FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

Data Science Day 2025

14 May 2025

Friedrich Schiller University Jena

Book of Abstracts

Olia Blacher (Ed.)
olia.blacher@uni-jena.de

Contents

EIFER: Electromyography-Informed Facial Expression Reconstruction via Unpaired Image-to-Image Translation

Author: Tim Büchner[1]

Co-authors: Christoph Anders[2]; Orlando Guntinas-Lichius[2]; Joachim Denzler[1]

[1] *Computer Vision Group, Friedrich Schiller University Jena*

[2] *Jena University Hospital*

Corresponding Authors: joachim.denzler@uni-jena.de, christoph.anders@med.uni-jena.de, orlando.guntinas@med.uni-jena.de, tim.buechner@uni-jena.de

Analyzing facial expressions through muscle activity, captured via surface electromyography (sEMG), offers rich insights for psychology, medicine, and animation. However, sEMG electrodes introduce significant occlusion, hindering accurate facial expression analysis, particularly for monocular 3D face reconstruction. Our method, **EIFER: Electromyography-Informed Facial Expression Reconstruction via Unpaired Image-to-Image Translation**, addresses this challenge. EIFER decouples facial geometry and appearance by leveraging unpaired image-to-image translation within a CycleGAN-like adversarial framework [3, 6]. We combine this with 3D Morphable Models (3DMMs), specifically FLAME [5] and BFM [2], to represent facial shape and expression parameters. EIFER performs monocular 3D face reconstruction using a neural renderer and learns to remove sEMG electrode occlusions via adversarial training against unpaired, occlusion-free reference images. Critically, EIFER establishes a bidirectional mapping between 3DMM expression parameters and sEMG muscle activity, enabling both physiologically-based expression synthesis and electrode-free facial electromyography. This approach overcomes the limitations of occlusion-sensitive photometric methods and integrates physiological information into 3DMMs [1, 4]. We validate EIFER on a novel dataset of synchronized sEMG and facial expressions, demonstrating faithful geometry and appearance reconstruction, expression synthesis from muscle activity, and muscle activity prediction from facial expressions. EIFER introduces a new paradigm for facial electromyography with potential extensions to other multi-modal face recordings.

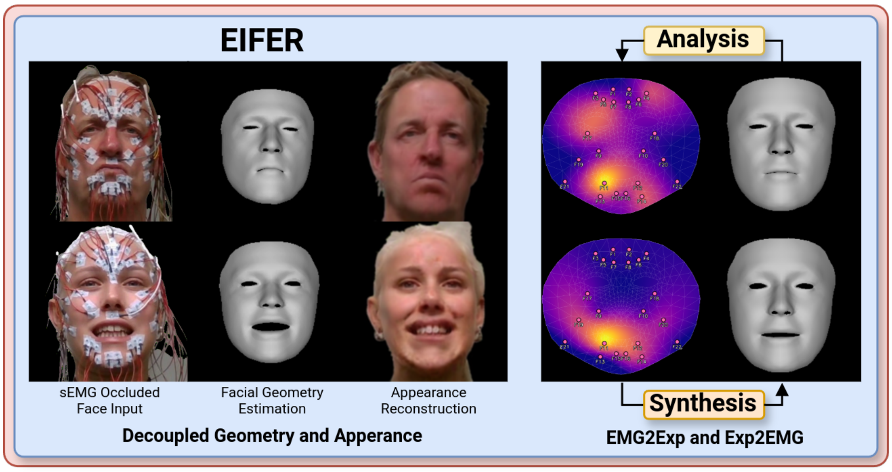


Figure 1: Bridging the gap between mimics and muscles: Our method EIFER utilizes neural unpaired image-to-image translation to decouple facial geometry and appearance for muscle-activity-based expression synthesis and electrode-free facial electromyography.

[1] Balconi, M., Bortolotti, A., & Gonzaga, L. (2011). Emotional face recognition, EMG response, and medial pre-frontal activity in empathic behaviour. Neuroscience Research, 71(3), 251-259.
[2] Blanz, V., & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. Proceedings of the 26th annual conference on Computer graphics and interactive techniques, 187-194.
[3] Büchner, T., Guntinas-Lichius, O., & Denzler, J. (2023). Improved obstructed facial feature reconstruction for emotion recognition with minimal change cyclegans. International Conference on Advanced Concepts for Intelligent Vision Systems, 262-274. Springer.
[4] Ekman, P., & Friesen, W. V. (1978). Facial Action Coding System: A technique for the measurement of facial movement. Consulting Psychologists Press.

[5] Li, T., Bolkart, T., Black, M. J., Li, H., & Romero, J. (2017). Learning a model of facial shape and expression from 4D scans. ACM Transactions on Graphics (TOG), 36(6), 1-17.
[6] Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of the IEEE international conference on computer vision, 2223-2232.

CausalRivers — Scaling up benchmarking of causal discovery for real-world time-series

Author: Gideon Henrik Stein

Co-authors: Maha Shadaydeh; Jan Blunk; Niklas Penzel; Joachim Denzler

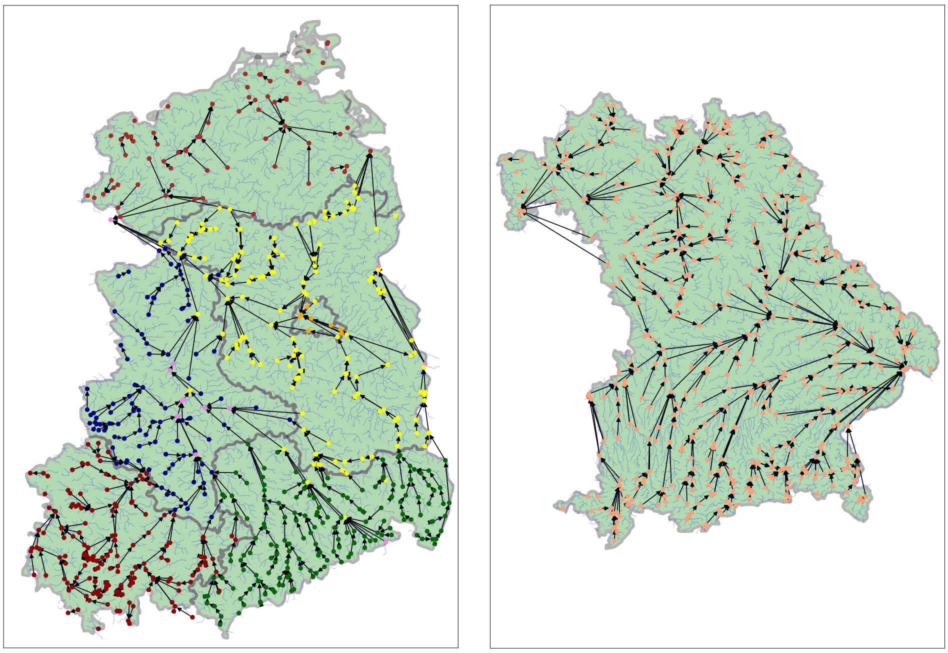
Computer Vision Group, Friedrich Schiller University Jena

Corresponding Authors: niklas.penzel@uni-jena.de, gideon.stein@uni-jena.de, jan.blunk@uni-jena.de, maha.shadaydeh@uni-jena.de, joachim.denzler@uni-jena.de

Causal discovery from observational data is a challenging problem, with many proposed methods lacking thorough real-world evaluation. Most studies rely on synthetic data or limited real-world examples under idealized assumptions, which do not accurately reflect the complexity of real-world systems.

To address this issue, we introduce CausalRivers, a comprehensive benchmarking kit for causal discovery in time series data. Our dataset consists of extensive river discharge measurements from 1,160 stations in eastern Germany and Bavaria, spanning 2019–2023 with 15-minute temporal resolution. We also include data from a recent flood event around the Elbe River, which exhibits a pronounced distributional shift. By leveraging multiple data sources and time-series metadata, we constructed two distinct causal ground truth graphs for Bavaria and eastern Germany, which can be sampled to generate thousands of subgraphs for benchmarking.

We demonstrate the utility of CausalRivers through multiple experiments and introduce effective baselines, highlighting areas for improvement in causal discovery methods. Our benchmarking kit has the potential to facilitate robust evaluations and comparisons of causal discovery approaches. We also anticipate its relevance to related areas such as time series forecasting and anomaly detection. With this, we hope to establish benchmark-driven method development in the field of causal discovery, as is the case for many other areas of machine learning.



(a) Eastern Germany (b) Bavaria

Figure 1: The causal ground truth graphs for river discharge measurement stations are provided with this benchmarking kit. Jointly, these two graphs hold over 1000 nodes.

Reinforcement Learning for Laser Alignment

Author: Matthew Schwab

Friedrich Schiller University Jena

Corresponding Author: matthew.schwab@uni-jena.de

Manual alignment of optical systems can be time consuming and the achieved performance of the system varies depending on the operator doing the alignment. A reinforcement learning approach using the PPO algorithm was used to train agents to align simple two-mirror optical setups, as well as a full regenerative laser amplifier. The goal is to produce agents that can reproducibly align the setup faster than a human and can correct long-term drifts in laser energy (time scale of approx. one hour) during operation. The work is still ongoing. Agents have been successfully implemented on hardware in the two-mirror setup, showing "super-human" performance in alignment time. The agents successfully "learn" to handle a significant amount of mechanical backlash in the used stepper motors and mirror mounts. Currently, the necessary hardware is being installed on a regenerative amplifier and agents are being further developed for this use case.

Data Literacy — More Than Just Data Science for Beginners

Authors: Volker Uwe Schwartze; Martin Kerntopf

Friedrich Schiller University Jena, Office for Digital Transformation, Project Data Literacy Jena (DaLiJe)

Corresponding Authors: martin.kerntopf@uni-jena.de, volker.schwartze@uni-jena.de

The increasing digitalization and datafication of society have made data competencies essential for all citizens, not just for specialists in computer science or data science. As data-driven decision-making becomes central to business, government, and everyday life, the ability to critically engage with data is emerging as a key qualification for the 21st century. Especially in light of recent advances in artificial intelligence (AI) and machine learning (ML), understanding how data is generated, collected, and used has become more important than ever. Thus, universities have to prepare students from all study programmes for critically interact with data, may it be in the context of their future research, jobs or in their private lives.
In this context, the concept of Data Literacy—commonly defined as the ability to read, work with, analyze, and communicate with data—has gained significant traction in education. However, it is often narrowly understood as a set of technical or methodological skills, such as statistics or programming. As a result, Data Literacy is frequently perceived as "Data Science for beginners." In this presentation, we challenge that view and argue that Data Literacy represents a broader, interdisciplinary set of competencies that extend beyond the traditional scope of Data Science.
We will explore three key dimensions:

i) How Data Literacy empowers students from all disciplines to critically engage with data and foundational data science concepts;

ii) How it expands and enriches the scope of Data Science, making it relevant for students in computer science and related fields; and

iii) How Data Literacy initiatives can act as bridges between disciplines, fostering interdisciplinary collaboration and shared understanding.

In this context we will present current initiatives for the implementation of Data Literacy at the Friedrich Schiller University Jena. We will also discuss potential future developments and strategies for further embedding Data Literacy across the university.

Assessing Segmentation Errors in Multiplexed Imaging for Tissue Property Inference

Author: Matthias Bruhns

Co-authors: Jan Schleicher; Manfred Claassen; Marcello Zago; Maximilian Wirth; Sepideh Babaei

University of Tübingen

Corresponding Authors: marcello.zago@uni-tuebingen.de, jan.schleicher@uni-tuebingen.de, sepideh.babaei@uni-tuebingen.de, maximilian.wirth@uni-tuebingen.de, manfred.claassen@uni-tuebingen.de, matthias.bruhns@uni-tuebingen.de

Multiplexed tissue imaging provides an unprecedented understanding of tissue biology by capturing the spatial relationships between cells and their protein expression profiles. Common approaches use highly multiplexed fluorescence microscopy to measure tissue slides as images where each channel represents the intensity of a particular protein marker. Single-cell features are generated by segmenting the cells and aggregating all pixel intensities per cell for all measured channels. The resulting feature table is then used for further analysis, including quality control, preprocessing, dimensionality reduction, and clustering. Cell segmentation masks play a critical role in the generation of feature tables. Thus, their accuracy is crucial for downstream analysis. However, segmentation remains challenging, with existing solutions being prone to errors. The performance of these segmentation methods is evaluated using metrics such as mean average precision, intersection-over-union, and F1 scores. In this work, we empirically examine how segmentation errors influence the outcomes of downstream analysis, providing insights into the epistemic uncertainty induced by segmentation errors. Specifically, we use a perturbation procedure to generate augmented segmentation masks, allowing for systematic investigations. Using this approach, we can quantify the propagated error in the subsequent pipeline, which includes neighborhood preservation, unsupervised clustering and phenotyping. We demonstrate that even small perturbations significantly impact cell type inference and that the commonly used IoU-F1 metric hides segmentation inaccuracies.

LLMs for Information Extraction from Clinical Trials

Authors: Jasvinder Singh; Diana Peters; Sirko Schindler

DLR – Institute for Data Science

Corresponding Authors: diana.peters@dlr.de, jasvinder.singh@dlr.de, sirko.schindler@dlr.de

Natural Language Processing plays a crucial role in the automated parsing of clinical data. This data encompasses a wide range of sources including clinical notes, trial documents, imaging and sensor data, and patient-reported outcomes. In the project Avatar, we explore clinical trial documentation. These describe the circumstances of a trial: which patient groups shall be in- or excluded, and which kind of treatment is considered. From these documents, we extract relevant details related to the patient accrual process with the goal of providing researchers with data that fits to their study.
Due to the often-unstructured nature of descriptions, inferring users' intent becomes a challenging task. To address this issue, we use Named Entity Recognition (NER) to extract and categorize information within clinical trial descriptions. Given the recent advances in pre-trained language models, we fine-tune a selection of such models on the Facebook NER (FBNER) dataset which consists of clinical trial descriptions. Our approach involves fine-tuning these language models using supervised token-level classification along with domain adaptation (from generic clinical as well as non-clinical models). We classify extracted information into categories such as "disease," "therapy," "drugs", "gender" etc. and evaluate the performance on German Clinical Trials.

Probabilistic Embeddings for Frozen Vision-Language Models

Authors: Aishwarya Venkataramanan; Joachim Denzler

Computer Vision Group, Friedrich Schiller University Jena

Corresponding Authors: joachim.denzler@uni-jena.de, aishwarya.venkataramanan@uni-jena.de

As deep learning systems are increasingly deployed in high-stakes domains such as medical diagnostics, climate modeling, and autonomous decision-making, their ability to express uncertainty in their predictions becomes crucial. Traditional neural networks, while powerful, often produce overconfident predictions, even when presented with out-of-distribution data, that are different from the training data or ambiguous inputs. This lack of calibrated uncertainty can undermine trust in AI systems.

In this work, we present recent advances in uncertainty quantification (UQ) methods for deep learning and demonstrate their practical relevance across both classification and regression tasks. These UQ methods are further applied to downstream tasks such as out-of-distribution detection, active learning, and model calibration. The results illustrate how uncertainty-aware models can more accurately reflect predictive confidence and contribute to the development of more robust and interpretable AI systems.

Spatial and Dynamics Informed Latent-Variable Estimation with Deep Gaussian Processes

Author: Marcello Zago[1]

Co-authors: Soham Mukherjee[2]; Manfred Claassen[1]

[1] *University of Tübingen*

[2] *University of Dortmund*

Corresponding Authors: manfred.claassen@uni-tuebingen.de, soham.mukherjee@tu-dortmund.de, marcello.zago@uni-tuebingen.de

Spatial transcriptomics is an emerging field enabling the study of gene expression within its spatial tissue context, offering critical insights into cellular function and organization. Incorporating spatial information is essential for accurate biological hypothesis generation, as it provides context often lost in traditional single-cell analyses. A key challenge in this area is the inference of underlying biological processes, which is central to many research domains. However, current deep learning-based methods often sacrifice interpretability, making it difficult to extract meaningful biological conclusions.

To address this, we propose a nested two-layer Deep Gaussian Process [1] model for the estimation of a latent time variable that captures the progression of a biological process. The first layer learns a latent representation that is explicitly informed by spatial location, embedding spatial correlation directly into the model. The second layer models gene expression as a function of this inferred latent time, which guides the latent time to produce temporally smooth gene expression changes, as observed in many biological processes.

This model estimates a latent time that characterizes an underlying biological process, directly informed by spatial context. By capturing smooth gene expression changes along this inferred trajectory, the model provides an interpretable representation of spatial biological dynamics. This interpretability supports the generation of meaningful hypotheses across a range of biological and medical research fields.

[1] Damianou, A., & Lawrence, N. D. (2013, April). Deep gaussian processes. In Artificial intelligence and statistics (pp. 207-215). PMLR.

Explainable Seismic Signal Discrimination: A Comparative Analysis of Convolutional Neural Networks and Vision Transformers

Author: Valentin Kasburg[1]

Co-authors: Marcel van Laaten[1]; Markus Zehner[2]; Jozef Müller[3]; Nina Kukowski[2]

[1] *Institute of Geosciences, Friedrich Schiller University Jena*

[2] *Institute of Geography, Friedrich Schiller University Jena*

[3] *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven*

Corresponding Author: valentin.kasburg@uni-jena.de

Traditionally, the discrimination of seismic signals has been carried out manually by expert analysts, who assign events – such as earthquakes, quarry blasts, and other anthropogenically induced events – to specific categories. This process often involves the use of cross-correlation techniques for pattern recognition. Although methods from the field of deep learning, particularly convolutional neural networks, have demonstrated significant potential in automating seismic event discrimination, their widespread application remains limited. Two main challenges hinder their further and frequent deployment: First, the difficulty of developing models that maintain high accuracy across different regional seismic networks. Second, the limited interpretability of such models, which often makes their internal decision-making processes non-transparent.
In this contribution, we investigate Vision Transformers as a novel deep learning approach for the discrimination of different types of seismic events. We present a comparative analysis between Convolutional Neural Networks and Vision Transformers with regard to their ability to distinguish between earthquakes, quarry blasts, and mining-related seismic events. Our results demonstrate that the tested vision transformer architectures can reach discrimination accuracies between 95% and 98%.
To enhance the interpretability of model outputs, we employ visualization methods that provide insight into the internal reasoning of the Vision Transformer models. Specifically, we apply Attention Rollout to track and aggregate attention scores across multiple transformer layers, and LeGrad, a gradient-based attribution technique that identifies input regions most relevant for the model's predictions. While Attention Rollout directly utilizes the architecture's inherent attention structure, LeGrad derives importance by analyzing the gradients of the output with respect to the input. Together, these approaches offer complementary perspectives on model behavior and contribute to a more comprehensible and transparent understanding of how seismic signals are discriminated by the deep learning models. This supports the potential of Vision Transformers as an interpretable and high-performing alternative to Convolutional Neural Networks in the context of automated seismic event discrimination.