助成対象研究の紹介文

多言語およびマルチモーダルな機械翻訳に関する研究

九州大学大学院システム情報科学研究院情報知能工学部門　助教　ZHAO YUTING

In recent years, with the advent of Transformer and high-quality image recognition models, many multimodal machine translation (MMT) models have been proposed and have shown advanced translation performance that exceeds conventional MT models that use only sentences in the source language. However, the practicability and efficiency of the MMT models are still limited to the weak correlation between text and image. This project will model a novel architecture of MMT to enforce the correlation between text and image concepts (words and objects) through learning a cross-modal shared embedding space on both semantic and spatial levels. The motivation behind this is to innovate previous MMT methods, which represent textual and visual features in different space independently, to represent multimodal information in a shared embedding space.
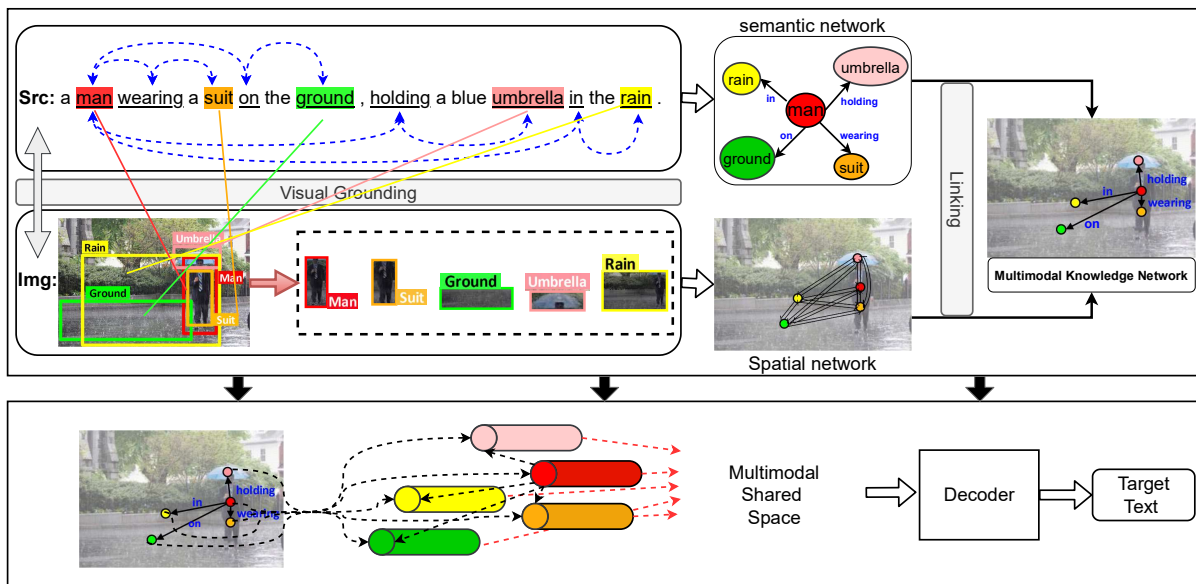
【実用化が期待される分野】

From a social perspective, this work comprehensively breaks through the existing barriers of multimodal translation, aiming to achieve high-performance and novel machine translation. Communication in the world is inseparable from language, and language communication is based on effective machine translation. Realizing a diversified machine translation system with multiple languages and diverse medium is the potential way for the next generation of machine translation applications. In the future, multilingual and multimodal machine translation systems will facilitate semantic communication between multiple languages with taking advantage of diverse modalities, it will become a popular way to allow people to communicate freely without having to learn multiple languages, such as travel, cultural exchange, and trade.