

報道関係各位

Press Release

2024年1月30日

すべての産業の新たな姿をつくる



オーダーメイド AI 開発
『カスタム AI』

株式会社 Laboro.AI

日本語話者の音声を収録した話者認識用データセット 「Laboro-ASV」を無償公開

株式会社 Laboro.AI

代表取締役 CEO 椎橋徹夫・代表取締役 COO 兼 CTO 藤原弘将

オーダーメイドの AI・人工知能ソリューション開発および AI 導入コンサルティング『カスタム AI』を展開する株式会社 Laboro.AI（ラボロエーアイ、東京都中央区、代表取締役 CEO 椎橋徹夫・代表取締役 COO 兼 CTO 藤原弘将。以下、当社）は、話者認識技術の発展を目的に、日本語話者の音声を収録した音声データセット「Laboro-ASV」の無償提供を開始いたしました。

Laboro-ASV は、B-CAS カードによるアクセス制限がないワンセグ放送を利用し、2022 年 2～7 月にかけて放送されたテレビ番組の録画データから構成された音声コーパスで、142 人の話者による約 95 時間分の音声が収録されております。

なお、Laboro-ASV については、2023 年 9 月 27 日に開催された、日本音響学会 第 150 回（2023 年秋季）研究発表会にて報告いたしました。予稿は下記よりご覧いただけます。

https://laboro.ai/wp-content/uploads/2024/01/Laboro-ASV_ASJ2023.pdf

A graphic featuring the text 'Laboro-ASV' in a large, bold, black sans-serif font. The text is centered and surrounded by a decorative border of colorful speech bubbles and human silhouettes in various colors (purple, orange, yellow, red, green, blue).

Laboro-ASV

背景

話者認識は、幅広い応用可能性を有している技術分野であり、例えば音声アシスタントやセキュリティーシステムなどさまざまな用途で使われております。中でも話者認識の精度を向上させるカギは、対象となる話者が使用する言語向けに作成された高品質な音声データセットが利用できることにあります。ですが、外国言語の著名な音声データセットとしては、VoxCeleb (※1) や CN-Celeb (※2) などがある一方、日本語の話者認識の分野では、以下の条件を満たす有用なデータセットがありませんでした。

- (1) 日本語話者に限定、または重きを置いている
- (2) 研究用途のみならず、商用利用も無償
- (3) "in the wild"のデータを収集し、それによってスタジオ収録の音声よりも実環境に近い音声データで構成されている

日本語を主たる言語としない話者による音声データセットでは、話者認識の精度を著しく低下させる可能性もあり、特定言語に特化したデータセットの必要性を背景に、当社では日本語話者に特化した音声データセットを作成いたしました。

(※1) 2017年に発表された英語話者向けデータセット (<https://arxiv.org/abs/1706.08612>)

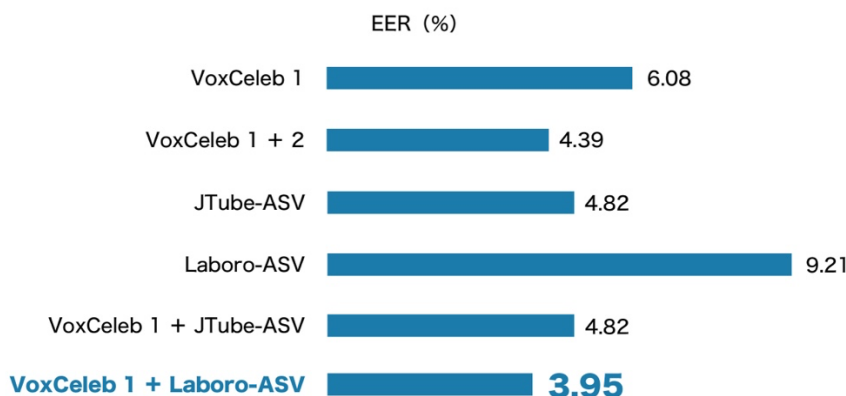
(※2) 2019年に発表された中国語話者向けデータセット (<https://arxiv.org/abs/1911.01799>)

Laboro-ASV の特徴

当社では「話者ごとの発話数」が話者認識のためのデータセットの有効性を決定付ける重要な要素であると考え、Laboro-ASV を、データセットとして充実させることを目的に、出演頻度に基づいて話者を選択し、話者ごとに十分な発話量を確保しております。話者1人当たりの発話数は、同種のデータセットで一般的に100~200であるところ、それをはるかに上回る450超の発話数で構成されております。

Laboro-ASV 比較実験について

Laboro-ASV の性能を評価するため、日本語話者照合タスクの観点から、JTubeSpeech (※3) データセットのトライアルセットをテストセットとして使用し、Equal Error Rate (EER) を評価指標として評価いたしました。なお、他のモデルとの公平な比較を保証するために、すべての実験は同じ設定で実施しております。また、speaker embedding の抽出と話者照合には X-vector/PLDA を使用しております。



Laboro-ASV の性能を日本語話者照合タスクの観点から評価した結果

EER は低い方が良いとされており、比較実験の結果、VoxCeleb1 と Laboro-ASV を組み合わせたデータセットで訓練されたモデルが最も良い性能を発揮できていることが上図からわかります。Laboro-ASV は、データ量の観点においては他の大規模データセットに比べて劣るものの、他のデータセットと併用することで日本語話者照合タスクの精度を向上させるために有効なデータセットであることを確認いたしました。

(※3) 音声認識と話者照合のために YouTube から構築される日本語音声コーパス

Laboro-ASV の利用について

Laboro-ASV に含まれる音声及びテキストデータの権利は、元のテレビ放送の著作権者に帰属していますが、著作権法 30 条の 4 に基づき、情報解析等の用途のために、商用利用および大学等の学術研究機関に対して無償で公開いたします。ただし、元のテレビ番組の音声を再構成し鑑賞する事を防ぐために、発話単位でランダムに並び替えられており、かつ番組名や放送局等の付加情報は含まれておりません。

ご利用にあたっては、当社 HP 内の特設ページ (<https://laboro.ai/activity/column/engineer/laboro-asv>) をご確認の上、お問い合わせフォーム (<https://laboro.ai/contact/other/>) よりお問い合わせください。

ご参考情報

■株式会社 Laboro.AI 会社概要

会社名：株式会社 Laboro.AI (ラボロ エーアイ)

所在地：〒104-0061 東京都中央区銀座八丁目 11-1

代表者：代表取締役 CEO 椎橋徹夫

代表取締役 COO 兼 CTO 藤原弘将

設立：2016 年 4 月 1 日

事業内容：機械学習を活用したオーダーメイド型 AI 『カスタム AI』の開発

カスタム AI 導入のためのコンサルティング

URL：<https://laboro.ai/>

株式会社 Laboro.AI は、オーダーメイドの AI ソリューション 『カスタム AI』の開発・提供を事業とし、アカデミア（学術分野）で研究される先端の AI・機械学習技術をビジネスへとつなぎ届け、すべての産業の新たな姿をつくることをミッションに掲げています。業界に隔たりなく様々な企業のコアビジネスの改革を支援しており、その専門性から支持を得る国内有数の AI スペシャリスト集団です。

以上

<本リリースに関するお問い合わせ>

株式会社 Laboro.AI マーケティング部 広報担当 中村麗奈

Mail：press@laboro.ai Tel：03-6280-6564（代表）