

Scalable Dataset Labeling and Contrastive Alignment for Cross-Modal Inference and Data Generation with Multimodal Large Foundation Models

A Project fully funded by YÖK-ADP, 2025-2027

Project No: 50192

Fully funded M.S. and Ph.D. positions are available!

Project Description: Large Foundation Models (LFMs) such as OpenAI's GPT and Google's Gemini now have significantly advanced the multi-modal data processing capabilities of AI systems, enabling coherent inference across diverse data modalities. By processing text, images, audio, and other data types, these models possess the potential to revolutionize numerous sectors. The proposed project aims to develop LFMs with multi-modal data processing capabilities to establish next-generation national technologies.

An LFM model capable of integrating and processing various data types, including text, images, audio, depth, infrared, GPS, and lidar, and performing real-time inference, will be created. Leveraging its cross-modal inference and prediction abilities, this model will be introduced as a fully domestic/national technology in diverse fields such as aviation, autonomous systems, robotics, healthcare, and security. The original contribution of this project lies in the development of a multi-tower architecture capable of simultaneously processing different data types. While current AI models typically focus on single or dual (e.g., language + vision) data modalities, this project will provide an innovative solution by offering multi-modal parallel data processing capabilities. Furthermore, the cost functions to be employed will incorporate several novel aspects alongside the increasing number and types of modalities. The project also includes the establishment of local HPC farms and storage system for domestic high-performance AI applications, facilitated by advanced GPU-based infrastructures.

The project will consist of three main phases: (1) data collection and labeling, (2) innovative techniques for the alignment of different data types into a common representation, and (3) development of cross-modal inference and future prediction capabilities for specific application fields (e.g. autonomous driving). These processes will utilize multi-layered modern deep learning algorithms and GPU acceleration techniques (through both hardware architecture – NVMe etc. – and innovative software approaches). The outcomes of the project will offer innovative and comprehensive solutions in autonomous vehicles, robotics, digital pathology, and security systems.

Our project will also elucidate, using concrete metrics, how the developed models will be impacted by adverse conditions specific to certain modalities. It will realize innovative ideas (perhaps from neuroscience) to prevent performance degradation under these conditions. By embedding these techniques into the architecture or integrating them into the training processes, the cross-modal data processing performance will be enhanced even in challenging conditions. The model's individual success across different modalities (such as language, computer vision, and audio) will be tested, and its modal tendencies and resulting model bias will be characterized. Our results are primarily intended to provide impressive accuracy performance and critical decision-making mechanisms in sectors where multi-modal data processing is essential, such as aviation and autonomous systems or human-level AI development (AGI). Fine-tuning mechanisms will be developed for the models obtained for these areas to enhance application-specific task completion performance.

Project Management:

1. Prof. Suayb S. Arslan (PI), Dept. of Comp. Engineering at Bogazici University and Research Affiliate, Brain and Cog. Sciences at MIT.
2. Assoc. Prof. Mehmet Turan (Co-PI), Dept. of Comp. Engineering, Bogazici University.
3. Asst. Prof. Mustafa Taha Kocyigit (Co-PI), Data Science and AI Institute, Bogazici University.
4. Dr. Yasin Almalıoğlu (Researcher), University of Oxford.

Applications and Compensations: Funding is available for M.S. and Ph.D. degrees.

1. Master of Science (M.S), 40,000TL + benefits
2. Philosophy of Doctorate (Ph.D), 50,000TL + benefits

Bonus: Opportunity to collaborate with a research group at MIT, with the possibility of a summer visit to the lab, contingent on the availability of external funding (e.g., TUBITAK, MIT MISTI, etc.). Guidance and help will be provided.

All applicants are encouraged to contact the PI and Co-PIs of the project should there be any interest to engage and contribute to this new exciting field. All questions about the logistics and official applications (A resume with previous project experience and publication record) should be directed to ***suayb.arslan@bogazici.edu.tr***. The available positions will be filled on a first-come, first-serve basis upon successful fulfilment of project requirements.