Research Statement

Nagur Shareef Shaik, shaiknagurshareef6@gmail.com
Master of Science in Computer Science, Georgia State University, Atlanta

I am deeply interested in Machine Learning research, with a specific focus on addressing challenges in computer vision and natural language processing. My passion extends to exploring interdisciplinary applications in biomedical sciences, where these challenges not only hold practical value but also stimulate the discovery of new problems. My research is driven by the goal of developing intelligent systems that excel in learning meaningful representations across diverse data modalities, including images and text. showcasing a robust ability to bridge theoretical exploration with practical applications. Initially, my research focused on extracting discriminative representations from various medical imaging modalities including retinal fundus, brain MRI, and chest CT scans. The goal was to build robust diagnostic systems for retinopathy severity classification, brain tumor recognition, and pulmonary infection detection.

Recent successes in machine learning can be attributed to the use of representations learned deep neural networks, spanning text embeddings from LSTMs or Transformers to image representations via CNNs. Nevertheless, traditional approaches encounter challenges in learning embeddings that capture diverse semantics or contexts from input data in medical domain.

- Diabetic Retinopathy Severity Classification: We initially designed a feature aggregation framework integrating CNN-learned features at various levels, creating a discriminative representation with low, mid, and high-level features from retinal images. Despite outperforming traditional CNN features, this approach had limitations in focusing on lesion-specific regions. To address this, the subsequent work introduced a composite deep neural network with gated attention, excelling in capturing lesion-specific features. However, handling higher-dimensional attention features posed a challenge. In the following phase, we introduced a joint neural network with a convolutional auto-encoder and a discriminator to compress these high-dimensional attention features. Additionally, a Neural Support Vector Machine was designed — a hybrid model combining the strengths of neural networks for representation learning and SVM for robust classification. All these efforts resulted in a 9% improvement in diagnostic accuracy for retinopathy severity grading when in comparision with existing reserach.
- Brain Tumor Recognition: Initial approach involved a joint model capable of learning distinct features from brain MR images, facilitating the identification of brain tumor types. The key strength lies in the model's ability to assimilate diverse features from two different channels. An extension to this work introduced Multi-modal Squeeze and Excitation based attention, enhancing the focus on tumor-specific regions across the features. Seeking further performance enhancement, a subsequent work resulted in multi-level attention network featuring various types of visual attentions. This design aimed to capture spatial and channel-specific attention scores in relation to tumored regions, resulting in a 6% improvement in precision.
- COVID-19 Detection: We introduced a novel ensembling approach centered on prediction confidence values. This innovative method aimed to construct a robust model capable of learning to rectify the errors inherent in the max-voting ensemble approach.

All above discussed approaches were designed to deal with limited annotated data, handling a crucial challenge in medical domain. Ongoing research is primarily focused on extending these approaches 3D medical imaging modalities such as 3D MRIs (sMRI & fMRIs) to develop intelligent diagnostic systems. I engaged in a real-time research project for diagnosing Schizophrenia using 3D sMRI images, despite having very limited training samples. My main contribution lies in the design of a unique attention mechanism called Spatial Sequence Attention, resulting in the highest diagnosis accuracy. Currently, my efforts are directed towards developing an intraoperative intelligent system for assisting surgeons in breast cancer surgery. A notable achievement in this project is the implementation of guided contextual gating attention, enhancing the interpretability of medical images. In a challenging intersection of NLP and CV, my thesis project centers on generating meaningful clinical descriptions from multi-modal retinal images. The key challenge lies in the careful integration of clinical keywords and visual features through an attention mechanism. To address this, I implemented a context-gating transformer network, focusing on lesion-specific regions. This approach establishes a vital correlation between expert-defined keyword embeddings and visual features, enhancing the precision of clinical descriptions during the generation process. Significant milestones in these research projects include the publication of methods and results in reputable journals (14) & prestigious conferences (3) including International Symposium on Biomedical Imaging (ISBI) and the International Conference on Image Processing (ICIP).

Research Publications

[1] T. K. Cherukuri, N. S. Shaik, and D. H. Ye, "Guided context gating: Learning to leverage salient lesions in retinal fundus images," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, Paper ID: 1605, 2024.

[2] N. S. Shaik and T. K. Cherukuri, "Gated contextual transformer network for multi-modal retinal image clinical description generation," *Image and Vision Computing*, vol. 143, no. C, 2024.

[3] N. S. Shaik, T. K. Cherukuri, V. Calhoun, and D. H. Ye, "Spatial sequence attention network for schizophrenia classification from structural brain mr images," in *Proceedings of the 21st IEEE International Symposium on Biomedical Imaging*, IEEE, 2024.

[4] N. S. Shaik, T. K. Cherukuri, N. Veeranjaneulu, and J. D. Bodapati, "Medtransnet: Advanced gating transformer network for medical image classification," *Machine Vision and Applications*, vol. 35, no. 4, p. 73, 2024.

[5] N. S. Shaik, T. K. Cherukuri, and D. H. Ye, "M3t: Multi-modal medical transformer to bridge clinical context with visual insights for retinal image medical description generation," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, Paper ID: 1604, 2024.

[6] N. S. Shaik and T. K. Cherukuri, "Visual attention based composite dense neural network for facial expression recognition," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 12, pp. 16 229–16 242, 2023.

[7] N. S. Shaik and T. K. Cherukuri, "Hinge attention network: A joint model for diabetic retinopathy severity grading," *Applied Intelligence*, vol. 52, no. 13, pp. 15 105–15 121, 2022.

[8] N. S. Shaik and T. K. Cherukuri, "Multi-level attention network: Application to brain tumor classification," *Signal, Image and Video Processing*, vol. 16, no. 3, pp. 817–824, 2022.

[9] N. S. Shaik and T. K. Cherukuri, "Transfer learning based novel ensemble classifier for covid-19 detection from chest ct-scans," *Computers in Biology and Medicine*, vol. 141, p. 105 127, 2022.

[10] J. D. Bodapati, N. S. Shaik, and V. Naralasetti, "Composite deep neural network with gated-attention mechanism for diabetic retinopathy severity classification," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 10, pp. 9825–9839, 2021.

[11] J. D. Bodapati, N. S. Shaik, and V. Naralasetti, "Deep convolution feature aggregation: An application to diabetic retinopathy severity level prediction," *Signal, Image and Video Processing*, vol. 15, pp. 923–930, 2021.

[12] J. D. Bodapati, N. S. Shaik, V. Naralasetti, and N. B. Mundukur, "Joint training of two-channel deep neural network for brain tumor classification," *Signal, Image and Video Processing*, vol. 15, no. 4, pp. 753–760, 2021.

[13] J. D. Bodapati, S. N. Shareef, V. Naralasetti, and N. B. Mundukur, "Msenet: Multi-modal squeeze-and-excitation network for brain tumor severity prediction," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 35, no. 07, p. 2 157 005, 2021.

[14] N. S. Shaik and T. K. Cherukuri, "Lesion-aware attention with neural support vector machine for retinopathy diagnosis," *Machine Vision and Applications*, vol. 32, no. 6, p. 126, 2021.

[15] J. D. Bodapati, V. Naralasetti, S. N. Shareef, S. Hakak, M. Bilal, P. K. R. Maddikunta, and O. Jo, "Blended multi-modal deep convnet features for diabetic retinopathy severity prediction," *Electronics*, vol. 9, no. 6, p. 914, 2020.

[16] V. Dondeti, J. D. Bodapati, S. N. Shareef, and N. Veeranjaneyulu, "Deep convolution features in non-linear embedding space for fundus image classification.," *Rev. d'Intelligence Artif.*, vol. 34, no. 3, pp. 307–313, 2020.

[17] J. D. Bodapati, N. Veeranjaneyulu, and S. N. Shareef, "Sentiment analysis from movie reviews using lstms.," *Ingénierie des Systèmes d Inf.*, vol. 24, no. 1, pp. 125–129, 2019.