Onderzoeksgroep Kernfusie / Research unit Nuclear Fusion

Onderzoeksstage / research internship

Zomer 2024 / Summer 2024

Automatic captioning of fusion physics figures with image-to-text AI

Promoter: Prof. Dr. Geert Verdoolaege
Supervisor: Jerome Alhage
Study programs: Master of Science in Engineering Physics, Master of Science in Physics and Astronomy, European Master of Science in Nuclear Fusion and Engineering Physics, Master of Science in Computer Science Engineering, Master of Science in Information Engineering Technology
Location: Technicum

Introduction

The research on controlled thermonuclear fusion aims at providing stable, baseload electric power by creating "a star on earth". Magnetic confinement fusion in tokamaks and stellarators is foreseen to result in clean and safe commercial power production by the second half of the century.

The research unit Nuclear Fusion (*infusion*) is active in the domain of *Fusion Data Science*, comprising the development of techniques in data science and machine learning with applications to the physics and technology of fusion devices. With the large volumes of complex data being generated at experimental fusion machines around the world, there is a strong need for automated analysis using data science and machine learning methods. The research of the group targets a broad array of applications in plasma control and plasma diagnosis, in increasing the understanding of the physics of magnetized fusion plasmas and in designing new fusion machines. This research combines two of the most topical and challenging issues of our time: sustainable energy supply and data science.

Problem statement

Fusion energy research is a long-term collaboration of specialists with various areas of experience, including technicians, engineers, physicists (experimentalists, theoreticians, modelers) and managers. The volume of published works, both internally and in research papers, is substantial. Graphical representations, such as plots, annotated images, or industry-standard diagrams, play a key role in communications regarding the operation of the machine and diagnostics, planning of experiments, measurements and models, physical interpretations, etc. However, uncovering specific information from the vast repositories that have been accumulated over many years, can be a considerable challenge.

Recent advances in large language models (LLMs) have shown an impressive capacity to generate acceptable responses to queries based on newly provided (in-context) information. A further step is the understanding of image data. Multi-modal systems utilize image-to-text neural networks to map figures to a common latent space, also known as *embeddings*. LLMs can then be used for retrieval tasks, classification, or even generation.

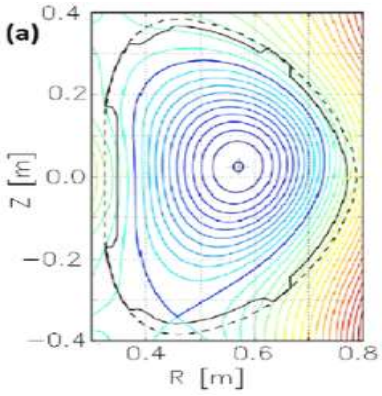
Goals of the internship

The goal of the internship is to create an expert system that can understand figures produced in fusion science, by applying existing image captioning techniques on a curated dataset. The internship is part of a larger project on conversational AI for querying intranet pages and documents of fusion institutes. The end user will be able to prompt the chatbot to look for a figure, understand its context, or answer questions about it. This will include the following research tasks:

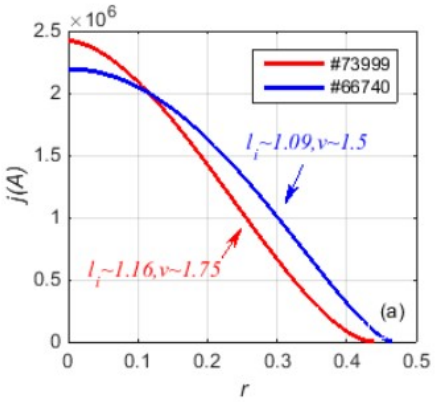
- Create a dataset of fusion physics images and their captions.
- Fine-tune a vision-language model to this dataset. Generate captions.
- Evaluate performance (e.g. sentence similarity).
- [optional] Improve generated output with image segmentation.
- [optional] Enrich generated output with OCR for annotations.

Depending on the student's background (emphasis on physics or computer science), the internship can focus on correct interpretation by the chatbot of physics results, or on the technical aspects of the tool.

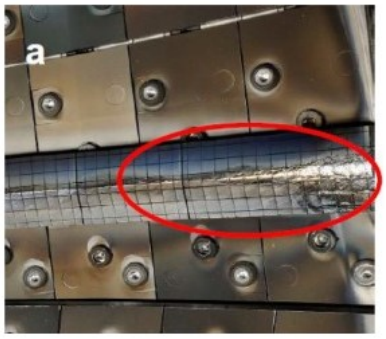
Below are a few examples of images with the original captions.



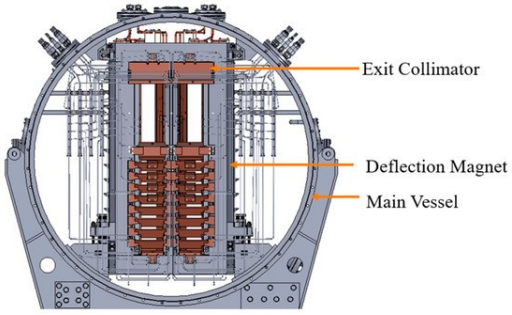
COMPASS poloidal cross section: (a) Equilibrium reconstruction of magnetic flux surfaces (discharge #6358 at t = 1080 ms). (Source: 10.1088/0741-3335/57/6/065006)



Current density profile of EAST #73999 at 50 s and #66740 at 14 s. (Source: 10.1109/TPS.2022.3179023)



Melt zones on beryllium plasma-facing components: (a) upper dump plates. (Source: 10.1016/j.fusengdes.2018.03.027)



Structure and layout of DM inside the second spool of the NBI system. (Source: 10.3390/en15051911).