

ディープ・ラーニングで 環境音認識

齋藤 睦巳



(a) 従来の調査法



(b) 新しい調査法

図1 鳴き声を手がかりとした生き物調査

MATLABなら ディープ・ラーニングも簡単

● ディープ・ラーニングで環境音が認識できる

本稿では、MATLABによる環境音認識アプリの作成を解説します。

MATLABで用意されている関数や機能を用いることで、音データの解析やディープ・ラーニングによる学習・認識の実行、認識結果を確認するGUIアプリ作成を簡単に行うことができます。そこで具体的な例として、環境音の録音データから抽出された鳥の鳴き声を種類別に認識（識別）するアプリを作成します。

● 現地調査員の負担をAIで軽減できる

このアプリは、自然環境での生き物調査への応用を考えて作成しています。

絶滅危惧種の保護などを目的として森林での野生生物の生息状況調査が求められることがあります。しかし、夜行性の生物のように目視で姿を見つけにくい生き物の調査は、これまでは調査員が現地に入って鳴き声を聞いて確認していました。この方法では調査員の

負担が大きいため限られた範囲しか調査できません。録音した環境音をAIで解析して鳴き声を認識できれば、より広い範囲を効率的に調査できるようになります（図1）。

音を認識する方法

● 音を画像に変換してディープ・ラーニングを適用

音を認識するにはさまざまな方法があります。例えば、人の声を認識して文字にしたり内容を理解したりするには、言語の情報まで含めた高度な認識モデルが必要になります。

一方で、特定の音を認識する場合には、もう少し簡単な方法も利用できます。

例えば、音をスペクトログラムと呼ばれる画像に変換して、ディープ・ラーニングを用いて認識するという手法が用いられることがあります。図2にその概要を示します。録音された音声は前処理（雑音除去/音声区間検出）を行った後に画像化し、その画像を学習・認識するという流れになります。

音は時系列信号ですが、周波数分析を用いてスペクトログラムとして画像化できます。

スペクトログラムは、グラフ上で横軸に時間、縦軸に周波数をとり、周波数成分の強さを濃淡で表示したもので、声紋とも呼ばれます。

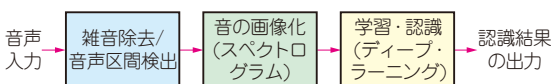


図2 音の認識の流れ