SUROJIT SAHA

surojit.saha@utah.edu • +1-(385)-419-6427 • surojit-utah.github.io • LinkedIn: surojit-saha-6258ab15

EDUCATION

Ph.D. in Computer Science, *Kahlert School of Computing, The University of Utah, USA* 2018 – 2024
***Research areas*:** Machine Learning and Computer Vision, ***Advisor*:** Dr. Ross Whitaker **3.98/4.0**

Master of Technology in Computer Science, *Homi Bhabha National Institute, India* 2009 – 2012
***Research area*:** Autonomous Navigation of Mobile Robots, ***Advisor*:** Dr. Prabir Kumar Pal **80.3/100**

Bachelor of Technology in Computer Science, *Maulana Abul Kalam Azad University of Technology, India* 2005 – 2009
***Research area*:** Database Management Systems, ***Advisor*:** Dr. Arup Kumar Bhaumik **9.08/10**

RESEARCH SUMMARY AND SKILLS

My research is focused on developing deep latent variable models (DLVMs), such as the VAE, with *provable* properties of the distribution in the latent space that find application in *outlier detection* and *disentanglement analysis*. I have developed a method to find the *relevant* latent variables in DLVMs that are sufficient to model a data distribution, representing the intrinsic dimensions of the dataset. In addition, I am interested in training deep neural networks with *limited* annotated data, and I have proposed methods to interpret biomedical and seismic images using *Gaussian processes* in the *few-shot* setup. Besides probabilistic modeling, I have worked on registering 3D scans (RGB-D data). During Ph.D., I had the opportunity to collaborate with researchers from **ExxonMobil, USA**, on multiple projects related to interpreting seismic images, such as few-shot segmentation, outlier detection, multitask learning, and Explainable AI.

Research Interests: Generative Models, Few-shot Learning, Metric Learning, Multitask Learning, Bayesian Statistics, and 3D Computer Vision.

Skills: TensorFlow, PyTorch, Python, C, C++, MATLAB.

PROJECTS

Generative Encoding Networks (*GEN*)

- Uses the *kernel density estimate* (KDE) to match the aggregate posterior distribution in the latent space to the prior.
- Properties of the latent distribution at convergence are derived analytically that are validated by empirical evidence.
- Demonstrates the use of the learned latent representations for *outlier detection* that favorably compares to the SOTA methods.

Aggregate Variational Autoencoders (*AVAE*)

- Matches the aggregate posterior to the prior in the VAE to alleviate the *posterior collapse* and *holes* observed in the latent space.
- Address the *bias* in the estimated bandwidth that encourages using KDE in high dimensional latent spaces (dimensions > 100).
- Latent factors (*not axis-aligned*) are more effective for *disentanglement* than VAEs, an advantage of matching aggregate distributions.

Automatic Relevancy Detection in the Variational Autoencoder (*ARD-VAE*)

- Discovers *relevant* axes in the VAE using a *hierarchical* prior (based on Bayesian statistics) without modification of the ELBO.
- *Robust* to the choice of the autoencoder architecture and optimization strategies across multiple datasets, unlike other methods.
- Empirical evaluations demonstrate the effectiveness in *modeling data distributions* and finding *latent factors explaining the data*.

Few-shot segmentation (FSS) methods using Gaussian processes (a Bayesian modeling framework**)**

- Proposed FSS methods for *seismic* and *microscopy* images as their annotations require especially skilled people, and thus, are costly.
- Developed a few-shot segmentation method for interpreting microscopy images, named the ***GP-UNet***.
- We present a few-shot semantic segmentation method for identifying seismic facies that can adapt to the *varying number of facies* across datasets, dubbed the ***AdaSemSeg***.

Multitask Learning

- Proposed a multitask learning method for regularizing deep neural networks designed to detect *geological features* in seismic images.
- Demonstrated the effectiveness of multitask learning in scenarios with *limited* labeled data to interpret complex seismic images.

3D computer vision

- We propose an alternative to pairwise *3D scan registration* that uses small n−cycle graphs from the pose graph to register scans.
- Computationally efficient solvers are designed for joint estimation of the camera poses in small loops, such as 3-, 4-, and 5-cycles, that recover 12, 18, and 24 degrees of transformation variables, resulting in lower drift error on loop closure.

PROFESSIONAL EXPERIENCE

The University of Utah, Graduate Research Assistant, ***Advisor*:** Dr. Ross Whitaker Jan 2019 – Present
Developed multiple DLVMs with provable properties of the latent space that find application in several downstream tasks. The proposed DLVMs address shortcomings in the existing methods and are not sensitive to the choice of neural network architectures or training strategies. The FSS methods using Gaussian processes on multiple non-natural image datasets solve many realistic problems where the availability of annotated data is a challenge. Self-supervised learning is used to initialize the backbone networks used in the FSS of seismic facies.

Ancestry.com, CV/NLP PhD Intern, ***Mentor***: Dr. Jack Reese May 2021 – Aug 2021
Developed regularization techniques for detecting paragraphs, tables, and other data layouts in records related to the ***Family History Book***.

The University of Utah, Graduate Assistant, ***Mentor*:** Dr. Srikumar Ramalingam Aug 2018 – Dec 2018
Developed a 3D scan registration method that jointly estimates the poses of multiple cameras using efficient solvers, resulting in lower error.

Department of Atomic Energy, Government of India, Scientific Officer Sept 2009 – Aug 2018
Developed accurate, robust, and efficient localization algorithms for indoor mobile robots based on the principle of Monte Carlo. Fortification of the security of IT assets against possible cyber-attacks. Setting up and maintaining the infrastructure for high-performance computing.

PUBLICATIONS

1. Surojit Saha, Sarang Joshi, and Ross Whitaker**,** *ARD-VAE: A Statistical Formulation to Find the Relevant Latent Dimensions of Variational Autoencoders*, 2024 (under review).

2. Surojit Saha, and Ross Whitaker**,** *AdaSemSeg: An Adaptive Few-shot Semantic Segmentation of Seismic Facies*, 2024 (under review).

3. Surojit Saha, Sarang Joshi, and Ross Whitaker**,** *Matching Aggregate Posteriors in the Variational Autoencoder*, 2024 (under review).

4. Surojit Saha*, Wasim Gazi*, Rehman Mohammed, Thomas Rapstine, Hayden Powers, and Ross Whitaker, *Multi-task Training as Regularization Strategy for Seismic Image Segmentation*, IEEE Geoscience and Remote Sensing Letters (**IF: 4.8**), 2023.

5. Xiwen Li, Tristalee Mangin, **Surojit Saha**, Rehman Mohammed, Evan Blanchard, Dillon Tang, Henry Poppe, Nathan Searle, Ouk Choi, Kerry Kelly, and Ross Whitaker, *Real-Time Idling Vehicles Detection Using Combined Audio-Visual Deep Learning,* International Conference on Intelligent Traffic and Transportation, 2023.

6. Surojit Saha, Shireen Elhabian, and Ross Whitaker**,** *GENs: Generative Encoding Networks*, Machine Learning (**IF: 7.5**), 2022.

7. Surojit Saha, Ouk Choi, and Ross Whitaker, *Few-Shot Segmentation of Microscopy Images Using Gaussian Process*, MOVI (a MICCAI workshop), 2022.

8. Pedro Miraldo, **Surojit Saha** and Srikumar Ramalingam, *Minimal Solvers for Mini-Loop Closures in 3D Multi-Scan Alignment*, IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.

ACHIEVEMENTS

1. Best paper award at MICCAI workshop, MOVI, 2022.
2. 99.16 percentile in GATE-2009, India (country-wide exam for pursuing Masters).

ACADEMIC SERVICES

Conferences: Served as a reviewer for the AAAI-23, ACML-23, AAAI-22, and ICVGIP-21.
Journals: Served as a reviewer for the IEEE TPAMI and IEEE TGRS.

TEACHING

1. *Teaching Mentor* for **Deep Learning**, Fall 2019, *Kahlert School of Computing, The University of Utah, USA*.
2. *Teaching Mentor* for **3D Computer Vision**, Spring 2020, *Kahlert School of Computing, The University of Utah, USA*.
3. *Delivered a tutorial on PyTorch* in **Image Processing**, Fall 2023, *Kahlert School of Computing, The University of Utah, USA*.
4. *Delivered a lecture on Introduction to Statistics* in **Probability and Statistics for Engineers**, Spring 2024, *Kahlert School of Computing, The University of Utah, USA*.

REFERENCES

1. Dr. Ross Whitaker
Professor, *Kahlert School of Computing, The University of Utah, USA.*
E-mail: whitaker@cs.utah.edu

2. Dr. Sarang Joshi
Professor, *Kahlert School of Computing, The University of Utah, USA.*
E-mail: sjoshi@sci.utah.edu

3. Dr. Shireen Elhabian
Associate Professor, *Kahlert School of Computing, The University of Utah, USA.*
E-mail: shireen@sci.utah.edu

4. Dr. Huseyin Denli
Research Engineer, Meta, USA.
E-mail: huseyindenli@meta.com