

■募集情報

募集番号	2023-2
部署名	ユニバーサルコミュニケーション研究所 先進的音声翻訳研究開発推進センター
インターンシップ課題名	Leveraging Diverse Segmentations For Neural Machine Translation
インターンシップ内容の概要	<p>Previous works have shown that dynamic segmentations that are semantically aware have a positive impact on neural machine translation. However, existing works do not leverage various sources of segmentation effectively. Based on this the research question is: how can we leverage diverse subword segmentation sources for neural machine translation.</p> <p>Specifically we want to:</p> <ol style="list-style-type: none"> <li>1. Develop an approach were diverse subword segmentations are effectively combined in an encoder. The simplest approach is to adopt multisource encoder approaches however, the main challenge is to ensure that the model uses information from all segmentations.</li> <li>2. Ensure that the method is efficient so that the final system is not much slower than one which leverages only one type of segmentation.</li> </ol>
課題に関する問い合わせ先	先進的翻訳技術研究室 研究統括 田中英輝 hideki.tanaka@nict.go.jp
応募条件	<ul style="list-style-type: none"> <li>・ 学年：大学院以上</li> <li>・ 専攻学科等：計算機科学</li> <li>・ 自然言語処理、機械翻訳に関する知見がある方を優先的に採択する。</li> </ul>
実施場所	ユニバーサルコミュニケーション研究所（けいはんな）、自宅
実施期間	2023年4月1日～2024年3月31日の間で60日間（休・祝日を除く）
受入予定人数	1名
選考課題	なし
備考	