Action-Aware Vision Language Navigation: A Novel Reinforcement Learning Framework for Dynamic Navigation in VR and Beyond

Liu, Jasmine (School: Shanghai American School - Pudong Campus)

Visually impaired individuals have dif?culty understanding their environment and the actions of others due to their limited visual abilities. Previous research has shown that approaches used in robotics and augmented reality can provide navigation assistance, but these approaches struggle to understand dynamic environments. Existing approaches such as SLAM-based navigation and VLN have limitations in understanding dynamic environments. In order to assist visually impaired individuals in navigating to their destinations, we propose a cross-modal transformer-based action-aware VLN system that understands natural language instructions and dynamic environments, including human actions, to aid navigation. Our project innovatively proposes a vision and language-based navigation framework that includes an Agent with scene navigation and action recognition algorithms, and a Simulator with human action. To train the Navigation Agent and Simulator, we will use reinforcement learning with a dynamic environment simulator that includes virtual human ?gures and their actions. Our framework is capable of: 1) using the cross-modal transformer structure to understand and reason about events and actions using vision, text, and audio, 2) navigating in environments with human actions, and 3) understanding natural language instructions. Experimental results demonstrate our framework outperforms other methods, and its benefits are evident in VLN navigation. Our ablation study demonstrates the bene?ts of leveraging visual and language modalities to understand human-like events. Our project has potential applications in ?elds such as robotics, augmented reality, and human-computer interactions, and could help advance the ?eld of arti?cial intelligence in the area of vision understanding.

Awards Won:

Fourth Award of $500

Association for Computing Machinery: Second Award of $3,000