

## 科学研究費助成事業 研究成果報告書

令和 4 年 6 月 22 日現在

機関番号：14301

研究種目：若手研究

研究期間：2019～2021

課題番号：19K20343

研究課題名（和文）マルチモーダルデータからの対訳資源の抽出によるニューラル機械翻訳

研究課題名（英文）Neural Machine Translation Based on Bilingual Resources Extracted from Multimodal Data

研究代表者

チョ シンキ（Chu, Chenhui）

京都大学・情報学研究科・特定准教授

研究者番号：70784891

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本プロジェクトでは、マルチモーダルニューラル機械翻訳（MNMT）について、主に以下の研究を行った：1). コンパラブル文を用いたMNMT。データセットを構築して第8回アジア翻訳ワークショップ（WAT 2021）でshared タスクを主催し、優勝した。2). セマンティック画像領域と単語領域アライメントを用いたMNMT。セマンティック画像領域と単語領域アライメントを用いたMNMTを研究し、NeurocomputingとTASLPという2つの著名な国際誌に論文を発表した。3). 映像付き機械翻訳。空間階層注意ネットワークを提案し、動詞と名詞両方の意味曖昧性解消が対応可能になった。

研究成果の学術的意義や社会的意義

機械翻訳における自然言語の意味曖昧性解消を目的として、マルチモーダルニューラル機械翻訳（MNMT）が主に研究されている。本プロジェクトでは、低資源な設定でコンパラブル文を用いたMNMTという新しい仕組みを考案し、画像を用いたMNMTにおいてはセマンティック画像領域と単語領域アライメントを用いたMNMTを提案し、映像を用いたMNMTにおいては空間階層注意ネットワークを提案し、機械翻訳における視覚情報の利用の有効性を示した。開発したMNMTシステムは映画、ドラマ、アニメやニュースなどの字幕の自動翻訳の精度向上に貢献できるし、大阪万博などの国際的イベントでの自動翻訳ニュースにも応えられる。

研究成果の概要（英文）：In this project, we mainly studied the following topics for multimodal neural machine translation (NNMT). 1). MNMT with comparable sentences. We constructed an MNMT with comparable sentences dataset and organized a shared task in the 8th Workshop on Asian Translation (WAT 2021). Our system achieved the best performance in this shared task. 2). MNMT with semantic image regions and word-region alignment. We studied MNMT with semantic image regions and word-region alignment and published them in two famous international journals Neurocomputing and TASLP. 3). Video-guided MT (VMT). We proposed VMT with a spatial hierarchical attention network, which can address both verb and noun sense disambiguation.

研究分野：自然言語処理

キーワード：機械翻訳 マルチモーダル

### 1. 研究開始当初の背景

国際化が進む中、翻訳の需要はさまざまな場面で急速に高まっており、すべてのテキストを手で翻訳することは不可能である。機械翻訳 (MT) は翻訳のコストを削減して効率を高めるための強力なツールとして、国際化を促進するのに重要な役割を果たしている。日本では、外国人観光客の急増 (2018 年は 3000 万人以上) や 2020 年東京オリンピック開催などのため、MT は不可欠である。

深層学習に基づくニューラル機械翻訳 (NMT) の発展により、MT の精度は大きく向上してきた (引用文献①)。NMT は大規模な対訳コーパス (文単位のバイリンガルテキスト) が入手できる場合に最先端の翻訳精度を達成したが、対訳コーパスが少量な場合、翻訳精度が低いことが知られている (引用文献②)。しかし、特定言語対や分野の対訳コーパスが大量に存在しない場面が数々ある。

### 2. 研究の目的

Web やソーシャルメディアの普及により、大量の画像とその説明文が簡単に収集できるようになってきた。例えば、スポーツサイトでは同じ話題のスポーツニュースが異なる言語で同じような画像を使って書かれたり、違う国の e コマースサイトでは同じような商品が画像と異なる言語の説明文で販売されたり、違う国の SNS やブログでは毎日数百万の写真と多くの異なる言語のコメントが投稿されている。これらの説明文は、異なる人が異なる言語で、類似した画像について書いているため、説明文の中には大量の対訳データが存在する可能性がある。しかし、このようなマルチモーダルデータは、これまで NMT のコミュニティで適切に利用されてこなかった。本プロジェクトはこのようなマルチモーダルデータを NMT での利活用を目指し、マルチモーダル NMT (MNMT) の精度向上を目的とし、MNMT の研究開発を行う。

### 3. 研究の方法

(1) コンパラブル文を用いた MNMT。スポーツ/e コマースサイトや各国の投稿などのマルチモーダルデータでは、対訳文はほとんど存在しないが、対訳辞書や対訳句を含むコンパラブル文が大量に存在する。コンパラブル文を含む MNMT のデータセットを構築し、MNMT の性能を測る。

(2) 画像を用いた MNMT システム開発。いかに画像情報を MNMT に最善に使うて翻訳性能を向上させるかを研究する。

(3) 映像を用いた MNMT システム開発。時系列情報も含む映像情報を MNMT いかに利用して翻訳性能を向上させるかを研究する。

### 4. 研究成果

本プロジェクトでは、MNMT について、主に以下の研究成果を得た：

#### (1) コンパラブル文を用いた MNMT

既存の画像キャプションデータセットからコンパイルな文からなる新しいマルチモーダルの大規模な (123k 画像) 英日コーパスを構築した。図 1 にコンパイルな文の例を示す。また、動詞に曖昧性を含む英語の画像キャプションを業者に日本語に翻訳してもらって、小規模な検証・テストセットも作った。コンパラブル文を用いた MNMT を評価するために、そのコンパラブルコーパスを用いて複数のベースライン MNMT モデルを学習し、英日翻訳の性能を比較した。さらに、構築したデータセットを用いて、第 8 回アジア翻訳ワークショップ (WAT 2021) で shared タスクを主催し、優勝した。



図 1 : コンパラブル文の例

#### (2) セマンティック画像領域と単語領域アライメントを用いた MNMT

Region-attentive MNMT (RA-NMT) と呼ばれる、セマンティック画像領域を用いた (MNMT) 手法を提案した。先行研究では、翻訳性能を向上させる

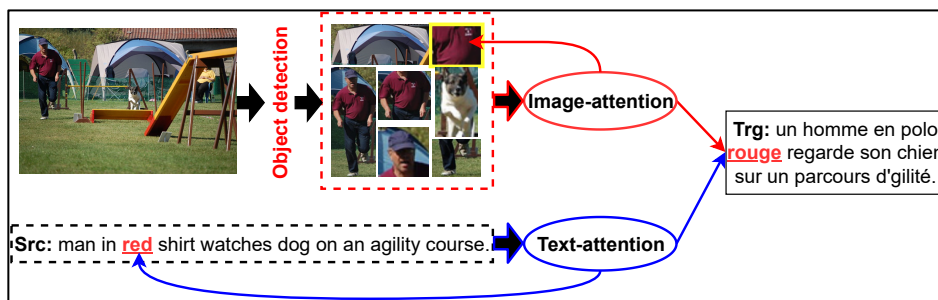


図 2 : Region-attentive MNMT

るために、畳み込みニューラルネットワークによって抽出されたグローバルな視覚特徴や等サイズグリッドの局所的な視覚特徴を用いることに主に焦点が当てられてきた。しかし、視覚的

微量に含まれるセマンティック情報の効果は無視されている。本研究では、物体検出により抽出されたセマンティック画像領域を MNMT に利用し、モダリティ依存の 2 つの注意メカニズムを用いて視覚特徴とテキスト特徴を統合する (図 2)。提案手法を NMT の 2 つのニューラルアーキテクチャ、リカレントニューラルネットワーク (RNN) とトランスフォーマーに実装し、検証を行った。Multi30k データセットの異なる言語対を用いた実験の結果、提案手法はベースラインよりも改善され、最先端の MNMT 手法のほとんどを凌駕することが示された。さらに、提案手法は視覚的特徴をうまく利用することで、より優れた翻訳性能を達成できることを示した。

### (3) 単語領域アライメントを用いた MNMT

単語領域アライメント (WRA) を用いてテキストと視覚モダリティの意味的相関を結びつける新しい MNMT モデルを提案した。先行研究は、主に視覚モダリティとテキストモダリティをいかに統合して効果を果たせるかに着目している。しかし、2 つのモダリティ間の意味的な関連性は考慮されていない。WRA をブリッジとして組み込むことで、MNMT における視覚とテキストモダリティ間の意味的な相関を学習させることを提案した (図 3)。本提案は、NMT の主流である RNN とトランスフォーマーという 2 つのアーキテクチャに実装されている。Multi30k データセットを用いた英独・英仏翻訳タスクと Flickr30kEnt-JP データセットを用いた英日翻訳タスクという 2 つのベンチマークを用いた実験により、提案手法は複数の評価基準においてベースラインより大幅に改善し、既存の MNMT モデルのほとんどを上回ることが示された。さらに分析から、提案手法が WRA を統合することでより良い翻訳性能を達成し、より良い視覚情報の利用につながるということがわかった。

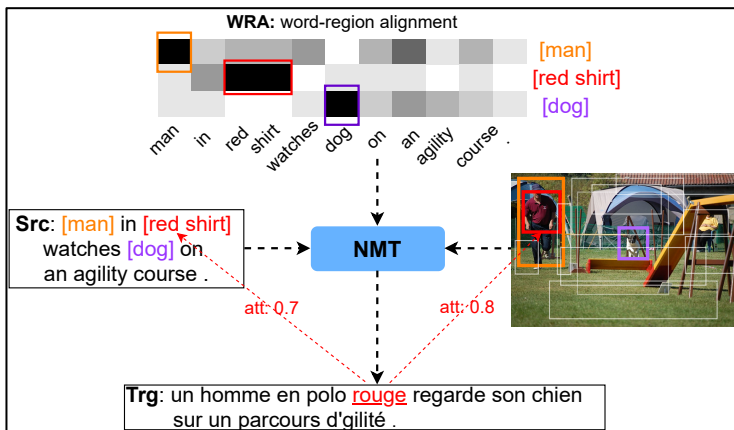


図 3 : 単語領域アライメントを用いた MNMT

Multi30k データセットを用いた英独・英仏翻訳タスクと Flickr30kEnt-JP データセットを用いた英日翻訳タスクという 2 つのベンチマークを用いた実験により、提案手法は複数の評価基準においてベースラインより大幅に改善し、既存の MNMT モデルのほとんどを上回ることが示された。さらに分析から、提案手法が WRA を統合することでより良い翻訳性能を達成し、より良い視覚情報の利用につながるということがわかった。

### (4) 映像付き MNMT

映像付き MNMT は、映像を補助情報として利用することによって、

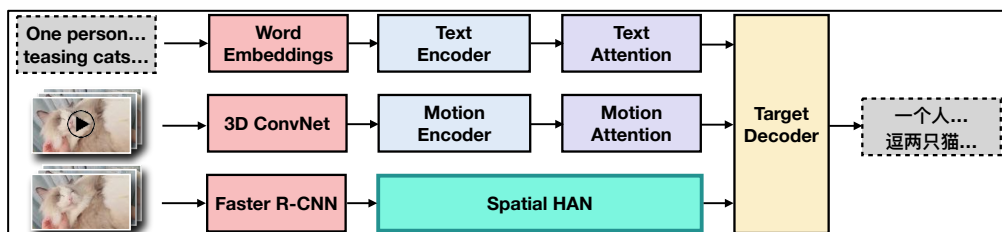


図 4 : 空間階層注意ネットワークを用いた映像付き MNMT

MT における語義曖昧性問題に対処することを目的としている。先行研究では、動詞の語義曖昧性を解決するために、事前に学習した動作検出モデルからの特徴量を MNMT システムに組み込む。しかし、名詞の語義曖昧性は課題として残っている。この課題を解決するために、映像中の空間表現と動作表現の両方を用いた映像付き MNMT システムを提案した。空間特徴については、物体レベルから映像レベルまでの空間情報をモデル化するために、空間階層注意ネットワークを提案した (図 4)。VATEX データセットを用いた実験の結果、提案手法はベースラインの性能を大きく上回ることがわかった。

#### <引用文献>

- ① Dzmitry Bahdanau, Kyunghyun Cho and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. In Proceedings of the 3rd International Conference on Learning Representations, (2015).
- ② Philipp Koehn and Rebecca Knowles. Six Challenges for Neural Machine Translation. In Proceedings of the First Workshop on Neural Machine Translation, pp. 28-39, (2017).

## 5. 主な発表論文等

〔雑誌論文〕 計12件（うち査読付論文 7件 / うち国際共著 3件 / うちオープンアクセス 12件）

1. 著者名 Nakao Ryota, Chu Chenhui, Kurohashi Sadao	4. 巻 28
2. 論文標題 Spoken-Written Japanese Conversion for Japanese-English University-Lecture Translation	5. 発行年 2021年
3. 雑誌名 Journal of Natural Language Processing	6. 最初と最後の頁 1034 ~ 1052
掲載論文のDOI (デジタルオブジェクト識別子) 10.5715/jnlp.28.1034	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Zhao Yuting, Komachi Mamoru, Kajiwara Tomoyuki, Chu Chenhui	4. 巻 30
2. 論文標題 Word-Region Alignment-Guided Multimodal Neural Machine Translation	5. 発行年 2022年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech, and Language Processing	6. 最初と最後の頁 244 ~ 259
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/taslp.2021.3138719	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Zhao Yuting, Komachi Mamoru, Kajiwara Tomoyuki, Chu Chenhui	4. 巻 476
2. 論文標題 Region-attentive multimodal neural machine translation	5. 発行年 2022年
3. 雑誌名 Neurocomputing	6. 最初と最後の頁 1 ~ 13
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.neucom.2021.12.076	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Kawara Yuki, Chu Chenhui, Arase Yuki	4. 巻 29
2. 論文標題 Preordering Encoding on Transformer for Translation	5. 発行年 2021年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech, and Language Processing	6. 最初と最後の頁 644 ~ 655
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/taslp.2020.3042001	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Dabre Raj, Chu Chenhui, Kunchukuttan Anoop	4. 巻 53
2. 論文標題 A Survey of Multilingual Neural Machine Translation	5. 発行年 2020年
3. 雑誌名 ACM Computing Surveys	6. 最初と最後の頁 1~38
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3406095	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Chu Chenhui, Wang Rui	4. 巻 28
2. 論文標題 A Survey of Domain Adaptation for Machine Translation	5. 発行年 2020年
3. 雑誌名 Journal of Information Processing	6. 最初と最後の頁 413~426
掲載論文のDOI (デジタルオブジェクト識別子) 10.2197/ipsjjip.28.413	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Andrew Merritt, Chenhui Chu, Yuki Arase	4. 巻 -
2. 論文標題 A Corpus for English-Japanese Multimodal Neural Machine Translation with Comparable Sentences	5. 発行年 2020年
3. 雑誌名 arXiv:2010.08725	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Vipul Mishra, Chenhui Chu, Yuki Arase	4. 巻 -
2. 論文標題 Lexically Cohesive Neural Machine Translation with Copy Mechanism	5. 発行年 2020年
3. 雑誌名 arXiv:2010.05193	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Chenhui Chu and Raj Dabre	4. 巻 -
2. 論文標題 Multilingual Multi-Domain Adaptation Approaches for Neural Machine Translation	5. 発行年 2019年
3. 雑誌名 arXiv:1906.07978	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Raj Dabre, Chenhui Chu, Anoop Kunchukuttan	4. 巻 -
2. 論文標題 A Survey of Multilingual Neural Machine Translation	5. 発行年 2019年
3. 雑誌名 arXiv:1905.05395	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 竹林佑斗, Chenhui Chu, 荒瀬由紀, 永田昌明	4. 巻 26
2. 論文標題 ニューラル機械翻訳における単語報酬モデルに基づく対訳辞書の利用	5. 発行年 2019年
3. 雑誌名 自然言語処理	6. 最初と最後の頁 711-731
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Raj Dabre, Chenhui Chu, Anoop Kunchukuttan	4. 巻 -
2. 論文標題 A Comprehensive Survey of Multilingual Neural Machine Translation	5. 発行年 2020年
3. 雑誌名 arXiv:2001.01115	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計21件（うち招待講演 1件 / うち国際学会 8件）

1. 発表者名 Haiyue Song, Raj Dabre, Zhuoyuan Mao, Chenhui Chu, Sadao Kurohashi
2. 発表標題 Representative Data Selection for Sequence-to-Sequence Pre-training
3. 学会等名 言語処理学会 第28回年次大会
4. 発表年 2022年

1. 発表者名 Yihang Li, 清水 周一郎, Chenhui Chu, 黒橋 禎夫
2. 発表標題 曖昧性を含む翻訳に着目したマルチモーダル機械翻訳データセットの構築方法の検討
3. 学会等名 言語処理学会 第28回年次大会
4. 発表年 2022年

1. 発表者名 Zhuoyuan Mao, Prakhar Gupta, Chenhui Chu, Martin Jaggi, Sadao Kurohashi
2. 発表標題 Lightweight Cross-Lingual Sentence Representation Learning
3. 学会等名 In Proceedings of ACL-IJCNLP 2021 (国際学会)
4. 発表年 2021年

1. 発表者名 Weiqi Gu, Haiyue Song, Chenhui Chu, Sadao Kurohashi
2. 発表標題 Video-guided Machine Translation with Spatial Hierarchical Attention Network
3. 学会等名 In Proceedings of the ACL-IJCNLP 2021 Student Research Workshop (国際学会)
4. 発表年 2021年

1 . 発表者名 Toshiaki Nakazawa, Hideki Nakayama, Chenchen Ding, Raj Dabre, Shohei Higashiyama, Hideya Mino, Isao Goto, Win Pa Pa, Anoop Kunchukuttan, Shantipriya Parida, Ondrej Bojar, Chenhui Chu, Akiko Eriguchi, Kaori Abe, Yusuke Oda, Sadao Kurohashi
2 . 発表標題 Overview of the 8th Workshop on Asian Translation
3 . 学会等名 In Proceedings of the 8th Workshop on Asian Translation (WAT2021) (国際学会)
4 . 発表年 2021年

1 . 発表者名 Yuting Zhao, Mamoru Komachi, Tomoyuki Kajiwara, Chenhui Chu
2 . 発表標題 TMEKU System for the WAT2021 Multimodal Translation Tas
3 . 学会等名 In Proceedings of the 8th Workshop on Asian Translation (WAT2021) (国際学会)
4 . 発表年 2021年

1 . 発表者名 Raj Dabre , Chenhui Chu , Anoop Kunchukuttan
2 . 発表標題 Multilingual Neural Machine Translation (Tutorial)
3 . 学会等名 In Proceedings of the 28th International Conference on Computational Linguistics (COLING 2020) (国際学会)
4 . 発表年 2020年

1 . 発表者名 Yuting Zhao, Mamoru Komachi, Tomoyuki Kajiwara, Chenhui Chu
2 . 発表標題 Double Attention-based Multimodal Neural Machine Translation with Semantic Image Regions
3 . 学会等名 In Proceedings of the 22nd Annual Conference of the European Association for Machine Translation (EAMT 2020) (国際学会)
4 . 発表年 2020年



1. 発表者名 Weiqi Gu, Haiyue Song, Chenhui Chu, Sadao Kurohashi
2. 発表標題 Video-guided Machine Translation with Spatial Hierarchical Attention Network Encoder
3. 学会等名 言語処理学会 第27回年次大会
4. 発表年 2020年

1. 発表者名 Haiyue Song, Raj Dabre, Chenhui Chu, Sadao Kurohashi, Eiichiro Sumita
2. 発表標題 Self-supervised Dynamic Programming Encoding for Neural Machine Translation
3. 学会等名 言語処理学会 第27回年次大会
4. 発表年 2020年

1. 発表者名 Zhuoyuan Mao, Prakhar Gupta, Chenhui Chu, Martin Jaggi, Sadao Kurohashi
2. 発表標題 Learning Cross-lingual Sentence Representations for Multilingual Document Classification with Token-level Reconstruction
3. 学会等名 言語処理学会 第27回年次大会
4. 発表年 2020年

1. 発表者名 瓦 祐希, Chenhui Chu, 荒瀬 由紀
2. 発表標題 Non-Autoregressive Translationモデルにおける事前並び替え適用手法の検討
3. 学会等名 言語処理学会 第27回年次大会
4. 発表年 2020年

1. 発表者名 Shuichiro Shimizu, Chenhui Chu, Sheng Li, Sadao Kurohashi
2. 発表標題 End-to-End Speech Translation with Cross-lingual Transfer Learning
3. 学会等名 言語処理学会 第27回年次大会
4. 発表年 2020年

1. 発表者名 Yuting Zhao, Mamoru Komachi, Tomoyuki Kajiwara, Chenhui Chu
2. 発表標題 Neural Machine Translation with Semantic Relevant Image Regions
3. 学会等名 言語処理学会 第27回年次大会
4. 発表年 2020年

1. 発表者名 中尾 亮太, Chenhui Chu, 黒橋 禎夫
2. 発表標題 日本語話し言葉書き言葉変換による大学講義の日英翻訳の精度向上
3. 学会等名 言語処理学会 第27回年次大会
4. 発表年 2020年

1. 発表者名 Yuting Zhao, Mamoru Komachi, Tomoyuki Kajiwara, Chenhui Chu
2. 発表標題 Double Attention-based Multimodal Neural Machine Translation with Semantic Image Regions
3. 学会等名 情報処理学会第241回自然言語処理研究会/NLP若手の会 (YANS) 第14回シンポジウム
4. 発表年 2019年

1. 発表者名 瓦祐希, Chenhui Chu, 荒瀬由紀
2. 発表標題 事前並び替え位置表現を用いたTransformerによる日英機械翻訳
3. 学会等名 言語処理学会第26回年次大会
4. 発表年 2020年

1. 発表者名 Mishra Vipul, Yuki Kawara, Chenhui Chu, Yuki Arase
2. 発表標題 Comparison and Analysis of 2-to-2 and Hierarchical RNN Models on Japanese-to-English Context-Aware Translation
3. 学会等名 NLP若手の会 (YANS) 第14回シンポジウム
4. 発表年 2019年

1. 発表者名 瓦祐希, Chenhui Chu, 荒瀬由紀
2. 発表標題 事前並び替えによる英日Transformerモデルの翻訳精度向上
3. 学会等名 NLP若手の会 (YANS) 第14回シンポジウム
4. 発表年 2019年

1. 発表者名 Raj Dabre, Atsushi Fujita, Chenhui Chu
2. 発表標題 Exploiting Multilingualism through Multistage Fine-Tuning for Low-Resource Neural Machine Translation
3. 学会等名 In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 Chenhui Chu and Rui Wang
2. 発表標題 Domain Adaptation for Neural Machine Translation
3. 学会等名 The 15th China Conference on Machine Translation (CCMT 2019) (招待講演) (国際学会)
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

<a href="https://researchmap.jp/chu/">https://researchmap.jp/chu/</a>
---

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	荒瀬 由紀 (Arase Yuki) (00747165)	大阪大学・大学院情報科学研究科・准教授  (14401)	
研究協力者	小町 守 (Komachi Mamoru) (60581329)	東京都立大学・システムデザイン学部・教授  (22604)	
研究協力者	梶原 智之 (Kajiwara Tomoyuki) (70824960)	愛媛大学・大学院理工学研究科・助教  (16301)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	黒橋 禎夫  (Kurohashi Sadao)  (50263108)	京都大学・情報学研究科・教授    (14301)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関			
中国	上海交通大学			
インド	Microsoft, Hyderabad			
米国	University of Georgia			
スイス	EPFL			