

## 令和元年度研究開発成果概要書

採 択 番 号 : 19701  
研究開発課題名 : 多言語音声翻訳高度化のためのディープラーニング技術の研究開発の研究開発  
副 題 : 深層学習によるマルチモーダル文脈理解と機械翻訳の高度化

## (1) 研究開発の目的

ビジネスシーンにおける外国人と日本人のコミュニケーションを円滑にするため、会議や社交における対話・雑談、およびニュース記事に対応した機械翻訳技術を研究・開発する。また、より日常生活に密着した機械翻訳に向け、マルチモーダル機械翻訳に取り組み、テキストと画像の統合的な理解に向けた研究を進める。全ての研究・開発において、深層学習を共通基盤として採用し、タスクやメディア、モーダルの垣根を超えた人工知能の実現を目指す。また、自然言語処理、画像処理、深層学習、報道などの分野で最先端の取り組みを継続しているグループでチームを結成し、人材やデータ、技術の交流を促進する。これにより、日本語に関する技術で世界トップレベルを維持するとともに、マルチモーダルな文脈理解など、世界に先駆けた研究に挑戦する。

深層学習に基づくアプローチでは、学習データの質と量が性能を左右する。そこで、本提案では研究資源（対話や翻訳、多言語画像技術文のコーパスなど）の開発にも注力する。本研究で得られた成果は、プロジェクト終了後も活用できるように検討し、研究コミュニティや社会に還元する。また、ニュース記事に対応した機械翻訳技術および自動要約技術などの成果は、ニュースの配信の実務と照らし合わせながら評価し、研究成果の社会実装に向けた実行可能性調査を行う。また、対話翻訳技術は NICT が開発している VoiceTra の翻訳エンジンや、スマートスピーカーへの応用を模索し、社会実装につなげていく。

## (2) 研究開発期間

平成 30 年度から令和 2 年度（3 年間）

## (3) 実施機関

国立大学法人東京工業大学 <研究代表者>  
国立大学法人東京大学  
国立大学法人愛媛大学  
日本放送協会  
一般財団法人 NHK エンジニアリングシステム  
株式会社時事通信社

## (4) 研究開発予算（契約額）

総額 780 百万円（令和元年度 240 百万円）  
※百万円未満切り上げ

## (5) 研究開発項目と担当

研究開発項目 1： 高度な文脈理解技術（インテリジェント翻訳技術）の研究開発

1. 対話単言語コーパスの構築 (東京工業大学)
2. 日英対話対訳コーパスの構築 (東京大学)
3. 文脈に応じて適切な訳を選択する技術の開発 (東京大学)
4. 訳語の一貫性を保持する技術の開発 (東京大学)
5. 翻訳に必要な情報を補完する技術の開発 (東京大学)

研究開発項目 2： 新語・新トピックへの即時対応技術（ニュース対応翻訳技術）

1. ニュース記事コーパスの整備 (時事通信社)
2. 人手によるニュース対訳コーパスの作成 (NHK エンジニアリングシステム)
3. 新語や新話題に対応した機械翻訳技術の構築 (東京大学)

- 4. ニュース翻訳システムの研究開発 (日本放送協会)
- 5. ニュース記事からの要約文生成の研究開発 (東京工業大学)
- 6. ニュースの長文翻訳および要約文生成の評価 (時事通信社)

研究開発項目 3: マルチモーダル翻訳技術の研究開発

- 1. マルチモーダル日英対訳コーパスの作成 (東京大学)
- 2. 機械翻訳のためのシーン画像認識手法の開発 (東京大学)
- 3. マルチモーダル翻訳手法の開発 (愛媛大学)

(6) 特許出願、外部発表等

		累計 (件)	当該年度 (件)
特許出願	国内出願	1	0
	外国出願	0	0
外部発表等	研究論文	1	0
	その他研究発表	54	34
	標準化提案	0	0
	プレスリリース・報道	1	0
	展示会	0	0
	受賞・表彰	8	6

(7) 具体的な実施内容と成果

研究開発項目 1:

1-1、1-2. ビジネスシーン日英対話対訳コーパスの構築

平成 30 年度に引き続きビジネスシーンにおける対話のシナリオを日本語および英語で創作し、それぞれ英語および日本語に翻訳することで、日英対話対訳コーパスを拡充した。令和元年度は昨年度の 4.5 万対訳文に加えて、新たに 4 万対訳文を追加し、合計 8.5 万対訳文となった。さらに令和元年度は共参照アノテーションが付与された英語のデータセットである OntoNotes5.0 のうち、Newswire カテゴリに属する 1.5 万文を日本語に翻訳することで対訳文を構築した。

1-3 文脈に応じて適切な訳を選択する技術の開発 および 1-4 訳語の一貫性を保持する技術の開発

項目 1-1、1-2 で構築した対訳コーパスの一部をテストデータとし、これまでに構築した全ての対訳コーパスを使って訓練した単文の機械翻訳エンジンで日英および英日翻訳を行い、翻訳結果をクラウドソーシングにより評価した。単文での評価において翻訳に問題がないと判定された文のうち、文脈付きで評価すると翻訳に誤りがあると判定されたものは日英で約 15%、英日で約 20%であった。次に 2-to-1 の方法で訓練した文脈付き機械翻訳で翻訳を行ったところ、単文での翻訳より多少 BLEU スコアが向上するが、オラクルのスコアよりは低い結果となった。文脈情報を利用することは有効であるが、文脈情報が不要な文まで文脈付きで翻訳すると、悪影響を与える可能性があり、全ての文において文脈情報を付与するのではなく、文脈情報が必要かどうかをあらかじめ判定する必要があることがわかった。また対話対訳では直前文が相槌などのため情報量が少ないものが多いため、数文前を文脈情報として利用する必要があることもわかった。これらの解決が今後の目標である。

1-5 翻訳に必要な情報を補完する技術の開発

BERT を用いて、翻訳文の文脈として適切な文を選択する手法を開発した。訓練済みの BERT に対し、与えられた 2 文が連続する 2 文か判定を行うタスクと、翻訳を行うタスクのマルチタスク

クでファインチューニングし、文脈として適切かどうかを判定する分類器として用いた。翻訳対象文と、その 1~5 文前をそれぞれペアにして入力し、分類器による連続文判定の確率が最も高い文を文脈文として利用し、2-to-1 翻訳を行った。実験の結果、単文で翻訳を行った場合および 1~5 文前をそれぞれ文脈文として用いて訓練した 2-to-1 翻訳機と比較して、文脈文選択を行った場合は日英翻訳で 0.10 から 0.67、英日翻訳で 0.15 から 0.94 ポイントの BLEU 値の向上が見られた。しかしながら BLEU 値が最も高くなる文脈文を選択した Oracle のスコアと比べるとまだ 7 ポイント程度の差があるため、改善の余地が十分に残されている。

#### 研究開発項目 2 :

##### 2-1、2-2 ニュース記事コーパスの整備、人手によるニュース対訳コーパスの作成

平成 30 年度に引き続き、人手翻訳による対訳コーパスの構築を進めた。令和元年度は、時事通信社日本語ニュースの日英翻訳と、日本語ニュースを英語ニュースの内容に合わせて修正する方式を検討して業者に発注し、人手翻訳による高品質な対訳として 17 万文対を構築。昨年度までに構築した人手翻訳による対訳データや、日英のニュース自動対応付けデータなどと合わせて合計 180 万文対の日英対訳コーパスを構築した。

##### 2-3、2-4 新語や新話題に対応した機械翻訳技術の構築、ニュース翻訳システムの研究開発

平成 30 年度までに構築した対訳コーパスを用いて日英・英日ニュース翻訳システムを構築し、評価型の国際ワークショップで 1 位を獲得した。主観評価では、ほとんどの重要情報が正確に伝達されているレベル (80%~) 以上の結果が得られた。また、複数の対訳コーパスを効果的に学習するための手法を考案し、その効果を実証した。さらに、翻訳における新語処理の調査、検討を進め、新語となる固有名詞は対訳辞書で訳を指定できる仕組みが有効であることを確認した。

##### 2-5 ニュース記事からの要約文生成の研究開発

要約・見出し生成のデモと API を公開した。また、元記事から逸脱した見出しが生成される問題を改善する手法を提案し、学習データから無理な事例を取り除くと見出しの忠実性が改善することを自動評価と人手評価で確認した。さらに、言語横断見出し生成などの新たな課題に取り組んだ。言語横断見出し生成では、機械翻訳と要約を組み合わせるマルチタスク学習の手法を提案し、中英言語横断見出し生成で最高性能を実現した。

##### 2-6 ニュースの長文翻訳および要約文生成の評価

時事通信社において、令和元年度に作成した「JLIQIM 翻訳品質評価ガイドライン」を使い、ニュース翻訳のテスト評価を実施した。また、時事通信社の短文ニュースを作成している担当者により、公開した見出し生成の主観評価を実施し、53.3%で「作業が楽になる」、26.1%で「ある程度メリットがある」という評価が得られた。

#### 研究開発項目 3 :

##### 3-1. マルチモーダル日英対訳コーパスの開発

平成 30 年度に引き続き、Flickr30K Entities データセットの英語キャプションを日本語翻訳し、マルチモーダル日英対訳コーパスの拡張を行った。本年度は新たに 63,566 文の翻訳を行った結果、これまでの合計は 127,132 文対となり、本項目で目標としていた 10 万文対の作成を達成した。本コーパスは、日英では現存する最大のマルチモーダル対訳コーパスであると共に、文中の複数のフレーズと画像中の領域が多対多で紐づいたアノテーションを有するという、他に類を見ない特徴を有している。また、平成 30 年度に作成したコーパスを用い、日本語フレーズによる物体検出タスクの実験を実施し、作成したデータの基礎的な評価を行った。この結果は LREC 2020 に採択され、当該論文の発表に合わせ、平成 30 年度版のコーパスを公開する予定である。

##### 3-2. シーン画像認識手法の開発

本年度は、画像認識の精度を向上させるための汎用的な基礎技術を二つ開発した。一つ目は、

データ拡張操作を微分可能なネットワークを構成することにより、データ拡張ポリシーを高速に自動探索する手法（Faster AutoAugment）である。本手法は、ImageNet の標準的な 1000 クラス識別タスクにおいて、ベースラインと比較して 0.6% 識別精度を向上させた。また、既存のデータ拡張探索手法と比較して数十倍学習が高速である。二つ目は、特徴マップをシャッフルさせることにより、大域的情報を畳み込みにおいて活用する Spatial Shuffling 法である。本手法も、ImageNet の識別タスクにおいて、さまざまな種類の畳み込みニューラルネットワークの識別精度 0.5% 程度向上させている。これらの手法を実装した結果、目標とする 100 クラスのシーン画像認識の精度が 75.0% から 77.2% まで向上した。

### 3-3. マルチモーダル翻訳手法の開発

マルチモーダル機械翻訳は、テキストだけを対象として翻訳するのではなく、そのテキストの情景を説明する画像も用いることで、より適切な翻訳を実現することを目的としている。本年度は、画像情報を参照するマルチモーダル機械翻訳システムの開発を行い、項目 3-1 で開発中のマルチモーダル日英対訳コーパスを用いて予備的な評価を行った。3-1 で開発中のマルチモーダル日英対訳コーパスは従来のマルチモーダル対訳コーパスと異なり、フレーズに対する言語間の対応関係を表すアノテーションとフレーズ-画像領域間の対応関係を表すアノテーションが付与されている。本研究はマルチモーダル機械翻訳の高精度化を目的として、これらのアノテーションをアテンション教師データとして用いるマルチモーダル機械翻訳手法を開発した。実験を行った結果、提案手法により翻訳精度 (BLEU) が 44.11% から 44.98% まで向上したことを確認した (+0.87% の向上)。