先端科学技術研究科　修士論文要旨

所属研究室 (主指導教員)	自然言語処理学 (渡辺　太郎 (教授))		
学籍番号	2211411	提出日	令和 7年 1月 21日
学生氏名	HANG JIANGNAN		
論文題目	Enhancing Low-Resource Language Document Translation through Pipeline Approach in LLM		

要旨

With the growing demand for document translation, advancements in this task have yet to match the success of sentence-level translation. This study identifies the limitations of existing traditional strategies in modeling long tokens and discourse information, as well as the opacity of large language models (LLMs) strategies. To address these challenges, we propose an innovative multi-stage approach that combines traditional translation models with the language understanding capabilities of LLMs, focusing on their advantages in handling low-resource languages. Our method constructs paragraphs by concatenating Transformer-based sentence-level translations and designs subtasks across sentence and context layers to address specific grammatical phenomena. These subtasks are executed in a specific sequence of LLM inferences, forming a pipeline to enhance translation capabilities. Our approach demonstrates success across language pairs on the COMET metric, which emphasizes semantic accuracy. Furthermore, the explicit task inference pipeline based on in-context learning effectively leverages the internal capabilities of general models, providing a transparent and efficient new paradigm for document translation task.