Research project for the Research Master Data Science

Project title: 1 Intelligent camera system for detecting drowning persons in swimming pools

Project overview

Number of students	1
Project Type	Pre-launch research with external practice partners from the bath operation
Project Owner	Supervision by Dr. Schultenkämper and Prof. Brandt-Pook in cooperation with application partners from the swimming pool and safety environment
Project Context	The project is located at the intersection of computer vision, AI-based security technology and applied research in public space. Typical pool and video scenarios, specialist expertise from the spa operation, methodical support in the field of machine learning as well as computational resources for training and evaluation are available in the project environment. From a perspective, cooperation with municipal or private bath operators and connection options for student support activities in the research environment are conceivable.

Abstract

The project investigates AI-supported camera systems for early detection of potential drowning situations in swimming pools. The aim is to develop and evaluate methods of image and video data analysis that robustly detect critical motion patterns and support the supervisory staff with reliable warnings. In addition to model quality, data protection, real-time capability and practical usability are considered.

Short description

Drowning accidents in swimming pools often occur in a short period of time and are not always immediately detected despite supervision, especially in the case of high pool occupancy, restricted visibility or parallel events during bathing operation. AI-supported camera systems promise an additional one

A layer of security by analyzing video streams in real time and automatically marking noticeable changes in movement or position in the water. The aim of the project is to investigate data-driven methods for the detection of potential emergency situations, to implement them prototypically and to evaluate them with regard to accuracy, false alarm rate, robustness and practical integrability in the swimming pool operation.

Task definition

The project analyzes the state of research on video-based detection of critical situations in water and derives a suitable methodological approach. Based on this, relevant event classes are defined, suitable image or video data is prepared, an evaluation procedure for the detection of potential drowning situations is developed and evaluated using comprehensible quality measures. Depending on the progress of the project, procedures for reducing false alarms, for explaining model evaluation, for data protection-compliant processing or for real-time notification can also be investigated.

Reference to the topic of data science

The project addresses key areas of data science, in particular data acquisition and processing, exploratory data analysis, feature extraction from image and video data, machine learning and the systematic evaluation of models. Methodologically , the focus is on computer vision and deep learning methods, such as object recognition, tracking, action and anomaly detection and the analysis of temporal patterns in video sequences. In addition, topics such as data quality, class imbalance, generalization capability, false alarm minimization, interpretability, and data protection-compliant data processing play an essential role.

Available resources

Methodological support in the field of data science and computer vision, common development tools for the analysis of image and video data as well as computational resources for training, experiments and evaluation are available for the project. Depending on the level of cooperation, sample video or simulation data from pool-related scenarios, domain knowledge from the safety and pool environment as well as feedback from practical applications can be incorporated. In addition, open source libraries and established frameworks can be used for deep learning, tracking and video interpretation.

Project plan

Iterative approach and early prototypes

The project deliberately does not follow a purely sequential approach, but combines scientific analysis with an iterative, prototypical development.

Simple, data-based prototypes (Minimum Viable Models) are to be developed at an early stage of the project in order to develop a practical understanding of the challenges of video data analysis. These early approaches can be based on simplified or publicly available sample data and will be gradually expanded and improved as the project progresses.

The aim is to continuously combine theoretical findings with practical experiments in order to identify potentials, limits and typical sources of error at an early stage.

First semester: Exposé & Basics

- Structured analysis of the state of research (computer vision / anomaly detection in videos)
- Overview of existing technical approaches (e.g. Object Detection, Tracking, Action Recognition)
- Market Overview:
  - Existing commercial systems
  - Relevant open source approaches
- First classification:
  - Which approaches are plausibly suitable for swimming pools?
- Coarse methodical concept for your own procedure
- First simple MVPs based on sample or open source data for early validation of basic assumptions

Target image: *"I understand problem space + solution space"*

Second semester: Methodical elaboration & first prototypes

- Definition of specific use cases / event classes (e.g. motionless in water, atypical movement patterns)
- Design of the database:
  - What data would be ideal?
  - What data is realistically available?
  - Annotations concept
  - Class imbalance
- Derivation of a suitable modeling approach based on data availability and use case, e.g.
  - Monitored learning (labeled data - "drowning" vs. "normal", vs. ...)
  - Anomaly detection (learns "normal behavior" and detects deviations)
- Coordinate the data concept with the practice partner (e.g. camera positions, viewing angles, usage scenarios) and derive data collection requirements (e.g. relevant scenarios, data quality) that can be implemented by the practice partner
- Baseline approach (e.g. simple heuristics or standard models)
- Further development of first prototypes on project-specific data
- Initial evaluation and analysis of typical error sources

Target image: *"I can translate the problem into data and models"*

Third semester: Iterative development & evaluation

- Development and comparison of several model approaches
  - Classic CV vs. Deep Learning
  - Different architectures or feature approaches
- Iterative improvement:
  - Hyperparameter tuning
  - Feature Engineering
  - if necessary Data augmentation

- Focus on:
    - False alarm rate (very important in context!)
    - Ruggedness (e.g. full vs. empty cymbals)
    - Generalizability
- Systematic evaluation:
    - Clean evaluation design
    - Appropriate metrics (e.g. Precision/Recall, False Positives)
    - Comparison of several iteratively developed model variants
    - Analysis of typical error sources
- Reflection:
    - Where do the approaches work well/poorly?

Target image: *"I understand what really works – and why"*

Fourth semester: Integration & Transfer

- Consolidate the best approach
- Transform an iteratively developed prototype into a consolidated solution
- Consideration of practical implementation:
    - Real-time capability
    - Integration with existing monitoring systems
- Rating:
    - Technical feasibility
    - Practical usability
- Optional:
    - Sketch a simple pipeline / system architecture
- Master thesis + colloquium
- Transfer of knowledge to the practice partner
    - Preparation of the project results for the practice partner as well as discussion of possible application scenarios and limits
    - Documented program code

Target image: *"I'm driving research toward application"*

Necessary competencies

Compulsory: Very good knowledge of Python as well as solid foundations in statistics, machine learning and data analysis; interest in computer vision, careful scientific work and readiness to familiarize with safety-critical application contexts.

Optional: Experience with deep learning frameworks such as PyTorch or TensorFlow , knowledge of image and video editing, experience with evaluation design, anomaly detection, edge or real-time systems, and a basic understanding of data protection and ethical issues in AI-based video analysis.

Acquirable competencies

Through the project , the student acquires skills in the scientific analysis of a current application field of data science, in the development and evaluation of computer vision and machine learning procedures as well as in the structured preparation of complex image and video data. In addition, he/she strengthens skills in research design, literature work, experimental validation, critical interpretation of results and target-oriented scientific communication. In addition, he/she develops an in-depth understanding of the requirements of safety-critical AI systems, in particular with regard to robustness, false alarms, data protection and practical feasibility.

Research project for the Research Master Data Science

Project title: **2 Learning digital twins**

Project overview

Number of students	1-3
Project Type	Project with external partners
Project Owner	Prof. Dr.-Ing. Christian Schwede
Project Context	Project in cooperation with the **Fraunhofer Institute for Software and Systems Engineering** (ISST) in Dortmund and in the context of the **Center for Applied Data Science** (CfADS) and **the Institute for Data Science Solutions** (IDaS); employment as a student assistant at Fraunhofer is possible

Abstract

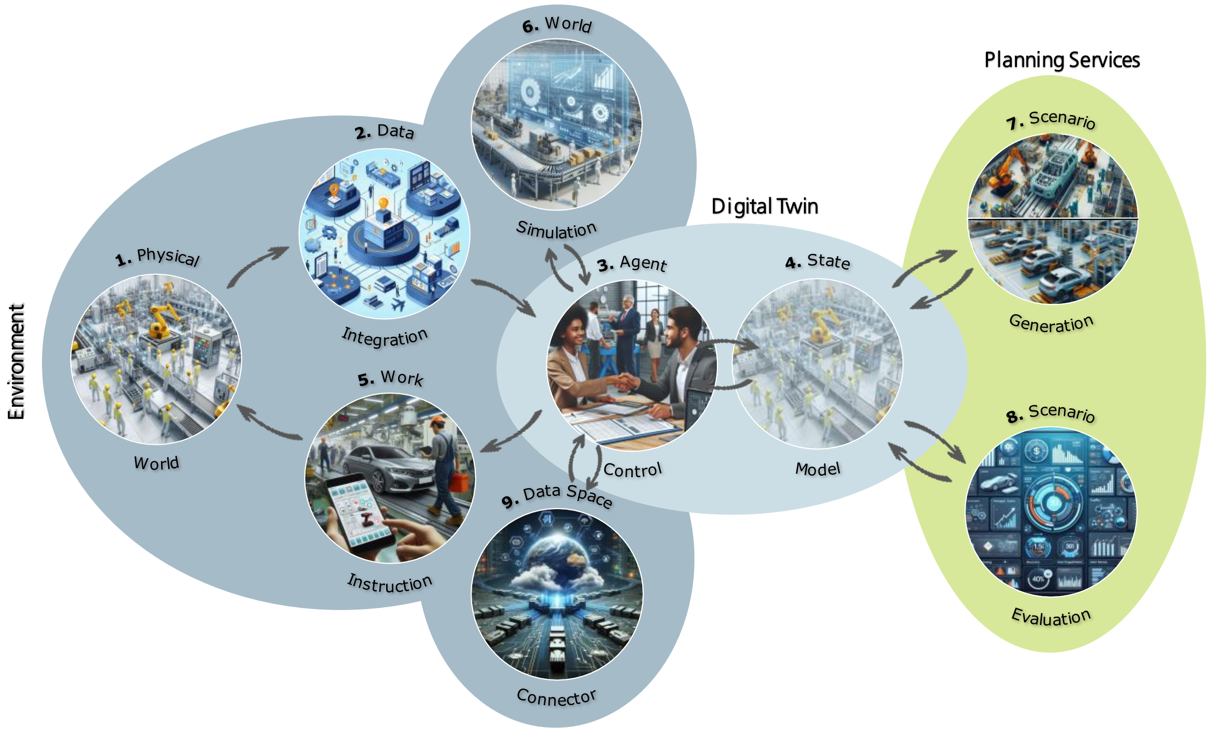
The aim of the project is to learn simulation-based digital twins for planning and controlling production and logistics environments from IoT data. It builds on and extends the open source framework OFacT developed by HSBI and Fraunhofer ISST. The data comes from Fraunhofer customer projects and the CfADS IoT Factory. In addition to process mining methods, generative models such as large language models (LLMs) are also used.

Short description

Digital twins are the key to the optimal design, planning and control of production and logistics environments. These software frameworks allow access to a constantly updated digital image of the system, enable the optimization and simulation of scenarios in this virtual environment and communicate changes directly back to the store floor by means of work instructions. In this way, decisions can be evaluated and optimized in all phases of the system's life cycle before they are made. Both long-term changes to the system and reactions to short-term disruptions can be optimally implemented. The potential for companies is estimated to be very high (forecast market volume of

USD 10 billion in 2028[1]). Fraunhofer ISST and HSBI have created the open source framework OFact (Open Factory Twin), which allows small and medium-sized companies in particular to digitize their production systems.

OFacT is based on a standardized state model that can be used to map any discrete material flow system. Any data interfaces can be addressed via the data integration module. The system is controlled by a multi-agent system so that control rules can be mapped and changed in a highly flexible manner and separately from the state model. The agents also control the material flow system by means of work instructions.



[1] Bain & Company: Global Machinery & Equipment Report 2024

As an environment, the framework can draw on a simulated reality to evaluate scenarios, for example. Planning services allow the creation and optimization of alternative courses of action and a dashboard allows evaluation based on standard key figures.

An important hurdle in the implementation of digital twins in small and medium-sized companies is the cost of manual modeling by experts. These costs are incurred not only during the creation, but also during regular updates. The aim of this project is therefore to learn digital twins from company data.

Task definition

In the project, the student should use methods from the field of machine learning to learn OFacT models based on data from production environments. The IoT data or event logs from the store floor of companies often already contain a lot of the graded information. Data from existing industrial projects from Fraunhofer ISST or the IoT Factory on the Gütersloh campus can be used. In addition to process mining methods, regression and classification methods, e.g. based on deep neural networks, will be used. It will also be investigated to what extent parts of the models can also be created with current Large Language Models (LLM). For this purpose, it may be useful to use Retrival-Augmented Generation (RAG) to provide the LLM with domain-specific information.

Reference to the topic of data science

The regression and classification methods used, as well as LLM and RAG, are part of the Research Master's course and core areas of Data Science.

Available resources

- Data is provided via Fraunhofer ISST and the IoT Factory
- The open source framework OFacT is being developed at HSBI and is available
- Hardware for machine learning is available via the Data Science Lab, CfADS and the AI computing cluster yourAI at HSBI

Project plan

First semester: Preparation of a research exposé as an examination. Familiarization with the task and the OFacT framework.

Second semester: Literature research on the generation of models using machine learning methods, in particular LLMs and RAG. Review of data sets and application of first simple methods of process mining and existing language models. Preparation of a paper that gives an overview of the research area is an examination achievement.

Third semester: Practical implementation of your own solution and generation of an OFacT model for a specific application. Publication of a paper with the first results is an examination achievement.

Fourth semester: Master's thesis and colloquium. Final evaluation by comparing different methods.

Necessary competencies

Mandatory:

- Programming skills (e.g. Python)

Optional:

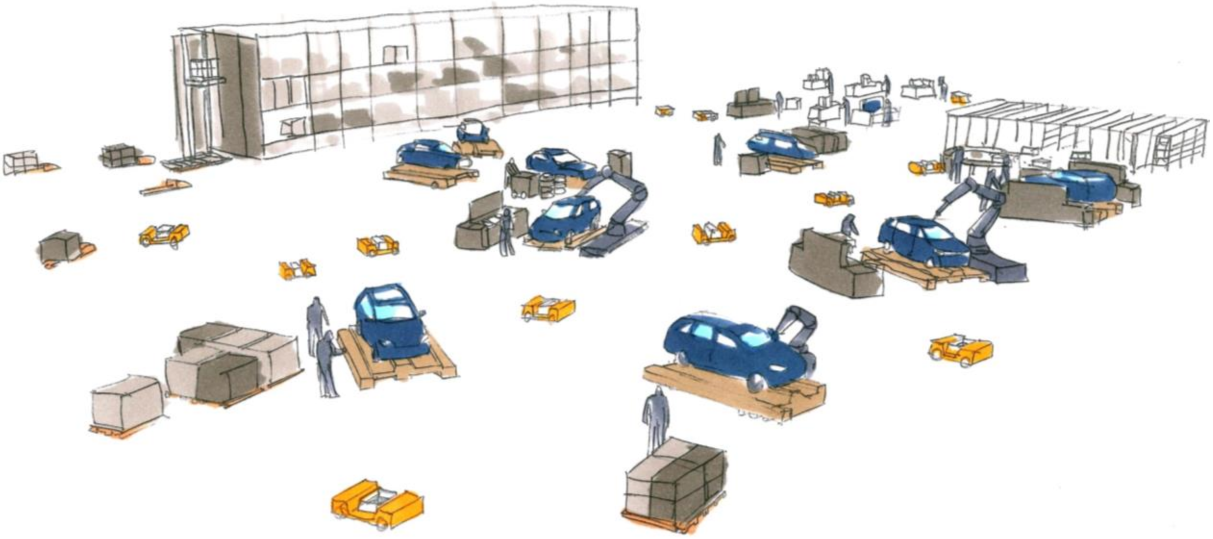
- Experience with machine learning
- Experience with LLMs

Acquirable competencies

- Development and use of digital twins
- Generation of models using LLMs and RAGs
- Publication of scientific papers at international conferences
- Competencies in working in a team
- Open source development

Research project for the Research Master Data Science

Project title: 3 AI for highly flexible, multi-variant mass production



Project overview

Number of students	1
Project Type	Project with external partners
Project Owner	Prof. Dr. Christian Schwede
Project Context	Project in cooperation with the **Fraunhofer Institute for Software and Systems Engineering (ISST)** in Dortmund and in the context of the **Center for Applied Data Science (CfADS)** and the **Institute for Data Science Solutions (IDaS)**; employment as a student assistant at Fraunhofer is possible;

Abstract

The constant individualization of products poses challenges not only for German automotive production, forcing a complete rethink of production organization. Away from clocked flow production towards decentralized matrix production systems. Here, swarms of autonomous robots coordinate the value creation processes and adapt flexibly to the current situation. While the basic idea of the concept is clearly outlined, there are still many open research questions in detail, particularly concerning the behavior of the individual robots. This is where the project comes in and explores different AI strategies in a virtual testbed.

Short description

In matrix production, production is coordinated decentrally by software agents in a multi-agent system (MAS). The resource agents negotiate with the order agent, for example, about the provision of value-added processes that are necessary for the realization of the order. The freedom of the sequence of production steps is only limited by the technical properties of the product. In this way, the entire system can adapt flexibly to unplanned incidents, such as a delivery bottleneck or a machine breakdown. While the concept was initially met with rejection and scepticism in the automotive industry, intensive efforts have been made in recent years to research and promote its operationalization. In principle, there are still many unanswered questions regarding its efficient use that need to be researched.

The open source OFacT (Open Factory Twin) software framework has been developed by the HSBI research group in cooperation with the Fraunhofer ISST in Dortmund for several years. Among other things, this tool allows highly flexible matrix production to be modeled using agent-based simulation and various strategies and design variants to be dynamically evaluated. The framework can also be used as a digital twin and can be connected directly to production via sensor data streams and used to control operational processes.

Task definition

The student has the opportunity to work on one of the following topics in order to explore the possibilities and limitations of matrix production and to contribute new solution concepts to the realization of the state of the art.

Agent control for decentralized production

One of the most crucial open questions is the optimal control of the agents. In particular, the synchronization of material flows with the main product is a decisive factor in preventing waiting times and maintaining flexibility at the same time. The development and implementation of efficient control strategies for the MAS is the core task here, which includes the question of whether and to what extent planning and reservations should be made in the future. Methods of optimization, search and planning algorithms from AI or reinforcement learning can be used here.

Layout design and optimization

Even if the term matrix production implies a grid-like arrangement of the production stations, the question of the optimal arrangement of the various elements of matrix production (stations, storage locations for parts, loading stations for transport vehicles, parking spaces for intermediate buffering of car bodies, etc.) has not yet been finally decided. Optimization algorithms, which are used in combination with simulation as an evaluation method, must be developed here in order to determine the optimal layout for a specific production setup. The question of how strongly the optimum layout depends on the specific order mix is relevant and could imply the need for regular re-planning of the layout.

Routing and congestion analysis

The formation of traffic jams in front of the production stations or on the various routes is a central aspect of the efficiency of the overall system. For this purpose, the simulation environment must be expanded to include a route planning element that also recognizes temporary route conflicts. AI search algorithms from the area of route finding must be implemented and various strategies tested. Finally, existing control strategies and layouts must be examined and evaluated against the background of congestion in the system.

Acceleration of negotiations by means of reinforcement learning

A weak point of multi-agent systems is the amount of communication that arises from the decentralized negotiations between the agents. Negotiation strategies need to be examined, implemented and compared and, in particular, unnecessary communication needs to be avoided. Reinforcement learning can be used to recognize when requests from agents are not expedient and can therefore be avoided in advance.

Agent-based value creation networks

The integration of logistical processes from the value creation network around production is another open topic. In addition to the development of suitable supply and disposal strategies, which must meet the needs of flexible matrix production, the visionary question of whether decentralized, autonomous value creation can also be extended beyond the factory boundaries should be answered. To this end, the agents must negotiate contracts with suppliers and service providers on the basis of money as a means of payment. The corresponding negotiation algorithms and scenarios must be implemented and evaluated.

Reference to the topic of data science

Software agents and multi-agent systems, search algorithms and reinforcement learning are core areas of AI and are addressed in various courses of the research master's program.

Available resources

- The modeling framework OFacT is provided by HSBI and Fraunhofer ISST
- Industry-related scenarios of individual bicycle production exist to validate the developments
- Scenarios from industrial customers of ISST and the IoT Factory in Gütersloh also exist for evaluating individual processes
- Hardware for more complex machine learning is available via the Data Science Lab or the CfADS of the

Project plan

First semester: Formulation of the research proposal, familiarization with the OFacT environment, selection of the topic.

Second semester: Modeling of an initial scenario, research on relevant work in the subject area and preparation of a report on the state of the art.

Third semester: Implementation and evaluation of initial procedures, evaluation of initial results based on the scenario, writing a paper for an (inter)national conference.

Fourth semester: Development of further procedures and optimization of the overall approach, evaluation of the quality and evaluation of the results in the Master's thesis

Necessary competencies

Mandatory:

- Programming skills (ideally in Python)

Optional:

- Experience with agent systems
- Experience with simulation and optimization
- Background knowledge of production logistics and production systems
- Experience with the implementation of reinforcement learning / ML methods

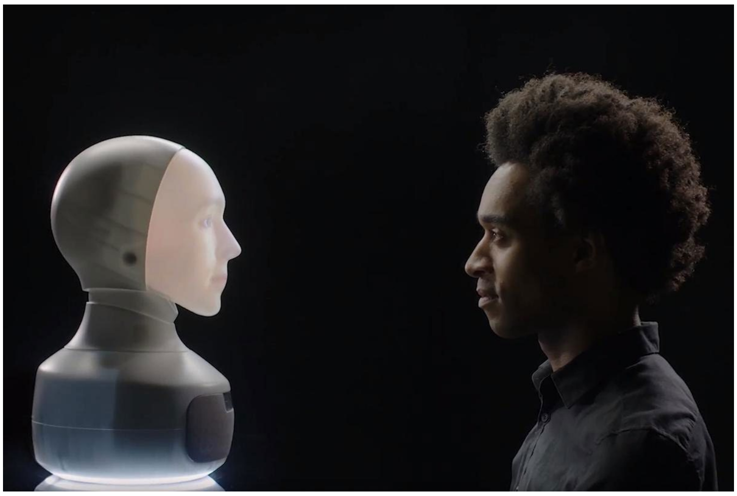
Acquirable competencies

- Use of reinforcement learning methods
- Development and evaluation of MAS
- Development of intelligent agents

- Working with agent simulations
- Planning of production and logistics systems, in particular market production systems and Industry 4.0

Research project for the Research Master Data Science

Project title: 4 Human-robot interaction: design, implementation and use of a humanoid robot as an individual conversation partner



Project overview

Number of students	1 to 2
Kind	Study project
Project responsibility	Prof. Dr. Thomas Süße and Dr. Maria Kobert
Project context	Possible employment as a student assistant depends on available financial resources. Cooperation with companies is also conceivable.

Abstract

The interaction between humans and social humanoid robots is central in current and future human-machine interaction. Social robots are able to communicate with people in a human-like way - by speaking, listening, showing emotions and maintaining eye contact. For example, they can be used to serve customers, provide companionship, train employees or teach a language. An example scenario with the Furhat robot, that is also available on campus Gütersloh is shown in this video:

https://www.youtube.com/watch?v=3lEQDf9Cv4s

Despite the many possible applications, the development and use of these promising technologies presents new challenges in terms of the fit and compatibility between the needs and requirements of users and humanoid robots. One challenge lies in configuring the robots in terms of speech, facial expressions and gestures so that they meet the individual needs of the users. This is where this project comes in. The overarching goal is to configure the humanoid robot Furhat, which is already available to the Department of Engineering and Mathematics at the Gütersloh campus, so that it can interact appropriately with humans in practical application scenarios to be defined. Furhat shall first be enabled to communicate and collaborate with people in a specific practical scenario. The practical scenario will be defined together with the project supervisors. In addition,

machine learning will be used to ensure that the robot learns from interactions with its human conversation partners and remembers their language, facial expressions, gestures, preferences and conversation content. Furhat should thus "develop its skills" and be perceived as an accepted, pleasant and helpful conversation partner by its human counterparts. Alongside the technical development of Furhat, the interactions between the humans and Furhat will also be examined from a human perspective. An experimental setting will be set up to investigate specific variants of configurations with regard to Furhat's "appearance" (in particular facial expressions, language, appearance) and their effect on the human counterpart.

Brief description

At present, the importance of interaction between humans and social robots is constantly increasing. The use of social robots is becoming increasingly widespread in the healthcare sector, in education, in hotels and restaurants, in shopping malls, in industry and at home. Possible fields of application include informing and activating older people, providing support in care-related services, supporting children and young people, e.g. with autistic conditions, supervising homework and generally imparting knowledge in the education and training sector, or receiving and accompanying visitors to companies, organizations or shopping centers. Despite the many potentials of the usage of these robots, there are also numerous challenges associated with their use. For example, initial studies with humans have shown a low level of acceptance of robots among people. In order to increase acceptance, a key challenge is to configure the robots in terms of speech, facial expressions, gestures and appearance in such a way that they meet the needs of the users. This is the overarching goal of this project. A learning algorithm is to be developed that enables Furhat to remember the characteristics of the human conversation partners and adapt to them.

The main research question of the project is: How can Furhat get to know people and remember them?

To this end, the social humanoid robot head Furhat, which was developed and is marketed by the Swedish company Furhat Robotics, will be used. Furhat is a robotic head without limbs with a face projected onto the inside, where the external appearance such as skin color, eyebrow position and make-up is adjustable. The Furhat robot thus realizes soft facial features and can participate directly in a conversation through natural movements such as shaking and nodding its head. Expressive characters are created with modern speech recognition, an advanced conversation system and automated lip synchronization.

Initially, the Furhat is available to students as a raw version. The first step is to configure it using a practical scenario (use case) and make it capable to interact with humans. In the second step, regular interactions with humans are to be carried out, e.g. weekly meetings. With the help of machine learning, Furhat should be made capable of building up a conversation history in the course of the joint meetings and thus remember past conversations with certain people. To this end, a learning algorithm is to be developed that will enable Furhat to remember and recall the language, facial expressions, gestures, preferences and content of conversations with the person in question. Furhat should therefore develop the ability, as a "conversation partner" for people, to link to information and results of past conversations and continue them. This can be done using a practical scenario (use case) in which a joint task between humans and Furhat is to be solved over several meetings. In addition, the human-robot interaction is examined from a human perspective with the help of interviews and observations.

Task

1. The raw version of the humanoid robot head Furhat has to be configured and made capable of interaction on the basis of a practical scenario (use case).
2. A learning algorithm has to be developed that enables Furhat to remember its human counterpart, his/her language, facial expressions, gestures, preferences and conversation content during regular interactions in the course of the use case developed under 1.
3. With the help of interviews with participating humans and observations of the interactions that take place, the human-robot interaction is examined from a human perspective.

Reference to the topic of data science

A machine learning algorithm will be developed.

Available resources

The Furhat robot head and the software required to program the robot are made available to the students.

Project plan

First semester: Preparation of a research exposé is an examination requirement.

Second semester: Preparation of a paper that provides an overview of the respective research area is an examination requirement.

Third semester: Preparation of a paper with initial quantitative results is an examination requirement.

Fourth semester: Master's thesis and colloquium

Suitability criteria

Mandatory: Programming skills, preferably in Python

Optional: Knowledge of conducting social science studies (e.g. conducting interviews)

Acquirable skills

- Testing a humanoid robot
- Artificial intelligence methods for problem solving
- Human-robot interaction

Research Project for the Research Master Data Science

Project title: 5 Machine learning for detecting anomalies and prediction of interactions based on movement information in a smart home environment

Project overview

Number of Students	1
Project Type	Project with external partners
Project Owner	Prof. Dr.-Ing. Thorsten Jungeblut
Project Context	Project in cooperation with **KogniHome - Technikunterstützes Wohnen für Menschen e.V.** in Bielefeld; Data sets and hardware resources are provided; Employment as a research assistant is possible; There is the possibility of close cooperation with members of KogniHome e.V., including Bethel, Steinel, Hettich, HUM Systems, C&S GmbH and many more.

Abstract

The aim of the project is the co-development of a system that uses AI methods to learn the usage behavior of the residents of a smart home in order to subsequently provide suitable recommendations for action or, if anomalies are detected, to initiate appropriate measures. The scientific challenge of the project is the application and evaluation of machine learning methods for learning interaction patterns based on the data provided by minimal-sensor technology, as well as the prediction of future interaction and anomaly detection.

Short description

Digitalization is penetrating more and more areas of life, and the associated networking of diverse components in our daily environment is leading to a situation in which constantly higher demands are being placed on everyday technology. Systems need to become smarter, automated and autonomous. Ideally, technology will anticipate the needs of people and they will no longer need to make corrections. Needs-based human-technology interaction requires the intelligent technical system (ITS) to adapt to the context of use and not the other way round. A requirement for intuitive interaction is therefore first and foremost the reliable recognition of the context of use, i.e. where the user is located and what action they are currently performing. Knowledge of regularly occurring interaction patterns enables the ITS to predict future interactions and predictively control assistance functions. Imaging sensor technology (e.g. cameras, high-resolution time-of-flight sensors) enables powerful recognition of the context of the action through person, object, gesture or even facial recognition, but brings with it the problem of collecting personal data. This may be undesirable for data protection reasons, particularly in private domestic environments, but also in office buildings or production facilities. Minimal sensors, such as motion or presence detectors, do not directly generate personal data, but only record selective information about the presence of people or objects in a spatially restricted area. For comprehensive detection of the context of use, the combination of many simple sensors promises sufficient detection of the context of use while at the same time preserving privacy. If more complex sensor technology is required for an assistance function (e.g. for voice or gesture recognition), it only needs to be activated (and only then) if the use of the assistance system is at least foreseeable. However, the continuous recording of the context of use also makes it possible to learn regular patterns of behavior. The ITS can derive expected future interactions from these learnt action patterns and make recommendations for the activation of assistance functions or prepare them (predictive control). In addition, deviations from the potential control state (anomalies) can also be recognized and responded to appropriately; the derivation of recommendations for action by the ITS can be rule-based in the simplest implementation (e.g. "If the motion detector in the hallway is activated, then the light should be switched on"). In systems with many sensors, however, this is very complex, cannot be flexibly customized and is difficult to scale to growing environments. In addition, different sensors provide different abstract information (e.g. local movement (PIR, ultrasound) or movement across several rooms (HF/microwaves), binary "on/off" or distances) or even more complex information from camera-based systems (e.g. number of recognized people/pets/objects (robot vacuum cleaner/transport platform)). It is therefore necessary to abstract the various pieces of information, anonymize them if necessary and use them in a holistic approach to learn the context of use, estimate future interaction and detect anomalies. The scientific challenge of the project is therefore the application and evaluation of machine learning methods for learning interaction patterns based on the data provided by the minimal-sensor technology, as well as the prediction of future interaction and anomaly detection.

Task definition

In this project, the student will co-develop an ITS that uses machine learning methods to learn the usage behavior of the residents of a smart home in order to subsequently provide suitable recommendations for action based on the real-time data or, if anomalies are detected, to initiate suitable measures. The development of the concrete application scenario, together with the KogniHome e.V. team, is part of the project. To illustrate this, the following scenario is described: "Sabine (74) gets up every morning between 7 and 8 a.m., goes to the toilet and then makes herself a coffee". Three activities can be recognized in this action sequence: Getting up, going to the toilet, making coffee. Each of these activities can now lead to recommended actions or anomalies. For example, the ITS could prepare for going to the toilet with suitable lighting or switch on the coffee machine. At the same time, the ITS monitors the activities and recognizes deviations from the usual behavior. For example, if Sabine falls on the way to the toilet and does not get up again on her own, the ITS should recognize this and, for example, call the emergency services or the nursing home.

Reference to the topic of data science

The evaluation and application of AI/ML methods for condition monitoring and prediction are a core topic of data science and are covered, for example, in the modules "Data Mining & Machine Learning" and "Artificial Intelligence". The recording of high-resolution measurement data (e.g. consumption data) from many sensors in complex living environments highly demands on the organization and processing of the data. This is the core of the "Big Data Architectures" module.

Available Resources

- Information required to create the scenario (system description, logistical processes, relevant key figures) is provided by KogniHome e.V.

- Access to the research apartment of KogniHome e.V.

- Extensive test data sets are available via the KogniHome e.V. research apartment

- The contact person at KogniHome e.V. will be available for the duration of the project

- Required materials will be provided by KogniHome e.V.

- Hardware for the more complex machine learning is available via the Data Science Lab, CfADS and the AI computing cluster yourAI at HSBI.

Project plan

First semester: Preparation of a research exposé as examination performance. Familiarization with the concepts and structures (IoT software, smart home protocols, interfaces, etc.) of the KogniHome e.V. research apartment.

Second semester: Creation of the system concept for anomaly detection for predictive control of assistance systems. Research on relevant work in the field of the use of AI methods for learning event sequences and for anomaly detection. Preparation of a paper that provides an overview of the respective research area as an examination performance.

Third semester: Practical implementation of various machine learning methods in the field of unsupervised and (semi-)supervised learning to optimize automated interaction between humans and machines, as well as their evaluation.

Fourth semester: Master's thesis and colloquium. Final evaluation by comparing the implemented strategies. Preparation of a paper with initial quantitative results as an examination.

Necessary Competencies

Mandatory:

- Programming skills (especially Python)

- Experience with the version control system "git"

Optional:

- Experience in electronics development

- Experience in the field of smart home technologies/IoT devices

- Programming of microcontrollers

Acquirable Competencies

- Artificial intelligence methods for problem solving

- Sensor-based information processing

- Predictive assistance and anomaly detection (transferability to industrial processes)

- Expertise in working in a team

Research Project for the Research Master Data Science

Project title: **6 Efficient Execution of Object Detection Algorithm on Edge Devices**

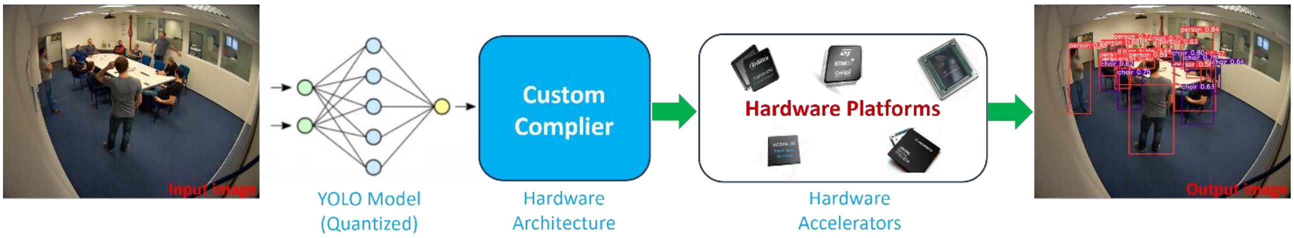


Figure 1: Efficient mapping of YOLO-Vx model on edge devices

Project Overview

Number of Students	1
Project Type	Study project in collaboration with an external partner
Project Owner	Prof. Dr.-Ing. habil. Thorsten Jungeblut Dr.-Ing. Qazi Arbab Ahmed
Project Context	**CareTech OWL** transfer project in collaboration with an external partner, **Steinel GmbH**, as part of the CareTech OWL research project. Parallel employment as a student research assistant (WHK) is possible.

Abstract

The artificial intelligence (AI) paradigm has replaced a handful of traditional computer vision techniques to intelligently predict the correct output of a complex system under tight constraints and varying conditions. However, many AI algorithms require huge amounts of hardware resources due to their massive computational requirements. The main task of this project is the development of a framework for the efficient implementation of object detection algorithms on highly resource-constrained hardware architectures (low-end FPGAs, embedded microcontrollers). Based on the framework, automated methods for exploring the design space of suitable combinations of object detection algorithms and hardware in terms of HW/AI-co-design will be explored. In particular, industrial applications with high latency requirements are targeted, such as real-time human presence detection.

Short description

For the efficient execution of AI algorithms, techniques such as federated learning (FL) and cognitive edge computing (CEC) have already shifted the load of training and inference of neural networks from the cloud to the edge. The partitioning of the execution of the application has a significant impact on the performance and resource efficiency of the overall system. For example, a distinction can be made here between fundamentally different approaches of decentralized feature extraction as close as possible to the sensor with subsequent fusion versus central processing in the cloud. The first approach requires powerful edge hardware for preprocessing and potentially offers advantages for high real-time requirements. The second approach places higher demands on the communication infrastructure but potentially enables the execution of more complex networks. The main goal of this project is therefore to explore automated methods for exploring the design space of suitable combinations of AI-based objection detection methods and resource-constrained hardware in terms of HW/AI codesign.

In the area of efficient execution of AI procedures on embedded systems (cognitive edge computing), great progress has been made in the past. At all levels of the different processing concepts (cloud, fog, edge, very edge) in the production chain, a variety of potential hardware architectures and AI accelerators can be found, which differ in the available system resources (e.g., performance or power consumption). Examples of relevant hardware architectures are embedded microcontrollers with integrated AI acceleration, embedded GPUs/FPGAs, dedicated AI hardware accelerators, or high-end GPUs/FPGAs from the HPC domain.

We consider the complete processing chain, starting with the pre-processing close to the sensor (very edge), by connecting the groups of sensors and their information in the edge, via local decentralized cloud instances (fog), to the central station of all relevant information (cloud). At each level in the processing chain, individual approaches exist to locally optimize the resource efficiency of AI methods, for example, by reducing numerical precision (e.g., from 32 to 16, or 8 bits) to enable more efficient execution on specialized hardware or to minimize local storage space. The goal of this project is to determine an optimal combination of hardware and object detection algorithms (e.g., YOLO) as part of a holistic design space exploration close to the sensor.

Task definition

In this project, students will develop a framework for the efficient execution of AI-based object detection algorithms on resource-constrained edge devices. In order to support the sustainable use of the developed methods and frameworks by the industrial partner, universally applicable models and automated design tools are to be made available in this project in the form of a standard development toolkit. The first main challenge is the compression of the latest object detection algorithm (e.g., YOLO) to reduce the model size, yet with acceptable accuracy, using AI approximation techniques. The next challenge is to develop resource-efficient techniques (inference compiler) to map the compressed model to a suitable hardware platform e.g., low-end FPGAs or embedded (AI-) microcontrollers, considering the resource utilization in terms of area, power/energy, and latency/throughput, as demonstrated in Figure 1.

Reference to the topic of data science

The evaluation and application of AI/ML methods in the field of "machine vision", e.g., the use of CNNs in object classification, are a core topic of data science and are dealt with, for example, in the modules "Data Mining & Machine Learning" and "Artificial Intelligence". AI-supported image processing places high demands on the organization and processing of data at all levels of IoT processing concepts (edge/fog/cloud). This is the core of the module "Big Data Architectures". The consideration of the entire system process from the imaging sensor to the cloud requires a holistic view of the complete data science process, which is covered in the module "Data Science".

Available Resources

Ensuring the availability of data, computing resources, hardware, application experts

- Information required to create the scenario (system description, interfaces, documentation, relevant key figures, etc.) is provided by the resource person

- Steinel GmbH provides extensive test data sets from real production environments
- The Steinel GmbH contact person will be available for the duration of the project
- The components required for the prototype development as well as other required materials will be provided by the Steinel GmbH
- Hardware for the more complex machine learning is available via the Data Science Lab, the CfADS, and the AI computing cluster, yourAI, at Bielefeld University of Applied Sciences

Project Plan

First semester: Preparation of a research exposé as an examination. Familiarization with the concept of AI-based object detection algorithms, in particular, YOLO, R-CNN, image recognition, neural network approximation techniques, hardware platforms, and design flow tools.

Second semester: Creation of the system concept for design space exploration of cognitive edge computing architectures. Research on relevant work in the field of AI-based object detection algorithms and model compression in the above-mentioned context. Preparation of a paper that gives an overview of the respective research area, as an examination.

Third semester: Development of a first demonstrator and proof-of-concept for hardware acceleration for human presence detector (HPD) using, for example, YOLO-V7. Comparison of the developed framework with a classic implementation.

Fourth semester: Master's thesis and colloquium. Implementation and comparison of different combinations of AI-based object detection algorithms and hardware accelerators. Systematic evaluation and exploration of the efficiency of the developed framework. Final evaluation by comparing the implemented strategies with the state-of-the-art methods. Preparation of a paper with the first quantitative results as an examination.

Necessary Competencies

Mandatory:

- Expertise in Python, Pytorch
- Good knowledge of C++

Optional:

- Experience with hardware design flow tools
- Programming of microcontrollers/FPGAs
- Basic knowledge of HDL (Verilog, VHDL)
- Experience with IoT devices
- Experience with the version control system "git"

Acquirable Competencies

- Resource-efficient information processing at edge (embedded microcontrollers, FPGAs) in line with the IoT processing concept
- Sensor-related information processing
- AI/ML-based object detection methods
- Leveraging embedded hardware to accelerate AI/ML processes

Research Project for the Research Master Data Science

Project title: 7 Hardware-AI Codesign for Efficient Execution of Neural Networks

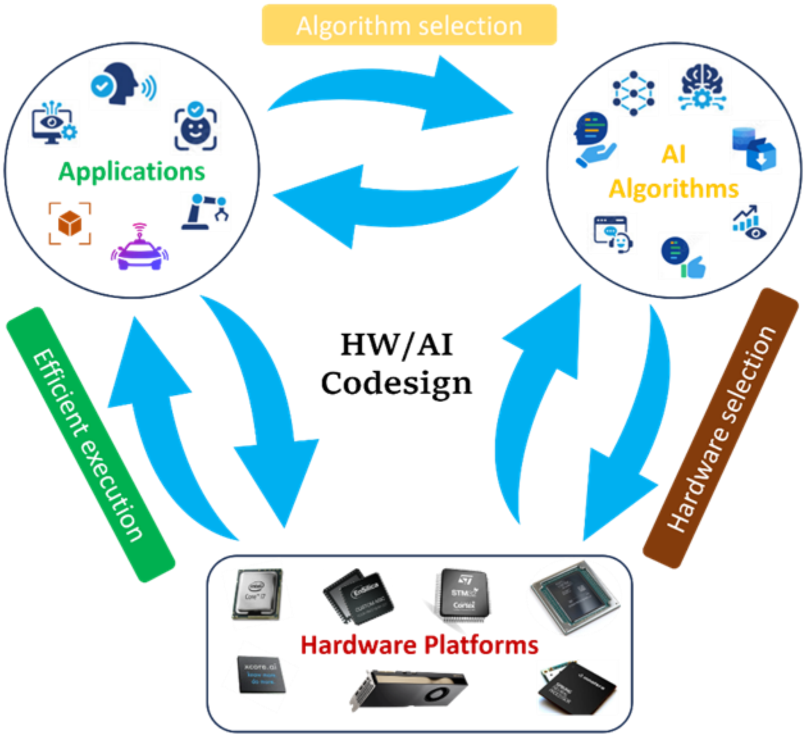


Figure 1: Design space exploration for Hardware-AI codesign

Project Overview

Number of Students	1
Project Type	Study project in collaboration with an external partner
Project Owner	Prof. Dr.-Ing. habil. Thorsten Jungeblut, Dr.-Ing. Qazi Arbab Ahmed
Project Context	**CareTech OWL** transfer project in collaboration with external partner, **Steinel GmbH**, as part of the CareTech OWL research project. Parallel employment as a student research assistant (WHK) is possible.

Abstract

The artificial intelligence (AI) paradigm has replaced a handful of traditional computer vision techniques to intelligently predict the correct output of a complex system under tight constraints and varying conditions. However, many AI algorithms require huge amounts of hardware resources due to their massive computational requirements. The goal of this project is the development of a framework for the efficient implementation of AI and ML algorithms on highly resource-constrained hardware architectures. Based on the framework, automated methods for exploring the design space of suitable combinations of AI algorithms and hardware in terms of HW/AI-co-design will be explored. In particular, industrial applications with high latency requirements are targeted. Not only the entire

chain as a linear process from model training to inference will be taken into account, but also the impact of the choice of possible hardware configurations on the original model development.

Short description

For the efficient execution of AI algorithms, techniques such as federated learning (FL) and cognitive edge computing (CEC) have already shifted the load of training and inference of neural networks from the cloud to the edge. The partitioning of the execution of the application has a significant impact on the performance and resource efficiency of the overall system. For example, a distinction can be made here between fundamentally different approaches of decentralized feature extraction as close as possible to the sensor with subsequent fusion versus central processing in the cloud. The first approach requires powerful edge hardware for preprocessing and potentially offers advantages for high real-time requirements. The second approach places higher demands on the communication infrastructure but potentially enables the execution of more complex networks. The main goal of this project is therefore to explore automated methods for exploring the design space of suitable combinations of AI methods and resource-constrained hardware in terms of HW/AI codesign.

In the area of efficient execution of AI procedures on embedded systems (cognitive edge computing), great progress has been made in the past. At all levels of the different processing concepts (cloud, fog, edge, very edge) in the production chain, a variety of potential hardware architectures and AI accelerators can be found, which differ in the available system resources (e.g., performance or power consumption). Examples of relevant hardware architectures are embedded microcontrollers with integrated AI acceleration, embedded GPUs/FPGAs, dedicated AI hardware accelerators, or high-end GPUs/FPGAs from the HPC domain.

We consider the complete processing chain, starting with the pre-processing close to the sensor (very edge), by connecting the groups of sensors and their information in the edge, via local decentralized cloud instances (fog), to the central station of all relevant information (cloud). At each level in the processing chain, individual approaches exist to locally optimize the resource efficiency of AI processes, for example, by reducing numerical precision (e.g., from 32 to 16, or 8 bits) to enable more efficient execution on specialized hardware or to minimize local storage space. The goal is to determine an optimal combination of hardware and software (i.e., AI algorithm) as part of a holistic design space exploration at all processing levels.

Task definition

In this project, students will develop an automated toolkit for design space exploration that can support the user in HW/AI codesign. In order to support the sustainable use of the developed methods and frameworks by the industrial partner, universally applicable models and automated design tools are to be made available in this project in the form of a development toolkit, which abstracts from the concrete use cases and make it possible to obtain recommendations for an optimal combination of AI methods and suitable (edge) hardware architecture based on a future application scenario. The goal is to provide the user with preliminary models and automated design tools that enable recommendations for a combination of AI methods and suitable (edge) hardware architecture based on a given application scenario and evaluation measures.

Figure 1 illustrates the general concept of design space exploration of hardware-AI codesign. The first main challenge is the compression/optimization of the neural network to reduce the size, yet with acceptable accuracy, using AI approximation techniques. The next challenge is to develop resource-efficient techniques to map the compressed (AI) model to a suitable hardware platform, such as GPUs, CPUs, ASICs, and FPGAs, considering the resource utilization in terms of area, power/energy, and latency/throughput.

Reference to the topic of data science

The evaluation and application of AI/ML methods in the field of "machine vision", e.g., the use of CNNs in object classification, are a core topic of data science and are dealt with, for example, in the modules "Data Mining & Machine Learning" and "Artificial Intelligence". AI-supported image processing places high demands on the organization and processing of data at all levels of IoT processing concepts (edge/fog/cloud). This is the core of the module "Big Data Architectures". The consideration of the entire system process from the imaging sensor to the cloud requires a holistic view of the complete data science process, which is covered in the module "Data Science".

Available Resources

Ensuring the availability of data, computing resources, hardware, application experts

- Information required to create the scenario (system description, interfaces, documentation, relevant key figures, etc.) is provided by the resource person
- Steinel GmbH provides extensive test data sets from real production environments
- The Steinel GmbH contact person will be available for the duration of the project
- The components required for the prototype development as well as other required materials will be provided by the Steinel GmbH
- Hardware for the more complex machine learning is available via the Data Science Lab, the CfADS, and the AI computing cluster, yourAI, at Bielefeld University of Applied Sciences

Project Plan

First semester: Preparation of a research exposé as an examination. Familiarization with the concept of AI algorithms (CNN, DNN), in particular, object detection and image recognition, hardware platforms, and design flow tools used.

Second semester: Creation of the system concept for design space exploration of cognitive edge computing architectures. Research on relevant work in the field of the use of AI/ML methods for data processing and model compression in the above-mentioned context. Preparation of a paper that gives an overview of the respective research area, as an examination.

Third semester: Development of a first demonstrator and proof-of-concept for hardware acceleration of relevant application. Comparison of an AI process with a classic implementation.

Fourth semester: Master's thesis and colloquium. Implementation and comparison of different combinations of AI/ML methods and hardware accelerators. Systematic evaluation and exploration of the efficiency of the combinations. Comparison of different processing concepts (embedded AI, edge, cloud). Final evaluation by comparing the implemented strategies. Preparation of a paper with the first quantitative results as an examination.

Necessary Competencies

Mandatory:

- Expertise in Python, Pytorch
- Good knowledge of C++

Optional:

- Experience with hardware design flow tools
- Programming of microcontrollers/FPGAs
- Basic knowledge of HDL (Verilog, VHDL)

- Experience with IoT devices
- Experience with the version control system "git"

Acquirable Competencies

- Resource-efficient information processing at the various levels (edge, fog, cloud) in line with the IoT processing concept
- Sensor-related information processing
- AI/ML methods
- Leveraging embedded hardware to accelerate AI/ML processes

Research Project for the Research Master Data Science

Project title: 8 Cognitive Edge Computing for AI/ML-based Surface Inspection

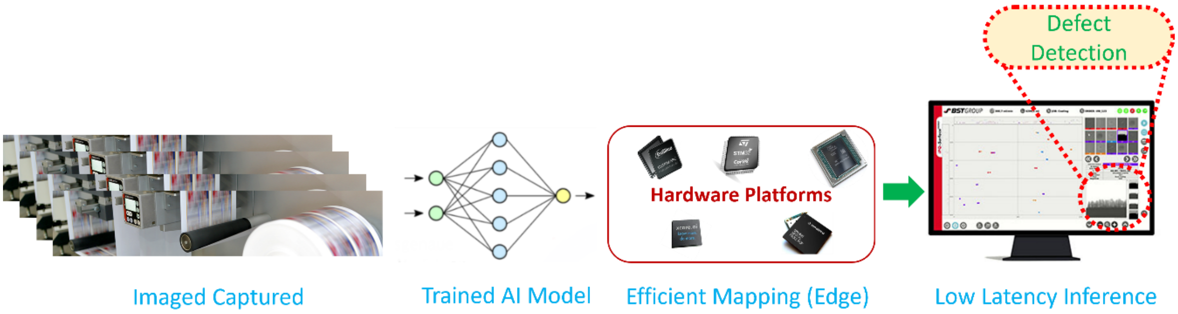


Figure 1: Efficient surface inspection using AI on edge

Project Overview

Number of Students	1
Project Type	Study project in collaboration with an external partner
Project Owner	Prof. Dr.-Ing. habil. Thorsten Jungeblut Dr.-Ing. Qazi Arbab Ahmed
Project Context	**CareTech OWL** transfer project in collaboration with an external partner, **BST Group**, as part of the CareTech OWL research project. The partner company **BST Group** provides extensive test data sets as well as a prototype hardware setup. Parallel employment as a student research assistant (WHK) is possible. The BST Group is very interested in the recruitment and long-term employment of junior staff.

Abstract

The artificial intelligence (AI) paradigm has replaced a handful of traditional computer vision techniques to intelligently predict the correct output of a complex system under tight constraints and varying conditions. The aim of the project is the design space exploration of AI/ML hardware accelerators in surface inspection. The focus is not only on learning the models on HPC systems but also on efficient execution (inference) on embedded hardware. The result of the design space exploration is the partitioning of the application, i.e. which AI methods can be evaluated on the sensor, which procedures can be accelerated via edge hardware (e.g. embedded GPU/FPGA), and which require powerful HPC hardware in the cloud.

Short description

Developments in the field of intelligent technical systems (ITS) are currently leading to a radical change in the entire value chain of industrial production. The increasing performance of information processing offers many new possibilities for the application area of industrial image processing, where humans and machines have so far reached their limits. In particular, the use of machine learning processes and artificial intelligence methods promises previously unknown possibilities, for example in object classification or visual quality control. Since around 2010, significant progress has been made with deep neural networks (DNN) and convolutional neural networks (CNN).

However, the performance of established AI/ML methods has so far mostly been based on the use of powerful decentralized computing resources (high-performance computing) in the cloud. The user is dependent on these powerful resources not only for learning the models but also for their execution (inference). However, the requirements arising in the field of industrial image processing deviate from the issues addressed by large providers of AI expertise due to high demands on low latency, real-time capability, or data locality. Issues such as maintainability, certifiability, or privacy also make it difficult to use popular models such as DNNs or CNNs in the cloud.

Nevertheless, great progress has also been made in the past in the area of efficient execution of AI/ML processes on embedded systems (**cognitive edge computing**). Suitable hardware accelerators can be found at all levels of the different processing concepts in networked production (edge/fog/cloud computing), which aim to achieve a suitable compromise between system resources such as required performance (e.g. classification accuracy), power consumption/energy requirements or data throughput/latency. Examples of relevant hardware architectures are embedded microcontrollers, embedded GPUs, embedded FPGAs, or dedicated AI hardware accelerators.

The BST Group develops systems for surface inspection (e.g., for battery cell production) that can reliably and immediately detect typical defects in the manufacturing process (e.g., coating discontinuities). BST systems can be perfectly adapted to a wide range of applications, due to their modular design. Image capture and defect detection can take place in real-time. The systems can be used on uniform, textured, and printed surfaces. The immediate and automatic detection and display of even the smallest defects and deviations enables the process to be adapted quickly and reliably to avoid rejects. The system consists of several optical sensors whose sensor data is suitably fused. The use of machine learning methods promises higher performance in defect detection and optimization of the entire process. However, with web speeds of several hundred meters per second, very high demands are required on the throughput and latency of the AI/ML processes used. Figure 1 shows the overview of the proposed methodology.

Task definition

The aim of the project is the design space exploration of AI/ML hardware accelerators for use in surface inspection. The focus is not only on learning the models of HPC systems but also on efficient execution (inference). Consideration of the entire system architecture, from intelligent sensors to edge gateways for local data pre-processing through to the cloud infrastructure, places high demands on the mapping of the AI network model to the hardware. The choice of target architecture, in turn, influences the choice and training of models. This results in an iterative cycle (model-to-inference-to-model), which includes both the selection of a suitable AI algorithm and the determination of the (hyper)parameters of the model. This approach of holistic design space exploration can also be referred to as HW/KI co-design, based on the established term HW/SW co-design. The design space exploration should consider numerous design goals, such as classification accuracy, latency, or resource requirements of the hardware. The result of the design space exploration is the partitioning of the application, i.e. which AI algorithm can be evaluated directly on the sensor (e.g. through dimension reduction or feature extraction), which algorithm can be accelerated via edge hardware (e.g. embedded GPU/FPGA or dedicated TPU accelerators) and which require powerful HPC hardware in the cloud.

Reference to the topic of data science

The evaluation and application of AI/ML methods in the field of "machine vision", e.g., the use of CNNs in object classification, are a core topic of data science and are dealt with, for example, in the modules "Data Mining & Machine Learning" and "Artificial Intelligence". AI-supported image processing places high demands on the organization and processing of data at all levels of IoT processing concepts (edge/fog/cloud). This is the core of the module "Big Data Architectures". The consideration of the entire system process from the imaging sensor to the cloud requires a holistic view of the complete data science process, which is covered in the module "Data Science".

Available Resources

Ensuring the availability of data, computing resources, hardware, application experts

- Information required to create the scenario (system description, interfaces, documentation, relevant key figures, etc.) is provided by the BST Group
- The BST Group provides extensive test data sets from real production environments
- The BST Group contact person will be available for the duration of the project
- The components required for the prototype development as well as other required materials will be provided by the BST Group
- Hardware for the more complex machine learning is available via the Data Science Lab, the CfADS, and the AI computing cluster, yourAI, at Bielefeld University of Applied Sciences

Project Plan

First semester: Preparation of a research exposé as an examination. Familiarization with the concept of BST's surface inspection systems, and the interfaces of the intelligent sensors provided.

Second semester: Creation of the system concept for design space exploration of cognitive edge computing architectures. Research on relevant work in the field of the use of AI/ML algorithms for sensor data processing in the above-mentioned context. Preparation of a paper that gives an overview of the respective field of research as an examination achievement.

Third semester: Development of a first demonstrator and proof-of-concept for hardware acceleration of an application for surface inspection of BST. Comparison of an AI-based inspection with a classic implementation.

Fourth semester: Master's thesis and colloquium. Implementation and comparison of different combinations of AI/ML methods and hardware accelerators. Systematic evaluation and exploration of the efficiency of the combinations. Comparison of different processing concepts (embedded AI, edge, cloud). Final evaluation by comparing the implemented strategies. Preparation of a paper with the first quantitative results as an examination achievement.

Necessary Competencies

Mandatory:

- Expertise in Python, Pytorch
- Good knowledge of C++

Optional:

- Basic knowledge of HDL (Verilog, VHDL)
- Experience with hardware design flow tools

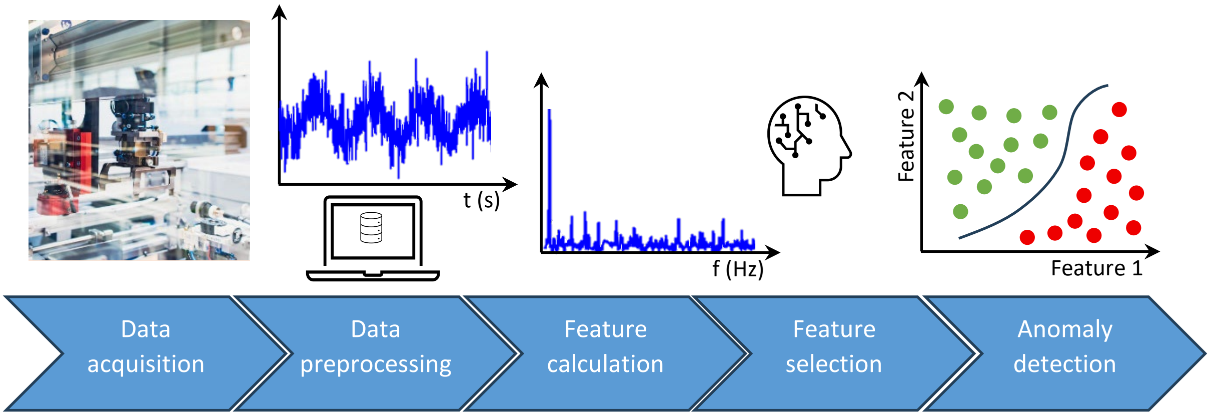
- Programming of microcontrollers/FPGAs
- Experience with IoT devices

Acquirable Competencies

- Resource-efficient information processing at edge (embedded microcontrollers, FPGAs) in line with the IoT processing concept
- Sensor-related information processing
- AI/ML-based object detection methods
- Leveraging embedded hardware to accelerate AI/ML processes

Research Project for the Research Master Data Science

Project title: 9 **Systematic Data and Feature Engineering for AI-Based Condition Monitoring in Industrial Automation**



Project overview

Number of Students	1
Project Type	Master student project
Project Owner	Dr.-Ing. Annika Junker Prof. Dr.-Ing Wolfram Schenck Dr.-Ing. Heiko Stichweh (Lenze)
Project Context	Project within the Center for Applied Data Science Gütersloh (CfADS) with the external partner Lenze in Groß Berkel

Abstract

This Master student project focuses on developing AI-driven methods for machine condition monitoring using data from frequency inverters, eliminating the need for additional sensors. Conducted in close collaboration with Lenze and embedded within the Center for Applied Data Science Gütersloh (CfADS), the project aims to develop innovative, physics-informed data processing and AI techniques for predictive maintenance.

Short description

The increasing digitalization in industrial environments and the rise of Industry 4.0 have created new opportunities for automated machine condition monitoring. Instead of adding costly sensors, modern drive systems already provide rich operational data. Lenze's advanced frequency inverters deliver high-quality signals that can be leveraged for sensorless monitoring, reducing hardware costs and simplifying integration. Combined with AI-based anomaly detection, this approach can significantly improve machine availability, reduce downtime, and optimize maintenance strategies. This research project aims to develop a physics-informed AI system for anomaly detection in storage and retrieval machines, using only inverter data. Key objectives include:

- Extracting time- and frequency-domain features with drivetrain knowledge.
- Preprocessing and segmenting data for comparable motion patterns.
- Building robust ML models for reliable anomaly detection.
- Development of robust machine learning models for reliable anomaly detection based on representative training datasets.

The expected outcome is a scalable solution for predictive maintenance in industrial automation.

Task definition

The objective of this research project is to design, implement, and critically evaluate a robust data and feature processing pipeline for AI-supported condition monitoring systems. The methods are designed to work with various machines and applications. The following work packages are planned:

Data Preprocessing and Feature Engineering: The student will systematically prepare operational data collected from frequency inverters, ensuring reliable performance across diverse operational motions, such as acceleration, constant speed, and deceleration. Preprocessing steps should guarantee that the resulting dataset allows for meaningful comparison and supports transfer to similar systems.

Feature Selection and Construction: Both data-driven features (e.g., statistical measures, time-series characteristics, and patterns learned directly from data) and physics-informed features (e.g., metrics derived from drivetrain models, domain knowledge, or physical laws) should be considered. Domain knowledge is to be explicitly integrated in the construction of physics-motivated features, leveraging expertise in drive technology and system dynamics.

Model Development and Robustness: Develop and train machine learning models for anomaly detection using the engineered feature sets. Robustness must be validated by applying the trained models to data from an additional, comparable machine. The student should ensure that both feature processing and model performance are transferable, not narrowly tailored to a single device or set of operating conditions. To avoid overfitting and improve model generalization, feature reduction techniques such as principal component analysis should be implemented, efficiently selecting the most relevant features for anomaly detection.

Comparative Analysis: The student will conduct a critical comparison between data-driven and physics-informed features, discussing their respective strengths and limitations. This includes:

- Assessing performance and robustness of models built on each approach
- Evaluating transferability to new machines or operational scenarios
- Considering factors such as interpretability, computational requirements, and resilience to changing system conditions

Documentation and Critical Reflection: The methodology, selection of algorithms, and all major decisions must be thoroughly documented. The student is expected to provide a reasoned discussion of the results, limitations, and practical implications, particularly with respect to the transferability and reliability of the developed system.

Reference to the topic of data science

This project addresses key aspects of data science, including data processing, particularly the preprocessing of raw data, and the targeted computation and selection of relevant features for subsequent AI-based condition monitoring. In addition, approaches from physics-informed machine learning will be explored, as the integration of physical domain knowledge into the machine learning pipeline represents an exciting and forward-looking research field.

Available Resources

- **Data access:** Availability of realistic operational data from modern frequency inverters across different machine types and applications. This data forms the basis for developing and training anomaly detection methods.
- **Expert support**: Guidance from specialists in automation and drive technology and data science available to advise on data interpretation, modeling, and implementation questions.

Project plan

First Semester:

- Conduct a literature review on condition monitoring and physics-informed machine learning.
- Familiarize with the operation of automation systems using frequency inverters.
- Perform data exploration and preprocessing for an exemplary system.
- Write a research exposé summarizing the findings.

Second Semester:

- Implement the calculation and selection of relevant features from both data-driven and physically motivated domains for the exemplary system of the first semester.
- Develop and evaluate a machine learning model to identify abnormal machine states.
- Write a paper of a literature review as the mandatory artefact.

Third Semester:

- Test the transferability of the developed model to at least two other systems: assess to what extent the model can be applied to comparable, yet previously unseen installations.
- Identify features critical for model quality reducing complexity and improving generalizability.
- Write a scientific paper with first results as a mandatory artifact.
- Define specific research questions to be addressed within the scope of the Master thesis.

Fourth Semester:

- Master thesis and colloquium.

Necessary Competencies

Mandatory:

- Bachelor’s degree in an engineering-related field.
- Interest in both data-driven and physically motivated modeling.
- Experience with the implementation of machine learning methods with Python.
- Strong passion for exploring complex technical challenges in depth.

Acquirable Competencies

- Application of machine learning methods to real industrially relevant problems.
- Application of physics-informed machine learning, which is a particularly exciting and forward-looking research field.

Research Project for the Research Master Data Science

Project title: 10 AI-Based Network Optimization and Strategic Internationalization for CTL

Project overview

Number of Students	1
Project Type	Project with external partner (Softward GmbH, 100% subsidiary of CTL AG)
Project Owner	Prof. Dr. Christian Schwede (Softward GmbH and CTL AG, Contact: Malte Konersmann, Softward GmbH, Tobias Jüngling, CTL AG)
Project Context	This project is conducted in close collaboration with Softward GmbH, a wholly-owned subsidiary of Cargo Trans Logistik (CTL) AG and the group's digital innovation specialist. Softward is responsible for the ongoing development of the CTL network's digital infrastructure and service optimization. The project aims to leverage AI-based methods to identify untapped optimization potential in CTL's network management and develop actionable strategies for international expansion. Students Assistant Job is included.

Abstract

CTL AG operates a complex hub-and-spoke groupage network with six main hubs across Germany. Transport routes and planning are currently based on rules derived from postal codes and static tables. With internationalization moving to the forefront of CTL's strategy, Softward GmbH seeks to harness

AI-driven optimization techniques to flexibly plan routes and depot utilization, simulate various expansion scenarios, and support strategic decision-making. The objective is to create a dynamic planning system that increases efficiency, scalability, and service quality—providing measurable value as the CTL network enters international markets.

Short description

Currently, CTL's route planning relies on a deterministic system based on postal code zoning and fixed product/service definitions (e.g., Express, Fixed Date, Classic). With growing cross-border business, an adaptive AI-driven routing framework is vital. This project, led by Softward GmbH, investigates the use of machine learning, reinforcement learning, and advanced optimization algorithms to dynamically analyze, predict and optimize network operations. Key questions to be addressed:

- Which AI methods unlock the most impactful network optimizations (e.g., lead time reduction, better resource allocation, increased depot utilization) in the CTL setting?
- How do various internationalization strategies (new international hubs/multihubs/national hubs as gateway) affect network performance, and how can AI-driven simulation steer this process?
- How can a dynamic planning platform be designed to integrate with CTL/Softward systems and scale with international growth requirements?

Task definition

- Analyze current routing and network planning (based on postal codes and fixed service rules)
- Curate and model relevant network datasets (shipment flows, depot capacities, service times, cross-border factors) e.g. in a digital twin.
- Design and implement AI-based optimization algorithms for flexible, scenario-driven network planning
- Simulate and benchmark several international expansion scenarios (adding new hubs, cross-border routing, dynamic time slot optimization)
- Evaluate efficiency gains versus the rule-based baseline; quantify benefits for internationalization
- Develop a prototype or proof-of-concept; propose roadmap for integration into CTL's operational landscape

Reference to the topic of data science

- Data exploration & preprocessing
- Graph/network modeling, combinatorial & stochastic optimization
- Predictive modeling, simulation (e.g., reinforcement learning, scenario analysis)
- Integration of AI models into operational logistics IT (APIs, microservices)

Available resources

- Access to comprehensive CTL routing, shipment, and depot data via Softward

- Mentorship and subject matter support from Softward and CTL AG (operational, technical, strategic perspectives)
- Modern data analysis and AI/ML toolkits, simulation capabilities
- API and system access for realistic prototyping
- Opportunities for workshops with logistics and IT experts at Softward and CTL

Project plan

- **First semester:** Network data exploration, evaluation of current system, identification of key gaps and AI opportunities. Research Expose.
- **Second semester:** Literature review, AI method benchmarking, initial model and simulation development, building prototype. Paper on literature review results.
- **Third semester:** Validation of optimization techniques and prototype, comparison of different internationalization strategies, interim results and recommendations. Paper on quantitive results.
- **Fourth semester:** Prototype refinement, documentation, integration planning, delivery and defense of Master thesis

Necessary competencies

Mandatory: Solid knowledge of data science, statistics, machine learning, React, MySQL, MSSQL, Azure, DOTnet

Optional: Experience in network/graph modeling, optimization algorithms, logistics or transport sector, software/API integration, Python

Acquirable competencies

- Applying and evaluating modern AI/optimization in a real logistics context
- Building and validating scalable simulation and planning models
- Cross-disciplinary collaboration (data science, logistics, digital transformation)
- Business communication and presentation of data-driven strategic impact

Research Project for the Research Master Data Science

Project title: 11 Multi-Agent Control of the IoT-Factory: Large Language Models and Reinforcement Learning on a Real-World Use

Project overview

Number of Students	1
Project Type	Student project
Project Owner	Prof. Dr. rer. oec. Pascal Reusch
Project Context	The project takes place in the IoT-Factory of the CfADS (https://www.hsbi.de/ium/cfads/projekte/iot-factory) and is supervised by laboratory engineer Roman Sliwinski and further staff of the CfADS. The Data Analytics Cluster of the CfADS is also available.

Abstract

For the IoT-Factory at CfADS, an alternative to the current central MES control is to be developed in the form of a multi-agent system. To this end, OptiFlow, an in-house Python-based control platform that already handles AGV fleet coordination and material flow control in the IoT-Factory, is to be evolved into an agentic system in which each physical component is modelled as an autonomous agent that cooperatively negotiates optimal order execution. Depending on the chosen focus, Large Language Model (LLM) based agent communication, learning agents using reinforcement learning, or agent economy approaches are to be investigated and evaluated on the real production line.

Short description

With the IoT-Factory, the Center for Applied Data Science (CfADS) at Bielefeld University of Applied Sciences and Arts offers a complete modular production line for the hands-on investigation of current research questions. Through a wide range of sensors within the individual modules and the data processing capabilities of the CfADS Data Analytics Cluster, the IoT-Factory opens up a wealth of new methods for addressing current research challenges. The current control of the plant is based on a central MES (Festo MES4.0) in conjunction with OptiFlow, a Python-based control layer developed at CfADS that consolidates material flow, AGV fleet logic, station coordination, and plant-wide procedures into a unified platform. Both systems operate deterministically: order prioritization, routing, and resource allocation follow static and inefficient rules and cannot autonomously react to changing conditions. The aim of the project is to evolve OptiFlow from a deterministic control platform into an agentic system. The currently static logic for material flow, AGV coordination, and station control is to be gradually replaced by cooperative agents that make their decisions autonomously and in a data-driven manner. The physical components of the IoT-Factory, robot cells, transport modules, storage, workpiece carriers, and the workpieces themselves, are to be represented as autonomous, communicative agents within OptiFlow and cooperatively negotiate optimal order execution.

Task definition

This research project pursues a fundamentally new approach: instead of central MES control, a multi-agent system (e.g. using OFacT) is to be applied in which each physical component of the IoT-Factory acts as an autonomous, communicative agent. These agents are to cooperatively negotiate how incoming production orders can be optimally processed in the IoT-Factory. Various research focuses can be defined within the project. One possibility is the use of LLMs in agent communication, where both a variant with natural-language LLM-to-LLM communication and a variant with structured communication that is "translated" by an LLM for human observers can be implemented. A second focus addresses reinforcement learning agents with an emphasis on optimizing production sequences, predictive resource planning, and AI-controlled AGV routing strategies. Further topics such as agent economy can also be agreed upon.

Reference to the topic of data science

The project addresses several core areas of the Research Master Data Science: development of autonomous software agents based on artificial intelligence (a study objective of the Master program), application of LLMs for reasoning and decision-making, data engineering in building the communication infrastructure and integrating heterogeneous data sources (MES, PLC, sensors), as well as the application of various AI approaches (e.g. LLMs or reinforcement learning). In addition, the project touches on topics of distributed systems, real-time communication, and human-machine interaction.

Available Resources

Ensuring the availability of data, computing resources, hardware, application experts

- IoT-Factory (23-module research and production plant with conveyor and AGV-based transport, assembly/disassembly robots, picking station, and assisted manual workstations for manual and combined manufacturing)
- MES access (full access to the Festo MES4 system including order management, workplan logic, and error handling)
- Open data basis (databases, machine sensor data, and parts of the software source code are openly available for research purposes)
- Computing resources (Data Analytics Cluster of the CfADS)
- Supervision by laboratory engineer M.Sc. Roman Sliwinski, laboratory director Prof. Pascal Reusch, and further CfADS staff

Project plan

First Semester: Familiarization with multi-agent systems, LLM-based agent frameworks, and the architecture of OptiFlow. Analysis of the existing MQTT communication infrastructure and conception of an agent-capable extension. Creation of an initial proof-of-concept with two communicating agents on the real plant. Writing a research exposé is the mandatory artefact.

Second Semester: Implementation of the multi-agent system as a prototype connected to the real plant modules via OptiFlow and MQTT. Depending on the focus: implementation and comparison of different communication variants (LLM-based vs. structured), development of initial RL policies, or design of economic allocation mechanisms. First production run on the IoT-Factory. Writing a paper providing an overview of the respective research field is the mandatory artefact.

Third Semester: Systematic evaluation of the agentic system across various scenarios (single orders, parallel orders, resource conflicts, disruption scenarios). Comparison with the central MES control based on defined KPIs such as throughput time, resource utilization, and decision quality. Writing a paper with first quantitative results is the mandatory artefact.

Fourth Semester: Master thesis and colloquium

Necessary Competencies

Mandatory:

- Willingness to work hands-on with an active and complex production plant
- Advanced programming skills in Python
- Interest in artificial intelligence, particularly LLMs and agent systems

Optional:

- Experience with data science projects, digital twins, intralogistics, and production systems
- Knowledge of reinforcement learning or natural language processing
- Experience with LLM frameworks (e.g. LangChain, CrewAI) or RL libraries (e.g. Stable Baselines3, RLlib)
- Basic knowledge of software architecture (e.g. Django, modular Python backends)

Acquirable Competencies

- Practical experience in developing and integrating software agents on real manufacturing infrastructure (PLC network, MES, MQTT, OPC UA, automated guided vehicles)
- Ability to comparatively evaluate different AI communication approaches
- Extensive experience in scientific work, including publication at international conferences
- Contribution to a real open-source software project (OptiFlow) with a Django backend and modular Python architecture
- Depending on the focus: in-depth knowledge of LLM engineering (prompting, RAG, agent design), deep reinforcement learning (policy optimization, reward design), or mechanism design (auctions, resource allocation)
- Broad skill set spanning data engineering, LLM engineering, IoT communication, and production logistics

Research Project for the Research Master Data Science

Project title: 12 Dialog-Based Materials Modeling and Simulation (Simulation Copilot)

Project overview

Number of Students	1 - 2
Project Type	Preliminary research as the basis for a project proposal
Project Owner	Prof. Dr. Christian Schröder
Project Context	Project in the **Computational Materials Science and Engineering Research Group** at the **Bielefeld Institute for Applied Materials Research (BIfAM)**. A position as a student assistant at HSBI is available.

Abstract

The goal is to develop a dialogue-based interface that uses a large language model (LLM) to specify material models in natural language, specifically request missing parameters, launch simulations via tool calls, and summarize results in a transparent manner. The study examines prompting, RAG, domain ontologies, and evaluation metrics through case studies. Outcome: a prototype plus quantitative evaluation.

Short description

Background: Material modeling and simulation are an essential part of modern product development. While product developers possess extensive domain knowledge, they often lack experience – or have only limited experience – in using the typically highly complex software tools. Each environment requires its own input language, parameters, convergence criteria, and post-processing. Many iterations arise from incomplete specifications, incorrect units, unsuitable boundary conditions, or unclear target values.

Motivation: LLMs enable natural language interaction, but without domain-specific validation, they can lead to hallucinations, incorrect units, or non-reproducible workflows. A "Holodeck-like" dialogue (for those familiar with Star Trek 😉) that systematically asks follow-up questions and correctly operates simulation tools could significantly improve accessibility, speed, and reproducibility. **Goal**: Research and prototyping of a dialogue-oriented “Simulation Copilot” that (i) accepts tasks in natural language, (ii) actively clarifies missing information, (iii) launches simulations via defined interfaces, (iv) summarizes/visualizes results, and (v) ensures reproducibility and traceability.

Task definition

You will develop and evaluate a prototype for a simulation tool (to be determined in consultation with your advisor) that meets the following minimum requirements:

1. **Domain model & dialog schema**
    - Definition of a structured schema for: material parameters, model type, geometry, boundary conditions, numerical methods, units, target variables, etc.
    - Design of a clarification policy: When must the system prompt for input, and when may it suggest default values?
2. **Tool integration**
    - Integration of at least one simulator (e.g., LAMMPS for MD or an FEM workflow via COMSOL interfaces).
    - Implementation of “Tool Calling”: Generating input files, starting runs, collecting logs/outputs.
3. **Knowledge Access & Validation**
    - Establishment of a domain-specific knowledge base (e.g., manual excerpts, material data sheets, best-practice guidelines) and RAG pipeline. RAG for source support (e.g., “Why this cutoff choice?” “Which boundary condition is appropriate?”).
4. **Evaluation**
    - Creation of a benchmark dialog for typical tasks (e.g., “Estimate elastic modulus → Configure simulation → Provide stress-strain curve”).
    - “Assumption Ledger”: List of all assumptions/defaults made during the dialogue.
    - Quantitative metrics: Success rate, number of dialogue turns, proportion of necessary corrections, simulation termination rate, time to result, reproducibility (run determinism/log quality)

Reference to the topic of data science

Subfields:

- Natural Language Processing / Conversational AI
- Information Retrieval (RAG), Knowledge Representation
- MLOps/LLMOps and Experiment Management
- Data-Driven Modeling & Surrogate Models (optional, but recommended)
- Evaluation/Statistics (quantitative studies, hypothesis testing)

Methods (selection, project-dependent):

- Prompt engineering, structured output (JSON schemas), guardrails
- Retrieval-augmented generation: embeddings, vector database, chunking/ranking
- Dialogue policy design: rule-based + data-driven (e.g., bandit/active demand heuristics)
- Validation: Constraints, unit checks, consistency rules, “self-check” with justification required

- Experiment design & evaluation: A/B tests (with/without RAG, with/without follow-up questions), confidence intervals, effect sizes
- Optional: Bayesian optimization/active learning for automated parameter search in the simulation space; surrogates (Gaussian processes/NN) for acceleration.

Available resources

A fully equipped workspace, access to computing resources (HPC cluster with over 5,000 CPU cores and high-performance GPUs), simulation software, domain and application experts, and a great team of people 😉

Project plan

1st Semester – Research Proposal

- State of the art: LLMs in engineering/simulation, tool calling, RAG, validation
- Precise research questions and hypotheses (e.g., "Demand strategy reduces dropout rate by x%")
- Methodology, data sources, planned metrics, risk analysis, ethics/compliance (licensing of tools/data)
- Selection of use cases in collaboration with the advisor, definition of metrics and baselines

2nd Semester – Survey/Overview Paper

- Systematic literature review: LLMs in engineering, tool calling, validation, reproducibility-by-design
- Taxonomy: input specification, demand strategies, knowledge sources, dialogue systems for engineering, simulator integration, verification/validation, reproducibility
- Identification of research gaps and derivation of own research questions and a reference workflow for "Dialogue → Setup → Run → Analysis"

3rd Semester – Paper with initial quantitative results

- Implemented prototype including simulator integration
- Benchmark evaluation and variant comparison (Baseline vs. +RAG vs. +Constraints vs. +Policy)
- Comparison of multiple variants (e.g., Baseline Prompting vs. RAG vs. RAG+Constraints)
- Discussion of error cases, limitations, generalizability

4th Semester – Master's Thesis & Colloquium

- Expansion into a robust research platform (logging, reproducibility)
- Extended evaluation (more tasks, possibly a user study with material/simulation users)
- Conclusion: Master's thesis (methodology, results, guidelines) + colloquium (demo + scientific classification)

Necessary competencies

Mandatory: Strong general computer science skills, programming skills (Python, etc.), interest in interdisciplinary work, curiosity, enthusiasm

Optional: Experience with LLMs and RAG

Acquirable competencies

- **Scientific expertise**: hypothesis formulation, experimental design, quantitative evaluation, clear reporting
- **Data science expertise**: RAG pipelines, NLP/LLM workflows, metrics, statistical analysis, benchmarking
- **Software engineering**: tool integration, interfaces/orchestration, testability, logging, reproducibility, deployment fundamentals
- **Domain expertise in material simulation**: parameterization, boundary conditions, units, interpretation of simulation results
- **Quality & safety**: handling hallucinations, validation strategies, documentation of assumptions/uncertainties
- **Communication**: structured dialogue designs, comprehensible summaries of results, scientific writing

Research project for the Research Master Data Science

Project title: 13 AI production control of electroplating systems based on a digital twin

Project overview

Number of students	1
Project Type	Project with external partner
Project Owner	Prof. Dr. Christian Schwede
Project Context	Project in cooperation with UBB GmbH Employment as a student assistant is planned.

Abstract

UBB GmbH operates two electroplating plants for the coating with zinc of previously manufactured steel components. The aim of this project is to develop an AI-based agent system that enables automated resource planning and control of the electroplating department. The introduction of this system aims to improve resource utilization, shorten throughput times and increase adherence to deadlines. As part of the project, the processes, customer requirements and resources are first mapped in a digital twin. Based on this, an AI agent system is developed that controls the processes and the use of resources, taking into account prioritized customer requirements and technical constraints.

Short description

UBB GmbH, headquartered in Burgbernheim, is a wholly-owned subsidiary of MÜPRO International GmbH The MÜPRO Group comprises 14 international subsidiaries, employs around 650 people and has its headquarters in Wiesbaden. With its approximately 90 employees, UBB produces fastening systems in 3-shift operation in various production areas as products of the core product range of the MÜPRO Group.

In the electroplating production area, metal parts are provided with a defined zinc layer to ensure effective corrosion protection for the subsequent application of the finished products. The two electroplating systems – a drum and a rack system – are an integral part of the production chain and are roughly planned and controlled via the ERP system (SAP) used. The detailed planning and concrete integration of the production orders into the electroplating process is currently carried out manually by specialist personnel. In addition to the order prioritization, which may be adjusted, galvanotechnical parameters such as the filling level of the drums and racks and the component geometries are also taken into account. This manual approach leads to blurring of planning and scheduling, which has a negative impact on adherence to deadlines and capacity utilization.

Task definition

The task is to design and implement an innovative AI agent system that automates and optimizes the existing manual and transaction-based planning and control processes in the electroplating sector of UBB GmbH. The focus is on the complete digitization of all relevant processes, including the acquisition and modeling of process parameters, resources and customer requirements in the digital twin. The system to be developed should be able to independently make decisions on resource planning and control based on real-time data and prioritized customer requirements, and be able to react flexibly to changes in the operational environment. The aim is to significantly increase the efficiency and quality of the electroplating processes, to increase the transparency in the flow of goods and to relieve or take over operational control. In addition, the agent system is designed to enable continuous adjustment and improvement of processes to meet the increasing demands of international and volatile customer projects.

Reference to the topic of data science

Develop an AI agent (Reinforcement Learning, Optimization, Agent AI) and create data-based digital twins.

Available resources

Access to the company's data, IT systems and assets is ensured. A specialist contact person is also available within the company. Compute resources are provided by the HSBI with the YOUR-AI cluster.

Project plan

First semester: Preparation of a research exposé as an examination. Instruction in the tasks in the company, review of the data and definition of the optimization criteria.

Second semester: Literature research on the process of AI-based production control. Implementation of the digital twin. Preparation of a paper that gives an overview of the research area is examination performance.

Third semester: Implementation of a first procedure and evaluation of the results. Publication of a paper with the first results is examination performance.

Fourth semester: Master's thesis and colloquium.

Necessary competencies

Mandatory:

- Programming knowledge (e.g. Python)

Optional:

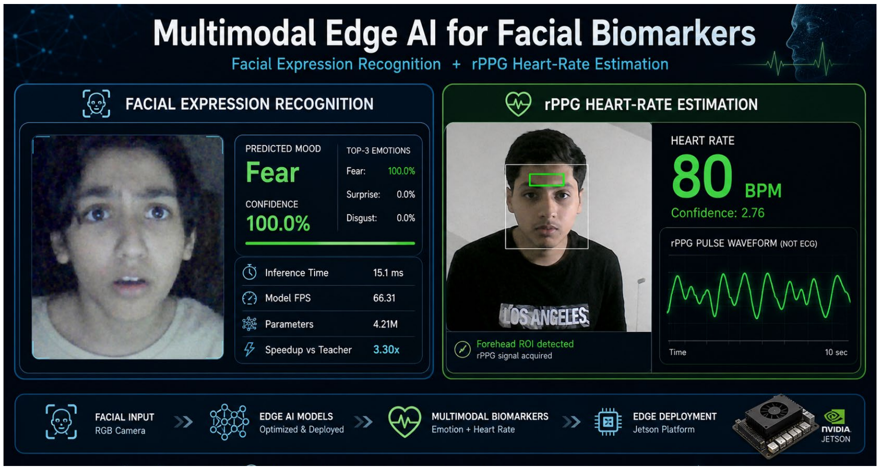
- Experience with machine learning
- Experience with optimization procedures
- Experience in production and logistics

Acquirable competencies

- AI process for production control
- Digital twin in production
- Publication of scientific publications at international conferences
- Competencies in working as a team

Research project for the Research Master Data Science

Project title: 14 Multimodal Edge-AI Biomarkers from Facial Expressions and rPPG for Affective and Cognitive State Screening



Project overview

Number of Students	1-2 students
Project Type	Preliminary research / applied AI prototype; possible extension with external clinical or care partners
Project Owner	Prof. Dr.-Ing. habil. Thorsten Jungeblut Dr. Hasina Attaullah
Project Context	The project is positioned at the intersection of computer vision, affective computing, digital health, physiological signal processing, and edge AI. It will use publicly available or request-based datasets related to facial expression analysis, mental-health screening, physiological video analysis, and cognitive-state research. The technical work can build on pretrained facial-expression and facial action-unit models, rPPG-based heart-rate estimation methods, and an NVIDIA Jetson edge device for real-time or near-real-time inference. If collaboration with a clinical, care-home, psychology, or gerontology partner is available, the project may include expert feedback, domain validation, and a small ethically approved pilot study. Without such a partner, the project will remain a public-dataset-based research and demonstrator project.

Abstract

This project develops an explainable edge-AI pipeline that analyses facial expressions and remote photoplethysmography (rPPG) from facial video to support non-contact screening of pain and affective/cog-nitive state changes [1-3]. It combines pretrained facial-action, expression, and video models with robust signal processing, low-label adaptation, and uncertainty estimation. A NVIDIA Jetson-based live demon-strator will show local inference while emphasizing privacy, robustness, and clinical safety limits.

Short description

The project investigates whether camera-based behavioral and physiological biomarkers can support early, non-contact screening in contexts where self-report is difficult or incomplete. The primary recommended focus is older adults and people with dementia, where verbal communication may be impaired and clinically relevant discomfort or distress can be under-recognized [6,7]. Secondary exploratory tasks may include depression severity estimation and dementia-related affective/cognitive change, framed strictly as research and decision support rather than automated diagnosis [4,5].

The technical idea is to combine facial expression recognition, facial action-unit dynamics, head and pose motion, and rPPG-based heart-rate estimation [1,3,4,6]. rPPG estimates pulse trends from tiny periodic skin-colour changes in facial video, typically from stable skin regions such as the cheeks or forehead [1,3]. It is useful for experiments and auxiliary physiological context, but it is not medical-grade and can fail under motion, poor lighting, dark or variable illumination, makeup, compression, unstable camera exposure, or occlusion [2].

Initial situation, motivation and objectives

- Clinical pain, depression, and cognitive decline are often assessed through self-report or observer-based scales, which can be subjective, intermittent, and difficult for patients with dementia or communication impairment.
- Facial expressions, facial action units, gaze, pose, and video-derived physiological cues can provide measurable behavioral markers, but models trained on controlled datasets often fail under real-world lighting, pose, age, cultural, and clinical-domain shifts.
- rPPG can add a non-contact heart-rate cue from facial video; however, it is sensitive to lighting and head movement and should be treated as an auxiliary experimental signal rather than a medical vital-sign monitor.
- Edge devices such as Jetson platforms make real-time, privacy-preserving local inference possible without streaming sensitive video to the cloud.
- The objective is to design, evaluate, and demonstrate a robust low-label adaptation pipeline for interpretable facial and rPPG biomarkers.

Task definition

- Conduct a systematic literature review on facial expression recognition, rPPG, and multimodal video analysis for pain, depression, and dementia-related affective/cognitive changes.
- Implement a video-processing pipeline: camera capture, face detection/tracking, landmark extraction, action-unit or expression features, rPPG signal extraction, temporal aggregation, and model inference.

- Compare pretrained models, classical machine-learning baselines, and temporal deep-learning approaches for selected tasks.
- Adapt centrality-guided maximum-diversity subset selection to select a small, diverse, representative subset of samples for annotation, calibration, or domain adaptation.
- Evaluate performance, robustness, calibration, fairness, and explainability across datasets, camera conditions, lighting changes, motion levels, and deployment settings.
- Deploy the trained inference pipeline on a Jetson device and build a live demonstrator that overlays facial landmarks, expression/action-unit cues, rPPG pulse waveform or BPM estimate, model confidence, and warning thresholds.

Concrete task of the students

- Select one primary target task focused on non-contact screening of clinically relevant facial and physiological state changes in older adults or people with impaired verbal communication. Depression-related or dementia-related affective/cognitive screening can be defined as optional secondary tasks.
- Prepare datasets and preprocessing scripts for public or request-based sources such as BioVid, UNBC-McMaster, DAIC-WOZ/E-DAIC, and suitable rPPG datasets where access is granted [8-10].
- Build a facial-video feature pipeline including face detection, region-of-interest tracking, facial expression estimation, facial action-unit analysis, head and pose motion features, and rPPG-based heart-rate signal processing.
- Train or fine-tune models using low-label subset selection and compare the results against random sampling, uncertainty sampling, k-center selection, diversity-only selection, and full-data baselines.
- Develop explainability outputs using action-unit importance, temporal saliency, rPPG signal-quality indicators, confidence intervals, and structured error analysis.
- Deploy inference on an NVIDIA Jetson edge device using PyTorch, ONNX, or TensorRT where appropriate, and measure latency, throughput, memory use, and energy profile.
- Document ethical and clinical limitations clearly: the system must not provide diagnosis and should be framed only as research, screening support, or caregiver/clinician decision support.

Reference to the topic of data science

The project covers core data science topics: data acquisition and preprocessing, representation learning, transfer learning, active learning/subset selection, time-series modelling, signal processing, multimodal fusion, model calibration, explainable AI, fairness and bias analysis, edge deployment, and reproducible evaluation. Methods may include CNN or ViT facial encoders, action-unit extraction, temporal CNN/LSTM/Transformer models, rPPG algorithms such as green-channel, CHROM or POS-style approaches, multimodal fusion, uncertainty estimation, and statistical validation.

Available Resources

- Hardware: NVIDIA Jetson Orin Nano, Orin NX, or comparable edge GPU device; USB/RGB camera or microscope/care-simulation camera setup for live demonstration.
- Software: Python, PyTorch, OpenCV, MediaPipe or OpenFace alternatives, NumPy/SciPy signal processing, ONNX Runtime/TensorRT, scikit-learn, experiment tracking tools, and containerized deployment.
- Data: public or request-based research datasets such as BioVid for pain, UNBC-McMaster for shoulder-pain expressions, DAIC-WOZ/E-DAIC for depression-related interviews, and rPPG datasets where available. Any clinical pilot data would require ethics approval and consent.

- Expertise: supervision in computer vision and data science; optional input from psychology, geriatrics, pain medicine, nursing, dementia care, or digital-health experts for labels and interpretation.

Project plan

First Semester - Research Expose:
Define use case and ethics boundary; complete literature review protocol; select datasets; implement basic camera capture, face/landmark/action-unit extraction, and first rPPG prototype; design evaluation plan.
Second Semester - Literature review paper:
Review FER, rPPG, multimodal pain/depression/dementia screening, and edge AI literature; reproduce baseline models; finalize dataset access and preprocessing; define low-label subset-selection experiments.
Third Semester - Paper with first results:
Implement CGMD-inspired subset selection for facial/video/rPPG data; train or fine-tune models; compare against baselines; run robustness tests for motion and illumination; create first Jetson demo. **Fourth Semester - Master thesis and colloquium:**
Complete robustness, fairness, calibration, explainability, edge-deployment, latency, and energy analyses; finalize demonstrator; write thesis and prepare colloquium presentation.

Necessary Competencies

Mandatory:

- Python programming and machine-learning basics.
- Computer vision fundamentals and experience with PyTorch or comparable deep-learning frameworks.
- Basic signal processing or willingness to learn filtering, spectral analysis, and signal-quality estimation for rPPG.
- Basic statistical evaluation, cross-validation, and reproducible experiment design.
- Willingness to work with sensitive health-related data under privacy and ethics constraints.

Optional:

- Experience with facial expression recognition, facial action units, time-series modelling, affective computing, or rPPG.
- Experience with edge AI, Jetson, ONNX, TensorRT, Docker, or Linux deployment.
- Background in psychology, medical informatics, gerontology, pain assessment, or dementia care.

Acquirable Competencies

- Building end-to-end video-based AI pipelines from raw camera input to explainable predictions.
- Remote physiological signal extraction from video, including rPPG preprocessing, filtering, and quality control.
- Transfer learning and low-label active/subset selection for domain adaptation.
- Clinical-AI evaluation, including robustness, calibration, bias, and interpretability.
- Privacy-aware edge deployment and real-time inference optimization.
- Scientific writing: expose, literature review, results paper, thesis, and colloquium presentation.

Ethical and clinical boundary

The project must be positioned as research and decision support, not as a diagnostic tool. Outputs should be phrased as risk scores, behavioral-cue summaries, physiological-signal-quality indicators, or prompts for human review. Any use with real patients requires ethics approval, informed consent, data minimization, privacy-preserving processing, and involvement of clinical experts. The system should log uncertainty and abstain when face visibility, lighting, motion, rPPG signal quality, or domain mismatch make predictions unreliable.

References

1. Debnath, U., & Kim, S. (2025). A comprehensive review of heart rate measurement using remote photoplethysmography and deep learning. *BioMedical Engineering OnLine*, *24*(1), 73.
2. Acharya, B., Saakyan, W., Hammer, B., & Drimalla, H. (2025). The reliability of remote photoplethysmography under low illumination and elevated heart rates. *npj Digital Medicine*.
3. Sakib, S., Hasan, Z., & Roy, N. (2025). A State-Of-The-Art Survey of Remote Photoplethysmography for Contactless Health Parameters Sensing. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *15*(3), e70039.
4. Li, S., & Deng, W. (2020). Deep facial expression recognition: A survey. *IEEE transactions on affective computing*, *13*(3), 1195-1215.
5. Takeshige-Amano, H., Oyama, G., Ogawa, M., Fusegi, K., Kambe, T., Shiina, K., ... & Hattori, N. (2024). Digital detection of Alzheimer's disease using smiles and conversations with a chatbot. *Scientific Reports*, *14*(1), 26309.
6. Cascella, M., Esposito, D., Muzio, M. R., Cascella, V., & Cerrone, V. (2025). Artificial intelligence for pain assessment via facial expression recognition (2015–2025): a systematic review. *Exploration of Medicine*, *6*, 1001370.
7. Smrke, U., Milošič, A., Mlakar, I., Kadiš, M., & Mulej Bratec, S. (2025). Pain Cues in People With Dementia: Scoping Review. JMIR Mental Health, 12, e75671.
8. Lucey, P., Cohn, J. F., Prkachin, K. M., Solomon, P. E., & Matthews, I. (2011, March). Painful data: The UNBC-McMaster shoulder pain expression archive database. In *2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG)* (pp. 57-64). IEEE.
9. Gratch, J., Artstein, R., Lucas, G. M., Stratou, G., Scherer, S., Nazarian, A., ... & Morency, L. P. (2014, May). The distress analysis interview corpus of human and computer interviews. In *Lrec* (Vol. 14, pp. 3123-3128).
10. Ringeval, F., Schuller, B., Valstar, M., Cummins, N., Cowie, R., Tavabi, L., ... & Pantic, M.(2019, October). AVEC 2019 workshop and challenge: state-of-mind, detecting depression with AI, and cross-cultural affect recognition. In *Proceedings of the 9th International on Audio/visual Emotion Challenge and Workshop* (pp. 3-12).