

## 綜 説

## 深層学習の基礎と腹部画像診断への応用

小 松 大 祐

信州大学医学部画像医学教室

## Basics of Deep Learning and Application to Abdominal Diagnostic Imaging

Daisuke KOMATSU

Department of Radiology, Shinshu University School of Medicine

**Key words:** diagnostic imaging, computed tomography, magnetic resonance imaging, artificial intelligence, machine learning, deep learning

画像診断, CT, MRI, 人工知能, 機械学習, 深層学習

## I はじめに

医用画像の分析と解釈は、放射線科医の基本的な業務とされてきたが、これらのタスクを効果的にコンピュータで自動化することは歴史的に困難であった。しかし最近、深層学習（ディープラーニング）と呼ばれる手法を使用したコンピュータサイエンスの研究者が、画期的なパフォーマンスの向上を実証した結果、画像分類、物体検出、言語解析、音声認識などの様々な領域で応用が進んでいる。

非医療領域の画像に対する畳み込みニューラルネットワーク（Convolutional Neural Network: CNN）を使用した深層学習の成功は、医療画像の自動分析に発展する希望をもたらしている。最近、これらの深層学習アルゴリズムは、マンモグラフィでの乳がんの検出<sup>1)</sup>、胸部CTによる間質性肺疾患の分類<sup>2)</sup>、磁気共鳴画像（MRI）による脳腫瘍セグメンテーション<sup>3)</sup>などのいくつかの医療画像を用いた臨床研究の報告がみられている。

本稿では、人工知能の主要な用語の解説から、従来のルールベースのシステムから深層学習に至る歴史的な経緯を確認するとともに、ニューラルネットワークの基本構造と畳み込みニューラルネットワークアーキテクチャ、深層学習のための技術およびデータの条件、その評価方法について説明する。また、深層学習を用

いた自験例をお示しする。

## II 人工知能

人工知能は、通常人間の知能を必要とするタスクを実行するシステムを作成するコンピュータサイエンスの分野である。これは、多種多様な領域や技術を包括する広義の用語となる。本稿では、機械学習の一種として、両者に共通する概念を解説するとともに深層学習に焦点を当てる（図1）。

## III 機械学習

機械学習は人工知能のサブフィールドであり、明示的なプログラミングではなくデータからパターンを学習することによってタスクを実行するようにトレーニングを行う。

従来の機械学習では、各領域の専門家が対象となるデータの特徴を識別し抽出を行った後に、統計的手法を使用して、これらの特徴量に基づいて対象データの分類や回帰を行う（図2）。

例えば、画像を分析する目的で、画像処理の専門家がアルゴリズムをプログラムして、入力画像をエッジ、グラデーション、テクスチャの基本要素に分解する。次に、特定の画像におけるこれらの特徴量の統計学的に分析して、画像の分類または解釈を行う。

ただし、多くの複雑な画像分類タスクの場合、機械学習アルゴリズムに最適な特徴量を抽出する方法は、専門家でも通常はわからない。したがって、コンピュータシステムが、望ましい出力結果を得るために

別刷請求先: 小松大祐 〒390-8621

松本市旭3-1-1 信州大学医学部画像医学教室

E-mail: kdaisuke1981@yahoo.co.jp

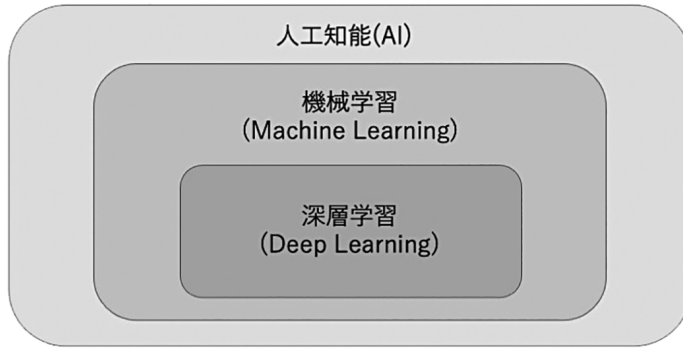


図1 人工知能と機械学習、  
深層学習の概念

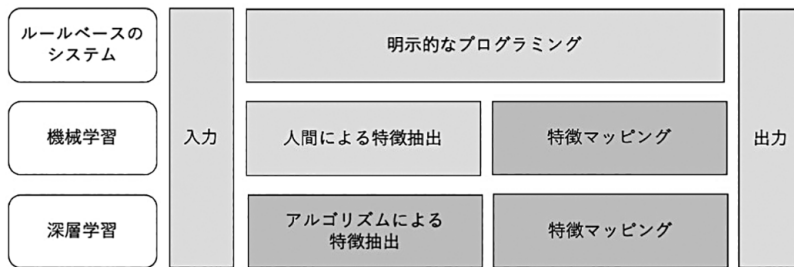


図2 ルールベースシステム、  
機械学習、深層学習の違い

は、人為的に選択された特徴量を学習するだけでなく、特徴量自体を抽出し、学習および最適化することが望ましい場合がある。

#### IV 深層学習

深層学習は、アルゴリズム自体がデータ構造の階層を反映する特徴を抽出し学習を行う。従って、形状、病変、臓器などのより複雑な特徴の構成要素として単純な特徴量（強度、エッジ、テクスチャなど）を抽出および学習することにより、end to endのアプローチおよび画像の構成的性質を活用することが可能となる。

#### V 歴史的経緯

深層学習は、人工ニューラルネットワークの構造を使用して特徴を符号化している。これは、生物学的ニューラルネットワークを模した構造である。このニューラルネットワークには、1950年代にまで遡る人工知能の長い歴史が存在している。1980年代に、誤差逆伝播法と呼ばれる方法に基づいてニューラルネットワークをトレーニングする体系的な方法が開発された<sup>4)</sup>。ただし、深い多層ニューラルネットワークのトレーニングの成功は、根本的な最適化問題の難しさ、およびその初期のコンピューティングハードウェアの限界によって制限されてしまった。

その後もニューラルネットワークは何十年にもわたって研究されてきたが、近年、安価で強力な並列計

算が可能となるコンピュータハードウェアの出現、インターネットなどのネットワークシステムの発達による大量のラベル付きのデータが利用可能となったこと、トレーニング技術とアーキテクチャの改善があり、大規模なニューラルネットワークの実装が可能となった。そして、画像を処理するために、畳み込みニューラルネットワークとして知られる深層学習のアーキテクチャが主流となった。

2012年の国際画像分類コンテスト（the ImageNet Large Scale Visual Recognition Challenge）で優勝したチームが、畳み込みニューラルネットワークを使用し、従来のコンピュータービジョン技術と比較して驚くべきパフォーマンスを示したことにより、大きな注目を集めた。この頃から多層ニューラルネットワークを指して、「深層学習」もしくは「ディープラーニング」と呼ぶようになるが、現在、深層学習というと、畳み込みネットワークを使ったものが大多数を占めている。

#### VI ニューラルネットワーク

脳では、ニューロンが化学的および電氣的シナプスを介して情報を交換する。電気化学信号はシナプス領域から樹状突起を通過して細胞体に向かって伝播し、特定の興奮閾値に達すると、細胞はその軸索を介して隣接するニューロンとのシナプスに向かって活性化信号を放出する。複雑な信号でもニューロンのネットワークによって表現することができる。例えば、視覚皮質

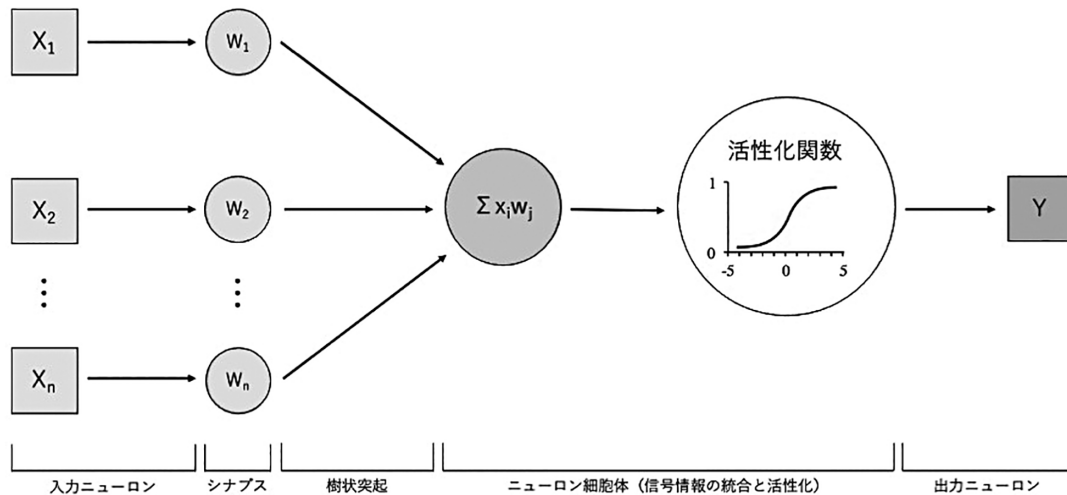


図3 人工ニューロンの構成

のニューロンの階層は、独立した視覚受容体からの信号を組み合わせることでエッジを検出でき、人工ニューラルネットワークは、この生物学的プロセスから着想を得ている。

人工ニューラルネットワークの基本単位である人工ニューロンまたはノードは、生物学的ニューロンに見られる基本的なメカニズムを模倣したモデルであるパーセプトロンというシンプルな基本単位で表される。このパーセプトロンにより、特徴のセットを入力し、それぞれに対応する重みを掛ける。重み付けされた特徴は合計され、非線形の活性化関数に渡されることとなる (図3)。

複数のパーセプトロンを重ねることにより、多層パーセプトロンと呼ばれるニューラルネットワークアーキテクチャを構成すると、非常に複雑な非線形関数を表現することが可能となる。これらの多層パーセプトロンは、通常、複数のニューロンを組み立てて層を形成し、これらの層を積み重ねて、1つの層の出力を次の層の入力へと接続することによって構築されている。深層学習の「深い」という部分は、多層パーセプトロンの多層アーキテクチャを指している (図4)。

最初の層は入力層と呼ばれ、個々のピクセル強度などの入力データを表す。出力層は分類結果などの出力結果を生成する。多層パーセプトロンの中間層は、直接的な結果を見ることはできないが、推論プロセスで役立つ入力フィーチャの中間表現を計算する。複数の層を積み重ねることにより、ディープアーキテクチャの構成力により、入力データを複雑な構成の階層として表現する。結果的にニューラルネットワークは抽象的な概念に基づいて決定を下すことができる (図5)。

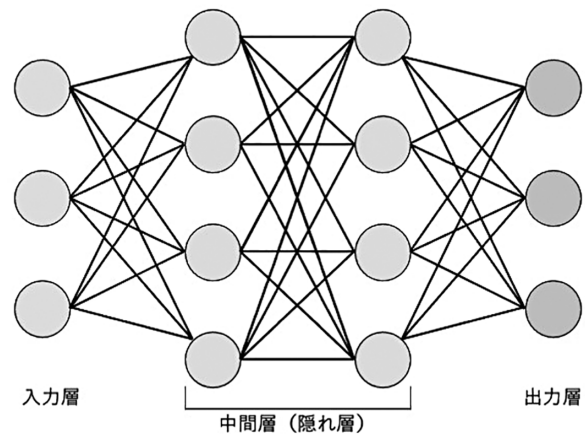


図4 人工ニューラルネットワーク

ニューラルネットワークは、各ノードの重みとバイアスで構成されるパラメータを調整することによってトレーニングされており、最新のニューラルネットワークには数百万のパラメータが含まれている。ニューラルネットワークを使用して、入力から予測を取得するプロセスは、順伝播と呼ばれる。入力層から出力層までの各層の各ノードにおいて、活性化関数を適用し、非線形変換を行うことにより、活性化値を順次計算する。分類タスクの設定では、通常、出力層の活性化において、実数のベクトルを確率分布にマップする正規化関数であるソフトマックス関数を使用する。

ランダムな初期構成から始めて、パラメータは勾配降下と呼ばれる最適化アルゴリズムを介して調整される。これは、トレーニングデータセットで適切に機能するパラメータのセットを見つけ出すアルゴリズムである (図6)。予測が特定のデータサンプルから計算されるたびに、予測の不正確さを定量的に測定する損

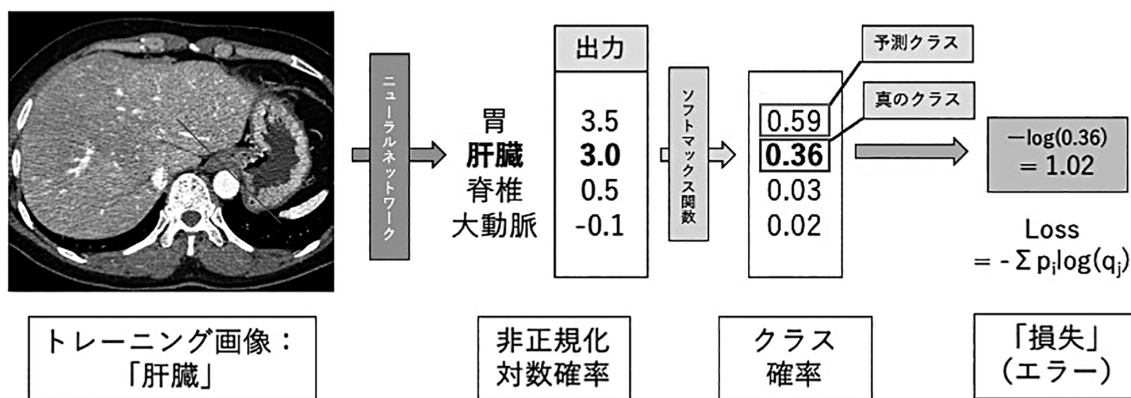


図5 ニューラルネットワークにおけるソフトマックス関数

- 学習アルゴリズム
1. 現在の設定に基づいて予測を評価するために、トレーニングサンプルのバッチ（一群）をネットワークに提示する
  2. 標的とするクラスの数（真の値）と予測値を比較し、損失を計算する
  3. 損失関数の勾配をモデルのパラメータごとに計算する
  4. モデルの重みを更新する
  5. 損失関数が最小となるまで、1-4 までの手順を繰り返す

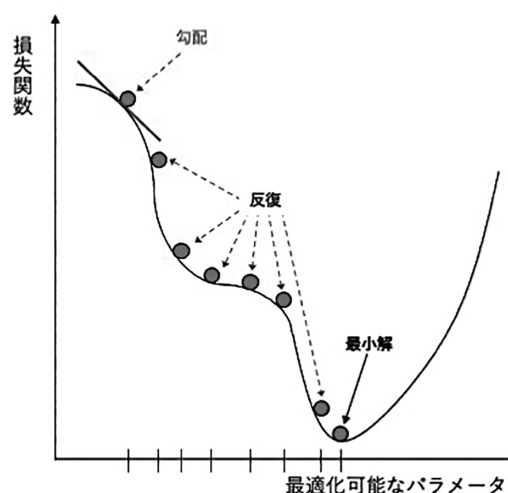


図6 学習アルゴリズムと勾配降下法

失関数によってネットワークのパフォーマンスが評価される。

次に、ネットワークの各パラメータは、損失を最小にする方向に少しずつ調整される。これは逆伝播と呼ばれるプロセスとなる。メモリの制限とアルゴリズム上の利点により、パラメータの更新は、各反復でランダムに選択されたトレーニングデータのサブセットから計算される。これが確率勾配降下法と呼ばれる最適化方法である。このトレーニング手順がトレーニングデータセットの各サンプルに対して複数回実行されると、パラメータはモデルの精度を最大化する値に近づくこととなる。

### Ⅶ 多層パーセプトロンと畳み込みニューラルネットワーク

多層パーセプトロンは、対象となるオブジェクトの形状、方向、および位置が変化する画像では、パフォーマンスが低下してしまう。一方、畳み込み

ニューラルネットワークは、畳み込み演算で画像のすべての部分に各特徴検出器を渡すことにより、これらの変動にロバスト性をもたらすことができる。そして、各特徴検出器は、その直接の入力で局所特徴を検出することに限定される。したがって、畳み込みニューラルネットワークは自然画像の構成構造を利用するため、画像内のオブジェクトのシフトや変形がネットワーク全体のパフォーマンスに大きな影響を与えることがなくなる。これらは、畳み込み、活性化関数、プーリング、ソフトマックス関数などのコンポーネントに基づく効率的なモデルアーキテクチャで、画像分類などの複雑なタスクに対処することができる。

### Ⅷ 畳み込み

入力イメージへの局所的な近傍フィルターおよびカーネルの線形適用の例を示す（図7）。これが画像の鮮鋭化や平滑化などの一般的な画像フィルターの仕組みである。

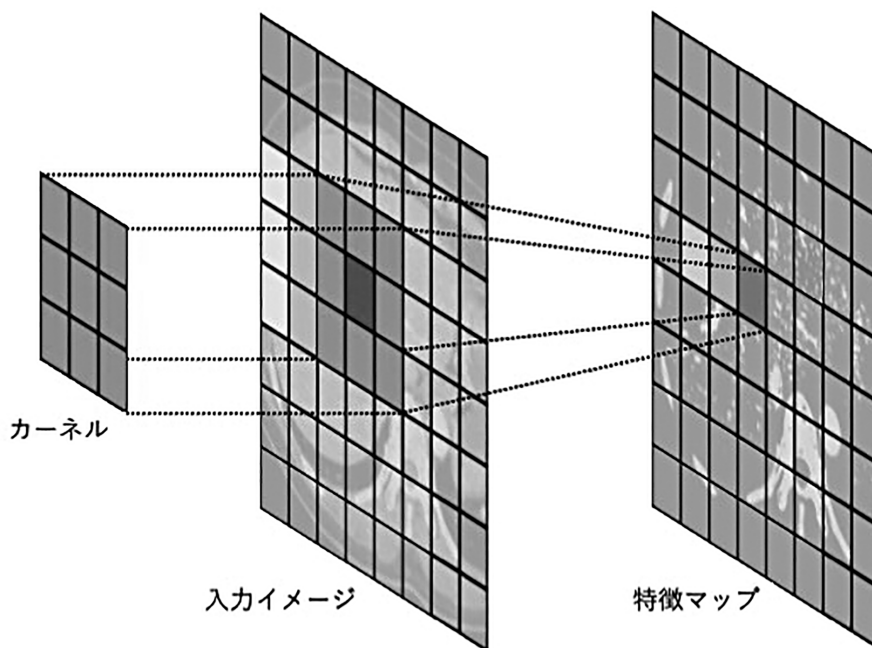


図7 畳み込み

特徴を表すフィルターは、通常、重みの小さなグリッド（ $3 \times 3$ など）によって定義される。入力に $n$ チャンネル（たとえば、異なるカラーチャンネル）がある場合、フィルターのサイズは $n \times 3 \times 3$ になる。

抽出される特徴は画像のどこにでも発生する可能性があるため、フィルターの重みはすべての画像位置で共有される。したがって、画像の特徴をより少ないパラメータでモデル化できるため、モデルの効率が向上する。通常、各層に対して複数の異なる畳み込みフィルターが学習され、多くの異なる特徴マップが生成される。

#### IX 活性化関数

ディープニューラルネットワークの主要な構成要素としては、活性化関数、つまりたたみ込みなどの線形演算の出力に適用される非線形関数が挙げられる。これらを積み重ねることにより、入力を線形分類器によって線形に分離可能な表現にマッピングすることが可能となる。活性化関数は、生物学的ニューロンの基本的な動作に関する観察から発想を得ている。ニューラルネットワーク層での活性化関数の役割は、一部の特徴を出力に渡すことができる選択関数と言える。初期には、シグモイド関数と双曲線正接関数が使用されていた<sup>5)</sup>。今日、殆どの畳み込みニューラルネットワークは中間層（隠れ層）に正規化線形ユニット（ReLU）を使用している。この活性化関数は、正の入力に対して完全に線形であるが、負の入力をゼロと

して評価する<sup>6)7)</sup>。

#### X プーリング

プーリング（またはダウンサンプリング）操作は、特徴として重要な情報を残しながらも、元画像を縮小する。特徴マップの活性化の結果が、より低解像度の特徴マップにグループ化され、情報が「削減」されることとなる。プーリングの一般的な形式は最大プーリングである。これは、プーリング領域内の最大の活性化部分を低解像度の機能マップに伝播する。連続したプーリング操作の結果、マップの解像度は徐々に低くなるが、対象の構造に関する情報は更に豊富となる。

#### XI 畳み込みニューラルネットワーク (Convolutional Neural Network : CNN)

初期の畳み込みニューラルネットワークは、脳の第1次視覚野に見られる単純型細胞（simple cell）と複雑型細胞（complex cell）に似た性質を持たせた、ニューラルネットワークであるネオコグニトロンに触発されたものであり、視覚的なパターン認識を可能とした。

畳み込みニューラルネットワークは、入力の近くでは、カーネルによってキャプチャされた局所的な特徴のみを考慮するが、ピクセル間の遠い相互作用は弱いように見える。また、表現の層が深いほど、プーリングにより特徴の空間位置の特性が粗くなってしまふ。したがって、これらのより深い層のカーネルは、更に

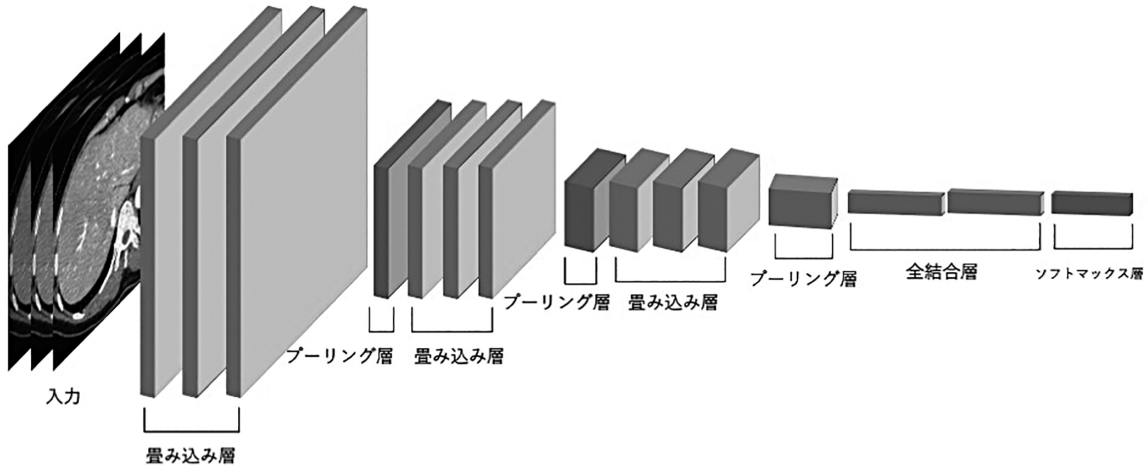


図8 畳み込みニューラルネットワーク

大きな空間スケールに対する特徴を考慮している。

畳み込み層と活性化関数は特徴マップを変換し、プーリング層は空間解像度を低下させる。分類のために訓練された典型的なネットワークでは、ネットワークの出力近くの粗い特徴表現は、全結合層を介してベクトル形式に変換されることとなる。全結合層では、各ニューロンは前の層のすべてのニューロンに接続される。全結合層により、画像のコンテンツ全体を推論することができる。分類タスクの場合、ニューラルネットワークの出力ノードは、各クラスの非正規化対数確率のベクトルと見なすことができる。最終層のソフトマックス関数を使用して、ニューラルネットワークの出力を正規化し、クラス確率のカテゴリ分布をパラメータ化できる (図8)。

### XII データ

機械学習アプローチには、トレーニングに必要なデータのタイプが異なる2つのタイプが存在し、教師あり学習と教師なし学習に分けられる。教師あり学習では、データセット内の各例にラベル付けが必要となる。例えば、肝腫瘍を分類するための機械学習システムでは、特定の腫瘍画像に「肝細胞癌」というラベル付けを行い、それを教師画像とする必要があるが、現状では多くの場合、人間による手動の作業となる。一方、教師なし学習では、ラベル付けの必要がないが、その代わりに、モデル自体が見出した規則性に基づいて画像をグループにクラスター化することを目的としている。

モデルを適切にトレーニングするための適切なデータセットサイズは可変であり、タスクの性質と複雑さに依存する。深層学習の方法は量に応じて適切にス

ケーリングされる。多くの場合、非常に大きなデータセットを活用してパフォーマンスを向上させることができる。大量のデータが望ましいが、高品質のラベル付きデータを取得するには、コストと時間がかかる点が大きな問題となる。

データ拡張は、小さなデータセットのサイズを人為的に拡大するために使用できる。ラベル割り当ての適切性を失わず、ランダム変換をデータに適用する。画像に適用できるランダム変換には、反転、回転、移動、ズーム、傾斜、変形などがある。したがって、データ拡張により、元のデータセットからの画像のバリエーションが作成され、深層学習モデルに提示される画像のトレーニングデータセットのサイズが拡大されることとなる。

利用可能なデータをトレーニングセット、検証セット、テストセットの3つのサブセットに分割することは、機械学習の標準的な方法である。トレーニングセットは、ニューラルネットワークのパラメータをトレーニングおよび最適化するために使用される。このトレーニングでは、トレーニングイメージを繰り返し実行し、エラーを使用してネットワーク接続の重みを調整する。検証セットは、トレーニングプロセス中にモデルのパフォーマンスを評価するために使用される。このデータセットは、モデルの選択を実行するためにも使用する必要がある。モデルのすべてのパラメータが決定したら、テストセットでそのパフォーマンスを測定できる。このセットは、最終モデルのパフォーマンスを報告するために、調査の最後にのみ使用される。

### XIII 学 習

ニューラルネットワークアーキテクチャの設計では、ニューラルネットワークの層の数やユニットの数、正

則化の係数など、人間が予め設定しなくてはならない多数のパラメータ（ハイパーパラメータ）を考慮する必要がある。トレーニングを実行し、全体的な精度を調べて、ハイパーパラメータを調整することとなる<sup>8)</sup>。

適切なデータセットとニューラルネットワークアーキテクチャを取得したら、モデルパラメータの学習に進むことができる。ここで重要となる機械学習の落とし穴は過剰適合（オーバーフィッティング）である。モデルは、特定の問題の一般化可能なパターンではなく、トレーニングセットの特異的な統計的変動を学習している。過剰適合は通常、トレーニングセットと検証セットのモデル精度の分析によって検出される。モデルがトレーニングセットでうまく機能し、検証セットでうまく機能しない場合、モデルはトレーニングデータを過剰適合していると言える。検証データに基づいてモデルを広範囲に評価すると、トレーニングデータと検証データの両方に適合しすぎる可能性がある。そのような状態のモデルはトレーニングセットと検証セットでは良いパフォーマンスを発揮するが、テストセットでは不十分となってしまう。また、パラメータが多すぎると、汎化性能が犠牲となることがある。モデルが過剰適合している場合は、パラメータの数を減らす、データを追加する、積極的なデータ拡張を適用する等を検討する必要がある。

#### XIV 転移学習

転移学習は、既に学習されたモデルを、他の類似タスクのモデルを学習するために使用する深層学習アプローチである。転移学習を用いたネットワークは、深層学習モデルをゼロから学習させるよりも迅速で簡便に使用することができるため、この転移学習のアプローチは一般的に、物体検出、画像認識、音声認識等で使用されている。

転移学習が利用される主な理由は、すでに大規模なデータセットで学習されているモデルを利用して、比較的少ないラベル付きのデータからモデルを学習することができる点、学習時間と計算リソースを大幅に削減できるため、新しいモデルの学習に必要な多くのエポック（全データセットを使用した1トレーニングサイクル）数が少なくて住む点が挙げられる。

転移学習は、物体認識などに適していて、AlexNet<sup>9)</sup>やGoogLeNet<sup>10)</sup>などの事前学習済みの様々なモデルをベースとして、一般に使用することができる。例えば、花の分類タスクにおいて、利用可能なデータが限

られている植物学のプロジェクトがある場合、画像を1,000カテゴリに分類する AlexNet モデルから重みや層を活用し、最終的な分類層だけを置き換えることで実現できる。

#### XV 評価方法

深層学習を含む機械学習モデルのパフォーマンスは、分類問題においては、すべての予測に対する正しく予測されたサンプルの比率である正解率（精度）を使用して評価される。そして、予測ラベルと真のラベルを報告する混合行列を生成することにより、視覚的に評価することができる。

一方、ほとんどの画像が病変（真陽性）もしくは正常（真陰性）である場合は、正解率に基づいてパフォーマンスを評価するのが困難となる。従って、再現率（感度、真陽性率）、適合率（陽性的中率）、F値などの真陰性を考慮しない評価指標の組み合わせにて評価を行う必要がある。

また、正解クラスと予測スコアが与えられたとき、各予測スコアを閾値としたときの偽陽性率を横軸、真陽性率を縦軸にプロットしたROC曲線の曲線下面積（ROC-AUC）の値に基づいて予測能・診断能を評価することもある。

#### XVI 深層学習の医療画像への適用

ディープラーニングは、写真などの自然画像に関連するタスクで、効果的なパフォーマンスを発揮した。この手法は、現在医療画像へと適用が進んでいる。

一般に、畳み込みニューラルネットワークを用いた実用的なシステムを構築するためには、大量の教師画像が必要となる。したがって、深層学習法による教師あり学習を行う場合、ラベル付けされた医用画像データセットを大量に取得する必要があり、さらに、バランスの取れた代表的なトレーニングデータセットを組み立てることは、臨床診療で遭遇する広範な病的状態を考えると困難な場合がある。

ラベル付けされた画像の不足を解決する一般的な戦略は、上述した転移学習を行うことである。そこで、我々の教室では、転移学習を利用することにより、腹部CTおよびMRI画像の分類タスクに深層学習を適用し、その精度を放射線科医の読影結果と比較を行った。

以下では、具体的に研究事例を紹介する。

### VII 腹部画像診断への応用

現在、肝腫瘤性病変の画像診断においては、超音波検査、CT、MRIなどの種々の画像検査が行われており、病理学的診断の経ずに、画像診断を診断根拠として治療方針が決定されることが多いという特徴がある。そして、客観的な診断根拠となるダイナミック造影CTや造影MRIはその中心的な役割を果たしている。一方、事前学習済みネットワークを利用した転移学習は、比較的少ない学習サンプルを用いて様々な対象に応用可能である。

そこで我々の教室では、事前学習済みの畳み込みニューラルネットワークモデルとダイナミック造影CTを組み合わせ、肝腫瘤性病変の診断能について検討を行った。また、ダイナミック造影CTの多時相画像をネットワークに学習させる際の位置合わせにおいても検討を行った。

当施設において肝切除が施行され、病理組織学的に肝原発悪性腫瘍と診断された連続215例の術前の3相ダイナミック造影CT (DCE-CT) を入力データとして使用した。呼吸運動を補正するために、3相ダイナミック造影CT画像 (造影前、造影早期相、造影後期相) を放射線科医による手動で厳密に位置合わせを行った。これらを入力画像とし、各々、青、赤、および緑の3つのカラーチャンネルに割り当てた合成RGB画像を作成した。次に、画像位置合わせの効果を評価するために任意の程度のピクセルシフト、回転、スキューによる3つのカラーチャンネルで意図的な位置ずれを加えた。作成した合成RGB画像を病理組織診断による組織型にて、ラベル付けを行って教師データとし、複数の事前学習済みの畳み込みニューラルネットワークモデル (CNN) (AlexNet, VGG-16, VGG-19, GoogLeNet, Inception-v3, ResNet-50, ResNet-101) を用いて転移学習を行った。組織型診断能については、事前学習済みCNNの種類および各種位置ずれの影響を検討した。またCNNの診断能は、一般放射線科医 (general radiologists: GRs) と腹部放射線科診断専門医 (experienced abdominal radiologists: ARs) と比較を行った。

GRs および ARs の DCE-CT における原発性悪性肝腫瘍の組織学的診断正診率は、それぞれ平均39.1% および47.9%であった。分散分析の結果、診断正診率は事前学習済みCNNのタイプの影響を大きく受けるが、スキュー変形以外の入力画像の不整合の影響は受

けないことが明らかとなった。

この結果から、呼吸性変動を避けることができない腹部画像診断において、位置ずれのない画像データを用意する煩雑な作業の必要性が明らかとなり、CNNの転移学習による診断能向上の可能性が示された。

なお本研究の詳細は、International Journal of Computer Assisted Radiology and Surgery (IJCARS) に掲載されている<sup>11)</sup>。

また、我々は、転移学習を腹部造影MRIに適用した研究も行っており、以下に示す。

肝細胞癌は多段階発癌をきたすこと知られており、今日では、早期肝細胞癌 (early hepatocellular carcinoma: eHCC) の病理学的診断基準のコンセンサスが得られている<sup>12)</sup>。進行肝細胞癌 (progressed hepatocellular carcinoma: pHCC) に対して、予後が良いことが知られているが、通常2 cm以下であることが多いeHCCに対する生検は困難であり、外科的切除による生存期間延長への寄与は乏しいとの報告<sup>13)</sup>もある。従って、画像診断によりeHCCとpHCCを区別することは臨床的に重要である。また、肝胆道造影剤を使用したMRIは、eHCCとpHCCを区別するための有用なモダリティではあるが、転移学習を用いた深層学習により、eHCCとpHCCを鑑別した報告はない。そこで、我々は、MRIでの小さなeHCCとpHCCの鑑別について、事前学習済みニューラルネットワークの診断能を検討した。

2008年から2016年の間に術前にダイナミック造影を含むMRIを撮像し、切除され病理学的に証明された2 cm以下の肝細胞癌 (eHCCで23人、pHCCで23人) 患者全46人を対象とした。MRIのシークエンスは、T1強調像同位相/逆位相、T2強調像HASTE/Fast Spin echo、拡散強調像 (b = 800/1000)、ダイナミック造影MRI (造影前、動脈相、門脈相、移行相、肝細胞造影相) の10種類を使用し、対象病変のクロップ画像をRGB画像として、最大3つのカラーチャンネルに割り当てた合成画像 (combined MR Images: cMRIs) を作成し学習及び検証として使用した (図9)。

事前学習済みネットワークであるVGG-16を転移学習用に使用した (図10)。また、全結合層のFC7までの学習結果から特徴抽出を行い、この特徴量をサポートベクターマシン (SVM) に学習させ分類器として使用し、eHCCとpHCCの分類を行った結果も検討を行った。

テストセットでのこれらの分類結果をROC-AUC



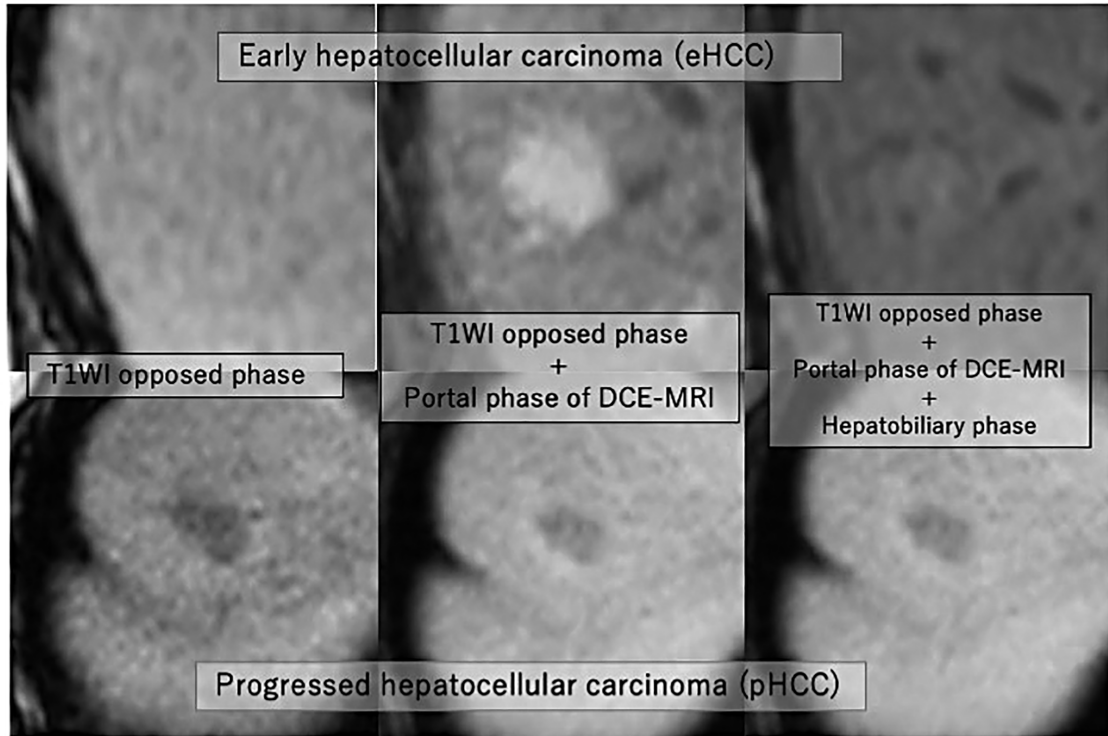


図9 合成MR画像 (cMRIs)

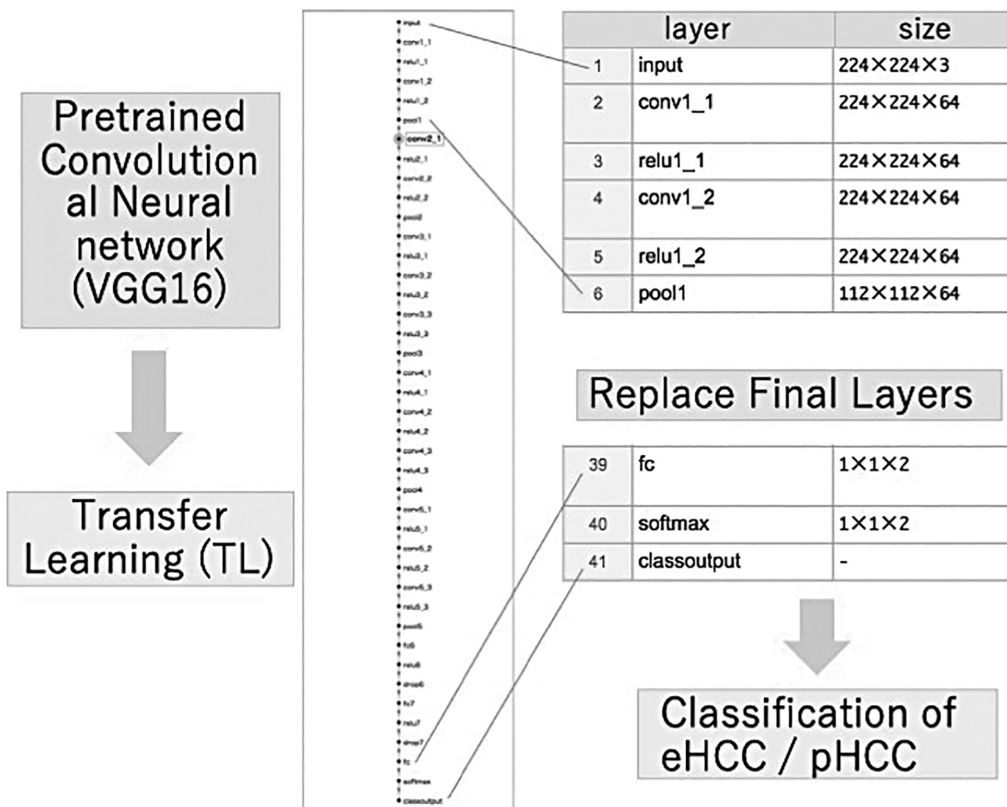


図10 転移学習 (VGG16)

を用いて、放射線科医の読影結果と比較した結果、最も高い AUC : 0.88を示したのは、転移学習に SVM を組み合わせた群の T1強調像同位相、T2強調像 (HASTE)、造影 MRI 動脈相 (T1I\_T2H\_Eap) の組み合わせであり、放射線科医の読影結果の AUC : 0.89とほぼ同等であった。

転移学習と cMRIs を使用することにより、放射線科診断医の診断能に匹敵するという結果が得られた。

### XIII おわりに

昨今の深層学習の発展と様々な領域への適用には目覚ましいものがあるが、深層学習がすべてのデータ分析問題に最適な機械学習手法ではない。データが適切に構造化されている場合など、従来の機械学習アルゴ

リズムが簡便かつ効果的な場合もある。そして、転移学習という手段を選択可能ではあるが、深層学習には大量のデータセットが必要となる点は、医療画像への適用には大きなハードルとなる。他にも、医療画像のプライバシーに関する懸念、複数の専門家または病理診断から正確なグラウンドトゥースラベルを取得するための困難性についても克服していく必要がある。しかし、これらの課題が解決された場合、放射線医学における深層学習技術は、放射線科医の行う様々な専門タスクを支援する可能性を秘めている。そのような時代が到来したとき、最適な放射線診断や放射線治療には、深層学習に基づくコンピュータ支援技術の概念、長所、限界に精通し、人工知能に使われるのではなく、人工知能を有効に使うことが重要となるであろう。

### 文 献

- 1) Becker AS, Marcon M, Ghafoor S, et al: Deep learning in mammography : diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer. *Invest Radiol* 52 : 434-440, 2017
- 2) Anthimopoulos M, Christodoulidis S, Ebner L, et al : Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE Trans Med Imaging* 35 : 1207-1216, 2016
- 3) Havaei M, Guizard N, Larochelle H, et al : Deep learning trends for focal brain pathology segmentation in MRI. Holzinger A, ed. *Machine learning for health informatics. Lecture Notes in Computer Science*, vol 9605. Cham, Switzerland : Springer International : 125-148, 2016
- 4) Rumelhart DE, Hinton GE, Williams RJ : Learning representations by back-propagating errors. *Cognitive Modeling* 5 : 1, 1988
- 5) Hochreiter S, Schmidhuber J : Long short-term memory. *Neural Computation* 9 : 1735-1780, 1997
- 6) Glorot X, Bordes A, Bengio Y : Deep sparse rectifier neural networks. *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics. J Mach Learn Res* 2011 : 315-323, 2017
- 7) Maas AL, Hannun AY, Ng AY : Rectifier nonlinearities improve neural network acoustic models. In *Proceedings of the 30th International Conference on Machine Learning 2013* : 30. Accessed October 24, 2020 at [https://ai.stanford.edu/~amaas/papers/relu\\_hybrid\\_icml2013\\_final.pdf](https://ai.stanford.edu/~amaas/papers/relu_hybrid_icml2013_final.pdf).
- 8) Kingma D, Rezende D, Mohamed S, et al : Semisupervised learning with deep generative models. *Adv Neural Inf Proc Syst* 27 : 3581-3589, 2014
- 9) Alec K, Ilya S, Geoffrey EH : Net Classification with Deep Convolutional Neural Networks. In *Proceedings of the Annual Conference on Neural Information Processing System 2012*
- 10) Christian S, Wei L, Yangqing J, et al : Going deeper with convolutions In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, 1-9, 2015
- 11) Yamada A, Oyama K, Fujita S, et al : Dynamic contrast-enhanced computed tomography diagnosis of primary liver cancers using transfer learning of pretrained convolutional neural networks : Is registration of multiphase images necessary? *Int J Comput Assist Radiol Surg* 14 : 1295-1301, 2019
- 12) International consensus group for hepatocellular neoplasia : Pathologic diagnosis of early hepatocellular carcinoma : a report of the international consensus group for hepatocellular neoplasia. *Hepatology* 49 : 658-663, 2009
- 13) Midorikawa Y, Takayama T, Shimada K, et al : Marginal survival benefit in the treatment of early hepatocellular carcinoma. *J Hepatol* 58 : 306-311, 2013

(R 2. 11. 4 受稿)