Sidong Zhang

sidongzhang@umass.edu linkedin.com/in/sidong-zhang-b25502143 github.com/stonezhng

Education

UMass Amherst's College of Information and Computer Sciences **Amherst, MA, USA**
Doctor of Philosophy *Sep. 2020 - present*

- Teaching assistant of graduate level CS 589 machine learning and CS 651 optimization class
- Research assistant in Information Fusion lab, currently funded by an NIH RO3

UMass Amherst's College of Information and Computer Sciences **Amherst, MA, USA**
Master of Science *Sep. 2018 - Jan. 2021*

- GPA: 3.82/4.0

Nanjing University, Software Institute **Nanjing, China**
Bachelor of Engineering *Sep. 2014 - Jun. 2018*

- GPA: 4.35/5.0

Project

Longitudinal Multimodal Modeling for Alzheimer's Early Detection in the Wild **Amherst, MA, USA**
Main contributor *Sep. 2020 - present*

- We work on datasets from ANDI and UK Biobank
- We preprocessed MRI for latent representation features via a mutual information maximizing encoder
- We designed training pipelines that improve validation stability and interpretability
- Our current model forecasts Alzheimer's 24 months in advance with an F1 score of 0.821 on a 3-label task
- We are experimenting multimodal models fusing time series patient records with brain image features
- We are introducing transfer learning and meta learning to build adaptive model working on multi-source datasets
- We are researching on time series models that can precisely capture the disease transition time
- Current result was presented in UMASS Initiative on Neurosciences Multidisciplinary Poster Conference for Neuroscience, Computer Science, and Engineering (Nov. 2021)
- Current result was presented in the poster session in New England Computer Vision Workshop (Dec. 2022)
- Current result is in submission to Journal of Neuroscience

FuseBox: Stage-Adjustable Multimodal Fusion **Amherst, MA, USA**
Main contributor *Apr. 2023 - present*

- The project originated from the observation that the choice of fusion point in multimodal research matters when utilizing unimodal pretrained models, i.e. early fusion, late fusion or intermediate layers fusion
- We work on a universal and model-agnostic approach to automatically select layers from unimodal pretrained models and compose informative fusion representations
- We selected layers based on an approximate maximization of conditional mutual information across modalities
- We are researching on other submodular or approximately submodular functions as approximate maximization target on selecting layers from unimodal pretrained pipelines

Multimodal Fusion for Multimedia Analysis **Amherst, MA, USA**
Collaborator *Jun. 2023 - present*

- This project is in collaboration with Dolby on multimodal speech separation
- We worked on a multimodal dataset of video and audio, The Grid Audio-Visual Speech Corpus
- We added a frequency based loss to identify male and female speakers besides the traditional time domain loss
- The new loss brings an improvement of 3.25% on SDRi

HyperFuse: Multimodal Fusion via Hypernetwork **Amherst, MA, USA**
Collaborator *Sep. 2022 - present*

- We designed a hypernet structure to add multimodal forward passing support to existing neural networks models consisting of linear layers and convolutional layers
- The HyperFuse structure gained performance improvements on CMU-MOSEI, AV-MNIST and M3A Financial data
- Current result is in submission to CVPR 2024

Information Bottleneck algorithm application on brain MR images **Amherst, MA, USA**
Main contributor *Oct. 2019 – Apr. 2020*

- The project worked on ANDI brain MR images
- The project experimented mutual information lower bound approach in information bottleneck algorithm with variational mutual information upper bound
- The result was reported as my Master Thesis

Clustered Vertical Attention for Irregular Time Series Modelling **Amherst, MA, USA**

Main contributor *Feb. 2019 - May. 2019*

- We worked on PhysioNet Challenge 2012 data set
- We improved the prediction accuracy of the in-hospital survival of ICU patients via existing imputation methods
- We ran Minimum Spanning Tree algorithm to determine the most correlated imputed clusters
- We trained separate Attention models for clusters and predicted on a Long Short-term Memory Attention model
- The model got accuracy improvements of 1.5%, 1.3% and 1.8% on 3 different imputation methods
- Results were submitted to ICML 2019 TimeSeriesWorkshop

Technical Skills

Languages: English (Proficient), Mandarin (Native)
Programming language: Java, Python, C, Lisp, Markdown, Latex