Department of Biomedical Informatics
Stony Brook University, New York, USA
(+1) 631-590-0604
saarthak.kapse@stonybrook.edu
My Webpage
Linkedin *Github*

SAARTHAK KAPSE

Currently, I am a PhD student at State University of New York, Stony Brook. My research work aims at designing AI-based systems for diagnosis and prognosis of cancer in digital pathology. I am actively working towards the direction of utilizing pathologists-driven concepts to enhance representation learning in this domain.

Research Interests

Machine Learning/Deep Learning (Self-supervised learning, Vision Transformers, Convolutional Neural Networks) for Medical imaging applications (computational pathology, multiple instance learning, radiology imaging, gigapixel image analysis, etc).

Education

2021 – present **Ph.D., Biomedical Informatics**, *Stony Brook University*, New York, USA.

2016 – 2020 **B.Tech., Electrical Engineering**, *Indian Institute of Technology*, Bombay, India.

Publications

2022 **Saarthak Kapse**, Luke Torre-Healy, Richard A Moffitt, Rajarsi Gupta, and Prateek Prasanna. Subtype-specific spatial descriptors of tumor-immune microenvironment are prognostic of survival in lung adenocarcinoma. In *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2022.

2022 **Saarthak Kapse**, Rajarsi Gupta, and Prateek Prasanna. Shape-based tumor microenvironment analysis to differentiate non-small cell lung cancer subtypes: a radio-pathomic study. In *Medical Imaging 2022: Digital and Computational Pathology*. SPIE, 2022.

2022 **Saarthak Kapse**, Srijan Das, Jingwei Zhang, Rajarsi R. Gupta, Joel Saltz, Dimitris Samaras, and Prateek Prasanna. Attention de-sparsification matters: Inducing diversity in digital pathology representation learning. *under Review at ICLR 2023 (OpenReview)*, 2022.

2022 **Saarthak Kapse**, Srijan Das, and Prateek Prasanna. Cd-net: Histopathology representation learning using pyramidal context-detail network. *under Review at ISBI 2023 (arxiv)*, 2022.

2021 Fan Wang, **Saarthak Kapse**, Steven Liu, Prateek Prasanna, and Chao Chen. Topotxr: A topological biomarker for predicting treatment response in breast cancer. In *International Conference on Information Processing in Medical Imaging*. Springer, 2021.

2021 **Saarthak Kapse***, Joseph Bae*, Gagandeep Singh, Rishabh Gattu, Syed Ali, Neal Shah, Colin Marshall, Jonathan Pierce, Tej Phatak, Amit Gupta, Jeremy Green, Nikhil Madan, and Prateek Prasanna (* equal contribution). Predicting mechanical ventilation and mortality in covid-19 using radiomics and deep learning on chest radiographs: A multi-institutional study. *Diagnostics*. MDPI, 2021.

Research Experience

Stony Brook University, New York, USA

Mar 2020 – present

Attention De-sparsification Matters: Inducing Diversity in Digital Pathology Representation Learning.

Demonstrated the requirement of tailoring self-supervised learning techniques (SSL) for digital pathology based on insightful observation about sparsity of attention. Proposed a framework *DiRL* to densely encode pathology characteristics and uses them in a dense matching objective for prior-guided pre-training based on SSL. Qualitatively we showed that *DiRL* de-sparsifies the attention map, while quantitatively showing consistent performance imporvement on multiple slide-level and patch-level classification tasks.

CD-Net: Histopathology Representation Learning using Context-Detail Network.

Developed an efficient architecture to jointly encode context and detail information through inputs from lower and higher magnification in whole slide imaging. Employed separate branch for each magnification, with residual connection to enable the communication between both input branches. CD-Net models larger field of view with context module, and concurrently the fine-grained details through detail module. Demonstrated its efficacy in classifying Lung Adenocarcinoma from Squamous cell carcinoma.

Subtype-specific spatial descriptors of tumor-immune microenvironment are prognostic of Survival in Lung Adenocarcinoma.

Implemented HoverNet to segment and classify cells into tumor and immune cells. We then characterized the tumor-immune microenvironment through two types of handcrafted features - graph-based and spatial nuclei heterogeneity based representations. This analysis was further augmented through integration of phenotypic information of the tumor.

Shape-based Tumor Microenvironment Analysis to Differentiate Non-Small Cell Lung Cancer Subtypes: a Radio-Pathomic Study.

Demonstrated that tumor-shape based features from radiology and pathology modalities are complementary to each other. The integrated representation was used to differentiate lung cancer subtypes. Explored tumor heterogeneity through clustering the tissue slide sub-regions followed by extracting cluster based features.

Outcome prediction in COVID-19 using chest radiographs.

Implemented Residual U-Net model for segmenting Lung and Image artifacts from Chest X-rays. From the lung region, we extracted Radiomics-based and Deep Learning–based representations and integrated both for Ventilation requirement and Mortality prediction in COVID-19 patients.

Automated Fazekas Scoring on Head CT Scans.

Implemented automatic brain extraction pipeline using Nipype library, followed by three separate slice-based binary classifiers using ResNet architecture. Aggregated slice-level to patient-level decisions using probabilistic majority voting.

Advisor : **Dr. Prateek Prasanna**, *Assistant Professor, Department of Biomedical Informatics*, Stony Brook University (*Web-page*)

Indian Institute of Technology, Bombay, India

Jan 2020 – Dec 2020

Gene Mutation Prediction.

Developed model for mutation prediction from non-small cell lung cancer (NSCLC) histopathology slides using convolutional neural networks. Implemented multiple instance learning to optimize the neural network using weak slides level labels.

Advisor : **Dr. Amit Sethi**, *Professor, Department of Electrical Engineering*, IIT Bombay (*Web-page*)

Philips Innovation Campus, Bangalore, India

May 2019 – Jul 2019

Liver Lesion Segmentation using Volumetric Approach.

Developed a volumetric medical image segmentation pipeline of Liver Lesions from CT Scans. Enhanced the segmentation mask prediction through post-processing using radiomics and ML Classifiers.

Advisor : **Dr. M.S Dinesh**, *Sr. Scientist II*, Philips

Computer skills

Programming Languages	Python, C++, Assembly-Language
ML Frameworks	PyTorch, Keras, Tensorflow
Software	MATLAB, Google Cloud, GitHub, Openslide

Position of Responsibility

2022 – present	Reviewer at AAAI, ISBI.
2021 – present	BMI department website team member, Stony Brook Univerisity
2022	Teaching Assistantship BMI 520: Data Analytics and Software Stacks, Stony Brook University