



A I は病気や障害のある人の助けになるか

進化を続ける人工知能（A I）による「認識」「生成」を生かし、障害のある人の意思伝達を図る取り組みも進められている。

脳性まひ 小樽の橋さん

「私に聞かれてくれたなまをへん3年間本当にありがたうございまして」。昨年3月、札幌・拓北養護学校で行われた卒業式。高等部卒業生の一人、橋脚聡さん（19）は、生徒、教員、父母らが居並ぶ体育館内に響いた。

脳性まひの橋さんは、体がこわばる「筋緊張」や、意図せずに体が動く「不随意運動」などの障害があるが、答辞の文章は視線で操作できるパソコンを使い自分で入力した。友だちや先生と話すのを楽しみ、自作の詩がコンテストで落選した時の悔しさも学校での思い出がつつらかった。

ただし、読み上げたのは、A Iが作り上げた橋さんらしい響きを持つ人工音声だ。橋さんは発声できるが、母・幸子さん（52）や拓北養護学校の担任教師だった夫野村さん（51）と日常的に接していないと即座に内容を理解するのは難しい。橋さんが作った文章をパソコンに入力し機械的に入音で読み上げるのは可能だが、橋さんらしきを感じてもらえない。

橋さんらしい響きのある人工音声は、IT企業のヒューマンテック



A I を使って作った自分の声質に似た人工音声で卒業式の答辞を述べる橋脚聡さん（中央、黒いマスク）＝昨年3月、橋幸子さん提供の動画から

「らしい響き」残し明瞭な声に

ノシステム（HITS、東京）と東北が開発した。HITSは2015年から特定個人の録音データから声質と話方を再現する人工音声を作り上げるソフトウェア「ボイスター」の製造と販売を手がけている。しかし、橋さんのように先天的に明瞭な発声ができな場合や、筋萎縮性側索硬化症（ALS）、嚥下が困難になった疾病や事故などで、喉の音が失われてしまうと、当人らしい声質と話し方の再現は困難だった。



A I を使って作った自分の声質に似た人工音声で卒業式の答辞を述べる橋脚聡さん（中央、黒いマスク）＝昨年3月、橋幸子さん提供の動画から

A I による音声作りのため、自分の声の収録に挑む橋さん（ヒューマンテックノシステム・渡辺聡さん提供）

を使った人工音声作りについて「明瞭性と（本人らしい）話感性、つまりコミュニケーションとアイデンティティのバランスを取ることに力を注いだ」と語る。思うように発声できない当事者の声と、健常者の明瞭な声を、それれあらかじめ定められたテキストを読み上げて収録。深層学習A Iの手法の一つ「敵対的生成ネットワーク」（GAN）を応用して、明瞭性と話感性を組み合わせた人工音声を作り、これをボイスターに発する仕組みだ。

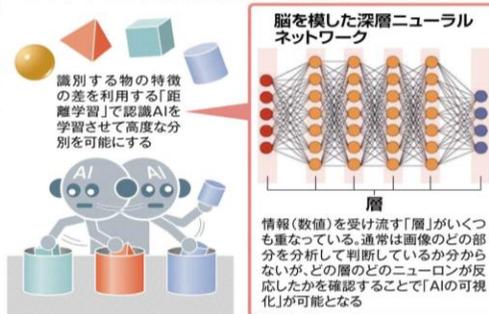
22年秋ごろから卒業式の前まで5回ほど声の収録が行われた。立ち会った教師の天野さんは「緊張し思うように言葉を出せない場面があったが、渡辺さんは辛抱強く対応してくれた」と振り返る。ただ、橋さんの場合、生まれつき発声に困難があったため、「明瞭性と響きは存在しない。正確性を上げれば、話感性、つまり橋さんらしさがなくなる。少し不明瞭な方がかえって魅力的でもある」と渡辺さん。

A Iによる自分の声の聞いた橋さんは「もう一人の自分がいるような気持ち。すべいことが起きるんだなと考えた。もっと広く使われることを願っています」と、自身の本音を一言ずつ振り絞るようにつぶやいた。HITSが支援したA Iによる音声合成技術開発は昨年3月で終了したが、渡辺さんは「当事者・関係者が、当者の声を基にした合成音声でコミュニケーションを取ることで、相手に対する好感度が上がったというデータが得られた。製品化に向けた準備を進めていきたい」と展望を示した。（編集委員 弓場敬夫）

廃棄物分別、顧客対応に応用

札幌のIT企業 独自に開発

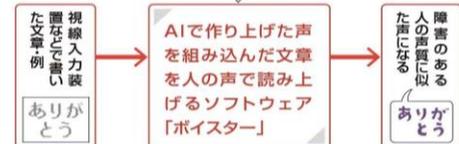
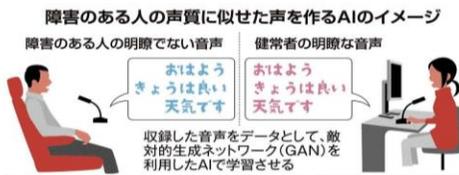
テクノフェイスが取り組む高度な分別AIの仕組み



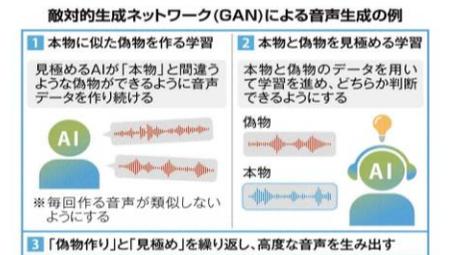
障害のある人の声質を持つ音声データつりに使われた敵対的生成ネットワーク（GAN）を、コンピュータで人間の脳を模した仕組み「深層ニューラルネットワーク（NN）」を利用した深層学習A Iの解は新たな手法が次々登場し、加速的な進化を遂げている。札幌のIT企業テクノフェイスも独自の認識A Iや、チャットGPTを応用した対話A Iを手がけている。

A I開発に携わる同社取締役の丸山哲太郎さん（46）は力を注いでいる分野の一つに、廃棄物を分別する作業を自動化するなどの分類A Iの高度化を挙げた。画像に何が写っているかを判断する技術開発は、この10年余りのA Iの進化に大きく貢献してきた。同社は物の特徴差による「距離学習」という方法を用いて、さらに高精度の分類が可能にしたとしている。新しい分類A Iは、道内の廃棄物処理業者で分別作業の実証実験が進められている。

さらに丸山さんは、チャットGPTを応用した製品「Techosearch」について「コトサチの特長は、ユーザーの意図を参照すべきデータがインターネット上にない場合は、的確な応答ができない。テクノサチは特定の企業が保有する製品情報などを取り込んで回答。これを利用して、コールセンターの顧客対応に役立っている」と語る。



※新エネルギー・産業技術総合開発機構（NEDO）のウェブサイトなどを基に作成



「生成」「識別」交互に学習 GAN

人工知能（A I）は、画像に何が写っているのか、どのような言葉が書かれたのかを識別する「認識A I」に加え、文章や画像、音などを生み出す「生成A I」が急速に発達している。いずれも、人間の脳を模した仕組みをコンピュータ上で再現する「深層ニューラルネットワーク（NN）」による深層学習A Iを使っている。人間の脳は、膨大な数の「ニューロン」と呼ばれる神経細胞が連結している。深層NNはコンピュータ上に仮想的なニューロンを多数置き、正解を導くようにニューロン間の情報（数値）のやりとりができるまで「学習」を繰り返す。

深層NNを使った生成A Iとして広く使われているのが「敵対的生成ネットワーク」で、英語表記の頭文字を取った「GAN（ジェネレーティブ・アドバーサリアル・ネットワーク）」と呼ばれる。GANは、A Iが本物と間違えるような文章や画像、音などを作り出し生成器と、作り出されたものが一定レベルに達して本物と見分けがつかない程度まで見極める識別器を交互に学習させることで、本物に近い文章や画像、音を生み出せるようにしている。障害のある人の声質を基にした「明瞭性の高い音声を作り上げた」のは、GANの応用技術だ。この技術の実用例としては、写真画像を特定の画像が描いた絵画のように変換する作業などが挙げられる。



人工音声などで円滑な意思伝達を支援します。