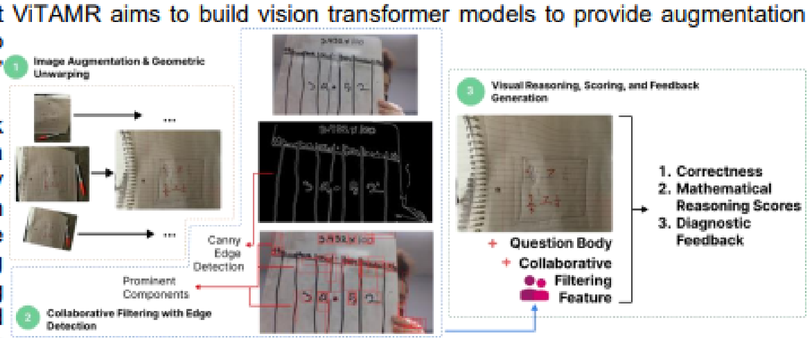
Project Summary

Overview: This project aims to support students' mathematical reasoning (MR) through the application of computer vision and diagnostic feedback. The development of students' abilities to reason mathematically is a central focus of theories and practices in mathematics education. Involving in MR would give students a powerful foundation for their understanding of many mathematical ideas, such as fractions and variables. Numerous studies have sought to investigate students' text and image artifacts. However, prior studies either manually examine students' learning artifacts, applied automatic analysis in impractical environment setups, or solely focused on text artifacts. Embedded in MathNet's multimodal data of students' artifacts in texts and images, our project ViTAMR aims to build vision transformer models to provide augmentation with diagnostic feedback to support students' mathematical reasoning.



Diagnostic feedback has been shown to be a valuable pedagogical strategy to invoke, assist, and train students' MR. Specifically, we will have a pipeline consisting of three components using computer vision (CV) and natural language processing (NLP) with deep neural networks: (1) an image augmentation and geometric unwarping component to synthesize training data (e.g., rotation, mirroring, blurring) and extract students' responses in the written paper while ignoring learning-irrelevant information; (2) a collaborative filtering system serving as a feature extractor to automatically group students based on their strategies (e.g., commonalities from image segmentations); and (3) a vision transformer that performs visual reasoning to determine if students' image artifacts match with problem text bodies and outputs scoring on students' MR with individualized feedback. To automatically score students' MR, we will utilize Amazon MTurk and the MR rubrics by Loong et al. (2018) to label 5,000 image responses. Findings from the project can be used to inform the design and development of automatic image grading systems for teachers and provide timely feedback on problem-solving for students.

Intellectual Merits: This project generates new understanding and insights into students' math learning strategies (e.g., through collaborative filtering) with hand-worked responses theorized with diagnostic feedback and mathematical reasoning (e.g., through visual reasoning). Second, this proposal creates a machine learning pipeline that utilizes multi-modal and multi-task learning to provide automatic diagnostic feedback on students' mathematical reasoning based on their hand-worked responses. To achieve the expected deliverable, we will extend current visual reasoning models by modifying their attention heads using shared- and self-attention to handle different sources of data (e.g., collaborative filtering features, texts, images). Our approach overcomes the previous challenges and the obstacles to accurately capturing and labeling the multimodal data in a pedagogically meaningful manner. The focus on using vision transformer models to perform visual reasoning is a cutting-edge application of computer vision to educational contexts. Finally, this project also demonstrates a methodological shift from the traditional predictive tasks with texts or images to both forms in a practical setting.

Broader Impacts: Our focus on mathematical reasoning and diagnostic feedback as a fundamental pedagogical framework to support mathematics education have important implications for developing transferable, scalable, and pedagogically meaningful tools for developing students' math competencies. First, we will open source the modeling process with detailed documentation and example codes to allow educational researchers to adopt or finetune our models. One of the biggest advantages of transformer models is their transferability, which democratizes the involved intellectual and computational resources. Second, this project can be adopted by math learning platforms such as ASSISTments to reduce teachers' workload on grading and provide students with automatic feedback. Further, our project outcome directly responds to the high-impact uses and capabilities to provide support for teachers to effectively and efficiently respond to students' mathematics work, and enhance prompt feedback and guidance provision to promote independent student learning. Lastly, we will publish our work with major conferences in learning analytics and artificial intelligence in education such as LAK, EDM, AIED, and L@S to document our techniques, marketize our results, and inspire further improvement.