

LINE WORKS、新技術「InterBiasing」を考案し、 音声・音響信号処理における世界最大規模の国際学会 「INTERSPEECH 2024」にて論文が採択

人名、業界特有の専門用語、社内用語などの、学習データに存在しない単語やフレーズを高精度で認識する新技術。今後提供するサービス内でも活用予定。

ビジネス現場のコミュニケーションツール「LINE WORKS」を提供する LINE WORKS 株式会社(本社:東京都渋谷区、代表取締役社長:共同代表:島岡 岳史、増田 隆一)は、音声・音響信号処理における世界最大規模の国際学会「INTERSPEECH 2024」にて新技術「InterBiasing」を考案した論文が採択されたことをお知らせいたします。

「INTERSPEECH」は、International Speech Communication Association (ISCA)が主催する音声・音響信号処理における世界最大規模の国際学会で、世界のトップレベルの研究者や専門家が集まり、最新の研究成果や技術の進展を発表・議論する権威のある場です。

今年で25回目の開催となり、採択された論文は9月1日から5日にかけて開催される「INTERSPEECH 2024」(ギリシャ・コス島)にて発表されます。

LINE WORKS 株式会社はAIの研究開発に積極的に取り組んでおり、本論文は、リサーチエンジニアの中込優、ヘンツェル・ミヒヤエルらによる研究成果です。今回の「INTERSPEECH 2024」での論文採択は、音声認識技術の可能性を広げ、製品への適用の期待も高まる有望な成果となりました。

▼論文の詳細については下記をご参照ください。

Y. Nakagome and M. Hentschel, "InterBiasing: Boost Unseen Word Recognition through Biasing Intermediate Predictions"

<https://arxiv.org/abs/2406.14890>

■論文の概要

追加の学習を必要とせず、事前にキーワードリストを与えるだけで、音声認識の精度を向上されることが可能に。製品への適用の期待も高まる有望な成果

本論文では、人名、業界特有の専門用語、社内用語といった、学習データに存在しない単語やフレーズを高精度で音声認識できる新技術「InterBiasing」を提案しました。

「InterBiasing」は、追加学習を必要とせず、事前にキーワードリストを与えるだけで、音声認識の精度を向上されることを可能にします。これは、LINE WORKS AI製品の「CLOVA Note β (AI音声記録管理サービス)※1」のような製品への適用も期待される有望な成果です。

※1: 高度な音声認識、話者分離、話者のフィラーや言い淀みを除去する機能が搭載されたAI音声記録管理サービス。

CLOVA Note β 公式サイト: <https://clovanote.line.me/>

LINE WORKS

新技術考案の背景<End-to-End 音声認識の課題>

音声認識の分野では、近年、深層学習モデルが音声から文字列を直接予測する End-to-End 型と呼ばれる方式が主流となり、優れた認識性能を発揮しています。しかし、End-to-End 音声認識が認識できる語彙は学習データに強く依存しており、人名、業界特有の専門用語、社内用語、最新の商品名や作品名といった固有名詞など、学習データに網羅しにくい言葉を正確に認識することが難しいという課題があります。

一般的には、認識させたい言葉を学習データに追加し、音声認識モデルを追加学習することが行われてきましたが、学習コストやユーザーが使えるようになるまで時間がかかる課題があります。

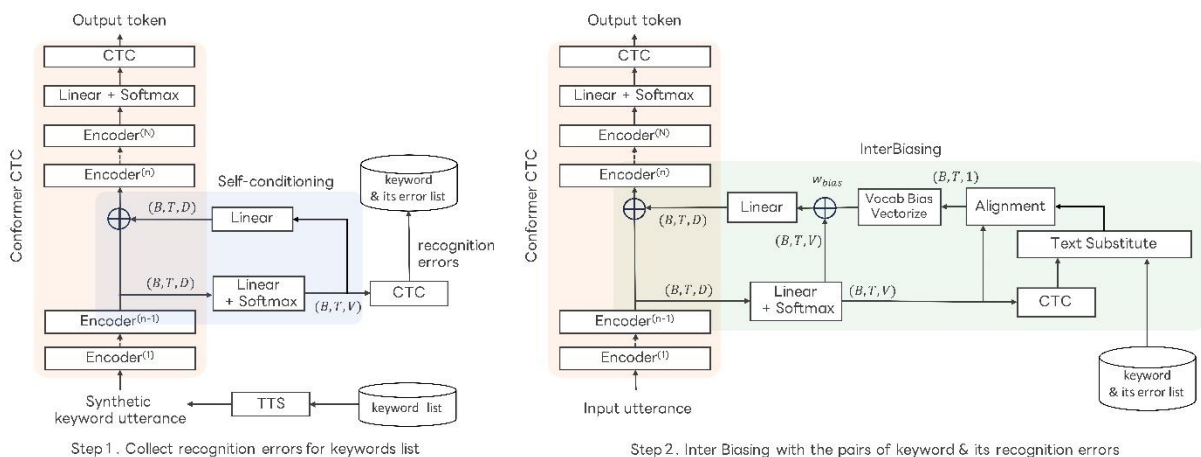
これらの課題を解決したいと考え、本論文では、音声認識モデルに事前に認識したいキーワードリストを与えておくだけで、一切の追加学習を必要とせずそのキーワードを正確に認識できるようになる音声認識手法の新技術「InterBiasing」を提案しました。

論文の詳細 「InterBiasing」について

本論文で提案した「InterBiasing」(図1参照)は、当社研究開発チームが以前から注力して研究している Self-Conditioned CTC※2 を発展させた新しい手法です。

この手法では、複数のニューラルネットワーク層からなる End-to-End 音声認識モデルの中間層に、認識させたいキーワードを条件付け、後続の層では与えられたキーワードを参考にしながら発話内容が推定されます。これにより、End-to-End 音声認識モデルが出力するトークンの事後確率分布を認識させたいキーワードが高い値を持つようにバイアスすることができ、追加学習を行わずにキーワードの認識精度を向上させることができます。

図1: 「InterBiasing」の概要図



<図1の説明: 提案手法は2つのステップで構成>

●ステップ1: キーワードの音声が発音合成(TTS)によって生成され、その音声を認識モデルに入力し、正解の文字列と認識誤りの文字列のペアが作成される。

●ステップ2: これらのペアは、中間予測における認識誤りの文字列を正解の文字列に置き換えるために利用される。後続の層はこのキーワードの正解文字列を参考にしながら認識仮説を推論する。

※2: J. Nozaki and T. Komatsu, "Relaxing the Conditional Independence Assumption of CTC-Based ASR by Conditioning on Intermediate Predictions," in Proc. Interspeech, 2021, pp. 3735-3739.

LINE WORKS

■LINE WORKS 株式会社が提供する AI サービス、研究開発について

LINE WORKS 株式会社では、ビジネス現場のコミュニケーションツール「LINE WORKS」に加え、LINE WORKS AI 製品として「LINE WORKS AiCall(電話対応 AI サービス)」や「LINE WORKS OCR(AI-OCR サービス)」などの高度な AI 技術を活用したサービスを提供しています。

また、サービスの提供だけでなく、AI 技術そのものの研究開発活動にも注力しています。2024 年 4 月に韓国・ソウルで開催された音声・音響信号処理の国際学会「ICASSP 2024」においても論文採択※3されており、研究の成果を積み上げています。

今後は、今回論文で提案した手法をさらに発展させると共に、LINE WORKS 株式会社のプロダクトへの適用、この活動で培った AI 技術を活用した新たな機能・サービスの創出に努めてまいります。

※3: 2024 年 2 月 1 日プレスリリース | LINE WORKS、効率的に LLM の知識が転送できる新しい手法を考案。音声・音響信号処理の国際学会「ICASSP 2024」にて論文が採択(<https://line-works.com/pr/20240201/>)

■会社概要

社名: LINE WORKS 株式会社

本社: 東京都渋谷区神宮前 1-5-8 神宮前タワービルディング 11F

設立: 2015 年 6 月

代表者: 共同代表: 島岡 岳史、増田 隆一

資本金: 55 億 2,000 万円

URL: <https://line-works.com/>

※記載の会社名、製品名は、それぞれ会社の商標または登録商標です。

※本プレスリリース記載の情報は発表日現在の情報です。予告なしに変更されることがありますので、予めご了承ください。