

AIってなんだっけ

技術部 製品開発グループ 近藤 佳秀

■はじめに

10月5～7日にポートメッセなごやで開催された「日本木工機械展2023」で気になる検査装置がありました。それは、カメラとAIモジュールが一体となったセンサーを本体とし、簡単な学習だけで検査対象の合否や文字の読み取りを判別できるという汎用性の高いもので、価格も比較的安いので、木製品の工場などでも導入しやすいのではないかと感じました。また、AIが一般的な技術として使われ始めていることを実感しました。

近年、様々な分野で話題となっているAIがどのようなしくみで実現されているか改めて調べてみたところ、とても奥が深い技術で簡単に説明できるものではないのですが、ここでは木材表面の節などを判別するような画像処理に用いるAIのしくみを中心にできるだけコンパクトに説明します。

■従来のプログラミングとAIの違い

AIとはArtificial Intelligenceの略で日本語では人工知能と訳されています。AIの歴史はさておき、現代のAIでは、機械学習、中でも深層学習と呼ばれる手法が利用されています。深層学習がなぜそう呼ばれるのかは後ほど述べるとして、まずは従来のコンピュータによるプログラミングと何が違うのかを説明します。

従来のプログラミングは入力をどのように計算すれば出力になるかを考え、それを直接プログラミング言語で記述します。これをコンピュータで計算できるように変換（コンパイル）し、アプリケーションとして登録します。

しかし、AIでは問題（入力）と答え（期待される出力）をコンピュータに入力し、正しい答えが導き出せるよう学習させます。計算方法は、AIが自ら導きます（図1）。

なぜ、こんなことができるのか。その秘密は深層学習を用いるAIの基礎となるニューラルネットワーク（以下、N.N.）という構造にあります。

■AIの基礎はニューラルネットワーク

N.N.は、生物の脳の構造、働きをモデル化してコ

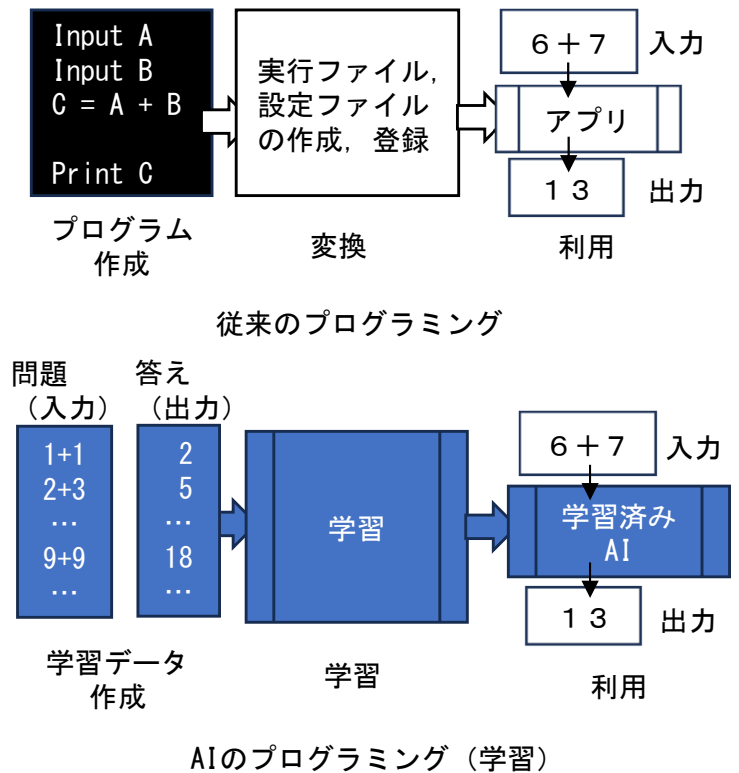


図1 従来のアプリとAIの利用法の違い

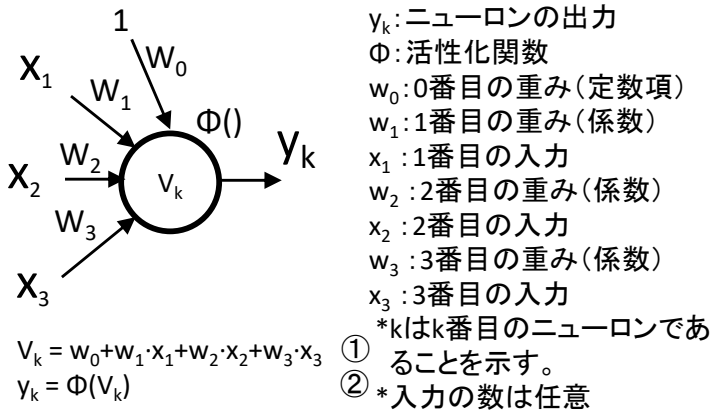
ンピュータの中に再現したものです。生物の脳は、ニューロン（脳細胞）が非常に数多く繋がってできています。ニューロンは周囲のニューロンから情報（刺激）を受け取り、情報を咀嚼（演算）して周囲の（他の）ニューロンに情報を伝えます。

ニューロンのネットワークが目で見えた映像や音、匂い、触感覚といった情報から外界の状況を記憶し、分析し、判断することで、生物は行動を起こします。

また行動の結果、失敗したときは学習し、行動を修正します。学習の結果は、ニューロンの繋がり方の変化として保存されます。

N.N.はニューロンをモデル化した人工ニューロン（図2）を数多くつなぎ合わせて構成します。

人工ニューロンは、他の多くの人工ニューロンから情報（数値）を受け取り、内部で演算して周囲の他の人工ニューロンに数値を伝えます。一つ一つの人工ニューロンは、N.N.の中でノードと呼んでいます。ノードのネットワークが例えば画像を分析し、判断し、結果を出力します。正しい結果を出せるように学習すると、ノードの繋がり方が変わります。



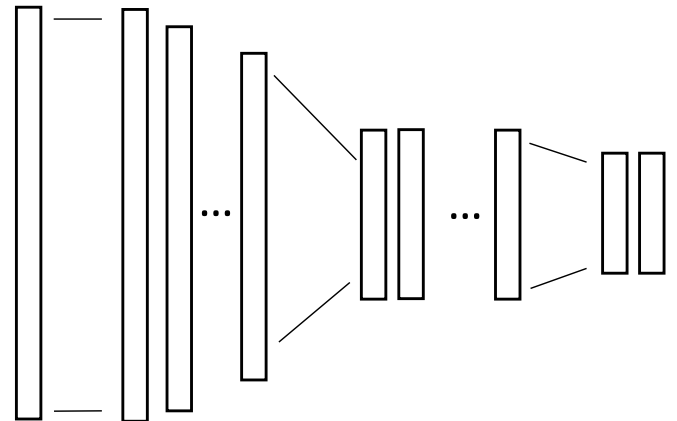
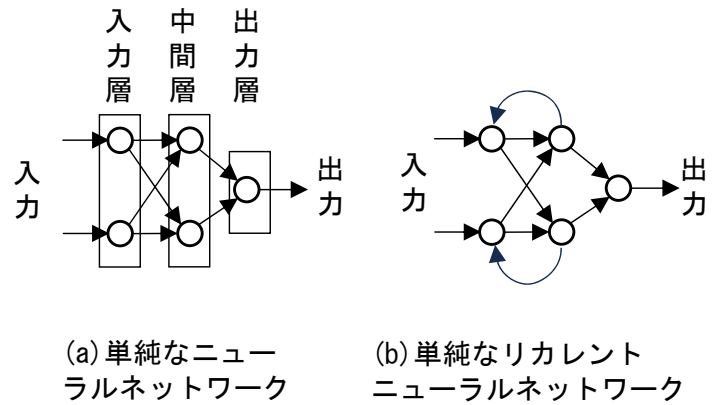
出力の計算方法

- ①. 入力ごとに異なる重み(係数 W)を掛けて足し合わせる。
- ②. ①の結果に活性化関数 Φ を作用させて出力とする。

図2 人工ニューロン(ノード)

生物の脳と違ってN.N.では学習前のノードの繋ぎ方を人間が決めます。画像処理用のAIでは、互いに接続しないノードの組を層と呼び、各層を入力から出力へ後戻りすることなく数値が伝わっていく構造でN.N.を構成するのが一般的です(図3(a), (c))。

参考までに、図3(b)のような後戻りする繋ぎ方もあります。こちらは、文章や音声など、少し前に書いてあったり、少し前に聞いた言葉が今の言葉の意味を理解するために必要な場合に用いられ、リカレントニューラルネットワークと呼んでいます。



入力層 コンボリューション層(4~5層) 全結合層(40~100層) 出力層

(c) ディープニューラルネットワーク(画像処理用)

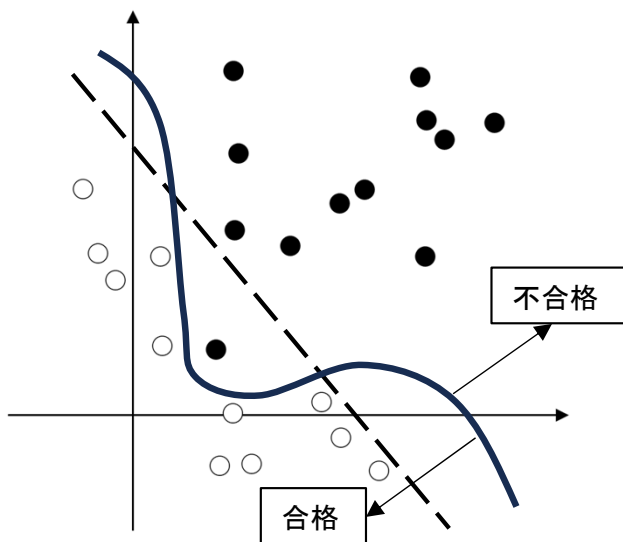
- ・ 四角の大きさは層内のノード数に相当する。
- ・ 入力層のノードの数は十万程度
- ・ コンボリューション層のノードは、前段の層の十個程度のノードからだけ入力を受けとることで、画像の特徴抽出を行う
- ・ 全結合層各層のノード数は数百程度。ある層は前段の層のノードすべてから入力を受けて後段のすべての層に出力を伝えることで、複雑な推論を行う。
- ・ 出力層は全結合層の推論結果をコンピュータが計算しやすい形式に整える

図3 ニューラルネットワークの構成例

現代のN.N.は、100層を超える多層構造が一般的ですので、これをディープニューラルネットワーク(深層ニューラルネットワーク)と呼び、その学習技術を深層学習(ディープラーニング)と呼んでいます(図3(c))。

■ニューラルネットワークはノードが記憶する

N.N.ではノードが重要な役割を果たします。ノードは、多数の入力(数値)を受け入れて内部で演算し、一つの出力(数値)を出します。ここで、重みと活性化関数が重要な役割をしますので、それぞれ



————— 活性化関数がある場合
 - - - - - 活性化関数がない場合
 ○: 合格, ●: 不合格
 活性化関数があると、複雑な曲線で合格の判別ができる。

図4 活性化関数の役割(イメージ)

について説明します。

ノードの演算手順は図2のとおり、

1. 入力ごとに異なる重み（係数W）を掛けて足し合わせる。
 2. 1に活性化関数Φを作用させて出力とする。
- です。

N.N.は学習してから使うと述べましたが、学習とは、重みを変化させることです。N.N.内のすべてのノードの重みが適切に変更され、期待される出力が出せるよう決定できた状態を『学習ができた』と呼んでいます。このようなしくみのため、学習に用いなかった入力に対しても概ね正しい出力が期待できます。つぎに、活性化関数は、N.N.に図4のような複雑な判断をさせることが目的です。活性化関数がないと、図4の例であれば合否判別区切りが単純な直線になってしまいます。詳細については、参考文献などを参照して下さい。ここでは、よく使われる活性化関数の名前と特徴を挙げておきます。

シグモイド関数：古くから使われている。N.N.の複雑さが増すにつれ学習に失敗することが多くなった。

ReLU関数：現在最も多く用いられている。欠点もあるので、新しい関数の研究は盛んに行われている。

■ニューラルネットワークの計算には時間が掛かる

N.N.はノード一つ一つの出力に複雑な計算を必要とするうえ、ノードの数が100万を超えることも現代では普通ですので一つの結果を出すためには膨大な計算が必要です。このような膨大な計算が可能となったのは、コンピュータの発展のおかげです。なかでも、並列演算技術の進歩が現代のAIの発展を支えてきました。その詳細を述べることは筆者の能力を超えていますので、ここでは並列演算装置であるGPGPUが行っている並列コンピューティングの考え方を示します。

通常のコピーは中央演算装置CPUのみで計算をするので、図5（上）の例のように図2の V_k にあたる3回の掛け算を順番に行った後、掛け算の結果を足し合わせるという計算に4ステップ必要です。

一方、AIの計算によく用いられるコンピュータのようにGPGPUがあると、掛け算はGPGPUが3個同時に計算してCPUに渡すため、CPU内では足し算だけすればよく、GPGPUと合わせても2ステップで計算が終わります。実際はもっと複雑ですが、基本の考え方はわかっていただけだと思います。

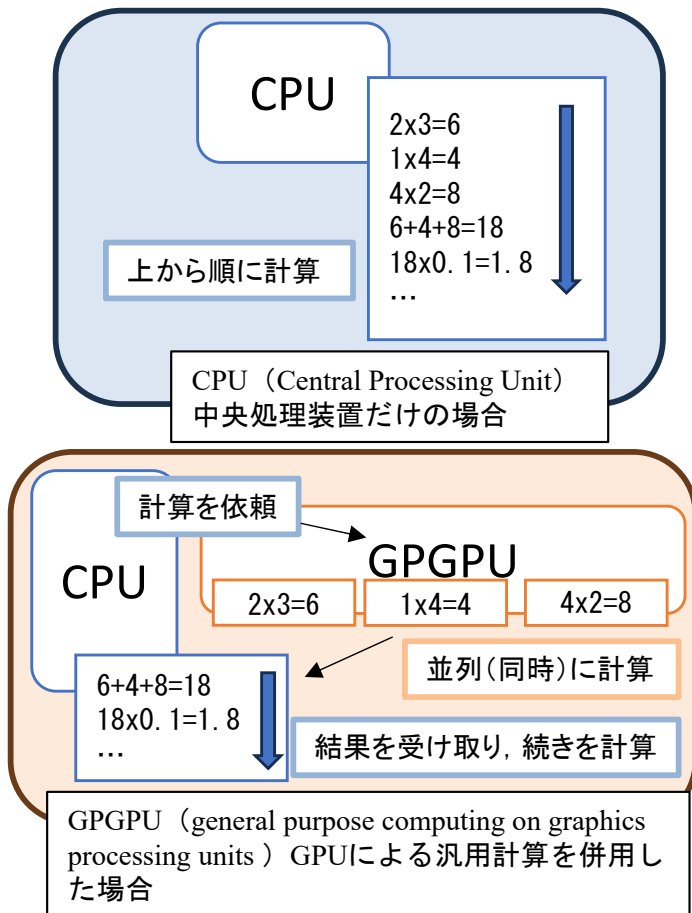
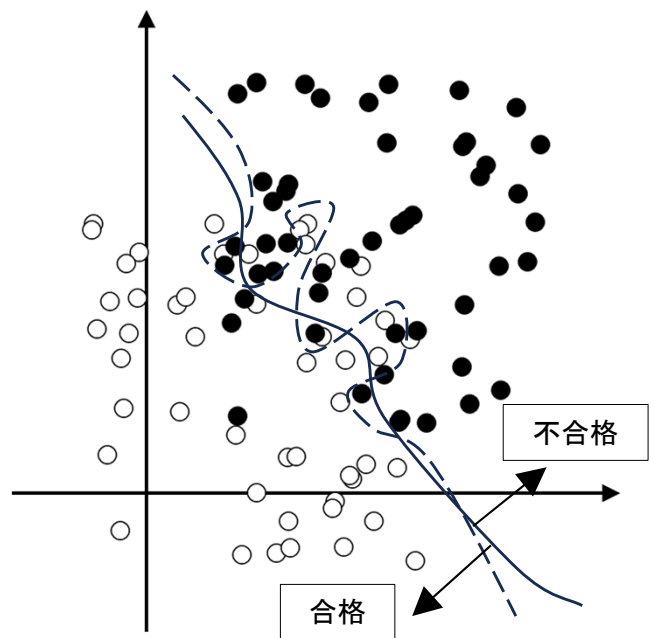


図5 GPGPUの動作イメージ



—— 期待する合否判別区切り
 - - - - 過学習による合否判別区切り
 学習データ ○：合格，●：不合格
 過学習したときは、判別条件が過敏なため判定結果が不安定となり、判別成績が下がる。

図6 過学習のイメージ

ちなみに一つ一つの計算に掛かる時間は、今のコンピュータではCPU、GPGPUともに1億分の1秒以下です。AIの推論に掛かる時間を概算すると、CPUだけだと1秒程度、GPGPUがあると0.1秒程度となり、冒頭のセンサーのような装置が実現できています。

■ニューラルネットワークの学習には落とし穴がある

ここでは、N.N.の課題の一つである過学習の捉え方について説明します。過学習とは、学習のために計算を繰り返すとある時点を超えて、学習に使ったデータの判別成績は良くなるものの、実際に判別させてみると不合格を合格としてしまう割合が増えてしまう現象です。この現象が起こる原因は、N.N.が学習データに過剰に適合することです。

図6に示したように、合否が入り組んだ結果を判別するためにN.N.を学習させたとき、ある時点では実線のような判別線で判断していたところが、学習のための計算を進めたところ、点線のような判別線を描いてしまうと考えてみてください。N.N.は学習データの判別成績を上げようとより複雑な判別線を見つけます。その結果学習データの判別成績は良くなるものの、入力値のわずかな差によって判別結果が大きく左右されることとなり、実際の判別作業では不合格品を合格とする割合が増えてしまうと考えられます。

このように考えると過学習の弊害を減らすためには、学習に用いるデータについて人間が判別しやすいものと人間が判別しにくいものの割合を変化させて学習させたり、学習を途中で打ち切って実際の判別結果が最も良い状態で使うなど、学習のさせ方に工夫が必要であると考えられます。

■ニューラルネットワークと他の機械学習の違い

ここまで、N.N.について説明してきましたが、今ほど話題になる前のAI研究では、深層学習以外の機械学習が主に使われてきました。その成果としては、カメラの顔検出等があります。

もともとN.N.と機械学習は全く別の技術として研究されていましたが、ノードが行っていることが機械学習の分野でよく知られている分類手法であると明らかになったことで研究が進み、N.N.による深層学習は機械学習の一分野であることがわかりました。

しかし、他の機械学習と大きく異なる深層学習の特徴が、特徴量を自ら見つけ出すという点です。

特徴量を説明するのは難しいのですが、例えば、

日本農林規格に従い製材の目視等級区分を決めることを考えると、節の大きさや位置、割れの長さ、腐れの程度等が特徴量に当たりますが、機械学習ではこれを数値にする必要があります。

また、目視等級区分を自動化したいときには、画像から節などの欠点をコンピュータが判別しなければなりません。しかし、例えば機械学習で画像から節を抽出する場合、節の特徴量をどのように数値化すれば良いのでしょうか。人の顔検出では、膨大な経験と試行錯誤、高度な数学的処理により特徴量が数値化されていますので、カメラなどでは写真の中の顔を非常に高い確率で認識できますが、節の特徴量を人の顔のように数値化するためには、これから多くの人員と時間を費やす必要があります。

一方、N.N.では節の判別を学習することができます。手順の一例は、

1. 節のある木材の写真を数十～数千枚用意する。
2. 写真一つ一つについて節の場所をマーキングし、“節”と名前をつける（この作業をアノテーションと呼んでいます）
3. 1を入力データ、2を期待する出力として、N.N.を学習する

となりますが、この手順の中に特徴量の数値化はありません。特徴量はN.N.が自ら抽出し、ネットワークの中にノードの重みとして保存されています。

■おわりに

AI (N.N.) がどのような計算方法で、結果を出しているか駆け足で説明しました。

最近のAIはパッケージ化が進み、学習する時の条件（活性化関数の指定や、学習率の指定などがあります）を細かく指定する必要は少なくなってきました。しかし、ここまで述べたような知識はパッケージ化されたAIを使う上でも有用であると思います。

林産試験場では、製材品の欠点判別やキノコの等級判別といった林産業で必要とされるAI技術の研究開発を行っています。

ご要望などがありましたら、技術部製品開発グループ（Tel. 0166-75-4264（ダイヤルイン））まで問合せ願います。

■参考文献

- 1) 例えば、赤石雅典：最短コースでわかるディープラーニングの数学、日経BP社（2019）等