Contribution ID: **41**   Type: **Poster**

Automatic translation of Japanese manga using MultiModal Large Language Models

Thursday 24 April 2025 10:45 (30 minutes)

Japanese manga has captivated readers worldwide with its vibrant and expressive form of art that provides compelling storytelling with intricate visuals. However, for many fans outside of Japan, language barriers often stand in the way of fully experiencing the depth of these stories. Traditional translation from language to language, while effective, can be a very time-consuming and labor-intensive process that requires teams of translators, editors, and cultural consultants to convey the essence of the original text accurately. The difficulty with Japanese is even greater, as it is a heavy-context-intensive language.

Machine translation is nothing new; we try and fail with basic translation from language A to language B, and Google Translate is its finest example. Large language models (LLMs) and with its new type, multimodal LLMs, have undergone substantial advancements, augmenting already powerful LLMs to support multimodal inputs or outputs via cost-effective training strategies. Those showed potential in many works such as coding, answering complex math questions, or understanding symbolism within texts. In this paper, we conduct a substential investigation on how well certain MMLLMs work against manga translation. We create benchmarks with text-only-wise; image-from1page-context; as well as whole volume thus far.

Although the results were sometimes satisfactory, they proved to be insufficient to be a standalone automatic translator, as translations were not understandable and/or too complex or too simple at times. However they prove to provide enough understanding of context to, with some human element, come up with accurate translation

Author: NITSCH, Tomasz (Warsaw University of Technology)

Presenter: NITSCH, Tomasz (Warsaw University of Technology)

Session Classification: Session B (Poster)