Project Proposals for Doctoral Research Positions 2020

ID01: Cooperative multi-agent reinforcement learning for next-generation cognitive robotics in laparoscopic surgery (Mathis-Ullrich, Wagner)

Doctorate at the faculty of Informatics

Laparoscopic surgery is a team effort. A surgeon and her assistant(s) collaborate to solve a shared task working individually and as a team for a successful surgical outcome. However, our society faces increasing shortage of skilled surgeons and assistants, especially in rural areas. This shortage may be resolved by providing cognitive surgical robots that automate certain tasks. Our project addresses the highly interdependent behavior of surgeon and assistant(s) as a multi-agent system problem of human and artificial (software) agents using methods of cooperative multi-agent reinforcement learning (cMARL). In contrast to previous work, this project aims to train multiple, decentralized artificial agents that cooperatively solve a shared, robot-assisted laparoscopic task. After training the artificial agents, a human may take over one or more of the hardware manipulators during critical phases. Critical phases are either identified by the human, or proactively by the artificial agents in situations of high policy uncertainty. As a result, we hope to improve surgical treatment quality by robotic assistance that allows for real human-robot-cooperation instead of simple automation of sequential tasks. Specific methodological challenges include but are not limited to policy quality, sample efficiency, and human-robot-cooperation.

ID02: Human Rules – AI brains: Automated CTV delineation for head and neck cancers (Frank, Jäkel)

Doctorate at the faculty of Mathematics

The project will investigate how best-in-class machine learning approaches cope with the most challenging task in radiation therapy: delineating the clinical target volumes. The primary goal is to establish a reliable, fully automated pipeline that is designed to respect the current expert guidelines. Comparison with straight-forward purely data-driven deep learning approaches will clarify the impact of such hybrid symbolic and non-symbolic methodologies. The adaptability of both approaches to modification in treatment guidelines – expected in evidence-based medicine – will also be clarified.

In the thesis, it will be investigated how accurate the straight forward data-driven best- in-class ML methods perform on the specific head and neck cancer cohort. Second, a simple divide-and-conquer strategy will be developed, which consists of a pipeline of single tasks (like localizing a certain bone in the CT scan) which will be arranged so that expert guidelines are satisfied. Third, on the methods side, we will investigate how constraints can be built into the learning routines. From a mathematical perspective, an ML algorithm solves a high-dimensional unconstrained global optimization problem. Including rules will be achieved by translating the constraints into either the target functional or the architecture of the neural network.

ID03: Scalable Cell-Tracking with Learnable Combinatorial Optimization (Rother, Nienhaus)

Doctorate at the faculty of Mathematics and Informatics

Cell segmentation and tracking is the problem of processing a time series of (3D) images showing development of an organism (e.g. drosophila) on the cellular level, that is, growth, movement, division and death of cells. Existing methods for this problem work reasonably well in simple cases with relatively few cells per time frame and relatively small temporal changes between them. The goal of the project is to develop a new cell-tracking method which is able to cope with the problem at later development stages, where existing approaches hit their limit. The new key components of the method are: (a) a scalable combinatorial solver which is able to efficiently deal with millions of variables and constraints, as cell-tracking can be seen as a large-scale combinatorial problem; and (b) a training technique to learn the parameters of the solver to fine-tune it to different types of input data.

ID04: Robust data-driven prediction of complications in minimally-invasive surgery (L. Maier-Hein, Kenngott)

Doctorate at the faculty of Mathematics and Informatics

Death within 30 days after surgery has recently been found to be the third-leading cause of death worldwide. In this context, anastomotic leakage has been recognized as one major cause for intraoperative or postoperative complication. Anastomoses in medicine refer to the surgical connection between two diverging tubular structures, such as blood vessels or intestines, while leaks may occur as late as days or even weeks postoperatively. Recent (yet unpublished) results suggest that the risk of anastomosis leakage varies crucially from hospital to hospital and can reach up to almost 20%. Reducing failure rates and understanding the reasons for failure could thus lead to clinical practice changing impact.

The OR Black Box® is an innovative, unique platform that allows healthcare professionals to identify, understand, and mitigate risks that impact patient safety. It combines input from video cameras, microphones, and other sensors with human and automated processing to produce insights that lead to improved efficiencies and reduced adverse events. It gathers such observational data from hospitals across the globe and has led to a valuable annotated surgical data pool.

The goal of this project is to develop the first fully-automatic video-based approach to anastomosis leakage prediction based on surgical black box data. Core data science challenges are related to heterogeneous data integration and scalability (generalization to different devices and hospitals).

ID05: ML based parameterization to simulate tissue and tumor development as emerging property from single cell events (Streit, Schug)

Doctorate at the faculty of Informatics

Despite decades of substantial research, cancer remains a ubiquitous scourge in the industrialized world. Effective treatments require a thorough understanding of macroscopic cancerous tumor growth out of individual cells in the tissue and microenvironment context. Clinical imaging methods only detect late-stage macroscopic tumors and provide low resolution imaging, while many quantitative experiments focus on small clusters of cancerous cells in microscopic detail but struggle to grow full tumors in-vitro.

Here, we aim to introduce the critical scale-bridging link between both these scopes by employing machine learning to drive model building between these scales. We want to simulate mm-sized virtual tissues such as embryogenetic brain tissue or cancerous tumors with more than a million µm-resolved individual cells by employing highly parallelized code on a supercomputer. Machine learning will be used to combine different scales of imaging as well as clinical and qualitative measurements to drive model generation and parametrization. This improved theoretical understanding of tissue growth as emerging behavior from single-cell events opens new avenues for different scientific fields, ranging from fundamental biology to testing treatment regimes in-silico for personalized medicine.

ID06: Integration of omics data to discover biomarker signatures for hypoxia and radioresistance (Brors, Baumann)

Doctorate at the Faculty of Biosciences

Resistance against radiotherapy is frequently observed in cancer and is tightly linked to the ability of cells to survive in an evironment with low oxygen concentration (hypoxia). Both phenomena are, however, highly heterogeneous both in time and space. To study hypoxia and radioresistance in more detail, model systems have been created from cell lines and animals, in particular for head-and-neck cancer. These have been profiled with respect to transcriptomic, epigenomic and genomic signatures. Furthermore, spatially resolved proteomics will generate data on spatially varying cellular composition during the course of this project. We aim to integrate all of this data and link them to regulatory pathways that control response to hypoxia and onset of radioresistance. To analyse spatial heterogeneity, the proteomic data will be used to create graph representations for each sample that model cellular composition and microenvironment, and compare those between resistant and sensitive samples. Finally, we will collaborate with an image-analysis group (Klaus Maier-Hein) to bring data from this project together with results from analysis of MRI scans. Methodological challenges include integration of data on different levels of regulation, development of methods to study and spatial heterogeneity, and data science methods to associate imaging data with various "omics" data types.

ID07: Explainable Artificial Intelligence in Life Science: An Application to Omics Data (XAI-Omics) (Sunyaev, Schlesner)

Doctorate at the Department of Economics and Management

As it is becoming progressively challenging to wholly analyse the ever-increasing amounts of generated biomedical data (e.g., CT scans, X-ray images, omics data) by means of conventional analysis techniques, researchers and practitioners are turning to artificial intelligence (AI) approaches to analyse their data. Yet, extant AI approaches often suffer from opacity; their sub-symbolic representation of state is mostly inaccessible and non-transparent to humans, thus limiting us in fully understanding and therefore trusting the produced outputs.

Explainable AI (XAI) describes a recent trend in AI research with the aim of addressing the opacity issue of contemporary AI approaches by producing (more) interpretable AI models whilst maintaining high levels of performance and accuracy

The objective of this doctoral thesis is the design, development and evaluation of an XAI approach to omics data. In particular, this doctoral thesis aims to identify omics use cases and current, viable approaches in the domain of XAI and apply and adapt them to the identified use cases.

ID08: Visualization for MRI-Based Psychiatric Diagnosis (Sadlo, Durstewitz)

Doctorate at the faculty of Mathematics and Informatics

Since reliable biomarkers for psychiatric diagnosis and treatment indication are lacking, such diagnostics and prognostics is mainly based on structural interviews with the patient. Although abnormalities in non-invasive measurements, including structural and functional magnetic resonance imaging (fMRI), are associated with psychiatric conditions, the effects are often small and very heterogeneous, and thus do not directly allow for medically reliable classification of psychiatric conditions or subtypes. The group of PI Durstewitz has suggested that derailed cortical attractor dynamics, rooted in known biophysical (synaptic) alterations, could be a hallmark feature of many psychiatric illnesses. Identification of attractor dynamics in brain recordings is, however, a highly challenging topic, and only recently progress in statistical machine and deep learning has enabled to extract crucial dynamical systems features from neural, including fMRI, recordings. Major challenges for translating these methods and observations into clinical practice include their extension to additional data sources for extended robustness, effective and efficient analysis of the resulting very high-dimensional state spaces, identification and extraction of those dynamical features most crucial to psychiatric diagnosis and treatment, and, last but not least, presentation of the results, together with context information, in a format that is accessible to clinicians. The goals of this project is to develop advanced visual data analysis techniques addressing these challenges. On the one hand, we develop techniques that reveal topological features, e.g., invariant manifolds in the high-dimensional phase space of the inferred dynamical systems, support navigation of these spaces, and bring them into context with additional data sources. On the other hand, we include structural magnetic resonance imaging data, develop respective local feature definitions, and investigate the integration of these spatial data with the high-dimensional phase space of the inferred dynamical systems. Last but not least, we combine the obtained techniques into an overall approach that enables effective application and interpretation by clinicians.

ID09: Data-driven Modeling and Uncertainty Quantification for Hereditary Cancer Research (Heuveline, Schlemmer, Frank)

Doctorate at the faculty of Mathematics & Computer Sciences

In the recent years, data generation and storage follow the Moore's law, leading to a quasi-exponential growth of the available data. This statement applies specifically for hereditary cancer research. In particular, the costs for omics data dropped dramatically including genome and exome sequencing. Moreover, histological images describing the molecular properties of tumors are an important source of high-dimensional data which can be beneficially used for diagnostic procedures and treatment optimization.

This tremendous amount of data needs to be transformed and investigated in such a way that a clinical knowledge on cancer evolution can be derived, which is crucial for adequate diagnosis, prevention and treatment strategies for different kind of tumors. Here, the innovative field of data-driven modeling (DDM) comes into play which is a concept to unify statistics, data analysis, machine learning and their related methods in order to understand actual phenomena with data.

Inherently, the considered data will be error prone especially due to measurement errors. In order to cope with this challenge, the data driven modeling will be coupled with methods in Uncertainty Quantification (UQ) to quantify the impact of uncertainties on data and mathematical models.

The goal of this project is to develop and analyze new and innovative computational approaches which couple DDM with UQ for solving current challenges in the field of hereditary cancer research. The derived models should enhance existing diagnostics and prevention strategies and allow for a better understanding of the interactions inherent to cancer development. By applying UQ-methods to molecular and mutational data, the impact of uncertainties both in the data and the model on methods needed by the clinicians will be analyzed.