

03

ディープラーニング (DL)
とは

ディープラーニングも機械学習の手法の1つですが、学習モデルが正解しやすいように人間がデータを加工するのではなく、学習モデルが自ら「特徴量」を抽出して学習していくという点に大きな違いがあります。

認識にすぐれたディープラーニング

Section02では触れませんでした。従来型の機械学習には大きな欠点がありました。それは、手元のデータをいきなり入力することができず、モデルが学習しやすいように、あらかじめ人間がデータを加工しなければならないという点です。ここでの加工とは、特徴の強弱を表す数値（**特徴量**）を数学的に算出することです。手書き文字の認識であれば、画像の「線の曲がり具合」「字の輪郭」「線のつながり方」、音声認識であれば「声の高さ」「声の大きさ」などを考えるとわかりやすいでしょう。

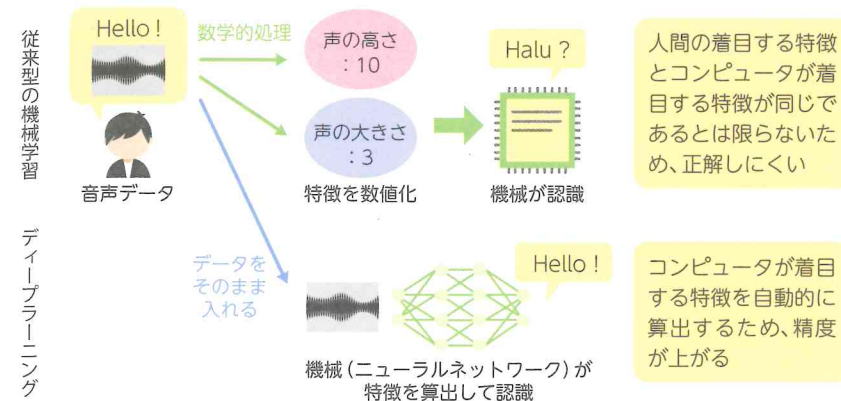
しかし、モデルが学習しやすいであろう特徴量を算出するのは、非常に難しいことでした。数式を駆使してどうにか数値化したとしても、モデルがその特徴を使ってうまく判別してくれるかどうかはわからないからです。一昔前の機械翻訳やカーナビの音声認識が使い物にならなかったのは、このような従来型の機械学習を使っていたためです。

そんな中、画期的な技術として台頭したのが**ディープラーニング**でした。ディープラーニングとは、脳の神経回路を模した**ニューラルネットワーク**と呼ばれる学習モデルを用いる機械学習のことです。入力層と出力層の間にある「**隠れ層**」が「深い」ことから、そう名付けられています。隠れ層とは「入力層から受け取った情報をさまざまな組み合わせで伝えていき、出力層に役立つ形に情報を変形して渡す」という役割を持つ層のことです。

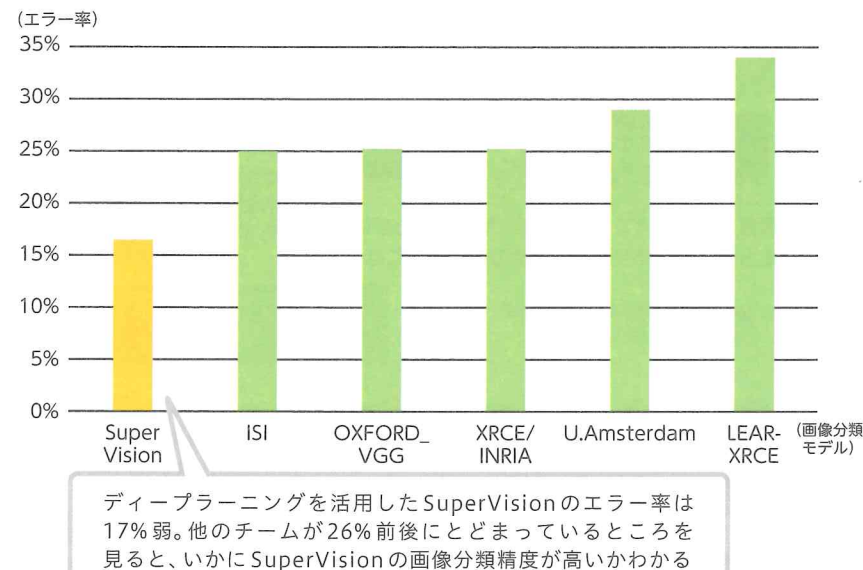
ディープラーニングが画期的だったのは、**最適な特徴量を自動的に抽出**するという点です。2011年に音声認識の分野で従来型の機械学習を大幅に上回る精度を実現したのをきっかけに、2012年の画像分類コンペティションILSVRC

(IMAGENET Large Scale Visual Recognition Challenge) でも大幅な性能改善が実現されました。さらに2015年には、ディープラーニングを用いて開発された画像認識プログラムが人間の誤認識率といわれる5%を切るなど、その性能に拍車がかかっています。

■ 特徴量の判別は難しい



■ ILSVRC-2012における画像分類モデルの比較



出典 : <http://image-net.org/challenges/LSVRC/2012/ilsvrc2012.pdf>