\Texify

Group 16

Hans Wu
Shukai Yin
12/7/2026

Our Project & Target Audience

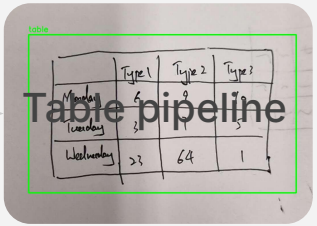
Our goal is to build a web app that converts handwritten math content, including symbols, equations, matrices, and tables, into LaTeX. The intended users are students, researchers, instructors, and professionals who frequently write formulas or sketches on paper or tablets and need high-quality LaTeX for homework, papers, or documentation.

Demo

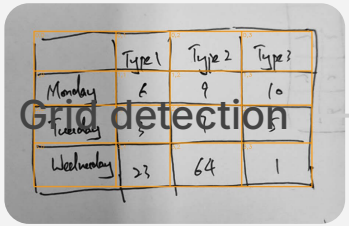
Main Pipeline

Content classifier

Table pipeline



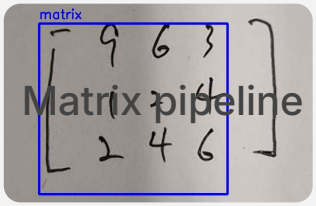
Grid detection



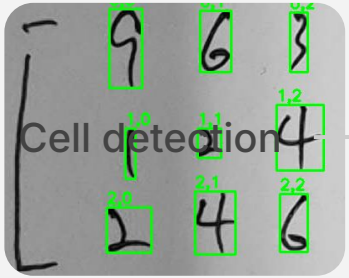
(base-stage1)

TrOCR → LaTeX Translator

Matrix pipeline



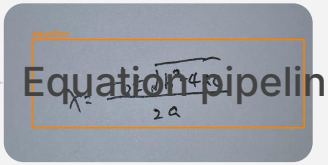
Cell detection



(base-handwritten)

TrOCR → LaTeX Translator

Equation pipeline



Fine-tuned pix2tex

Table Pipeline

1. Vertical / horizontal edge detection
2. Cell segmentation
3. Cell extraction
4. Microsoft TrOCR (base-handwritten)
5. LaTeX formatter

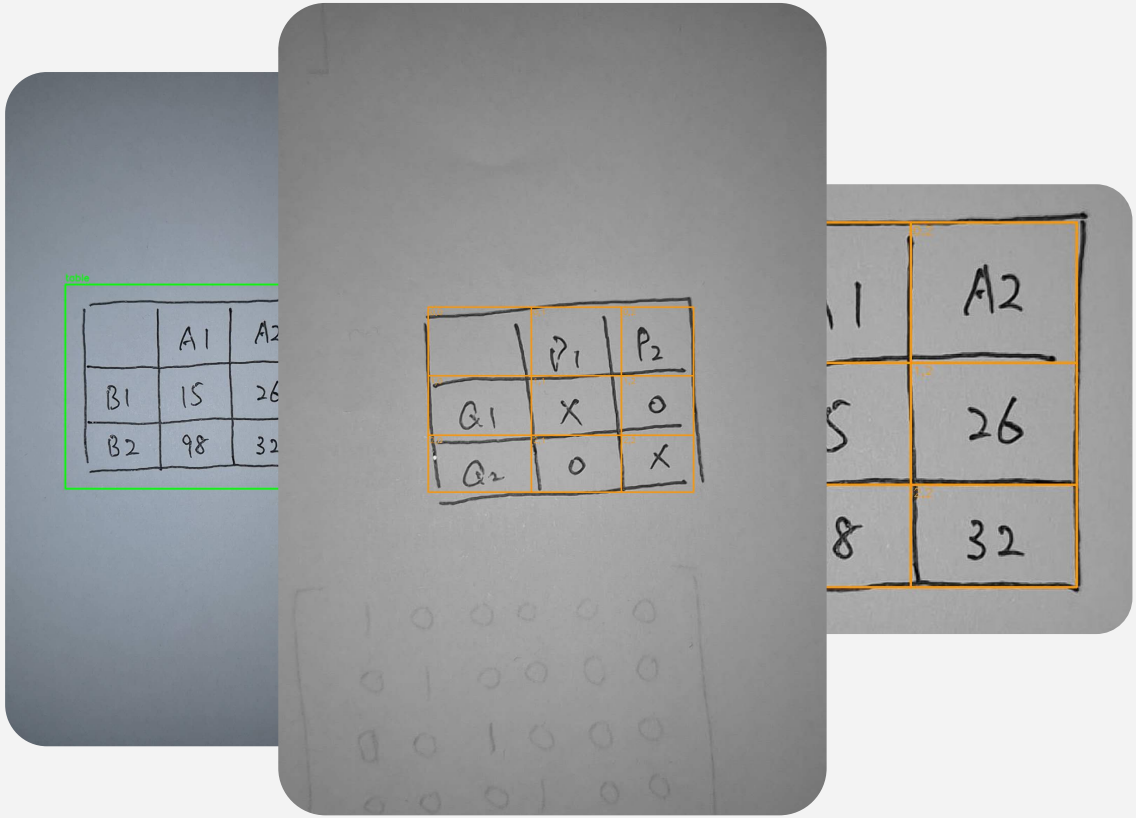
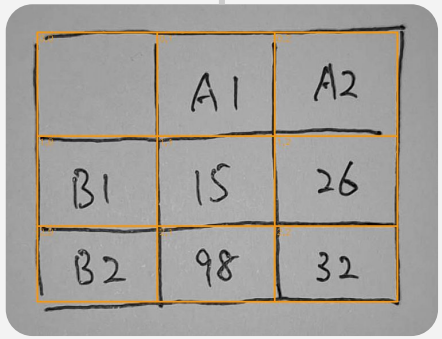
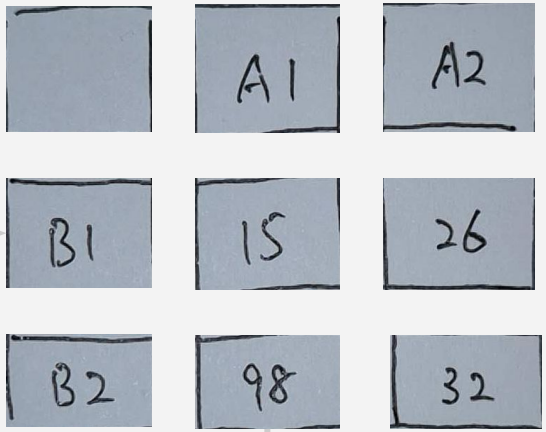


Table Pipeline

1. Vertical / horizontal edge detection
2. Cell segmentation
3. Cell extraction
4. Microsoft TrOCR (base-handwritten)
5. LaTeX formatter



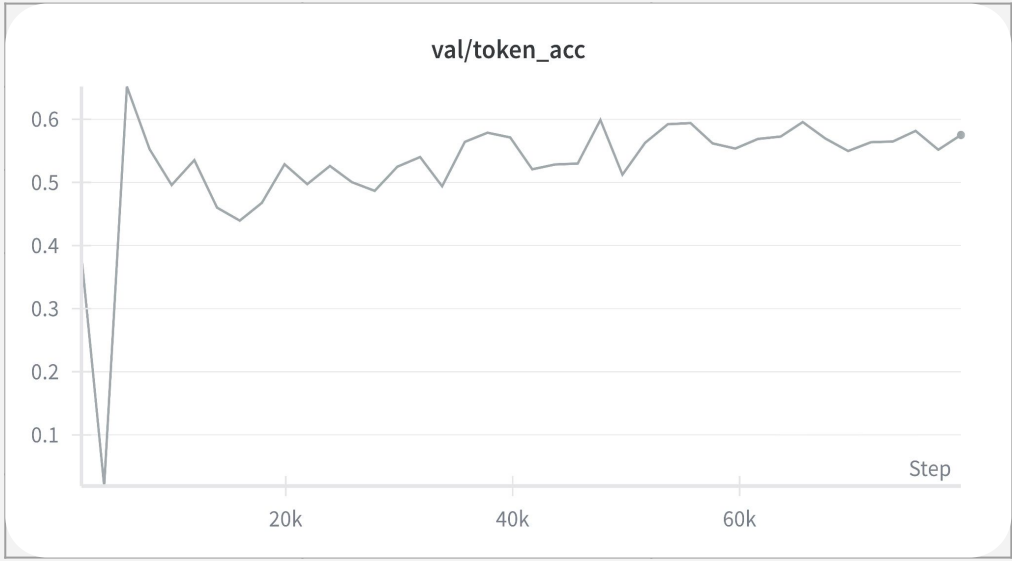
Matrix segmenter



Microsoft TrOCR
+
LaTeX Translator

```
\begin{tabular}{|c|c|c|}
\hline
- & A1 & A2 \\
\hline
BI & 15 & 26 \\
\hline
B2 & 98 & 32 \\
\hline
\end{tabular}
```

Equation Pipeline



Final Pipeline
CROHME
Fine-tuned pix2tex
• Preprocessing • ViT image encoder • Transformer decoder • LaTeX sequence generation
0.5752

\ Thank you 🫶

Questions?

Checkout our github repo:
https://github.com/Hans-Wu5/Texify

or email us at:
zihanw@seas.upenn.edu
sky1122@seas.upenn.edu