

TYPE OF
INDUSTRY

ニューラル機械翻訳は、ニューラルネットワークを用いて翻訳を実現する仕組みである。そのモデルは、特定の言語対のデータを使用して訓練される。

情報通信研究機構

NICT 先端研究

(201)

ニューラル機械翻訳 多段階訓練で品質向上

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 ダブレ・ラジ
先進的翻訳技術研究室 有期研究員

18年京大院博士後期課程修了。同年NICTに着任、現在に至る。以来、低資源言語対の機械翻訳、多言語機械翻訳の品質向上に関する研究に従事。博士(情報学)。



科学技術・大学

しかし、高品質なニューラル機械翻訳の訓練時モデルを段階的に訓練を実現するには数百万に複数の言語対のデータを必要とすることによってこの文の翻訳の実例データを複数使用して翻訳品質を向上させる多言語翻訳が必要であり、翻訳の例が小規模（例えば、数千文）しか存在しない言語間の翻訳は非常に難しい。

この低資源言語対の問題に対して、ニューラル機械翻訳の

しかし、高品質なニューラル機械翻訳の訓練時モデルを段階的に訓練を実現するには数百万に複数の言語対のデータを必要とすることによってこの文の翻訳の実例データを複数使用して翻訳品質を向上させる多言語翻訳が必要であり、翻訳の例が小規模（例えば、数千文）しか存在しない言語間の翻訳は非常に難しい。

この低資源言語対の問題に対して、ニューラル機械翻訳の

1. 事前訓練 英語→中国語

2. 多言語対の洗練

英語→中国語
英語→日本語
英語→クメール語
英語→マレー語
英語→インドネシア語
英語→ベンガル語
英語→タガログ語
英語→ベトナム語

3. 言語対ごとの洗練 英語→クメール語

複数の言語対のデータを使用する
多言語多段階学習(NICT提供)

載)

(火曜日に掲

問題に対処するよりも、複数の言語を使用しながら、より小規模で管理しやすい工程を通じて対処する方が良いことを示している。