Image to LaTeX

Chu Meng, Hu Songbo, Zhang Zhengyang

May 2025

1 Participants

- Chu Meng 2023010819
- Hu Songbo 2023010544
- Zhang Zhengyang 2024011339

2 Introduction

As students who study computer sciences, writing LaTeXis a common task. However, it is not easy to write LaTeXcode for beginners. The complex grammar and syntax of LaTeXcan be overwhelming, actually we felt the difficulty of become familiar with it when we were freshmen.

However, in some scholar cases, we need to write LaTeXcode for all kinds of tasks, which may be a barrier for those who are not familiar with it. In this project, we aim to solve this problem by using a neural network to convert images of formulas into LaTeXcode.

3 Pipeline

3.1 SFT

There are already some multimodel models that can convert images to text, such as Gemma by google and they are actually doing a good job in converting images to text. So, we can use supervised fine tuning to train a model that can convert images of formulas into LaTeXcode based on these multimodel models.

3.2 RL

Reinforcement learning is a powerful tool for training models to make decisions based on feedback from the environment. Actually, writing LaTeXcode is a decision-making process, where the model needs to decide which formula to write based on the former formulas. In this project, we can use reinforcement learning to train the model to write LaTeXcode based on the feedback from the environment. The environment can be a model that can evaluate the quality of the LaTeXcode generated by the model.