

(第3種郵便物認可)

サイ・テク こらむ 知と技の発信

埼玉大学・理工学研究の現場

【420】

近年、人工知能(AI)や深層学習は、その文字を目にしない日はないほど私たちにとって身近な技術となってきました。実際、深層学習の技術は、工業製品の検査など私たちの労働の一部を肩代わりするタスクにも応用されています。

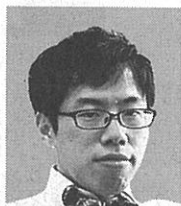
皆さんの中には、この急速な深層学習の普及の先に、社会がAIに乗っ取られてしまうと、と感じる方もいるでしょう。ただし情報科を専門とする私から言つと、今のところ深層学習が社会基盤に深く入り込むことはありません。その大きな理由が「ブラックボックス問題」です。ブラックボックス問題とは、深層学習が「なぜ」この結果を出したのか、中をのぞい

てもその根拠がまるで分からないというものです。このため、医療現場や大型の工場など判断一つで重大な損害が発生する分野では、深層学習の導入を躊躇(ちゆうちゆう)する人も多いです。

私の研究では、このブラックボックス問題を解決するため、人間に近い意思決定ができるように深層学習を誘導する仕組みの開発を行っています。この仕組みについて、私のメインテーマである音声の雑音除去を例に説明します。

音声の雑音除去とは、騒音に埋もれる音声から騒音のみを抑制して、目的の音声のみを取り出す技術です。これは音声通信や音声認識で使われる技術です。この雑音除去に深層学習を用いる研究がな

見える、深層学習を作る 杉浦 陽介 助教

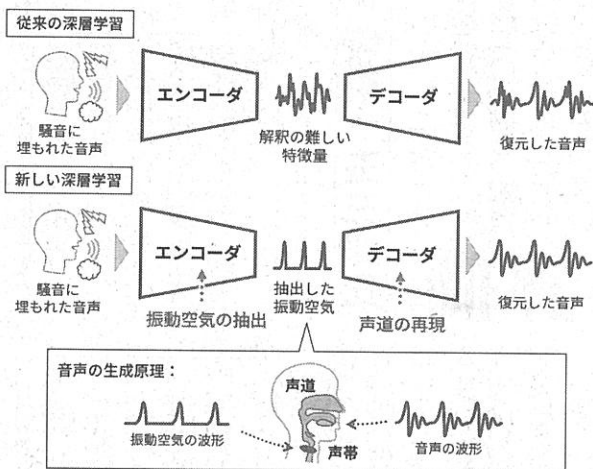


すぎらひら・よすけ 1986年生まれ。2013年大阪大学大学院修士(工学)。東京理科大学助教を経て15年4月から現職。専門は適応信号処理や最適化アルゴリズム。

されています。ここで、深層学習は騒音を含む音声を入力として音声のみを出力するように学習します。ただし、内部で何を基準に音声と騒音を区別している分かりません。そこで、音声が発生される原理を深層学習に教えることで、音声らしさとは何かを効率的に獲得するように誘導します。図は深層学習内の処理の様子を表しています。図解。

この深層学習はエンコーダとデコーダと呼ばれる二つのニューラルネットワークで構成されます。人間の音声は、声帯から発する振動空気が喉や口、鼻などの声道を通って声になります。この仕組みは騒音にはありません。そこで、

エンコーダでは入力から振動空気を取り出す役割を持たせ、デコー



ダでは空気から音声を変換する。声道の役割を持たせます。明確な役割・意味を持たせることで、深層学習は効果的な音声の選択を行い、より高品質な音声を生成できるようになります。

今は人間の物理的な仕組みを雑音除去用の深層学習に適用している段階ですが、今後もし人間の意思決定をうまくモデル化できれば、画像識別などの他のタスクにも応用し、深層学習のブラックボックス問題を解決できると考えています。そうなれば、深層学習は安心して使える技術として、さらに普及すると期待しています。