



Feature

# AI活用最前線

——ここまできた画像認識

Googleの「AlphaGo」が囲碁のプロ棋士を打ち負かし、インテリジェントな音声認識機能を備えたスマートスピーカー「Amazon Echo」「Google Home」といった製品が相次いで登場するなど、さまざまな話題を振りまいている人工知能(AI)。その中であって、自動運転技術にも用いられ、AI応用の典型例ともされているのが、ディープラーニングを用いた画像認識です。本稿では、ディープラーニングによって飛躍的な発展を遂げつつある画像認識の現在、そしてこれからについてお伝えします。

青木 功介(あおき こうすけ)

株式会社インテック  
先端技術研究所  
シニアスペシャリスト/博士(工学)

1992年 富山大学大学院理学研究科修士課程修了、インテックに入社。同年 インテック・システム研究所に転出。現在まで、コンピュータグラフィックス、画像処理、ヒューマンインタフェースの研究開発に従事。2010年 色情報を利用した画像処理の研究により、博士(工学)取得。2016年 アスベスト迅速検査により環境賞(環境大臣賞)を受賞。



## ディープラーニングによる 画像認識のブレークスルー

画像認識分野における機械学習技術の導入は古く、10年以上前にさかのぼります。

ただ、機械学習と言っても、かつての画像処理に用いられていた技術は、今日のディープラーニング（深層学習）とは異なるもので、その処理能力には限界がありました。

画像認識では基本的に画像としてとらえた物体の特徴（特徴量）を抽出し、解析することによって「それが何か」を識別します。これまでの機械学習を用いた画像認識のしくみでは、認識に用いられる特徴量を人が設計していました。例えば、「トラ」と「ネコ」との差異（特徴量の差異）を見つけて識別のルールを決め、プログラムを組むところまでを人が行っていたわけです。そしてコンピュータは、多くの画像から、どの程度の特徴があつたら「これはネコ」「これはトラ」と判断するか、その「しきい値」を割り出し、「トラ」か「ネコ」かを識別していたのです（図1）。

このような状況をガラリと変容させたのが、ディープラーニングです。この技術によって、コンピュータが物体の特徴や識別のルールまでを自動で学習するようになりました。ディープラーニングの技術を使うことで、コンピュータに画像データを読み込ませるだけで、「ネコ」や「トラ」の特徴を自動で学習させ、識別させることが可能になったのです（図2）。

ディープラーニングの進化により、画像認識の領域においてこれまで困難とされてきた「一般物体認識」が可能になると注目されています。

一般物体認識とは、いわゆる普通の写真に写る、人や椅子、机、窓といった物体のすべてを識別・認識する処理のことです。人であれば、写真を見て、そこに写り込んでいるモノが何か、すぐに理解できます。しかし、コンピュータに同じことをさせるのは非常に困難でした。というのも、コンピュータにはモノに対する「概念」がないからです。

図1：旧来型 AI（機械学習）を使った画像認識のしくみ

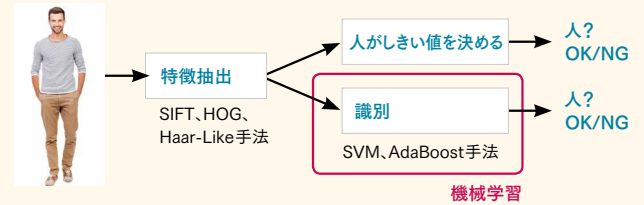
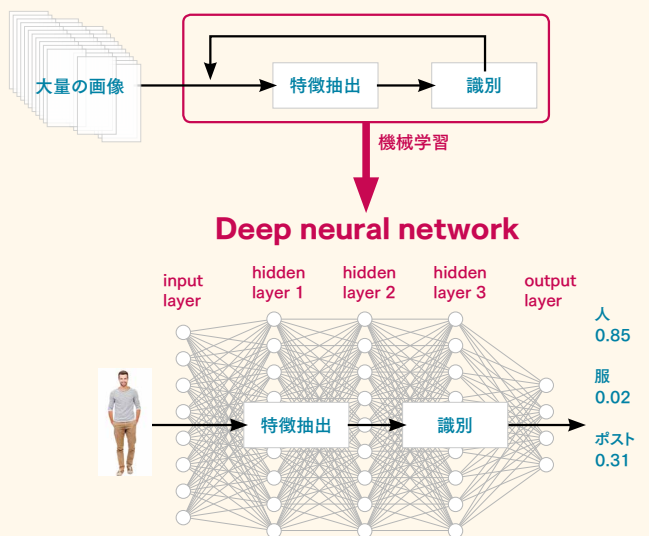


図2：ディープラーニングによる画像認識のイメージ



人は「椅子」という概念を持ち、さまざまな形状の「椅子」を「椅子」として認識できます。ところが、コンピュータには「椅子」の概念はありません。写真の中にある形状の異なる椅子を「椅子」として認識させるには、「椅子は人が座るもの」「椅子の前にはテーブルや机がある」といった常識をコンピュータに教え込む必要がありました。要するに、コンピュータに「一般物体」を認識させるには、「教え込む」という作業を実世界のあらゆるモノについて行う必要があるのです。そのため、ディープラーニングの登場以前は、コンピュータによる一般物体認識は困難とされていたのです。

ところが、ディープラーニングによって、コンピュータは数多くの画像からモノの特徴を自動的に学習し、それが何かを認識できるようになりました。つまり、大量の画像データから、「人とはこういうもの」「机とは

こういうもの」ということを学習し、それぞれを識別できるようになったのです。これにより、画像認識の長年の懸案であった、「コンピュータによる一般物体認識」の可能性が大きく広がっているのです。

## 人を超える認識精度

ディープラーニングの威力は、画像認識の世界的なコンテスト「ILSVRC(ImageNet Large Scale Visual Recognition Challenge)」でも実証されつつあります。

このコンテストのひとつが、数万点の画像をおおよそ200程度のカテゴリーに分類し、それぞれの認識率を競い合うというものです。図3のグラフは、そのコンテストの年度ごとの結果（誤認識率）を示したものです。2012年にディープラーニングによる画像認識のブレイクスルーがあり、ご覧のとおり、この年から画像の誤認識率が大きく下がり始めました。そして2015年にはコンピュータの精度が人(5%)を上回り、2016年には誤認識率が3%に達しています。

画像認識・画像処理は、AIが注目される以前から研究開発が進んでいます。工場のラインで不良ロットを発見するといった、あるルールの中において特定の画像を認識・処理するスピードは、以前から人の能力を超えています。一方、一般物体認識のように人が自然に行っている画像認識はコンピュータでは実現することができませんでした。しかし、そうした状況を、ディープラーニングが打破しつつあるのです。

## 進む画像認識AIの実践活用

ディープラーニングを用いた画像認識は、さまざまな分野へ適用することのできる技術です。そのため、この技術を使った画像認識は、さまざまな領域で実用化が始まっています。その中でも研究が盛んなのが医療分野で、画像解析によって腫瘍が悪性か否かを判断する研究が進展しています。

ディープラーニングの登場以前は、これをコンピュー

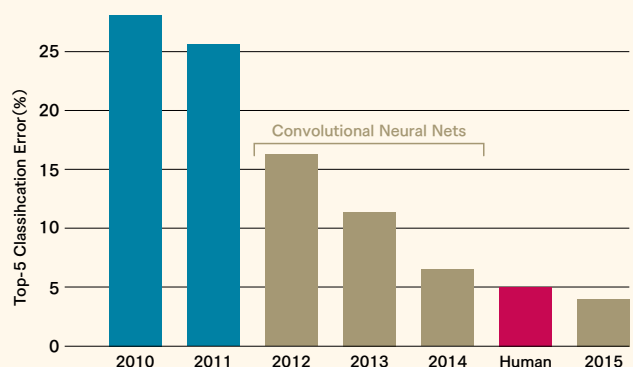
タに判断させるのは至難でした。しかし今日では、腫瘍に関する大量の画像をコンピュータに読み込ませることで、コンピュータが「悪性かどうかを識別するために、どこを見ればいいのか」までを判断できるようになってきたのです。

また、「顔認識」におけるディープラーニングの応用も進んでいます。顔認識技術も古くから研究開発が進み、ディープラーニングを使わずとも、かなりの精度で個人の識別を行うことができます。

しかし、さらに精度の高い認識や、年齢変化、環境変化に対応した認識には、ディープラーニングが使われるケースが増えています。ディープラーニングは学習により精度を向上させることができ、特徴をとらえることができるからです。さらには顔認識に使われるセンサーも発達しており、「iPhone X」に搭載されているセンサーは、立体イメージとして人の顔をとらえ、顔の凹凸まで計測（センシング）することができます。こうしたセンサーの発達により、顔認識の精度はより一層向上すると考えられます。

もうひとつ、AIによる画像認識の有名な応用例と言えるのが、Googleの囲碁プログラム「AlphaGo」です。AlphaGoは、膨大な数の棋譜パターンを画像として学習し、対局を勝利へと持ち込むしくみになっています。さらに、コンピュータ同士の対戦によって、大量の新しい棋譜を猛スピードで生み出し、自己を強化してい

図3：  
画像認識技術による誤認識率の推移  
ILSVRC(ImageNet Large Scale Visual Recognition Challenge)の結果



くことができます。つまり、AlphaGoは数千年に及ぶ囲碁の歴史が生んだ棋譜を超える量をきわめて短い時間で生成し、学習していくことができるのです。

もちろん、AIを使った今日の囲碁AIにも限界があります。例えば、囲碁のルールがいきなり変更になったら、対応できません。現在のAIは、過去のデータに基づいて、決められたルールの中で最適解を導出する機能しかないのです。そのため、現状では自ら課題を設定したり、新しいルールを作ったりすることができません。AIがそうした人と同等の能力を身に付けるまでには、一定の時間が必要であると考えられています。

## センサーの高度化で 画像認識がさらに進化

前述のとおり、ディープラーニングなどのAI技術は、画像認識技術の進化・発展に大きく貢献しています。もっとも、画像認識技術を高度化させているのは、AIだけではありません。解析用のデータを収集するセンサーもまた進化し、それが画像認識・画像処理の可能性を押し広げています。

今日の高速カメラを用いれば、1/1000秒のスピードで映像を撮影できます。それに対応するように、画像処理のスピードも増しており、2020年の東京オリンピック・パラリンピック競技大会では、高速カメラで撮影した画像を基にしたまったく新しい映像を観ることができるかもしれません。また、高速カメラで撮影した映像をAIで解析し、勝つための戦略に役立てようとするチームも出てくる可能性があります。

一方ハイパースペクトルカメラも、画像認識・処理の適用範囲を押し広げるものです。これまでのカメラは、三原色（赤・緑・青）のスペクトルしかとらえていませんでした。しかし、ハイパースペクトルカメラは、それをはるかに超える数の色をとらえ、記録することができます。そのカメラで撮影した画像を分析すると、モノの材質を見分けたり、モノに海水が染み込

んでいるかどうかを判断したり、金属の老朽化度合を判定したりすることができるようになります。ですから、建物の非破壊検査から橋梁の老朽化検査、さらには宝石の真贋検査に至るまで、実に多様な場面での活用が考えられるのです。

## インテックが目指すもの

約30年の長きにわたり、インテックではこれまで説明してきたような画像認識・画像処理の技術研究・開発に取り組み、すでに多くの実績を積み上げています。例えば、アスベストの画像認識技術の開発では、環境大臣賞を受賞しています。また、特定の細胞を見つけるための画像処理技術や、病理診断・農作物発育診断に向けた画像解析技術も開発・提供しています。これらの研究開発では、常に最新の技術を取り入れてきましたが、そのスタンスに変化はなく、今日ではディープラーニングを用いた画像認識ソリューションの研究開発に意欲的に取り組んでいます。

その一例が、視覚障がい者による買い物支援のソリューションです。これは、ディープラーニングによる画像認識を応用したしくみで、スマートフォン（以下、スマホ）用のアプリと、バックエンドのシステムから成ります。

このしくみを使うことで、例えば、商店を訪れた視覚障がい者は、スマホのカメラを店内へ向けるだけで、アプリが商品棚の位置を認識し、音声でその位置を知ることができます。また、その棚へスマホのカメラを向けると、そこにどのような商品があるか、さらに棚の商品を手に取りスマホをかざすと商品名・賞味期限・含有添加物などを音声で案内します。もちろん、視覚障がい者にとって、スマホカメラのフレームを商品に合わせることも簡単ではありません。そこで、手持ちの商品をカメラのフレームにスムーズに合わせるための技術も独自に開発しています。

インテックは、お客さまのニーズにいち早くお応えするために、より一層の力を注いでいきます。