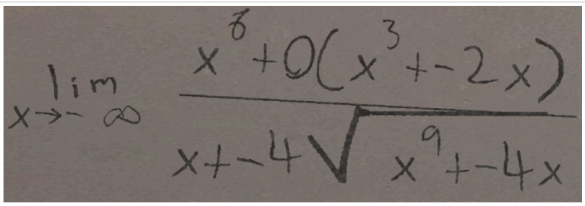
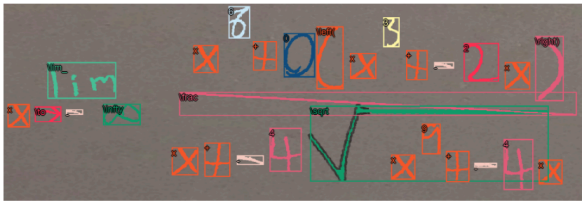
CSC671 Deep Learning Project

Team members:

Alpha Diallo, Miguel Logarta, Eric Kunzel, Issac Moreno

For our deep learning project, we will be developing an image classification model that can convert handwritten math equations into valid LaTeX syntax. The purpose of this project is to make it easier for users to convert their handwritten homework into digital paper. The benefits are that users do not have to concern themselves with learning LaTeX syntax to type out math equations. For experienced users, this model can be used as a time-saving tool for writing out equations on their papers.

For our training data, we'll use the **Aida Calculus Math Handwriting Recognition Dataset**: https://www.kaggle.com/datasets/aidapearson/ocr-data. This dataset contains around 100,000 synthetically generated images of handwritten calculus math expressions specifically focused on limits. It is split into 10 batches, each containing corresponding images, bounding boxes, and LaTeX expression data.

Training Data	Expected Output
 	LaTeX Expression `$\lim_{x\to-\infty}\frac{x^6+0(x^3+-2x)}{x+-4\sqrt{x^9+-4x}}$` PDF Output $\lim_{x\to-\infty}\frac{x^6+0(x^3+-2x)}{x+-4\sqrt{x^9+-4x}}$