Exhibit D

Research Project Requirement Template

Exploring Top-Down Visual Attention for Transportation Behavior Analysis

Recipient/Grant (Contract) Number: The University of Texas at Austin/Grant # 69A3552344815 and 69A3552348320

Center Name: Center for Understanding Future Travel Behavior and Demand (TBD)

Research Priority: Improving Mobility of People and Goods

Principal Investigator(s): Zhigang Zhu

Project Partners: N/A

Research Project Funding: $245,046 (Federal + non-Federal funding)

Project Start and End Date: 6/1/2024 - 5/31/2025

Project Description: This project stands at the intersection of cognitive psychology, AI and computer vision, and transportation safety and efficiency. By focusing on the nuanced ways in which humans allocate their visual attention, and how this can inform the development of artificial intelligence (AI) and machine learning (ML) to aid in self-driving cars, transportation safety automation, and transportation planning and scheduling in general, this project promises to contribute significantly to the field, ensuring safer, more intuitive driving experiences, and smoother traveling experiences for the traveling public.

By performing human behavior analysis with visual attention, we aim to develop best practices for safe and efficient interaction of automated roadway vehicles with existing vehicles, roadside hardware, pedestrians, cyclists, and motorcyclists. The advantages of a top-down attention approach include: prioritizing relevance, improving accuracy, enhancing machine learning efficiency, adapting models to scenarios, and enabling better human interaction. By exploring top-down visual attention, we aim to build machine learning models to achieve the following objectives that are coherently connected with each other, where the first two will be the objectives in the base phase of this proposal and the last two would be in a second phase for a follow-on effort:

(1) Develop human behavior analysis machine learning architectures that allow autonomous driving and other transportation systems to anticipate the attention and reaction patterns of both human drivers and pedestrians, thereby preventing accidents. These include the human behavior analysis of the interaction between a driver and their vehicle, driver and pedestrians, humans with the existing vehicles and roadside hardware. The ML architectures explored will be CNN for image encoding for improving accuracy and reducing computation, GCN for relation reasoning focusing on human interaction and actions, and transformers for self-attention and feedback.

(2) Investigate the potential of using visual attention models to improve autonomous and/or automated vehicle navigation and decision-making processes in complex environments. The visual attention mechanisms will be driven by both data and knowledge, including dynamic transportation information, roadside hardware information, location-based information (maps, events, tasks). As a start point, we will leverage the state-of-the-art (SOTA) model such as *Analysis-by-Synthesis Vision Transformer (AbSViT)* to encode feature selection, higher-level feedback and top-down input, added on the typical bottom-up process in deep models.

(3) Develop multimodal human-machine interface dashboards in self-driving cars and vehicle safety automation system, making them more intuitive for human users. These include audio, visual and haptic features as well as accessibility functions that the team has studied for helping the navigation of people who

are blind or have low vision. Supported by the AI/ML-based architectures and attention models, the interface as dashboards will also allow developers, engineers and users to access the intelligent transportation systems for interaction, interpretation and diagnosis.

(4) Furthermore, collaborative opportunities may arise with existing projects, especially in applying the findings to enhance the travel pattern analysis and other safety features of self-driving and/or existing vehicles and pedestrians. Collaboration could involve sharing data, methodologies, and insights to refine autonomous driving technologies' perception and decision-making capabilities.

US DOT Priorities: This proposal addresses the following aspects of the transformation research priorities: 1. Integrated System of Systems (p. 50) – Attributes of a Desired Transportation System of the Future (People-Centered, Data Driven and Intelligent) (p. 52); Transportation Grand Challenge (Automation, AI and Machine Learning) (p. 54, p. 57); 2. Data-driven Insight – Data Science (p. 50), especially to conduct exploratory research on transformational mobility data analytics (p. 59); 3. New and Novel Technologies – Automation (p. 50), especially to develop best practices for safe interaction of automated roadway vehicles with existing vehicles, roadside hardware, emergency responders, pedestrians, cyclists, and motorcyclists. (p. 60). The technology transfer/deployment plan of this project including the following: 1. Evaluating the performance in real-world applications (p. 64), with the TBD Center (p. 68), and through the US DOT T2 program (Item 2) on p. 65); 2. Publish technical papers and guides (p. 64); 3. Present webinars and deliver presentations to stakeholders (p. 64); 4. Create websites (p. 64), by making the code and data available in GitHub; 5. Explore pilot and demonstration projects (p. 65), in collaborating with the CUNY spin-off Nearabl, Inc. (a location-based tech startup) (Item 3) on p. 65, Table 7 on p. 67).

Outputs: The proposed project will include *a basic phase effort of this proposal and a possible follow-on effort in a second phase proposal.* The outputs of *the first phase* include: (1) A comprehensive survey on the topic by one PhD student and the PI/Co-PI, and submit to a top journal, tentatively Computer Vision and Image Understanding; (2) A second PhD student work on data collection for driver behavior analysis; (3) A basic core attention model developed by the PhD students with PI/Co-PI on Analysis by Synthesis Using Visual Transformer for Driver Behavior Analysis. A GitHub page will be created and a paper to be submitted to the premier vision conference CVPR or transportation meeting TRB, while posting on arXiv. ***The second phase*** of a follow-on effort will be devoted to a general framework for human behavior analysis, multimodal dashboards, performance evaluation in real-world, and more publications and open access data and code.

Outcomes/Impacts: The comprehensive survey will provide insights to transportation researchers for the state-of-the-art AI/ML models and algorithms for human action and behavior analysis using computer vision. This will provide a foundation and increase in the body of knowledge for applied researchers in transportation for real-world problems. The collaboration of PI (AI/ML) and Co-PI (transportation) will ensure the relevance of the research to transportation to design new AI/ML attention models to increase the understanding and awareness of self-driving vehicles and transportation automation in relation to other entities (existing vehicles, pedestrians and hardware).

In summary, the approach will be a top-down approach that integrates transportation-domain knowledge to the data-driven AI models. The aforementioned advantages of a top-down attention approach will lead to significant impacts on transportation in transportation automation/safety and society in general: (1) improving efficiency by prioritizing relevance and enabling task-specific attention; (2) reducing fatalities by improving accuracy of detection and understanding; (3) decreasing capital or operation costs in computing facilities by enhancing machine learning efficiency especially with deep learning models; (4) promoting wide deployment by adapting models to various scenarios; and (5) improving demand management by enabling better human interaction.

Final Research Report: A URL link to the final report will be provided upon completion of the project.