VISHNU SASHANK DORBALA

◇ **Website**: https://vdorbala.github.io ◇ **Phone**: (+1) 202-322-2098 ◇ **Email**: vdorbala@umd.edu

EDUCATION

University of Maryland, College Park *July 2021 - Present*
Doctor of Philosophy (PhD), Computer Science

University of Maryland, College Park *July 2019 - May 2021*
Masters of Engineering, Robotics

Symbiosis International University, Pune *June 2013 - July 2017*
Bachelor's of Engineering, Electronics and Telecommunication

RESEARCH INTERESTS

Generative AI, Embodied AI, Large Language Models, Human-Robot Interaction, Social Robotics

WORK EXPERIENCE

GAMMA Lab, University of Maryland January 2020 - Present
Advised by ***Prof. Dinesh Manocha***

Generative AI for Navigation

- Conducting research tackling *embodied* problems in generalized, zero-shot settings.
- Developed a novel zero-shot Object Navigation scheme that queries Large Language Models (LLMs) for commonsense knowledge to explore unseen environments (Accepted at RA-L 2023).
- Currently exploring the use of LLMs with *long-term memory* to learn user habits for efficiently performing *portable* object navigation in a dynamic environment.

Social Robot Navigation in Human-Centric Environments

- Conceptualized and developed a novel *Reinforcement Learning* based robot navigation scheme, that utilizes ***affective* cues** to gauge *trust* from human interactions (Published at ICRA, 2021).

Amazon Alexa AI May 2023 - Sept 2023
Applied Scientist Intern under ***Dr. Reza Ghanadan***

Large Language Models for Generating Subjective Queries

- Proposed and tackled a challenging *Embodied Question Answering* problem with subjective queries. Our work released a useful human-verified LLM-generated dataset and established a quantitative **benchmark** for EQA with subjective queries. (Under Review at ICRA, 2024.)

Amazon Alexa AI May 2022 - Sept 2022
Applied Scientist Intern under ***Prof. Gaurav Sukhatme***

CLIP-Nav: Using CLIP for Zero-Shot Vision-and-Language Navigation

- Tackled a challenging **vision-and-language navigation** problem called REVERIE. Our results show improved generalizability over dominant supervised learning approaches, and established a **zero-shot baseline**. (Published at CoRL LangRob Workshop, 2022.)

Nokia Bell Labs June 2020 - December 2020
Robotics Research Intern/Co-op

Robust Navigation on an Indoor Robot

- Implemented indoor **Visual SLAM** on a Loomo robot, in a complex office environment.

Center for Visual Information Technology, IIIT, Hyderabad Sept 2017 - May 2019
Research Fellow under ***Prof. C.V. Jawahar***

Visual Navigation for an Assistive Robot

- Worked on developing SARA, an assistive ***robot wheelchair*** platform, capable of interacting with people.
- Performed ***autonomous corridor following*** on SARA via **Visual Servoing** (Published at IROS, 2019).

PUBLICATIONS

- ***V.S. Dorbala***, et. al, "Right Place, Right Time! Towards ObjectNav for Non-Stationary Goals", *Under Review* at Double Blind Computer Vision Conference, 2024 *(Link to Paper)*
- ***V.S. Dorbala***, et al., "S-EQA: Tackling Situational Queries in Embodied Question Answering", *Under Review* at the IEEE Int. Conf. on Intelligent Robots and Systems (**IROS**), 2024 *(Link to Paper, Video)*
- ***V.S. Dorbala***, et al., "Can LLM's Generate Human-Like Wayfinding Instructions? Towards Platform-Agnostic Embodied Instruction Synthesis", *Accepted* at the Annual Conf. of the North American Chapter of the Association for Computational Linguistics (**NAACL**), 2023 *(Link to Paper)*
- ***V.S. Dorbala***, et al., "Can an Embodied Agent Find Your "Cat-shaped Mug"? LLM-based Zero-Shot Object Navigation.", IEEE Robots and Automation Letters (**RA-L**), 2023 *(Link to Paper)*
- ***V.S. Dorbala***, et al., "*CLIP-Nav: Using CLIP for Zero-Shot Visual-Language Navigation*" Conference on Robot Learning (**CoRL**), 2022 *(Link to Paper)*
- ***V.S. Dorbala***, et al., "*Can a Robot Trust You? A DRL-Based Approach to Trust-Driven Human-Guided Navigation*" IEEE Int. Conf. on Robotics and Automation (**ICRA**), 2021 *(Link to Paper, Video)*
- *V. Narayanan*, ***V.S. Dorbala***, et al., "*ProxEmo: Gait-based Emotion Learning and Multi-view Proxemic Fusion for Socially-Aware Robot Navigation*" (**IROS**), 2020 *(Link to Paper, Video)*
- ***V.S. Dorbala***, et al., "*A Deep Learning Approach for Robust Corridor Following*" IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (**IROS**), 2019 *(Link to Paper, Video)*

SELECTED PROJECTS

- ***Do Face Generation Models Generate Good Faces?*** Oct 2023
  - Used GPT-4 to generate queries that *indirectly* ask a person about their facial appearance.
  - Interviewed 10 participants about in a '*Wizard-of-Oz*' setup, using the generated questions. Prompted DALLE with the responses to generate faces.
  - **Inference *(Report)***: Observed unusual annotations, a beard bias and exaggerated facial markings.
- ***Did Subtitles Ruin Your Movie?*** Nov 2022
  - Used Neural Machine Translation (NMT) to translate movie subtitles of 58 movies from French to English and vice versa.
  - Defined a Key Scenes of Disagreement (KSD) metric across the bilingual movie scripts and used it for correlation with movie ratings. Performed a human evaluation with the machine translated and original subtitles.
  - **Inference *(Report)*** - In some crucial cases, the NMT subtitles convey scene sentiments far better than human-written ones.
- ***Interpretability of Video Classifiers*** May 2022
  - Tackled the challenging problem of interpreting videos (finding keyframes), employing saliency techniques.
  - Compared an LSTM for detecting saliency with and without input cell attention on the KTH dataset.
  - **Inference (*Report*)**: As there is no quantitative benchmark, it proved hard to objectively validate saliency. However, a correlation of temporal with manually chosen frames was observed.
- ***Classifying Glucose Levels with EEG Data*** Dec 2021
  - Used a Muse headband to capture EEG brain signals non-intrusively for approximating glucose levels.
  - Utilized EEGLab to perform dimensionality reduction via Independent Component Analysis (ICA) and then trained a CNN classifier on the collected brain-wave data.
  - **Inference (*Report*)**: As the data was incredibly noisy given the sheer temporal volume (collected in milliseconds), training a CNN based classifier did not give favourable results. Classical hand-crafted feature extraction proved superior.

ACHIEVEMENTS

- Co-Chaired at a conference session on Human-Robot Interaction: Robot Navigation at ICRA 2021.

- Peer reviewer for ICRA 2019-2022, IROS 2018-2021 and EMNLP 2023.
- Received the Dean's fellowship for PhD students awarded by the University of Maryland, College Park.