**Progress**

Document layout analysis: layout parser, Azure document intelligence

Data collection: contacted CUHK Digital Repository, received the data sets, special thanks to the editor, 魏鵬展

Test out image generation with LLM

**DLA**

Layout parser: [layout\_parser](https://www.kaggle.com/code/jeremyipks/layout-parser)



Azure document intelligence:

Result

A screenshot of a computer screen

Description automatically generated

Problems

1. Recognizing one modern verse as several segments
2. Recognized segment not aligning   
   A close up of a paper

   Description automatically generated

Will test custom models

TODO:

Try azure document intelligence custom models

Try other DLA tools, e.g. paddleOCR

Last resort: use manual laboring

**Testing image generation**

According to the methodology from the paper “Automatic Generation of Multimedia Teaching Materials Based on Generative AI: Taking Tang Poetry as an Example”, we conducted several image generations.

Main ideas: get keywords, make a prompt, generate image, evaluate by comparing the semantics similarity between the original poem and descriptive text from the generated image

<https://poe.com/s/hTjDoX5pL8rb5CumJBg3>

<https://poe.com/s/NnfHJOgrsjifkg7h4iRc>

TODO: understand the evaluation process specifics

TODO: formalize the process and make it automatic

TODO: maybe can use the evaluation to re-prompt