Statistical Methods in Imaging (SMI) 2026 – Program

Poster and Mixer Event

4:30–6:30 PM | Michigan League Ballroom

P2

Haozheng Xu

Vanderbilt University Medical Center

A Curated Meta-analytic Neuroimaging Dataset from Neurovault

Co-author(s): Simon Vandekar

Abstract

Replicability has been a major challenge in neuroimaging research, driven by limited sample sizes and substantial heterogeneity across data collection, preprocessing, and analysis pipelines. Meta-anlaysis is a critical tool to integrate information across studies, however most neuroimaging meta-analysis is coordinate-based and does not directly use the images. Neurovault (neurovault.org) is an open-access, continually updated repository that includes group-level statistical images across thousands of studies, but is highly heterogeneous and noisy making it difficulty to standardize image values and alignment. We develop an automated, scalable quality-control and inference pipeline for neuroimaging meta-analysis for the Neurovault repository. Study-level maps are retrieved automatically from NeuroVault and rigidly registered to MNI 2 mm standard space, with systematic exclusion criteria and information retrieval implemented at each processing step. Test-statistic images are converted to effect-size maps and implausible images are automatically identified based on the expected variability of these summaries as a function of study sample size and voxel coverage. We perform voxelwise inference based on semantic-embeddings derived from an LLM to contrasts meta-analytic differences in task activation of fMRI data. is performed using linear mixed-effects models with random study effects and heteroskedastic variance weights, allowing principled aggregation across studies. The pipeline is fully reproducible, implemented in R, and designed for automated monthly execution with cloud-based archiving through Zenodo's API. The curated dataset serves as a valuable resource evidence synthesis and increases the utility of the Neurovault dataset for image-level meta-analysis.

P3

Chen Mu

Florida State University

Statistical Analysis and Characterization of Gold Nanoparticle Morphology from Tomographic Image Stacks

Abstract

Electron tomography provides high-resolution volumetric imaging of nanoscale materials, yet extracting reproducible quantitative morphology from tomographic reconstructions remains a fundamentally statistical problem due to reconstruction artifacts, segmentation uncertainty, and projection variability. This work develops a statistical imaging framework for inferring nanoparticle morphology from electron tomography image stacks.

We formulate nanoparticle detection and measurement as a multi-stage statistical pipeline combining image preprocessing, segmentation, geometric feature extraction, and uncertainty quantification. Individual particles are identified from reconstructed volumes, and their morphology is summarized through shape descriptors including equivalent diameter, aspect ratio, and anisotropy measures derived from second-moment tensors. Population-level morphology is modeled using distributional summaries and hierarchical modeling to characterize heterogeneity across projections and reconstructions.

Spatial organization is analyzed through spatial point-process methodology, including nearest-neighbor distance distributions and clustering diagnostics, enabling inference on particle dispersion and aggregation. To quantify measurement reliability, we evaluate variability induced by projection angle, reconstruction settings, and segmentation thresholds, yielding empirical uncertainty estimates for key morphological descriptors.

This study highlights the role of statistical methodology in transforming tomographic imaging data into reproducible quantitative measurements. The proposed framework provides a principled foundation for uncertainty-aware morphological analysis in nanoscale imaging and offers a generalizable approach for statistical inference from tomographic datasets.

P4

Pascal Tyrrell

University of Toronto

Restoring Radiomic Feature Interpretability After Principal Component Analysis in Prostate MRI Cancer Risk Models

Co-author(s): Khashayar Namdar, Hamed Zakeri, Saeidehsadat Mirjalili, Pascal Tyrrell, Leo Anthony Celi

Abstract

Description of purpose: To enhance prostate cancer risk stratification through explainable artificial intelligence (XAI) by developing a pipeline that integrates principal component analysis (PCA) for dimensionality reduction and restores interpretability in radiomics-based classification using multiparametric MRI.

Methods: We used the 2021 PI-CAI dataset from RadMLBench, comprising 3,045 radiomics features extracted from T2-weighted, apparent diffusion coefficient (ADC) maps, and diffusion-weighted imaging (DWI) MRI sequences of 969 patients. The task involved binary classification of csPCa (Gleason >3+3, n=637) vs. non-csPCa (n=332). We applied 10-fold nested cross-validation using LightGBM and compared baseline performance with a dimensionality-reduced variant utilizing PCA preserving 95% of variance. Grid search was conducted on the validation set, and model performance was evaluated using the area under the receiver operating characteristic curve (AUC). The preprocessing workflow included removal of highly correlated features (correlation >0.95), elimination of near-zero variance features (threshold <0.05), and Min-Max normalization. All preprocessing steps were derived from the training set and applied to validation and test sets to prevent leakage. To overcome PCA's limitation in post-hoc interpretability, a feed-forward neural network (NN) with a single hidden layer (256 units) was trained to reconstruct the full radiomics space. Feature importance scores from LightGBM were projected through this mapping and transformed into the original feature domain using the stored eigenvector matrix derived during PCA.

Results: The baseline 10-fold AUC without dimensionality reduction was 0.701 (95% CI: 0.680–0.722). Incorporating PCA improved performance, achieving an AUC of 0.753 (95% CI: 0.728–0.778). On a representative 60/20/20 train/validation/test split, the test set yielded an AUC of 0.735, accuracy of 0.727, and F1-score of 0.819. Post-hoc interpretability analysis identified top contributing features in the original radiomics space, including dwi_original_glcm_Imc1 and dwi_wavelet-LLH_firstorder_Minimum.

Conclusions: Dimensionality reduction via PCA improved model performance but introduced a loss of interpretability. By integrating an NN-based reconstruction step and leveraging the PCA eigenstructure, our pipeline restored feature-level attribution, enabling meaningful clinical interpretation.

P5

Khashayar Namdar

University of Toronto

AI-Powered Endometriosis MRI Reporting: Automated #Enzian Scoring in Pelvic MRI Enabled by Large Language Models

Co-author(s): Dominik Deniffel, Joana Kostova, Jonas Stief, Gustav Andreisek, Amelie Lutz, Eliane Pauli, Christian Houbois, Khashayar Namdar

Abstract

Background: The #Enzian classification provides a comprehensive 14-compartment framework for describing the anatomical extent of endometriosis on imaging. Despite its clinical value, routine implementation in radiology reporting remains limited due to the complexity of manual extraction and scoring from narrative MRI reports. Large language models (LLMs) may enable automated extraction of structured #Enzian scores from free-text reports and potentially support clinical training and standardized reporting.

Purpose: To evaluate the ability of online and locally deployed LLMs to automatically extract #Enzian scores from pelvic MRI reports and compare their performance with radiology trainees.

Methods: A single-center retrospective study was conducted using 186 pelvic MRI reports from patients with suspected or confirmed endometriosis. Ground-truth #Enzian scores were established through manual review by a fellowship-trained uroradiologist. Twelve radiology trainees without prior #Enzian experience independently scored a subset of 50 cases. Six online LLMs and six local LLMs were evaluated using one-shot prompting, and classification accuracy was compared across the 14 compartments of the #Enzian system.

Results: Online LLMs demonstrated consistently high performance with pooled multiclass accuracies ranging from 89.1% to 92.4%, with Gemini 2.5 Pro achieving the highest accuracy (92.4%) and outperforming trainees in 11 of 14 categories. Local models showed performance strongly dependent on model size; the best-performing local model (GPT-OSS 120B) achieved 87.1% average accuracy, whereas smaller models showed substantially lower performance. Cost analysis revealed considerable variation among online APIs (ranging from $3.82 to $42.26 for 186 cases), while local deployment provided a privacy-preserving alternative.

Conclusion: LLMs can reliably extract structured #Enzian scores from pelvic MRI reports. Online models currently demonstrate the highest performance and may support training and reporting workflows, whereas large local models provide a viable privacy-preserving alternative with slightly lower accuracy.

P7

Susan Glenn

Los Alamos National Lab

Bootstrap-based Hypothesis Test of 2D Contours using Elastic Shape Analysis

Co-author(s): Justin Strait, Kelly Moran, Chris Danly, Matthew P Selwood

Abstract

Shapes of objects in images are often complex, high-dimensional, and vary in ways not captured by standard Euclidean geometry and statistics. Statistical shape analysis encompasses methods for flexible and interpretable measurement of intrinsic shape and shape variability in geometric objects. Elastic Shape Analysis (ESA) is one such method that measures shape differences between objects in a way that is invariant to rotation, scale, translation, and parameterization. Although ESA is useful for quantifying shape of objects in many image applications, formal methods for statistical inference in image-based ESA remain limited. This work introduces an inference procedure for the construction of empirical confidence intervals of the elastic shape distance (ESD) between a proposed underlying true shape and an estimated shape. The confidence intervals are created using the a bootstrap procedure for non-smooth functionals, which accounts for the non-differentiability of the ESD. The effectiveness of the method is illustrated through both numerical studies and real world image examples from inertial confinement fusion (ICF).

P8

Robert J Bishop

JointSpace and Operating Rhythms

Abstract

We live in the JointSpaces our bodies create with the world. JointSpace is the hidden engine of the animate world creating relations between our nerves and machines. We employ do-calculus to mathematize the relationship of abiotic to biotic interfaces.

P9

Stefan Eng

University of Michigan

Modeling Differences in Tumor-Immune Co-Localization to Predict Immune Checkpoint Inhibitor Response in Metastatic NSCLC

Co-author(s): Jean Morrison, Nicholas Lesniak, Timothy Frankel, Veera Baladandayuthapani

Abstract

We jointly analyze two cohorts of metastatic non-small cell lung cancer (mNSCLC) patients: one cohort (Qin et al. 2022) in which patients received immune checkpoint inhibitors (ICI) following prior lines of therapy, and a second in which ICI was administered as first-line treatment. Using multiplex fluorescent immunohistochemistry (mfIHC) on pre-treatment tumor tissue, we perform cell segmentation and classification to characterize the tumor immune microenvironment. Specifically, we used antibodies for CD3, CD8, CD163, PD-L1, pancytokeratin, and FOXP3, resulting in cell phenotype of epithelial tumor cells, cytotoxic T cells, helper T cells, regulatory T cells, antigen presenting cells (APC) with PD-L1 positivity status for APCs and tumor cells. We analyze the resulting cell populations as spatial marked point processes. First, we quantify first-order differences in cell proportions, counts, and density between ICI responders and non-responders. We then model the local microenvironment surrounding reference cell populations using hierarchical frameworks for two- and three-way co-localization, explicitly accounting for target cell prevalence and spatial correlation structure. Specifically, we analyze the interactions between the tumor-immune cells, and immune-immune cells followed by three way interactions between tumor and two other immune cells. Finally, we introduce preliminary methods for jointly modeling local cell-cell interactions using latent Gaussian models.

P10

Xiaoyu Qiu

University of Michigan

Benchmarking Uncertainty Quantification of Plug-and-Play Diffusion Priors for Inverse Problems Solving

Co-author(s): Taewon Yang, Zhanhao Liu, Guanyang Wang, Liyue Shen

Abstract

Plug-and-play diffusion priors (PnPDP) have become a powerful paradigm for solving inverse problems in scientific and engineering domains. Yet, current evaluations of reconstruction quality emphasize point-estimate accuracy metrics on a single sample, which do not reflect the stochastic nature of PnPDP solvers and the intrinsic uncertainty of inverse problems, critical for scientific tasks. This creates a fundamental mismatch: in inverse problems, the desired output is typically a posterior distribution and most PnPDP solvers induce a distribution over reconstructions, but existing benchmarks only evaluate a single reconstruction, ignoring distributional characterization such as uncertainty. To address this gap, we conduct a systematic study to benchmark the uncertainty quantification (UQ) of existing diffusion inverse solvers. Specifically, we design a rigorous toy model simulation to evaluate the uncertainty behavior of various PnPDP solvers, and propose a UQ-driven categorization. Through extensive experiments on toy simulations and diverse real-world scientific inverse problems, we observe uncertainty behaviors consistent with our taxonomy and theoretical justification, providing new insights for evaluating and understanding the uncertainty for PnPDPs.

P11

Martin Salgado-Flores

Southern Methodist University

Relational Persistent Homology of the Tumor Microenvironment for Survival Analysis of Lung Adenocarcinoma

Co-author(s): Chul Moon

Abstract

Topological data analysis (TDA) offers tools to quantify the shape and relational structure of spatial data. A popular tool used to quantify topological information that has been found useful in the analysis of medical images is persistent homology. Stolz et al. (2024), demonstrated through simulation that relation filtrations, such as Dowker and Multispecies witness filtrations, can capture informative persistent homology that may be missed by classical methods such as Vietoris-Rips filtration. This study applies these approaches to lung adenocarcinoma (LUAD) data from The Cancer Genome Atlas (TCGA) publicly accessible via The Cancer Imaging Archive (TCIA). Using the TDA framework proposed by Stolz et al. (2024), we seek to find relational topological features between tumor, lymphocyte, and stromal cells in the TME of LUAD images. These topological features are incorporated into survival analysis models to investigate their association with patient survival outcomes.

P12

Shashipraba Rajakaruna

Texas Tech University

Topology guided conditional generative adversarial network for medical image simulation

Co-author(s): Asim K. Dey, Suprateek Kundu

Abstract

Medical image analysis faces enduring challenges from data scarcity, privacy restrictions, and the under-representation of diverse populations. While generative adversarial networks (GANs) have shown promise in producing synthetic medical images, conventional approaches often overlook higher-order structural features and statistical uncertainties, limiting their reliability. In this study, we develop a topology-guided conditional GAN (TDA-cGAN) that incorporates persistent homology into the generative process to preserve multi-scale geometric and structural information of medical images. Experiments on malaria cell images and Alzheimer's brain MRI images demonstrate that TDA-cGAN consistently outperforms standard cGAN, achieving lower Fréchet Inception Distances and Kernel Inception Distance, which reflect superior image quality and diversity. To further validate these results, we conduct statistical analyses of pixel intensity distributions and higher-order moments, which confirmed that our synthetic images closely match real images. Additionally, we employ topological data analysis based hypothesis tests to evaluate the differences between the distributions of real and simulated images. The results indicate no statistically significant discrepancies between real images and those simulated by the TDA-cGAN.

P13

Sanya Kejriwal

University of Pennsylvania

Evaluating the Performance of the SPICE Test Under Heterogeneous Spatial Autocorrelation

Co-author(s): Yiyan Hao, Brian R. White

Abstract

Comparing the spatial patterns of brain maps across modalities (e.g., structural, functional, genomic) is increasingly common in neuroimaging research. However, inference is complicated by spatial autocorrelation. Intermodal correspondence tests typically assume that spatial autocorrelation is uniform across the brain. Recent work has shown that nonstationarity can result in Type I error for the Spin Test, a popular method for intermodal correspondence testing. In this study, we evaluate the performance of another popular method, the Simple Permutation-based Intermodal Correspondence (SPICE) test, under varying spatial autocorrelation structures.

Using simulated spatial signals with controlled patterns of heterogeneous smoothing, we generate pairs of maps under both null and correlated conditions. We then apply the SPICE framework to assess correspondence between maps across repeated simulations. Our analysis quantifies the impact of spatial heterogeneity on both Type I and II error, providing insight into the robustness of permutation-based spatial correspondence tests for neuroimaging applications.

P14

Jessica Aldous

University of Michigan

MoSAIC: Multi-Resolution Spatial Regression Analysis of Cellular Colocalizations in Cancer Imaging

Co-author(s): Michele Peruzzi, Maria Masotti, Aaron Udager, Allison May, Evan Keller, Veera Baladandayuthapani

Abstract

Hierarchical multiplex imaging approaches generate spatially resolved single-cell measurements across multiple, spatially organized fields of view (FOVs) within patient tumor specimens, thereby enabling systematic investigation of how the organization of the tumor microenvironment varies along biologically meaningful intratumoral gradients. Existing approaches fail to jointly address this multi-resolution data structure needed to recover true biological signals. We propose MoSAIC: multi-resolution spatial regression analysis of cell colocalizations, a hierarchical Bayesian spatial regression model designed for multi-resolution spatial data. MoSAIC decomposes the joint variation into three model components: (i) global tumor-gradient effects, (ii) patient-specific effects to capture inter-patient variability, and (iii) Gaussian process models to account for spatial dependence between FOVs within each patient tumor tissue. Simulations demonstrate MoSAIC has improved prediction and model fit compared to existing spatial and non-spatial model alternatives. Our method is motivated by and applied to a renal cell carcinoma multiplex imaging cohort to investigate immune–tumor colocalization patterns across the epithelial-to-mesenchymal transition (EMT) gradient. MoSAIC identifies increased macrophage–tumor colocalization and decreased cytotoxic T–tumor colocalization progressing across the increasing EMT gradient, consistent with EMT-associated immune suppression and spatially varying immune engagement. Overall, MoSAIC provides an interpretable, multi-resolution framework for quantifying spatial tumor-gradient effects in cancer imaging studies. Software is available on Github at jcaldous/MoSAIC.

P15

Zihang Wang

Emory University

Flexible Normative Brain Modeling Using Quantile Super Learning with Motion Control

Co-author(s): Benjamin B Risk

Abstract

Normative brain modeling provides a statistical framework for characterizing patterns of brain development. These models are trained on large cohorts of healthy individuals, enabling the identification of abnormal individuals whose brain deviates from typical patterns. Such approaches can identify developmental concerns. However, existing normative modeling relies on parametric regression models that may perform poorly for extreme quantiles. In addition, data quality is often overlooked. Head motion during MRI acquisition can lead to spurious cortical thinning estimates and mischaracterize developmental trajectories. In this work, we propose a flexible normative modeling framework based on a quantile Super Learner that aggregates a diverse library of semi-parametric and machine learning quantile estimators. We apply this method to brain morphology data from the Reproducible Brain Charts to construct normative models. The Euler number is incorporated as a covariate to account for motion-related artifacts. Results show that our ensemble normative model improves conditional quantile estimation and enhances the identification of individuals with abnormal brain structure.

P16

Yuting Duan

University of Michigan

Robust Image-on-scalar Regression with Simultaneous Activation Region and Extreme Signal Detection

Co-author(s): Jian Kang, Hui Jiang

Abstract

Brain imaging modalities are widely used to investigate the association between brain activity and clinical variables. However, extreme signals frequently arise in brain imaging data and must be carefully addressed to ensure reliable data analysis. In this paper, we propose a robust image-on-scalar regression approach (RISE-X) that mitigates the influence of extreme signals by formulating the problem as a penalized regression model with truncated convex function minimization. This approach effectively reduces the impact of extreme signals while preserving meaningful spatial patterns. We demonstrate the advantages of RISE-X through simulations and the application to the Adolescent Brain Cognitive Development (ABCD) study.

P17

Yuhan Geng

University of Michigan

Threshold Spatial Attention Transformer for Efficient Image Generation

Co-author(s): Jian Kang, Wei Hao

Abstract

We propose the Threshold Spatial Attention Transformer (TSAT), a novel model for efficient and high-quality image generation in neuroimaging. Existing approaches such as GANs and Diffusion Probabilistic Models often suffer from data inefficiency, heavy computational demands, and strong reliance on labeled datasets. TSAT addresses these limitations through block-wise feature sampling, low-rank spatial attention kernels, and nested thresholding, reducing trainable parameters while preserving fine-grained spatial structure. Using both supervised and semi-supervised training, TSAT adapts to limited data settings. On MNIST, TSAT achieves competitive fidelity and diversity, with semi-supervised learning leveraging large unlabeled sets to overcome small labeled sample sizes. On Human Connectome Project fMRI data, TSAT generates plausible activation maps conditioned on gender, IQ, and working memory, consistently outperforming GANs and DDPMs in PSNR, SSIM, classification accuracy, and FID, while requiring substantially less training time. These results demonstrate TSAT as a scalable and efficient framework for medical image synthesis under real-world data constraints.

P18

Yiyan Hao

University of Pennsylvania

Applying Multiple Imputation On Neuroimaging Data While Maintaining Spatial Autocorrelation

Co-author(s): Ofer Harel, Simon N. Vandekar, Russell T. Shinohara, Brian R. White

Abstract

Neuroimaging datasets (e.g., magnetic resonance imaging, MRI; diffuse optical tomography, DOT) often present with variability in the field-of-view (FOV) across subjects. In group-level analysis, voxels with incomplete data are commonly excluded (complete case analysis), reducing FOV. Multiple imputation (MI) is a rigorous statistical method for missing data. Despite its desirable properties, the use of MI in neuroimaging is underexplored. One complicating factor is that imputation may not preserve spatial autocorrelation (SA) in the imputed images. Proper accounting for spatial autocorrelation and the spatial extent of neuroimaging features is essential to proper inference (e.g., statistical parametric mapping). Traditional imputation models do not utilize spatial information.

Our long-term goal is to develop an MI pipeline suitable for high-dimensionality neuroimaging data. As a first, we characterized existing imputation models in their ability to reproduce the original covariance structure. We conducted simulation studies to systematically evaluate imputation strategies for neuroimaging data with variable fields-of-view, assessing their ability to preserve spatial autocorrelation and support accurate group-level inference.

Our findings demonstrate that the choice of imputation model crucially affects both the fidelity of reconstructed data and the validity of statistical inference. MICE (R package for MI) with a LASSO linear regression model provides a promising foundation for a spatially informed multiple imputation pipeline for neuroimaging data, which will enhance reliability and reproducibility of neuroimaging studies.

P19

Rosa Fallahpour

University of Toronto

A Spatial Entropy Framework for Characterizing Radiomic Heterogeneity: Application to Healthy and Pediatric Tumor Brain MRI

Co-author(s): Khashayar Namdar, Pascal. N. Tyrrell

Abstract

Quantifying spatial heterogeneity in medical imaging is important for understanding structural variability and improving imaging-derived biomarkers, which can be beneficial for machine learning-based modeling. In this work, we introduced a Shannon entropy-based methodology designed to quantify spatial heterogeneity in brain MRI using radiomics-derived traits. In this approach, each brain image was partitioned into regions, and radiomics features were extracted from each region using a standardized radiomics pipeline. Continuous feature values were then transformed into categorical traits, allowing the construction of trait distributions across spatial regions. The proposed methodology summarizes the overall spatial heterogeneity of MRI images by aggregating entropy-based heterogeneity measures derived from these trait distributions. This formulation provides an interpretable metric that captures the diversity of radiomic patterns across brain regions. The proposed framework offers a systematic and reproducible approach for translating radiomics features into quantitative measures of spatial complexity, enabling comparative heterogeneity analysis across medical imaging datasets and supporting future studies in developing heterogeneity-aware machine learning models trained by MRI images. The framework was evaluated on two brain MRI datasets: the IXI healthy adult dataset (n = 270) and the BraTS-PED pediatric brain tumor dataset (n = 251). Using first-order radiomic entropy as an example feature and dividing each brain into four spatial quadrants, the proposed method produced heterogeneity average index (HTA) of 0.820 for the IXI dataset and 0.884 for the BraTS-PED dataset. Individual quadrant-level indices (HTI) exhibited substantial variance (0.618 to 0.995), reflecting the tool's sensitivity to localized structural diversity. These results indicate that the framework provides a standardized scale for comparing spatial complexity across diverse cohorts, offering a necessary precursor for integrating heterogeneity-aware features into predictive machine learning models.

P20

John Bodenschatz

Marquette University

Distributionally Accurate Phase Activation in FMRI

Co-author(s): Daniel Rowe

Abstract

Functional magnetic resonance imaging (fMRI) is used to study how the brain works by measuring changes in signals across different parts of the brain. Researchers want fMRI scans that are both fast and detailed. However, increasing temporal and spatial resolution also increases noise, which can make the data harder to analyze. Complex-valued fMRI time series can be deconstructed into two parts: magnitude and phase. While most studies focus on magnitude images, research shows that the phase component can also contain important biological information. In this work we focus on analyzing phase-only activation using maximum likelihood estimation with a mathematically accurate model. This approach allows phase information to be modeled more accurately, particularly in cases where the signal-to-noise ratio is low, resulting in the detection of task-related signal changes in the phase component of fMRI time series.

P21

Ziyu Liu

University of Michigan

Bayesian Image-on-Image Regression for Linking Resting-State Connectivity to Task fMRI Activation

Co-author(s): Jian Kang, Timothy Johnson

Abstract

Functional MRI (fMRI) data provide rich insights into the brain's functional organization. Understanding how resting-state functional connectivity is associated with task-evoked activity is an important problem in neuroimaging. To study this association, we propose an image-on-image regression (IIR) model, where predictor and outcome images are represented with fixed spatial bases and link them through low-dimensional basis-to-basis coefficient matrices that provide interpretable and spatially coherent mappings between modalities. For cortical surface outcome images, spherical harmonics capture smooth spatial variation. Posterior inference is performed using an efficient Markov chain Monte Carlo algorithm enabling scalable computation for high-resolution imaging data. Applied to the Adolescent Brain Cognitive Development (ABCD) Study, our method reveals coherent spatial associations between resting-state functional connectivity and the activation from a working memory task fMRI study with interpretable estimation and uncertainty quantification.

P22

Longhao Pang

University of Michigan

Bayesian Joint Singular Value Decomposition with Application to Image Analysis

Co-author(s): Jane E. Huggins, Jian Kang

Abstract

Modern imaging technologies, including electroencephalography (EEG) and functional magnetic resonance imaging (fMRI), produce high-dimensional, structured data that can be naturally organized as matrices — such as channel-by-time signals or region-by-region connectivity matrices. Analyzing these matrix-variate data presents two fundamental challenges: the inherent low signal-to-noise ratio in imaging measurements, which demands effective information pooling across subjects, and substantial inter-subject variability in the latent signal patterns, which complicates group-level inference and individual-level prediction. Existing approaches either ignore the matrix structure by vectorizing the data, or struggle to scale to high-dimensional settings due to the large number of parameters relative to available sample sizes.

To address these challenges, we propose Bayesian Joint Singular Value Decomposition (BJSVD). The framework extracts shared latent bases and subject-specific latent features from observed matrix-variate data, which can be used for downstream predictive models that link these features to scalar outcomes. For posterior computation, we employ a Gibbs sampler with a fast von-Mises Fisher distribution sampling for the basis based on Ulrich-Wood rejection sampling. In simulations, BJSVD outperforms standard methods and other state-of-art methods in classification accuracy. We apply BJSVD to EEG data analysis in brain–computer interface (BCI) involving participants with amyotrophic lateral sclerosis (ALS), recovering meaningful latent spatial-temporal patterns from EEG signals and achieves high accuracy in P300 speller letter prediction.

P23

Yuxiao Nie

University of Michigan

Sparse Dynamic Latent Factor Models for Multi-Channel EEG based Brain Computer Interfaces

Co-author(s): Jian Kang, Ji Zhu

Abstract

Electroencephalography (EEG)–based brain–computer interfaces (BCIs) aim to translate neural activity into control signals for communication and assistive technologies. A key task in these systems is to extract informative features from high-dimensional EEG recordings, including latent neural source signals and interactions across multiple channels, for downstream classification. This problem is challenging due to the low signal-to-noise ratios in EEG measurements and the complex spatial–temporal dependence among channels.

To address these challenges, we propose a sparse dynamic latent factor modeling framework for EEG signals in BCI. The proposed model represents observed EEG signals through a set of latent dynamic factors that capture underlying neural source activity while introducing structured sparsity to identify interactions among multiple channels. In particular, the framework is designed to capture higher-order channel interactions that may arise from coordinated neural activity across spatially distributed brain regions. This formulation provides a flexible approach for modeling temporal dynamics and complex dependence structures in high-dimensional neural signals while maintaining interpretability of channel-level relationships.

We develop an efficient Gibbs sampler for posterior inferences. The proposed framework is evaluated through simulation studies and applications to real BCI-EEG datasets . The results demonstrate that the model can effectively capture structured dependencies among EEG channels and provide informative latent representations for BCI applications.

P24

Grant Carr

University of Michigan

SPARCL: Spatially Adaptive Regression for Covariate Driven Clustering

Co-author(s): Veera Baladandayuthapani, Jian Kang, Tim Frankel

Abstract

The tumor microenvironment (TME) is a complex system of tumor cells, immune cells, and surrounding tissue architecture. Recent work highlights the role of the TME in the evolution and prognosis of cancer, frequently highlighting the spatial arrangement of immune cells as predictive of disease characteristics. Here we present SPARCL, a spatial clustering algorithm to aid in the identification and interpretation of cell aggregates in cancer biopsy imaging data. We demonstrate the ability of SPARCL to identify clusters linked with patient and cell-level covariates in simulation settings, and we apply SPARCL to a lung cancer imaging dataset to demonstrate the identification of spatial clusters in a real data setting.

P25

Razmin Bari

University of Pittsburgh

Extracting Bipolar-Associated Functional Connectivity Subgraphs in Adolescent rs-fMRI Using Adaptive Dense Subgraph Discovery

Co-author(s): Michele Bertocci, Qiong Wu

Abstract

Bipolar disorder (BD) may involve distributed disruptions in large-scale brain networks rather than isolated regional abnormalities. Using resting-state fMRI data from the Inpatient Child and Adolescent Bipolar Spectrum (InCABS) Imaging Study, we constructed ROI-based functional connectivity matrices for adolescents with BD, other psychopathology (OP), healthy controls (HC), and a high-risk group (HR). Group differences were first assessed through mass-univariate two-sample t-tests across all ROI pairs, producing BD-specific contrast matrices.

Because edgewise approaches overlook coordinated multiregional effects, we applied Adaptive Dense Subgraph Discovery (ADSD) to identify coherent subnetworks where BD-related disruptions are collectively strong. ADSD combines greedy adaptive density search, integrated-likelihood selection of the tuning parameter λ, and a node-label permutation test to evaluate subgraph-level significance. This framework detects dense, interpretable connectivity alterations even when individual edges fail to survive multiple-comparison correction.

By integrating connectome-wide testing with subgraph discovery, this work highlights network-level signatures of BD in youth and demonstrates the utility of graph-based statistical methods for psychiatric neuroimaging.

P26

Siyan Wen

Penn Statistics in Imaging and Visualization Endeavor (PennSIVE), Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania Center for AI and Data Science for Integrated Diagnostics (AI2D), Perelman School of Medicine, University of Pennsylvania

Out-of-Sample CovBat for Multi-Site Harmonization of Resting-State Functional Connectivity

Co-author(s): Zheng Ren, Dhivya Srinivasan, Russell T. Shinohara, Haochang Shou

Abstract

Multi-site neuroimaging studies increase statistical power but introduce scanner- or site-related batch effects that can obscure biological signals. While ComBat-based harmonization methods remove for mean and variance differences across sites, they do not handle covariance differences, which are critical for connectivity analyses. CovBat extends ComBat by harmonizing covariance structure; however, its performance in out-of-sample (OOS) settings, where new sites are harmonized using a previously trained model without retraining, remains underexplored. In this work, we develop the OOS framework for CovBat harmonization by leveraging eigenvector derived from existing harmonization model and projecting new out-of-sample subjects directly into the same principal component space and evaluated its performance in terms of preserving disease signals while removing batch effects.

We evaluated OOS CovBat using resting-state fMRI connectivity data from 2,525 subjects in the iSTAGING consortium for connectivity features derived from a 17x17 personalized functional network parcellation via pNet. We simulated varying degrees of disease effects by introducing 25% reduction of the mean connectivity over 10 connectivity features among varying proportions of subjects per site. We compared four scenarios: unharmonized data, in-sample CovBat where all sites are included in training, control-only OOS CovBat trained on controls from training sites and applied to unseen sites, and full OOS CovBat trained on all subjects from training sites. Disease effects were assessed using feature-wise regression models adjusting for age and sex. We examined detection sensitivity and specificity, stability of effect size estimates, and preservation of age-related trajectories. Global batch effects were evaluated using multivariate distance matrix regression and feature-level site effects were assessed using ANOVA, Kruskal-Wallis, and variance-based tests.

Both OOS harmonization scenarios removed batch effects to a degree comparable with in-sample harmonization, with post-harmonization multivariate distance matrix regression yielding non-significant p-values under the control-only OOS scenarios, in contrast to the highly significant batch effects observed in unharmonized data. Disease-effect estimates remained stable across conditions, with sensitivity and specificity maintained at a level comparable to the gold-standard

in-sample approach. Age-related associations were preserved across all harmonization conditions. Performance degradation was observed when training data lacked age distribution overlap with test sites or when the training sample size was insufficient to reliably estimate the underlying covariance structure. In summary, these results support OOS CovBat as a reliable harmonization strategy for growing multi-site connectivity studies, enabling disease effect preservation without model retraining under appropriate training data conditions.

P27

Mara Sherlin Talento

King Abdullah University of Science and Technology

KenCoh: A Rank-Based Estimator of Canonical Coherence

Co-author(s): Sarbojit Roy, Hernando Ombao

Abstract

This work is inspired by the problem of characterizing a dependence measure between two cortical regions of the brain where each region contains multiple signal recordings from several neurons or channels (e.g., inhibitory and excitatory neurons). The goal is to identify differences in the structure of brain functional connectivity between known brain states. An exploratory tool for studying the dependence between two random vectors is via canonical correlation analysis. However, these are limited to only capturing linear associations and are sensitive to outlier observations. Mitigating these limitations is crucial because brain functional connectivity is likely to be more complex than linear, and brain signals may exhibit heavy-tailed properties. To overcome these limitations, we develop a robust method, Kendall's tau-based canonical coherence (KenCoh), to learn connectivity structure among neuronal signals filtered at given frequency bands. Our simulation study demonstrates that KenCoh is competitive with the moment-based estimator and outperforms the latter when the underlying distributions are heavy-tailed. We apply the proposed KenCoh method to EEG recordings from a virtual-reality driving experiment and to calcium imaging recordings in inhibitory and excitatory neurons of the auditory cortex in mice subjected to sound stimuli. Our findings reveal distinct regional dependencies across frequency bands and brain states.

P28

Frank Fazekas

University of Maryland

Centrosome Positioning and Nuclear Architecture: Implications for T Cell Activation

Co-author(s): Matthew Connell, Aashli Pathni, Ivan Rey-Suarez, Mikayla Greiner, Arpita Upadhyaya

Abstract

T cells, crucial players in the adaptive immune response, become activated when receptors on their surface recognize and bind to pathogenic peptides on antigen-presenting cells (APCs). Activation triggers a series of changes in cellular morphology, including notable alterations to nuclear shape and chromatin organization. Upon activation, the centrosome (the microtubule organizing center) repositions toward the immune synapse, the contact site between the T cell and the APC, where it plays a critical role in establishing cell polarity and directing intracellular transport mechanisms. Using confocal microscopy, we find that in both the model CD4+ Jurkat T cell line and in primary cytotoxic T lymphocytes (CTLs), the centrosome is tightly associated with a deep invagination of the nucleus. To characterize this association, we developed a 3D image analysis pipeline for nuclear surface reconstruction, curvature mapping, and tracking of centrosome–nucleus dynamics. Using this approach, we show that upon activation, the centrosome and nucleus reorient as a cohesive structural unit to direct synapse organization and effector function. This dynamic association is maintained by a balance of forces generated by the actin and microtubule cytoskeletons. Furthermore, this centrosome-proximal region of the nucleus constitutes a distinctive chromatin microenvironment depleted of heterochromatin. These findings deepen our understanding of nuclear organization and dynamics during T cell activation and establish a foundation for examining how nuclear architecture and interactions with the centrosome regulate immune function.

P29

Minjin Lee

University of Illinois Chicago

Monte Carlo–Based Assessment and Spatial Smoothing for CTRW Parameter Estimation in DWI

Co-author(s): Muge Karaman, Joe X. Zhou, Ping-Shou Zhong

Abstract

Diffusion-weighted imaging (DWI) plays a critical role in probing the microstructural properties of biological tissues. Classical DWI methods assume Gaussian diffusion in a homogeneous medium, leading to a mono-exponential signal model and the estimation of apparent diffusion coefficient (ADC) maps. However, the intrinsic heterogeneity of biological tissues often gives rise to non-Gaussian diffusion behavior. To better characterize tissue microstructure, the continuous-time random walk (CTRW) model has been proposed to account for both temporal and spatial diffusion heterogeneity and to capture intravoxel structural variability. Despite its potential, statistical estimation and inference for the CTRW model in DWI have not been systematically studied. In this work, we first investigate statistical uncertainty and estimation accuracy through Monte Carlo simulations. We derive the theoretical MRI signal model under the CTRW diffusion framework, enabling the generation of DWI signals with noise distributions implied by the CTRW model rather than relying on artificially imposed noise structures. Second, we extend the classical CTRW framework by incorporating spatial smoothing to address the limitations of conventional voxel-wise estimation. By borrowing information from neighboring voxels, the proposed method improves estimation stability and efficiency. We develop an iterative smooth-CTRW estimator for the generalized diffusion coefficient and the temporal and spatial heterogeneity parameters. Simulation studies and real data analyses demonstrate that the proposed smooth CTRW approach improves parameter estimation accuracy compared with conventional voxel-wise methods.

P30

Ralph Jiang

University of Michigan

Scalable Image-on-Scalar Regression for High-Dimensional Neuroimaging Data

Co-author(s): Jian Kang, Hui Jiang, Thomas Nichols

Abstract

Modern neuroimaging studies often seek to identify associations between subject-level variables and high-dimensional brain images. In image-on-scalar regression, the image is treated as the outcome and scalar variables act as predictors, but existing methods can become computationally challenging at large sample sizes or large image sizes while also yielding coefficient estimates that are difficult to interpret spatially. To address this setting, we develop the Scalable Image-on-Scalar Regression Algorithm (SIRA), a region-based method for image-on-scalar regression that assumes the coefficient image is both sparse and spatially homogeneous. SIRA estimates the coefficient image through a set of local region-updating operations together with pre-computed summary statistics, allowing it to remain computationally feasible in large imaging studies. In simulation studies, SIRA showed strong support recovery and coefficient estimation performance while remaining more tractable than several existing spatially aware alternatives. We then applied SIRA to 38,639 task fMRI contrast maps from the UK Biobank facial recognition task. In this application, SIRA identified interpretable age-related activation patterns concentrated in occipital and fusiform regions with known roles in visual and face-processing tasks. These results suggest that SIRA provides a practical framework for scalable and interpretable image-on-scalar regression in modern neuroimaging settings.