

2024年8月29日

## 「音声認識 AI のための自己教師学習とデータ拡張技術」に関する論文が、 音声・言語処理技術の国際会議「INTERSPEECH 2024」で採択 ～効果的・効率的な学習手法を開発し、音声認識 AI の性能向上を確認～

株式会社リコー（社長執行役員：大山 晃）は、「音声認識 AI のための自己教師学習とデータ拡張技術」に関する論文が、音声言語処理分野の国際会議「INTERSPEECH 2024」に採択されたことをお知らせいたします。リコーの論文が、本会議で採択されるのは初めてとなります。

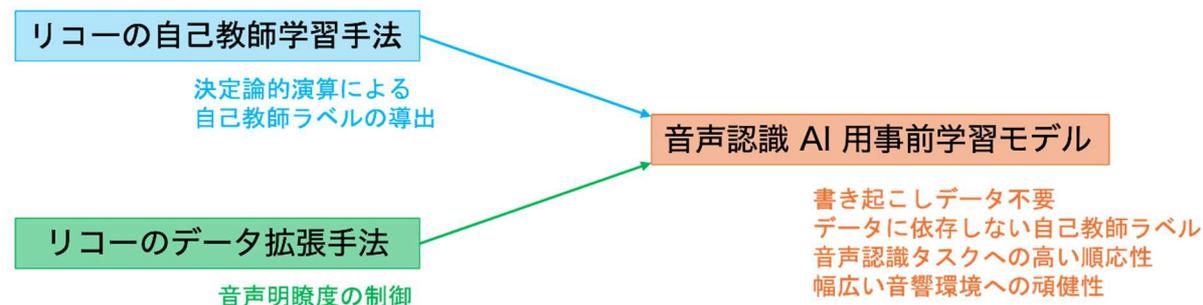
今回採択された論文では、音声認識 AI モデルに、書き起こしテキストの無い音声データのみを用いて効果的かつ効率的に学習させる手法の技術開発について発表しました。

従来の音声認識の学習過程では、入力する音声とそれに対応する書き起こしテキストのペアを用いて、それらの対応関係を学習させる『教師あり学習』を行うのが一般的です。その学習方法では、大量の書き起こし付き音声が必要となるため、学習データの入手に多大なコストがかかることが課題でした。

また、実際の使用環境で得られる音声は、用途・場所等により音声品質が異なるため、さまざまな環境で活用できるようにするために、音響ノイズへの耐性の強化が必要となります。

今回、リコーが開発した独自の自己教師学習手法と、音響ノイズ耐性を強化するデータ拡張技術により、従来の学習方法と比べて低コストで、より精度の高い音声認識性能を実現しました。

本技術についての詳細については、技術ページをご覧ください。  
技術ページ「音声認識 AI のための自己教師学習とデータ拡張技術」  
[https://jp.ricoh.com/technology/tech/134\\_speech\\_recognition\\_AI](https://jp.ricoh.com/technology/tech/134_speech_recognition_AI)



音声認識 AI は、人が発した言葉や声、会話を AI が認識・解析し、テキストデータへ変換して出力する技術です。昨今、会議中の字幕表示、議事録や報告書の作成等、ビジネスの現場で活用が広がっています。音声認識 AI を使えば、人がキーボードを使って音声をテキスト化する場合に比べて、早くテキスト化・システムへのデータ入力が可能となるため、業務効率化につながる非常に有効な入力方式として期待されています。

リコーの開発する音声認識 AI は、マイクから離れたノイズや残響を含む音声、人同士のくだけた話し言葉も高い精度で認識することができる独自の技術を有しており、正確さが求められるビジネスの現場での活用に適しています。また、この特徴はリコーが開発中の AI エージェントにも活かされています。リコーの AI エージェントは、音声会話の内容を AI が素早く認識・解析し、動的に次の質問を生成して対話を重ねていくことで、会話の中でお客様のニーズを捉え、的確な提案を行います。さらに、リコーはこのような音声認識 AI と他の AI 技術とを組み合わせ、お客様の働く場所での会話・音声データの利活用を支援し、新たな価値に変え、お客様の経営課題解決につながるデジタルサービスをご提供します。



「INTER\_SPEECH」は International Speech Communication Association (ISCA) が主催する音声言語処理分野における世界最大規模の国際会議で、今回が 25 回目の開催となります。採択された論文は、2024 年 9 月 2 日からギリシャ・コス島で開催される「INTER\_SPEECH 2024」で発表される予定です。

リコーは中期展望として、変容するワークプレイスにおいて一貫したサービスをグローバルに提供する「ワークプレイスサービスプロバイダー」を目指しています。これまでのオフィスに限らず、働く人がいるあらゆる場所や空間を「ワークプレイス」と捉え、AI やデータなどのデジタルの力でワークプレイスにおける生産性を向上するとともに、創造的なコラボレーションやイノベーションを創出し、はたらく人の創造力の発揮を支援します。これからも、デジタルサービスの提供を支える音声・言語・画像領域での AI の研究開発を推進し、業種業務に合わせて利用できる AI サービスの提供により、お客様のデジタルトランスフォーメーション (DX) を支援してまいります。

## ■関連リンク

技術ページ「音声認識AIのための自己教師学習とデータ拡張技術」

[https://jp.ricoh.com/technology/tech/134\\_speech\\_recognition\\_AI](https://jp.ricoh.com/technology/tech/134_speech_recognition_AI)

INTER\_SPEECH 2024のウェブサイト(英語)

<https://interspeech2024.org/>

※社名、製品名は、各社の商標または登録商標です。

---

## ｜ リコーグループについて ｜

リコーグループは、お客様のDXを支援し、そのビジネスを成功に導くデジタルサービス、印刷および画像ソリューションなどを世界約200の国と地域で提供しています(2024年3月期グループ連結売上高2兆3,489億円)。

”はたらく”に歓びを 創業以来85年以上にわたり、お客様の“はたらく”に寄り添ってきた私たちは、これからもリーディングカンパニーとして、“はたらく”の未来を想像し、ワークプレイスの変革を通じて、人ならではの創造力の発揮を支え、さらには持続可能な社会の実現に貢献してまいります。

詳しい情報は、こちらをご覧ください。

<https://jp.ricoh.com/>