

BrainCreators

Research Internship 2021-2022: Symbolic & Subsymbolic Hybrid Ai

[General Information](#)

[Contact and Interviews](#)

[Introduction and overview](#)

[Research internships in our applied vertical teams](#)

[Research internships on other activities](#)

[What we offer, what we expect](#)

[The Research Internship Position: Symbolic-Subsymbolic Hybrid Ai](#)

[Summary](#)

[Product](#)

[Technology](#)

[Neural Logic Programming](#)

[Research Questions](#)

[Engineering & MLOPs](#)

[Sources](#)

General Information

Contact and Interviews

Before you read on, .. **we encourage interested candidates to contact us as soon as possible for an intake interview!**

Please contact our head of research, Maarten Stol: maarten.stol@braincreators.com

Or visit our website: <https://www.braincreators.com/contact>

Like previous years, we have a limited number of available positions, and expect another rise in the number of candidates. Interviews will take place in September & October, and decisions to hire will be made for a timely start in November 2021.

To some degree, and given equal skills, this will be a first-come-first-serve selection process. (there is a chance of new positions opening up later in the year, so if you read this after November 2021, the door is not fully closed yet)

Introduction and overview

Welcome! You are looking at the research internships BrainCreators has on offer in 2021-2022.

BrainCreators is at the forefront of applied AI, with many years of successful research internship projects that combine cutting edge science with the challenges of applying AI in the real world. Located at Amsterdam's Prinsengracht and Science Park, we are a growing team of AI experts, software developers, MLOps & DevOps specialists and researchers.

Research internships in our applied vertical teams

The 4 business verticals that offer a research internship position this year are:

- **Road surface inspection**, combining Deep Learning Object Detection with Geo-information (and possibly 3D data).
- **Conveyor belt applications**: recognition, localization, and manipulation by robot of objects on a conveyor belt. Challenges concern high variance of object shape and visuals, and detection of out-of-distribution imagery.
- **Video surveillance**, based on, and extending our anonymization tooling. The focus is on understanding person and crowd behavior, anomaly detection, and video retrieval, all based on video representation Deep Learning and self-supervision.
- **Fashion & Retail**: this year with a focus on generative models for Virtual Try-on of clothing items.

Research internships on other activities

In addition to our business verticals, there are research topics that are more general, or concern pure research which is not immediately related to our commercial activities.

If you would like more information on topics like these, please contact our head of research, Maarten Stol: maarten.stol@braincreators.com

- **MLOps** is an essential part of every product we roll out live. Topics include data unit tests, live evaluations, deployment monitoring, handling shifting data, containerization, building KubeFlow pipelines, and scaling deployments.
- **Symbolic/Subsymbolic Hybrid Ai** In particular we are interested in compensating a lack of annotated training data with symbolically encoded background knowledge about the application domain. If valuable explicit background knowledge is available in the form of

rule-based information, then we are interested in e.g., imposing this knowledge as regularizers on our object detection models, or in other ways to exploit relational information.

- **Astronomy** A position working in tandem with our partners on the Cortex Consortium in the field of astronomy. BrainCreators is an industrial partner in this 6 year project, providing research and development with a focus on topics like neural network compression and autotuning of real-time ML pipelines. For a general impression see:
 - <https://www.uva.nl/en/shared-content/faculiteiten/en/faculteit-der-natuurwetenschappen-n-wiskunde-en-informatica/news/2019/06/self-learning-machines-hunt-for-explosions-in-the-universe.html?cb>
 - <https://www.esciencecenter.nl/projects/cortex/>

What we offer, what we expect

We offer:

- Be part of a growing company with a proven track record in applied Ai
- A research internship position on one of our vertical teams
- Interaction with research interns from our other vertical teams, in a science oriented horizontal research team.
- A protected environment for your research, without distraction by commercial deadlines of the team
- Opportunities to contribute to the team by developing dual-use software: for your own research and the team's products.

- Weekly supervision on scientific progress, experimental design, and thesis text
- Weekly supervision on software development and code reviews
- Daily contact with the vertical team, and morning stand-up meetings
- Weekly participation in internal ML workshops, sharing ideas with others
- Access to compute resources (in addition to University resources)
- Opportunity to work from home, or work from our HQs at Prinsengracht or Science Park Amsterdam.
- A financial compensation of 300 euros per month

- Learn all the essential things a Master program typically does not offer, e.g.,
 - onboarding with software development skills,
 - MLOps skills,
 - optimal use of compute resources,
 - versioning of ML and datasets,
 - collaboration software,
 - and communication skills.
- Be the eyes and ears of your team, looking for promising academic developments that might be relevant to the vertical
- Opportunities to become a permanent team member, and join as ML engineer after the research internship.

We expect:

- Workload contribution of 40h per week, 6-8 months (all activities related to your MSc program are included in this 40h, other jobs and classes are not)
- Capable to work independently on your own research questions and experiments
- Active participation in team effort when needed
- Solid control of spoken and written English language

- A strong opinion on ML research and how to apply it in practice
- Solid fundamental knowledge of ML theory and practice
- Overall knowledge level of a graduating Ai MSc student
- Good PyTorch skills
- Good understanding of the required mathematics
- Good software development skills

- Active participation in internal workshops, presenting your progress, and discussing your experimental design choices with your team and other verticals in the company
- Willingness to rewrite the thesis as a publishable paper
- Co-authorship for your thesis supervisors on publications derived from the thesis.

Please contact our head of research, Maarten Stol: maarten.stol@braincreators.com

Or visit our website: <https://www.braincreators.com/contact>

The Research Internship Position: Symbolic-Subsymbolic Hybrid Ai

Summary

Join our science team working on long term research questions and help us explore the big bets of academic research as it applies to our business activities. Move between teams to find out how Hybrid Ai might help each set of products. The end-products of your projects will become part of our commercial R&D in the future. Describe your scientific results in a publishable paper.

Product

Symbolic-SubSymbolic Hybrid Ai can be relevant to all of our products. Labeled data is always a bottleneck. Rule based background knowledge, and reasoning, can be another way to provide the learning systems with domain specific knowledge.

BrainCreators is willing to spend 1 year of R&D effort on the topic of Symbolic-SubSymbolic Hybrid Ai. You will join our internal science team, and perform experimental research in order to improve our understanding of relevant Ai technologies. The resulting software or insights serve as the basis for

commercial work in following years. A publishable paper should also be the aim of the research intern and their internal/external supervisors.

Technology

Examples of interesting technology include the following topics. These are examples we know we are interested in. An internship does not need to cover all of these, but combinations of topics like this are perhaps of higher scientific and practical relevance.

Neural Logic Programming

The primary reason BrainCreators is interested in symbolic hybrid ML is the potential to decrease the dependency on labeled data. Other approaches, like self-supervised learning have their place, but when valuable background knowledge is available in the form of symbolic, rule based knowledge, it seems like a waste not to use it. However, getting rule based knowledge into the ML pipeline remains a challenge, just like getting symbolic information out of the pipeline for the purpose of ML explainability. For a good introduction and overview of the field, see [1].

The previous two years, BrainCreators has explored ways to approach this problem. One was based around the use of graph based knowledge for the purpose of ML regularization in object detection. Our work together with the University of Groningen was based on [2]. Other work, together with Vrije Universiteit Amsterdam and Insight Centre Dublin, concerned the extraction of rules and graphs from trained neural networks, based on [3]. Both projects have delivered valuable results and insights for our application of this technology.

An example of other technologies that we are interested in this year is the combination of Deep Learning with probabilistic logic programming, found in [4]. The DeepProbLog framework seems to combine DL with ways to exploit rule based background knowledge in some of our use cases. More recent work that seems to overcome certain challenges is [5].

While classical experiments with DeepProgLog showed e.g., how to do multi-integer arithmetic with MNIST digits, or even solve Sudoku games with MNIST digits, we would like to take this a step further. In particular, we are interested in extending the frameworks from image classification to object detection. This is not a trivial challenge. Getting a differentiable signal to pass between the ML part and the reasoning part will require some fundamental changes to the architecture.

Perhaps a simpler approach is possible. If fully adapting DeepProbLog to object detection is too daunting, then at least we can work with crops of the defects and use their positional meta-data in the reasoning modules.

As an example, in road defect detection we may know that defect A always occurs at the edges of the road, while defect B occurs in the middle. We would like to impose this knowledge on the ML pipeline. In crowd analysis by video, we may want to describe typical trajectories for pedestrians. When recognizing and manipulating objects on conveyor belts, we may want to reason about subparts of objects, or about colors, or any other aspect of an object that is typical for a class but varies between individual objects and their orientations.

Hybrid Ai is a hot topic. There are many other approaches currently available in the literature besides Neural Logic Programming. The candidate is encouraged to bring to our attention any approach they deem promising for our products or that has their particular interest

Research Questions

A research internship could focus on questions like:

- Is increased performance in object detection possible using a framework like DeepProbLog?
- What is the price we pay? How much knowledge engineering for a single application domain is needed, and how much labeling effort can we avoid this way?
- When keeping the existing DeepProgLog architecture for classification, which compromises might enable us to apply the architecture to object detection challenges, e.g, like working with object image crops?
- If hybrid integration with object detection is possible, which types of rules work best? What is the influence of the size of the knowledge base on system performance? E.g., are many shallow rules better than a small but deep knowledge base?
- How much time is spent in logical inference vs. backpropagation? How is the CPU/GPU hardware best utilized during training?
- For each of our business verticals, what are the pros and cons of a hybrid approach? How do verticals differ, and what are their common properties wrt Hybrid Ai?

Engineering & MLOPs

The research intern will be partly responsible for integration of developed technologies into our product stacks, to facilitate deployment and scaling of the solutions with MLOps.

While this requires a substantial amount of skills that are often different from typical Ai research, we hope to provide the research intern the opportunity to learn as much as possible, and implement the solution together with our team.

Support for working with data from image capture hardware will be provided by our partners and us. We also envision a modular character for the internship, where most of our topics of interest can be developed independently from the engineering questions related to image capture.

Sources

[1] Modular Design Patterns for Hybrid Learning and Reasoning Systems: a taxonomy, patterns and use cases <https://arxiv.org/abs/2102.11965>

[2] Hybrid Knowledge Routed Modules for Large-scale Object Detection
<https://arxiv.org/abs/1810.12681>

[3] Extracting knowledge from Deep Neural Networks through graph analysis
<https://www.sciencedirect.com/science/article/pii/S0167739X21000613>

[4] Neural Probabilistic Logic Programming in DeepProbLog <https://arxiv.org/abs/1907.08194>

[5] DeepStochLog: Neural Stochastic Logic Programming <https://arxiv.org/abs/2106.12574>

For more information, please contact our head of research,
Maarten Stol: maarten.stol@braincreators.com
Or visit our website: <https://www.braincreators.com/contact>