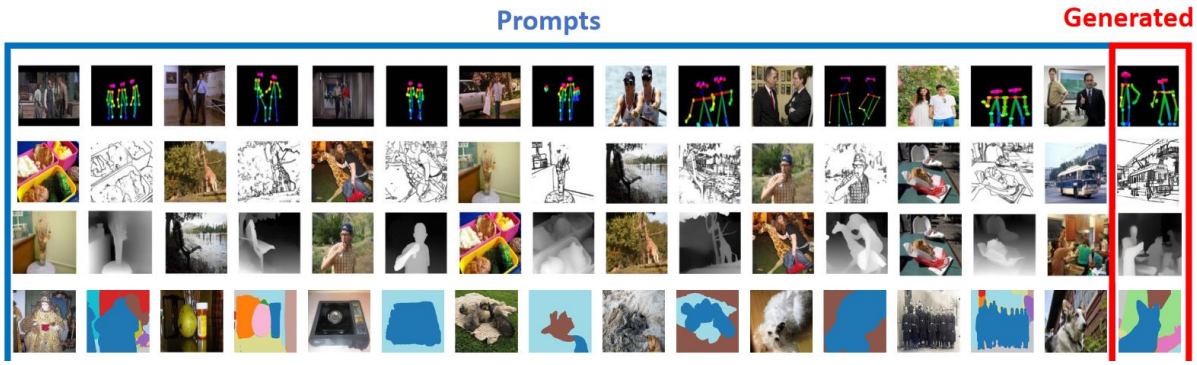

Figure 1: In-Context Learner is prompted with a task and generates the corresponding output for the last image in the prompt (image credits [1]).

Vision models that are capable to, given a „visual description" of a task (prompt), solve this task for a new incomming image without being explicitly re-trained on this task are so called in-Context Learners [1,2,3].

In this master's thesis, I would like to explore together with you what the limits of such models are and how their capabilities really transfer to data from different imaging domains, that they were not explicitly trained on. This evaluation might include microscopy imagery, satellite images, medical images and more as well as a broad variety of tasks in these domains. Based on the observed transfer capabilities, adaptation strategies for training in-Context Learners will be explored to address identified shortcommings.

Experience and Knowledge:

- Interest in the topic of computer vision and deep learning.
- Having visited lectures such as "Deep Learning for Computer Vision", "Neural Networks" or the like is helpful for fast on-boarding.
- Python programming skills and knowledge of PyTorch are desirable.

Contact:

- Send me an e-mail at [simon.reiss(at)kit.edu]!

If you are excited by these topics as I am, please feel free to send me your CV and Transcript of Records. I welcome any interesting ideas and look forward to publication at international conferences with you.

References:

[1] Bai, Yutong, et al. "Sequential Modeling Enables Scalable Learning for Large Vision Models." *arXiv preprint arXiv:2312.00785* (2023).

[2] Bar, Amir, et al. "Visual prompting via image inpainting." *Advances in Neural Information Processing Systems* 35 (2022): 25005-25017.

[3] Wang, Xinlong, et al. "Images speak in images: A generalist painter for in-context visual learning." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023.