

総 説

画像診断とAI(人工知能)

画像診断科

川 本 誠 一

Radiology and Artificial Intelligence

Seiichi Kawamoto

Key words: 人工知能, 深層学習, コンピュータ支援検出／診断システム (CAD), 画像診断, 精密放射線医学
Artificial Intelligence, Deep Learning, Computer aided/assisted
detection/diagnosis (CAD), Radiology, Radiomics, Radiogenomics

1. は じ め に

人工知能 (Artificial Intelligence: 略して AI) はすでにわれわれの日々の生活を変えつつある。Apple の Siri, Google の音声認識, 天気予報, 迷惑メールスパム判定などがその例である。AI とは何か。名付け親である John McCarthy は, *"It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable"*¹⁾ と言っている。AI とは「もの」ではなく, 知的な処理をする技術である。

2016 年 3 月, Google 傘下の DeepMind 社が開発した AlphaGo が, 囲碁におけるトップ棋士である李世石九段に 4 勝 1 敗のスコアで勝利した。このことで, AI の進歩のスピードが関係者たちの予想よりもはるかに早いことが実証されたかたちになり, 世界に大きな衝撃をもたらした²⁾。チェスや将棋に比べて盤面がより広くて対局のパターン数が桁違いに多い囲碁において, 機械が人間を打ち負かすまでには何年もかかるだろうと思われていたからである。

AlphaGo で使われた手法は深層学習 (Deep Learning) という機械学習 (machine learning: 略して ML) の手法の一つである。機械学習とは, サンプルデータや過去の経験を使

用してパフォーマンス条件を最適化するためのコンピューター・プログラミングである³⁾。現在実用化されつつある AI は, この ML の手法を用いて, 例えば監視カメラで不審者を見つけ出すといったある目的に特化した AI が主流である。これ以外に, AI は自然言語処理 (natural language processing: NLP), 医療診断, bioinformatics (生命情報学, 生物情報科学), 財務データ分析などに応用されつつある。

一方, 画像診断は, 過去数十年間にコンピュータ断層撮影 (CT), 超音波, 磁気共鳴イメージング (MRI), および陽電子放出断層撮影 (PET) など様々な方法が開発された。画像診断の日常診療では, これらの医用画像が放射線科医によって読影されるが, 現代の医療技術の進歩により, 画像データの量は急速に増加して読影医に大きな負担が生じている⁴⁾。

画像診断の目的の 1 つは, 構造異常を検出し, それらを疾患に結び付け分類することである。機械学習はこのタスクに有用で, 放射線科医の負担軽減につながるということで, 1980 年代以降, 様々なアルゴリズムを実装したシステムが開発されてきた。それらのうち, いくつかのコンピュータ支援検出／診断システム (Computer aided/assisted detection/diagnosis: 略して CAD; CADe / CADx) が 2000 年代初期の臨床ワークフローに導入されたが, これらは人間よりも時間を要したものや, 誤検出率が高く, 生検が増えるといっ

た悪影響を与えたものもあり⁵⁾、多くは臨床的には評価されず^{5,6)}、普及しなかった。深層学習技術は、これら CAD システムの限界を克服できるものとして期待されている。

深層学習は、膨大なデータから有益な feature (特徴 [量]) をコンピュータ自らが学習し作り出す。人間が特徴量を設計しなくても、コンピュータ自らが高次の特徴量を獲得し (表現学習 representation learning)、それをもとに画像や音声进行分类できるようになる。松尾は、著書：人工知能は人間を超えるか (角川 E P U B 選書) の中で、深層学習を、「人工知能研究における 50 年来のブレイクスルー」と言っている⁷⁾。この新しい AI 技術は医療においては、病理診断や画像診断など画像を扱う分野で最も活用される可能性がある。

AI の先駆者でありこの深層学習のブームを作り出した研究者の Geoffrey Hinton は 2016 年の Machine Learning and the Market for Intelligence conference で「5 年、おそらく 10 年で放射線科医は AI に取って代わられるだろう。」と発言⁸⁾して物議をかもした。彼の予言した期間内に AI が放射線科医そのものにとって代わるかどうかはさておき、AI 技術の急速な進歩は、好むと好まざるにかかわらず、近い将来に放射線科医の在り方そのものに影響を及ぼすのは間違いない。そこで本稿では、機械学習の原理・概念を簡単に紹介し、深層学習の一般的応用、さらに医学的、および放射線学的応用についての現状を述べ、問題点などを考察する。

2. 機械学習について

2.1. 機械学習の分類

機械学習は、人工知能 (AI) における研究課題のひとつとして、1960 年頃から研究が始まり、2 度のブームとその後の氷河期を経験したが、その間も様々な ML アルゴリズムが考え出されて活用が進んでいた。機械学習は統計学的手法を基礎とする統計学的機械学習が主流である。そのアルゴリズムを訓練する機械学習の方法は、大きく「教師あり学習 (Supervised learning)」と「教師なし学習 (Unsupervised learning)」に分けられる。もうひとつ「強化学習 (Reinforcement Learning)」を加えて 3 つに分ける考え方もあるが、これはどちらかといえば教師付き学習の特殊な問題ととらえることができる。

2.1.1. 教師あり学習

「教師あり学習」は、あらかじめ用意された「入力」と「正解出力」がセットになった複数の訓練データをもとに、新しい入力データが来たときに、それに対する正しい出力をするような関数を導き出すことが目的である。表現形式は、一連の方程式と係数または数値化した重みである。通常、1 つの

「入力」データは、そのデータの特徴を表したもので、特徴ベクトル (あるいは素性ベクトル) などと呼ばれる、ある次元数をもった実数値ベクトルとして表現される。この多次元空間は特徴空間と呼ばれ、特徴ベクトルは特徴空間の 1 点であらわされる。「出力」は、離散的な値であったり、連続数 (実数) であったりする。前者の場合は分類プロセスであり、**クラス分類や識別 (ラベル付け)**といわれる。後者の場合、訓練プロセスは一般に**回帰**と呼ばれ、データの予測をおこなう。教師あり学習は最も一般的な機械学習である。

2.1.2. 教師なし学習

教師なし学習では入力データのみが与えられる。正しい出力は与えられないため、何らかの基準を設けてそれを最適にするような出力の割り当てを求めることになる。その目的はデータの分布などから、データの特徴的なパターンやルールを見つけ出すことである。言い換えれば、サンプル生成の背後にある状態などを表現するのに用いられる潜在変数 (latent variable) を明らかにすることである。教師あり学習と異なり、予測は何らかの主観や視点に基づいているため、基準の設計の観点によって結果は異なってくる。教師なし学習の代表例としては、**クラスタリングや頻出パターンマイニング**がある。深層学習 (ディープラーニング) の登場で、その応用が拡がりつつある⁹⁾。

教師あり学習と教師なし学習の中間的な「半教師あり学習 (Semi-supervised Learning)¹⁰⁾」と呼ばれる問題設定もあり、これは訓練データのなかに、少数の正解付きデータ (labeled data) と大部分は正解出力が与えられないようなデータ (unlabeled data) が混在するような場合に、より予測精度の高いクラス分類を実現するのが目標。主にデータのラベル付けが非常に高価であるか、または一部のアプリケーションで不可能であるために開発された。生物のほとんどの学習が、このような半教師あり学習問題だと考えられる^[1]。半教師あり学習の例には、半教師ありクラス分類と情報推薦システム¹¹⁾が含まれる。近年深層学習を利用することで、従来の半教師あり学習と比べ、はるかに高い学習性能を発揮するアルゴリズムも考えられている^{12,13)}。

2.1.3. 強化学習 (reinforcement learning)

試行錯誤をしながら、目的に合った結果を得られると成功や成果に対してスコアを与える。これが「報酬」で、最も報酬が得られるように学習していく方法。強化学習は、AlphaGo で深層学習とともに用いられて注目されている¹⁴⁾が、教師あり学習や教師なし学習と比べると実装例がまだまだ少ない。ロボットの制御や囲碁・チェスなどの対戦ゲームで利用され

[1] 人間は生まれてから様々な物体を、正解を与えられずにただ観察する。そしてある日、その物体のラベル (例えば「これは猫」「あれは犬」など) が与えられると、その教師事例を途端に認識できるようになる。こうした生物の学習の過程では、教師なしのデータから何らかの学習を行っていると考えられるからである。

る。今まさに大きく発展を遂げつつある分野である⁹⁾。

2.2. 機械学習の主な手法

カテゴリー別に主な機械学習の手法の概要を紹介する。

2.2.1. 回帰 (Regression)

売上予測などのような、実績値から未知の数字を予測する教師あり学習の手法。線形回帰はデータ空間で最適な直線（回帰直線）を表す方程式の相関係数と切片を最小二乗法や最尤推定で推定する。予測したい値を算出する式を連続する多項式として表し、各係数を最小二乗法や最尤推定法で求めることでモデル^[2]となる式を決定する。

ロジスティック回帰は、線形回帰の出力をロジット関数^[3]に入力することで、二項分布の分類問題に対応した非線形の2分類回帰モデル。

サポートベクターマシン (Support Vector Machine: SVM) も回帰によって分類する手法。マージン最大化を取り入れることで、少ないデータでも汎化 (generalization) 性能が高い2分類回帰モデルを実現できるのが特徴である。初期のSVMは線形分類器にしか適応できなかった¹⁵⁾が、カーネルトリックと呼ばれるカーネル関数を取り入れた手法で、パターンを特徴空間に写像し、特徴空間上で線形分離を行うことで非線形分類問題にも優れた性能を発揮する。ただし、大きなデータを対象としたとき学習時間は長くなる。

2.2.2. 決定木 (decision tree)

木構造のモデルによって分類する手法。上から1つの説明変数とその閾値によってデータを2つに分け、さらに枝先（ノード）で同様に別基準でデータを分けることによって、分類するモデル。非線形モデルでありながら可読性が高い。ランダムフォレストなどの手法が使用されるが、分類（分類

木 classification tree) だけでなく、回帰問題などに利用することができる（回帰木 regression tree）。勾配ブースティング木 (Gradient Tree Boosting) と呼ばれる Gradient Boosting と Random Forest のアルゴリズムを組み合わせたアンサンブル学習^[4]は、決定木を大量に生成し、各決定木の結果を集計して予測する手法である。

2.2.3. ナイブベイズ法 (Naive Bayes)

確率論における重要定理であるベイズの定理を活用し、最もその分類が発生する確率が高いものを予測とする手法。主に文書分類タスクに利用される。ベイズの定理は、機械学習の様々な分野で多用される。

2.2.4. 最近傍法 (Nearest neighbor)

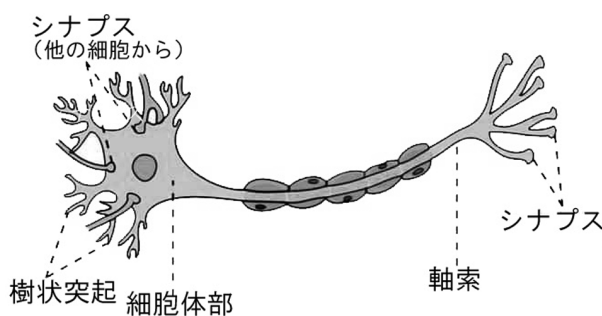
k近傍法 (k NN) とも言い、特別な空間での最も近いトレーニング例をベースとする統計分類の手法で、多くの機械学習アルゴリズムの中では最も単純で、その分ノイズに影響されやすい。古典的パターン認識として重要な方法である。

2.2.5. クラスタリング (clustering)

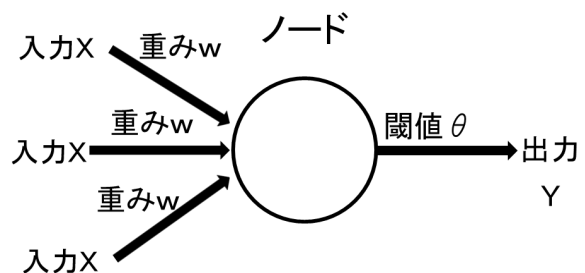
値やデータの類似性をもとに、データを自動的にグループ分けする手法。分類対象の集合を、内的結合 (internal cohesion) と外的分離 (external isolation) が達成されるような部分集合に分割すること。k平均法 (k means) などの手法がある。クラスタリングは探索的 (exploratory) なデータ解析手法であって、分割は必ず何らかの主観や視点に基づいている。よって、クラスタリングした結果は、データの要約などの知見を得るために用いられる。

2.2.6. 次元削減 (dimensionality reduction)

多変量解析分野では、主成分分析などで、データ変動の構



(a) 人間のニューロン



(b) 人工ニューロンモデル

図1

[2] 対象を記述するための枠組み・言語。入力される予測対象から、予測結果への対象対応を記述する写像、もしくはそれらの候補集合で、数学的関数か論理式を用いて表現する。機械学習での学習とは訓練データを参照して、モデル集合から正解を最もよく近似すると思われるモデルを選択することである。

[3] ロジスティック関数、シグモイド関数（図2）の逆関数である。

[4] 弱学習器 (weak learner= 性能の低い学習器) を組み合わせて高精度の学習器 (強学習器) を構成する方法。

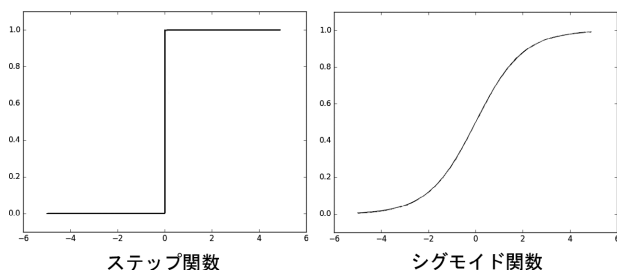


図2 左) ステップ関数 右) シグモイド関数

造を説明する少数の変数を描出することが重要で、こうしたデータの低次元化には、計算量削減やデータを人間が見やすいように可視化する目的もあった。これは観測データの次元削減と呼ばれる。観測データから本質的な情報を抽出する内部表現 Internal representation^[5] の手法の1つである。

2.2.7. 人工ニューロン (Artificial neuron)

人間など生物の脳神経系の学習能力にヒントを得て、2択の予測において要因となる入力値に重みを付けて合算し、それが閾値 (バイアス) を超えると活性と判定する手法。脳のニューロンの伝達の仕組みを模した人口ニューロンである。図1に、その概念を示す。

パーセプトロンは1957年 Rosenblatt により考案された¹⁶⁾。入力層と出力層のみからなる最も基本的な順伝播型のニューラルネットワーク (feedforward neural network) で、入力ベクトルの各ニューロンに入力された値に重みを乗算し、それらを足し合わせたものにバイアス項を足し算し、その結果に対してステップ関数 (図2左) という閾値活性化関数を用いる人工ニューロンの特殊な場合である。線形分離可能なパターンしか分離できない。

2.2.8. 多層ニューラルネットワーク

(Multilayer neural network)

人工ニューロンを入力層・1つ以上の中間層 (隠れ層)・出力層と階層型に構築したモデルである (図3)。相互接続された人工ニューロン (ノード) で構成される。各ノードがシナプスを模している。各ノードは、受け渡された信号の重み付き和を活性化関数として最もよく使われるシグモイド関数^[6] (図2右) にかけて決定信号を出力する単純な分類器モデルを実装する。このつなぎ部分の重み付けの調整は誤差逆伝播 (Back Propagation) などの学習アルゴリズムによって訓練される。誤差逆伝播は、出力の誤差を、入力伝播するのと逆方向に伝播させ学習することで誤差を最小にする教師あり学習の手法で、1986年のRumelhartら¹⁷⁾により発明された。一つのブレイクスルーである。のちの現在に続く第3次ニューラルネットワークブームのコアとなるアルゴリ

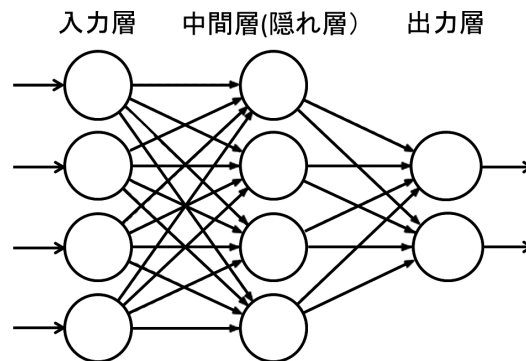


図3 ニューラルネットワーク

ズムである。誤差逆伝播学習によって、多層にわたる順伝播型ニューラルネットワーク (Feedforward Neural network: 略して FFNN) における要素間の結合重みを効率的に学習することが可能となり、観測データからタスクに適した内部表現 (特徴) を学習により獲得できることが示された。パーセプトロンが第1次のNNブームで、この階層型NNが第2のNNブームであったが、何十層にもわたるような深層NN (DNN) は2010年代の計算環境の発達を待たなければならなかった。

2.3. フィーチャー・エンジニアリング Feature engineering

MLで入力データとして使用される変数はfeature (特徴 [量], 素性) と呼ばれる。その特徴を定量的に表したものである。統計学的手法では予測対象を記述する特徴をベクトルの形式であらわす (特徴ベクトル)。どのような変数を読み込ませて、観測データに内在する本質的な情報表現すなわち内部表現 Internal representation を抽出するかが、ML研究では重要なプロセスである。これをフィーチャー・エンジニアリング Feature engineering というが、featureを何にするかで予測精度が大きく変わってしまう。良いfeatureが生成できれば、それらのベクトルを並べた特徴空間で簡単な線形識別器でクラス分類が可能になり、教師あり学習も容易になる。以前のMLでは、人間が考えるしかなかった。直感やドメイン専門家の知識に基づきさまざまな評価手法を適用して特徴量を生成しようとした。パフォーマンスの改善を得るために、さまざまなデータクリーニングや特徴量選択方法が開発されてきた。機械学習における本質的な問題である。

2.4. 過学習 Over training

機械学習で、訓練データに合わせ過ぎたモデルを選択すると、学習サンプルとの誤差すなわち経験誤差は小さいが、正解との差である汎化誤差が大きくなり、本来の学習の目的を達成できていない状態に陥る。統計学では過剰適合 (over

[5] 「2.3. フィーチャー・エンジニアリング」参照

[6] ロジット関数の逆関数。 (標準) ロジスティック関数=シグモイド関数と言う。

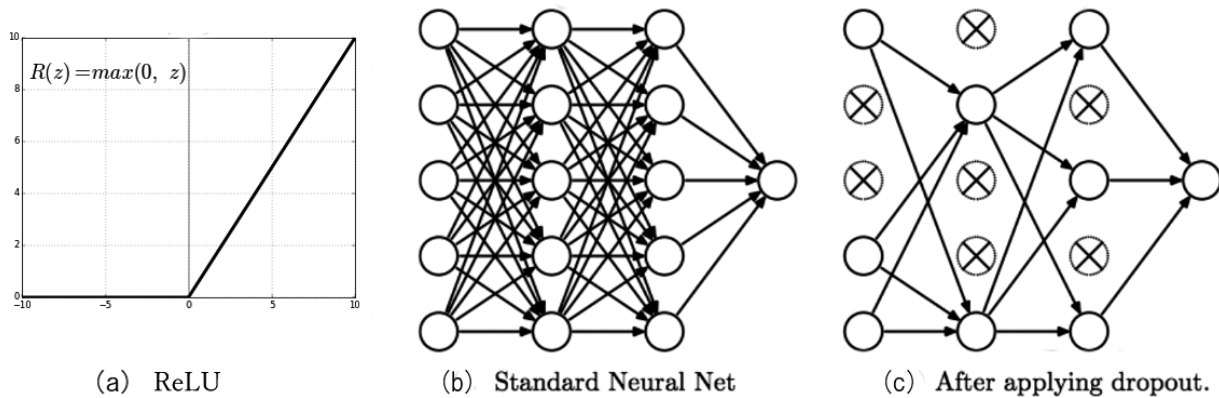


図 4

- (a) : ReLU 最も現在使われることの多い活性化関数. x が負のとき 0, それ以外は恒等関数. シグモイドなどより計算が早いことが特徴. x が正のとき活性化関数の微分は常に 1 になるので, Relu は勾配消失が起きにくい.
- (b),(c) : DropOut 学習時にネットワークを構成するパラメータを強制的に削減し, 自由度を下げることで過学習を抑制する効果がある.
- 出典: 文献 [18]. 一部修正.

fitting) という. この汎化誤差をいかに小さくするかが機械学習の目標である. 正則化や交差検証といった過学習を防ぐ手法がある. 深層学習では, DropOut (図 4) と呼ばれる手法が過学習を防ぐ高い能力を有する¹⁹⁾.

3. 深層学習

深層学習 (Deep Learning) は, 深い, 多くの層を持ったニューラルネットワーク (Deep Neural Network : DNN) の総称である. ニューラルネットワークは 20 年ほど氷河期にあったが, その間も研究は地道に続けられていた. Hinton ら²⁰⁾ は 2006 年頃, 制限ボルツマンマシン (Restricted Boltzmann Machine : RBM) というニューラルネットワークをベースとして用いた. 多層のネットワークを層ごとに教師なしで事前学習させてその結果をもとに教師あり学習を行うというもの. 局所最適解 (local optimum) に収束してしまい, 最適解に至らない問題点 (最適化問題, 勾配消失問題^{[7])} や過学習などの深層学習の限界が克服された. その後, 確定的に動作するニューラルネットワーク・ベースの深層学習が誤差逆伝播学習アルゴリズムの組み合わせで, 盛んに研究されている.

深層学習の本質的な意義は, 深く階層化されたニューラルネットワークによって, 観測データに内在する, 本質的情報表現 (内部表現 Internal representation) すなわち feature (特徴 [量]) をデータからの学習によって自ら獲得することにある. 汎用性の高い内部表現が得られれば, 多くの現実問題に対して良い性能を示す識別器などが作れる可能性がある. ニューラルネットワークの深層化には, 近年の ICT 化を通じたいわゆるビッグ・データの蓄積と, その膨大な計算を可能にするだけのマシンパワーが必要になる

が, GPU (Graphic Processing Unit) という高速の画像処理に適したプロセッサを汎用のプロセッサとした GPGPU (General-purpose computing on Graphics Processing Units) や, HPC (High-performance computing) などの計算環境の発達がこれを可能にした. これらのブレイクスルーにより, 音声認識・画像認識の分野では, 深層学習が従来の方法を駆逐した. 自然言語処理でも, 従来手法と同等の性能を達成している. 従来のほぼすべての認識技術は, 2010 年代の深層学習アルゴリズムによって洗い流されてしまった. 深層学習は, あらゆる分野において, 共通の基盤技術となる可能性を秘めている²¹⁾.

深層学習には多くのタイプがあるが, 現在は畳み込みニューラルネットワーク (Convolutional neural network : CNN) と 回帰結合ニューラルネットワーク (Recurrent Neural network : RNN) が盛んに研究され, 実際, 活躍している. 勾配消失問題に対処できる ReLU (図 4 左) や, より複雑な関数に対処できる MaxOut などの活性化関数, 過学習対策の DropOut (図 4 右), Adam などの最適化手法といった技術改良がなされている.

3.1. 畳み込みニューラルネットワーク

畳み込みニューラルネットワーク (Convolutional neural network : CNN) は人間の視覚信号の処理を参考にし, ある範囲の情報をまとめていく畳み込み構造を特徴とする深層のニューラルネットワークである. 2012 年, 世界的な画像認識のコンテスト ILSVRC (ImageNet Large Scale Visual Recognition Challenge)」で, 初参加の Hinton 率いるカナダ・トロント大学²²⁾ が開発した SuperVision が突出した成果をしめし, 他を圧倒した. この時使われたのがこの CNN で,

[7] ニューラルネットワークの誤差逆伝播法の過程で, 活性化関数の微係数が何度も乗算されることによって指数的に変化し, 勾配法による最適化が困難になる問題.

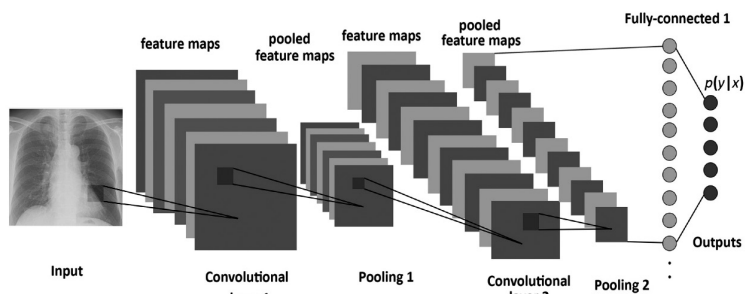


図5 畳み込みニューラルネットワークの構造。畳み層、プーリング層、全結合層からなる。
出典：文献 23)。一部修正。

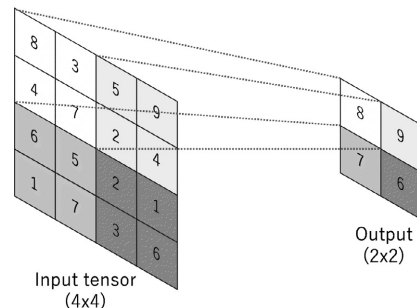


図6 最大プーリングの例。入力画像から2 x 2のパッチを抽出し、各パッチの最大値のみを出力し、画像を縮小させる。
出典：文献 24)。一部修正。

事前学習なしの誤差逆伝播学習アルゴリズムとの組み合わせで確定的に動作するニューラルネットワーク・ベースの深層学習である。ほぼすべてのニューラルネットワークは、この誤差逆伝播学習アルゴリズムによって訓練されている。この後、深層学習は画像認識タスクの標準的な手法と考えられるまでになった。様々な深層学習が研究されているがその中心はCNNである。

図5のように、CNNの構造は畳み込み層とプーリング層が交互に結合したニューラルネットワークである。出力層付近には隣接層間のノードすべてを結合した層（全結合層 fully connected layer）が配される。基本的構造は、順伝播型ニューラルネットであり、ある層のノードはその直前の層の複数のノードからの入力の重み和にバイアスを加算して活性化関数に入力した変数がこのノードの出力となり、次の層へ伝播する。deepな構造に適したReLU活性化関数（図4左）がよくつかわれる。

畳み込み層 convolution layer の主な目的は、画像認識において特徴的な局所モチーフのようなエッジ、線、その他の視覚的要素を検出することである。画像上にフィルタ、あるいはカーネルと呼ばれる正方形の小領域を設定し、この入力（画像）に複数のフィルタを畳み込む計算（畳み込み演算）を行う。このフィルタの係数が畳み込み層の重みであり、これらと出力ノードのバイアスは学習により決定される。各フィルタは、与えられた画像上をスライドして各層の特徴マップ（feature map）を生成する。畳み込み層では、複数のフィルタが適用され、同数の特徴マップが出力される（図5）。モノクロならばチャンネルは1つだが、カラーならばRGB3チャンネルすなわちフィルタの数は3倍になる。畳み込みはCNNの重要なコンポーネントであり、画像の分類やセグメンテーションなどの画像処理タスクの成功に不可欠である。プーリング層 pooling layer では、特徴として重要な情報を残しながら元の画像を縮小する。畳み込み層で生成された特徴マップがプーリング層の入力となる。プーリングの処理は画像を縦横方向に間引いて一つの値に集約（プーリン

グ）される。ノードからの入力の平均を圧縮されたノードの出力とする方法は平均プーリングであり、最大値を出力とする場合は最大プーリングである。初期のCNNの研究では平均プーリングが主に使われていたが、現在は最大プーリング（max pooling）が一般的である（図6）。また、プーリングには、特徴の位置感度を低下することで、位置に対するロバスト性（頑強性）を高める効果もある。これらの畳み込み層およびプーリング層は、複数回繰り返される。全結合（FC）層は、こうして得られた画像全体からのすべての特徴応答を統合し、最終結果を出力する。

CNNが画像から取り出している特徴（内部表現）にはネットワークの構造と対応した層構造があることが示されているが、これは生物の脳の視覚野にみられる階層構造と形態的にも機能的にも類似性が報告されている²⁵⁾。注目すべきは、この特徴（内部表現）の階層構造は、より下層の特徴はより普遍的であり、異なるタスクでも共通していることである。つまり、ある認識タスクで学習された特徴を他の認識タスクに転用することができる（転移学習 transfer learning）²⁶⁾。

DNNの扱いを難しくしている要素のひとつにそのパラメータの多さがある。畳み込みの段数をどれくらいにするか、あるいは畳み込みのパッチの大きさをどれくらいにするか、収束計算の際の学習率をどれくらいに設定するかなど、かなりの数のパラメータが存在する。DNNでは、特にこうしたパラメータの最適化は重要で、近年ではベイズ最適化（Bayesian Optimization）と言われる方法が使われることが増えている。こうした最適化の手法を用いて、最適なパラメータを効率よく、自動的に探索するような手法も試されるようになってきている。

3.2. 回帰結合ニューラルネットワーク

音声認識の領域でもニューラルネットワークを利用するための試みは多数行われてきた。その中でも回帰結合ニューラルネットワーク（Recurrent Neural network；RNN）は、ユニッ

ト間の接続が一方方向の方向を有するループを形成する時間的データに特化したNNである。これにより、時間変化の情報を示すことができる内部表現が生成される。順伝播型のニューラルネットワークとは対照的に、RNNは、記憶を保持することができるようになっているが、実際にはその記憶を学習することが難しいという欠点があり、これらを解消するためのLSTM ; long shot-term memory 法がある。LSTMは時系列データに対する強力なRNN予測モデルであり、勾配消失問題を解決し、長期の時系列を学習できるようになった²⁷⁾。このような深層学習は自然言語認識、手書き認識、音声認識、機械翻訳の大幅なパフォーマンス向上をもたらすことが期待されている²⁸⁾。

4. 非医療用応用

深層学習の使用は、画像認識の分野を急速に進化させた。前述の一般物体認識 (general object recognition) のコンテスト ILSVRC2012 でエラー率が 16%²⁹⁾ であったのが、2016 年には 3% を下回った³⁰⁾。もはや物体認識 (クラス分類) タスクにおける人間のパフォーマンス (5.1%) を上回っている。

音声認識は、深層学習以前にも多くの研究がなされてきたが、近年、音声認識の分野は、深層学習の進歩とビッグ・データのために長足の進歩を遂げている。これは、日常使われている Google, Apple, Microsoft, Amazon などの多くの企業の音声認識システムを見れば一目瞭然である。

DNN の効果的な学習には大規模の学習データセットが必要であるが、クラウドソーシングを基盤としたアノテーション技術の向上により、大規模な教師付きデータセット

が次々と作られるようになって、研究者自らがこのようなデータセットを収集することなく競争し、協力し合う新しいオープンな競争の場が生み出された。近年、大学などの研究機関が実績のある ML アルゴリズムをオープンソース化 (OSS) している。中でも California 大学が提供する「Caffe³¹⁾」や、Montreal 大学の「Theano/Pylearn2³²⁾」などが研究者の間では広く良く利用されてきた。さらに 2015 年には日本の Preferred Networks が「Chainer³³⁾」を、Google が「TensorFlow³⁴⁾」というフレームワークをそれぞれ公開した。そのほか、「Torch³⁵⁾」、「Scikit-learn³⁶⁾」など、この OSS のフレームワークが導火線となり、機械学習は一気に実用化段階に入っている。2014 年末に IBM Watson が ML の API^[8] を公開すると、2015 年からは主たるパブリッククラウドで続々と実用的な ML アルゴリズムがライブラリ^[9] としてサポートされるようになった。Amazon Web Services の「Amazon Machine Learning³⁷⁾」、Google の「Cloud Machine Learning³⁸⁾」、IBM の「IBM Cloud (Watson)³⁹⁾」、Microsoft の「Microsoft Azure Machine Learning」などが挙げられるが、これらのサービスや API を利用すれば、アルゴリズムと計算環境が提供されるので、誰でも容易に機械学習を活用できるようになった。

深層学習の分類における技術革新は、セグメンテーション (semantic segmentation) や、リアルタイムの物体検出 (detection) などにも応用が広がっている。画像をピクセルレベルでクラスラベル (「人物」、「道路」、「空」、「海」、「自動車」など) に関連付けるセグメンテーションの例としては、SegNet⁴⁰⁾ や U-Net⁴¹⁾ (図 7) などのアルゴリズムが開発

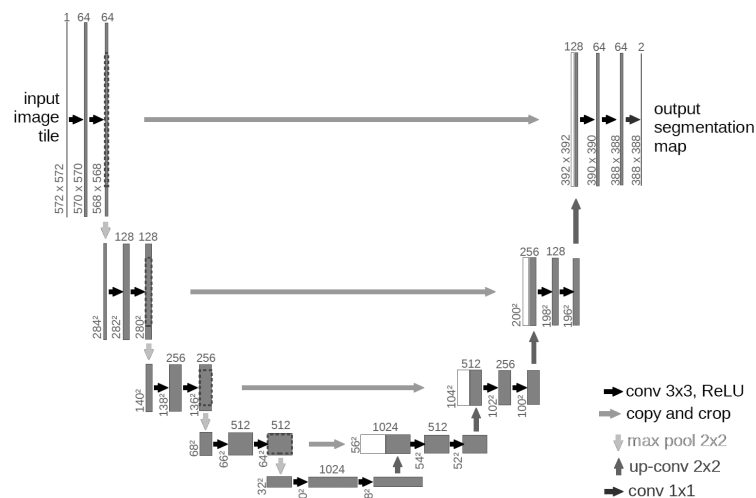


図 7 U-Net .

出典 : U-Net : Convolutional Networks for Biomedical Image Segmentation.
(<https://lmb.informatik.uni-freiburg.de/people/ronneber/u-net/>.)

[8] Application Programming Interface の略。ソフトウェアから OS の機能を利用するための仕様、またはインターフェースのこと。

[9] 汎用性の高い複数のプログラムを再利用可能な形でひとまとまりにしたものである。機械学習ライブラリはオープンソース化 (OSS) されて誰でも利用できる場合が多い。

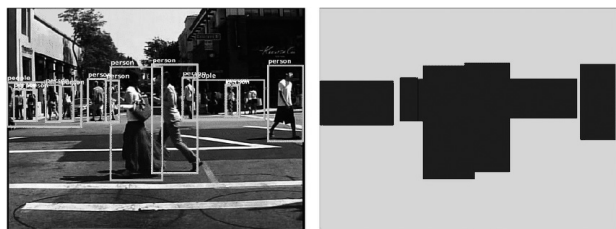


図8 右側の画像では、黒の値は1であり、白は0に設定されている。出典：文献64)。

されている。物体検出は、図8のように、バウンディングボックスと呼ばれる四角形の枠取りで画像内の物体の場所を指し示すタスクである。YOLO⁴²⁾、Faster R-CNN⁴³⁾、SSD⁴⁴⁾などのアルゴリズムが考えられている。

またCNNベースの画像認識フレームワークとRNNベースの言語モデルを統合し、画像内のオブジェクトやアクションに基づいて画像から説明文を生成するプロセスである**画像記述 (Image captioning)**⁴⁵⁾や、ある画像とその画像に関する質問を提示されたときに、正しい答えを導き出すタスクである**Visual Question Answering (VQA)**^{46,47)}など画像だけでなく、音声、自然言語なども扱うマルチモーダルな手法も盛んに研究されている。

5. 医用診断画像への応用

近年、医用画像解析にも深層学習技術が急速に導入されてきた。医療画像関連の機械学習の論文は、2015年から増え始め、2017年から一気に発表されるようになった。そのほとんどがCNNを用いた研究である。

CADe / CADx とは、医用画像をコンピュータで解析し、病変候補の検出や病変の質的診断を行うものである。含まれるタスクは病変の分類、描出、検出、そして、所見の音声入力、自動レポート作成である。

そのほか、画像処理にも深層学習技術が応用されている。主なタスク別にその概要を紹介する。

5.1. 分類 Classification

放射線科の領域では、例えば、異常の有無、良性か悪性かの判断、病理組織やゲノム情報（「5.6 Radiomics」を参照）による癌の悪性度の判定、予後、あるいは画像データの組織化の目的など、さまざまな分類タスクがある。

このわずか5年間に深層学習技術の発展をもたらしたImageNetのような自然画像データセットと比較して、利用可能な医用画像データは限られている。したがって、医用画像分類における深層学習の多くのアプリケーションは、事前学習された既成のCNN分類器の特徴量を利用する転移学習に依存している⁴⁸⁾。Antropovaら⁴⁹⁾は、事前学習された既

成のVGG19ネットワークを転移学習させて抽出された特徴量を、マンモグラフィー、超音波、およびMRIの乳房病変の良悪の鑑別に用い、従来の手作りの特徴量と比較して、深層学習でも同等の精度を示したとしている。異なるデータセットで事前学習されたネットワークを微調整する転移学習のやり方は、MRIによる前立腺病変の良悪の鑑別⁵⁰⁾、X線単純写真での変形性股関節症同定⁵¹⁾などで用いられている。

転移学習の派生手法として既成のネットワークをあらかじめ、目的のデータセットで学習させて、よりタスクスペシフィックな特徴量（内部表現）を得ようとする方法も、乳癌のMRI画像に関する研究など、転移学習よりも優れた結果を得ている⁵²⁾。

十分なデータが利用できる場合は、深層ニューラルネットワーク全体を一から訓練することができる。訓練されるネットワークのサイズは、タスクとデータセットの特性に依存するが、多くの医用画像研究^{53,54)}で使用されている既成ネットワークは、AlexNet（8層）やVGG（16層）で、層はそれほど深くはない。深ければ深いほど、より高度で複雑な特徴を抽出していると考えられており⁵⁵⁾、Residual Networks (ResNets)⁵⁶⁾（152層）や異なるサイズのフィルタの畳み込み層を複数横に並べて、それを結合するインセプション構造といわれるGoogLeNetやInception⁵⁷⁾などのネットワークなども医用画像解析で用いられている⁵⁸⁾。

CNNとは別に、**自己符号化器 Auto-encoder**^[10]（AE）を深層化した深層自己符号化器やその汎化能力を高めたスパース・オートエンコーダ（sparse auto-encoder : SAE）などを用いて特徴表現を得る方法が、MRIによる多発性硬化症の病変検出⁵⁹⁾などで用いられている。

5.2. セグメンテーション（領域抽出、領域分割）

Image segmentation

画像の分類（画像認識）では、位置や歪みなどの影響をあまり受けない、頑強なパターン認識をめざしている（画像としては解像度の減ったボケた画像になる）が、領域抽出では、図9のように、物体の特徴に加えて全体での位置情報も元の入力画像上で特定しなければならない。つまり、プーリング層でぼかされた局所の特徴の位置情報を元画像上にpixel単位で精細に復元する必要がある。深層学習を用いた医用画像セグメンテーションのための最近の突破口は、細胞膜のセグメンテーションで、Ciresanら⁶⁰⁾は専用の特徴抽出ステップなしで畳み込みニューラルネットワーク（CNN）を用いて小さな画像パッチの中心を直接分類することを提案した。MRI脳腫瘍描出⁶¹⁾などに用いられている。しかし、この方法はパッチの重複で、計算効率が悪いなどの欠点があるため、最近、パッチではなく全体イメージを用いる異なるCNN

[10] ニューラルネットワークを使用した次元圧縮のためのアルゴリズム。3層ニューラルネットにおいて、入力層と出力層に同じデータを用いて教師あり学習させたものである。

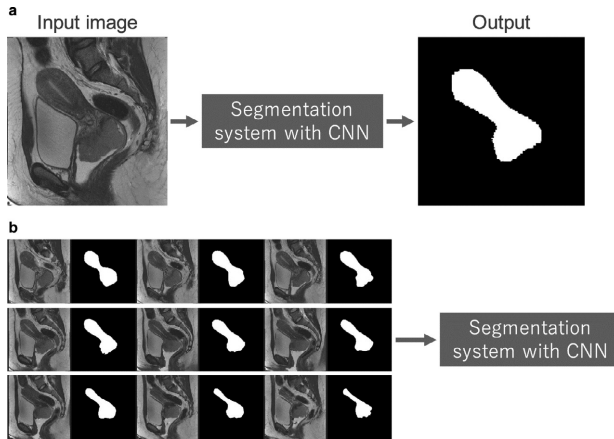


図9 A. CNNを用いたセグメンテーション。B. 学習段階で使用されるトレーニングデータ。オリジナルイメージとそれに対応する手書きセグメンテーションは、お互いに並んで配置されている。
出典：文献 24)。

アーキテクチャが提案されている。これにより、冗長な計算が不要になり、より効率的に、より高画像解像度が実現できる。Fully Convolutional Network (FCN, FCNN, Fully-CNN) と呼ばれる⁶²⁾。領域抽出では画像が出力になるため、全結合層が必要ない。畳み込み層だけのネットワーク = 完全畳み込みである。しかしながら、FCNは、畳み込みおよびプールの連続使用のために、これらの両方が次元性を低下させ、元画像と比較してセグメント化の解像度が低下してしまう。入力画像と同じ解像度のセグメンテーション画像を得るために、「物体の局所の特徴と全体的位置情報」の両方を統合して学習させる Encoder-Decoder 構造を有する SegNet⁴⁰⁾ や U-Net⁴¹⁾ などが開発された。U-Net (図7) は畳み込みおよび非畳み込み層を U の字に配した NN の組み合わせにより、下向きパス (エンコーダー) は、プーリング層を用いて空間的な次元を徐々に減少させる。一方、上向きパス (デコーダー) は、畳み込みを行わず特徴を保持したまま、空間的な次元を復元することができる (upsampling)。両方のパスにおいて、空間的な次元が同じものをそれぞれショートカット接続することによって、局所の特徴を保持したまま全体的位置情報の復元を行うことができる。しかしこの上向きの upsampling は解像度を完全に補間するものではないため Brosch ら⁶³⁾ は最近、多発性硬化症の病変抽出のために新たな FCN 構造を提案して、入力画像と同じ解像度のセグメンテーションを生成できるとしている。

5.3. 検出 Detection

画像診断領域においてさまざまな検出タスクに使用される最も一般的なアーキテクチャは、2 相のプロセスを組み合わせた方法で、第 1 段階では、感度をあげて、疑わしい領域をすべて特定する (領域提案 Region proposal)。すなわち偽

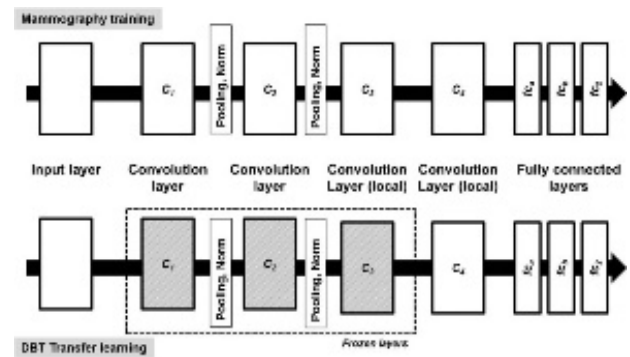


図10 DCNN による学習：上段は乳房撮影による事前学習。下段は事前学習させたネットワークを乳房トモシンセシス (DBT) に利用した転移学習。C1~C3 の feature layer はパラメータを“凍結”させて学習させる。
出典：文献 67)

陽性が多く含まれる状態で、この相では、分類に使用されるアーキテクチャに基づく回帰ネットワーク⁶⁵⁾ が用いられ、第 2 段階では、前の段階で抽出された候補領域の画像の単純な分類が行われる。深層学習は通常、転移学習が用いられる。CT 肺血管造影画像の肺塞栓検出⁶⁶⁾ では自然画像を用いて事前学習している。事前学習として他の医用画像データセットを使用するもの (図10) もみられる⁶⁷⁾。第 2 相の分類器は VGG⁶⁸⁾、GoogLeNet⁶⁹⁾、Inception⁷⁰⁾、ResNet⁷¹⁾ などの既成ネットワークを用いる。

2 相法ではモデルは各相で別々個別に訓練されるが、両方の相を End-to-End で学習する方法がある。単一の CNN ネットワーク構造からなる Faster Region-based Convolutional Neural Network (faster R-CNN)⁴³⁾ で、最近医療画像化に適用された。他に、領域提案の第 1 段階を排除し、画像認識を回帰問題に落とし込み、「画像の領域推定」と「分類」を同時に行うことを実現し、従来の方法より速く、精度も上がった You Only Look Once (YOLO)⁴²⁾、Single Shot MultiBox Detector (SSD)⁴⁴⁾ を用いた研究もみられる。乳房撮影における関心領域の検出⁷²⁾ や超音波画像での乳房腫瘍検出がある⁷³⁾。有用な結果が得られている。

5.4. 病変の自動記述 captioning

画像の内容を自動で記述する画像記述では、特定の疾患や臓器の自動認識や位置同定のためには、まず対象病変や解剖学的な位置などはっきりとラベル付されたデータセットをそろえる地道な努力を必要とする⁷⁴⁾。

Hoo-Chang Shin ら⁷⁵⁾ は公開されている胸部 X 線写真とその所見レポートのデータセットを使用し、CNN と RNN に学習させ、胸部 X 線写真に疾病の状況や病気の説明 (部位、重症度、罹患臓器) を自動的に記述する方法を示した。

Chen ら⁷⁶⁾ は造影胸部 CT で肺塞栓 (PE) の検出とタイプ (慢性、急性)、部位、中枢性か、末梢性かを自動で記述

する方法として、深層学習と従来の自然言語処理（NLP）法と比較して、同等またはそれ以上の正確さで、診断し、記述することができたとしている。ICU 患者の頭部 CT の重症度分類と、自動記述において、深層学習と従来の NLP モデルと比較した研究⁷⁷⁾でも、深層学習のほうが良好な成績を示した。

International Coding of Diseases（ICD）を用いて放射線診断レポートをコード（ICD9）化して自動記述する試み⁷⁸⁾もある。

5.5. 画像処理への応用

放射線科領域での機械学習は CAD のための、分類、セグメンテーション、および検出が基本となるが、画像処理にも深層学習が応用されている。主なものについて述べる。

5.5.1. レジストレーション（画像重ね合わせ）

Image registration

2つの画像間で、片方の画像を移動、回転、変形させることにより、もう片方の画像に合わせこむ手法を画像重ね合わせ（Registration）と言うが、近年、放射線診断や放射線治療において、重要な技術になってきている。医用画像位置合わせタスクでは、例えば X 線透視やエコーガイド下の IVR 手技や外科手術におけるナビゲーションなど、非常に高い精度と同時性が要求される。

いくつかの方法があるが、深層学習の一つの方法としては、CNN をグランドトゥルース^[11]な変形情報を用いて訓練する教師ありの回帰技法を用いて予測する方法である⁷⁹⁾。

教師なしの方法を用いた研究⁸⁰⁾もある。CNN は、位置合わせされていない画像対の相違度を最小にする空間変換を生成するために End-to-End で学習するやりかたである。

5.5.2. 画像生成 Image generation と

画像エンハンスメント Image enhancement :

医用診断画像における画像生成は、深層学習を用いて、同じモダリティで画質を向上させる試みや、違うモダリティ画像を生成する方法などが考えられている。

画像エンハンスメントは、超解像やノイズ除去などの様々な手法を用いて解像度、S/N 比の改善、不要な情報の抑制などの画像処理を行うことである。これらの画像処理タスクはお互いに重なる部分も多い。

まず、同じモダリティで画質を向上させた画像を合成するアプリケーションについては、図11のように、低線量の CT 画像（通常の線量画像と比較して画質が劣る）からの「通常線量」CT 画像の再構成は、深層 CNN を用いる方法⁸¹⁾や、U-Net の改変版⁸²⁾を使用する方法などが発表されている。

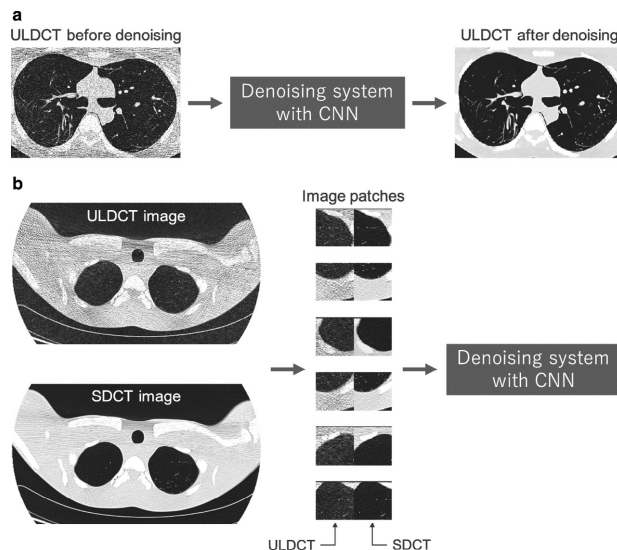


図11 ファントムの超低線量 CT（ULDCT）画像のノイズ除去システムとその訓練データの代表例のシェーマ。(a). CNN によるノイズ除去システム。(b). 学習段階で使用されるトレーニングデータ。SDCT：標準線量 CT
出典：文献 [24].

異なるモダリティの画像を合成する例としては、骨盤 MRI から CT 画像を生成するために FCN を用いた研究⁸³⁾、生成モデルの一種 GAN⁸⁴⁾を用いた MRI の「圧縮センシング⁸⁵⁾」といわれる、少ない収集データから画質を保持したままりアルタイムに再合成する高速撮像技術が提案されている⁸⁶⁾。この Generative Adversarial Network : GAN（敵対的生成ネットワーク）は generator と discriminator という 2つのネットワークからなる教師なし学習のアルゴリズムで、Generator は訓練データと同じようなデータを生成しようとし、discriminator はデータが訓練データから来たものか、それとも生成モデルから来たものかを識別する。学習が進んでいき、最終的には、generator は訓練データと同じようなデータを生成できるようになることが期待される。GAN は足りないデータを補完できるため、本物と同様に AI のトレーニングに役立つ。利用できるデータの量が限られる医療などの分野で重要になるとみられている。完全に合成された患者のデータセットを生成することが可能になる。GAN は、新たなブレイクスルーになる可能性があるということで注目されている。

画像の超解像技術の例としては、Umehara⁸⁷⁾らが公開されている胸部 CT 画像データセットを用いて、CNN に基づく画像超解像手法を適用して、間質性肺疾患の検出に適した高解像度 CT 画像を生成している。

胸部 X 線写真における肺構造描出を向上させる試みとし

[11] ground truth：衛星データとの比較評価用に地上で測定すること。転じて、機械学習では学習に使われる真の正解データを意味する。



図12 骨消去の例。

出典：Schalekamp, S., van Ginneken, B., Koedam, E. et al. Computer-aided detection improves detection of pulmonary nodules in chest radiographs beyond the support by bone-suppressed images. *Radiology*. 2014; 272: pp. 252-261. 左) 処理前, 右) 処理後. (文献 88) とは無関係.)

て、骨構造（肋骨および鎖骨）を抑制（図12）した画像生成においても深層学習の有用性が示されている⁸⁸⁾。機械学習で胸部CT画像の血管影を消すことにより肺野結節影を検出しやすくするシステムも製品化されている⁸⁹⁾。

5.6. Radiomics・Radiogenomics

Radiomicsとは放射線医学で扱われるMRI、CT、超音波、PETなど多量の情報を系統的に扱う科学という意味合いの造語である。具体的にはこれらの医用画像データをもとに、ラベル付けされた画像の特徴を用いて解析、学習を行い、疾病や予後を推論、推定する。Radiogenomics^[12]は、*Radiology*とgenomeが含まれているように、医用画像の特徴パターンと遺伝子レベルでの相関関係を明らかにして予防および治療に結び付けて行こうとする研究である。精密医学Precision Medicineは、ゲノム配列を含む個々の患者情報、医用画像、その他大量のデータを評価することによって、個体の表現型に基づいた予防および治療を行おうとする手法であるが、精密医学の成功は、良好な定量的バイオマーカーに大きく依存している⁹⁰⁾。中でも医用画像は非侵襲的^[13]かつ臨床診療で日常的に実施されており、定量的イメージングバイオマーカーを計算するために使用することができる⁹¹⁾。

深層学習によって、ルーチンのCTやMRIから、現在は見られないゲノムや診断特性を反映した画像上の特徴を視覚化する人間では見つけられない遺伝子マーカーを予測できる。

深層学習は、人間の入力なしに新しいバイオマーカーを発見することができ、この目標を達成するための理想的なツールであると考えられている⁹²⁾。Radiogenomicsでは、侵襲を伴う生検等を行わなくてもルーチンのCTやMRI画像だけで、AIがゲノム配列の変異を予測して疾病の予後や適切な治療方法を提案できるようになる可能性がある。

6. 画像診断領域におけるAIの現状と問題点

深層学習が、将来、放射線科領域において大きな役割を果たすということはもはや疑う余地はない。近い将来、AI技術により、指摘すべき画像が抽出され、疾患候補名も表示されるようになれば、診断医の負担が大幅に軽減される。さらに病変の検出から予備レポート作成までAIが行うようになれば、例えば左右を書き間違えたとか、骨折や肺結節を見逸したなどといった事象は、過去のものとなるだろう。そうして最終的に、AIが（少なくとも画像読影能力において）放射線医にとって代わると予測する者もいる。

しかし現実には、導入の前に解決すべき複数の課題がある。最優先の課題はやはり技術的課題である。最近多くの研究が放射線科専門医に匹敵する性能を示している^{93,94,95)}が、それらは未だ放射線科医が実行するタスクのごく一部でしかない。まだまだAIが画像診断の全領域において放射線科医に取って代わるにはほど遠い。その潜在的能力からみると未だ搖籃期に過ぎない。

AIは学習に用いるデータの量と質に依存する。AIの利活用においては、きちんと診断のついたデータセットの集積が最も重要である。放射線画像は統一規格のDICOM形式で保存されたデジタルデータであり、データベースが十分に大きいといった要因は、医療領域の出発点として画像診断が選択される理由ともなっているが、初期の時代の電子カルテの文化が災いして、さまざまなデータが存在するにもかかわらず、各々がサイロ化しているために、総括的なデータベースの作成を妨げている。しかも、ほぼすべての医療機関が画像アーカイブに膨大な量のデータを保存しているが、それらは手つかずの原材料というべきもので、機械学習の前に「抽出」、「精製」といった作業が必要となる。

世界の多くの企業や政府機関が必要なデータを取得し、有用なデータベースの構築を急いでいる。我が国でも、厚生労働省が、国立研究開発法人日本医療研究開発機構（AMED）を通じて、関係学会を中心としたデータベースの構築を行っている⁹⁶⁾。さらにこれらデータベースを、二次利用しやすくするためのオプションとして「次世代医療基盤法」を策定、2018年5月に施行した。国が医療情報を匿名化する業者を認定することで、医療情報を用いた研究・開発を加速させるのが狙いである。

診療情報を活用する仕組みづくりはプライバシーを配慮して展開することが絶対条件⁹⁷⁾であるし、Radiomicsでは、ゲノム解析、AIの使用判断等と患者への情報共有のための一定の判断基準作りなど、AIの利活用についての法的整備も必要である。

AIはあくまでclosed-endな教師付きデータに基づく学習によって結果を予測しているのであって、データにないよ

[12] Radiogenomicsには放射線が遺伝子改変に与える有害性に着目し、これを研究しようというアプローチのことをいう場合もある。

[13] もちろんX線を使用する限り最低限の侵襲性はある。

うな新しい、あるいはまれな事象には予測の精度は著しく低下する。人種や地域、罹病率、診断機器、撮像条件等の違いを考慮した場合、開発された方法が広く有用であることをどのように確認できるのだろうか？

現在の深層学習技術のブラックボックス性も考慮されなければならない。深層学習が優れた結果を示しているが、システムの技術的および論理的根拠を説明することはほとんど不可能である。エビデンスが重視される今日の医療においては致命的な欠点といえる。

このようなブラックボックス性や、学習に用いたデータセットの信頼性に加えて、CADの実装に関しては、学習によってソフトウェアの性能が継続的に変化していく点など、米FDA（食品医薬品局）や我が国のPMDA（医薬品医療機器総合機構）はどのようにして審査し^[14]評価するのだろうか？

さらにAIを採用する場合、患者の被害につながるような事象があった場合、誰が責任を取るのか？

倫理的・法的・社会的課題（ELSI: Ethical, legal, and social issues）は山積みである。

もう一つの問題は、機械学習アプリケーションはクラウドを使用して提供され、調整されると考えられるが、ほとんどの画像診断部門では、クラウドを信頼して使用するには時間を要するであろうし、患者個人情報のセキュリティを重視する病院医療情報部門がこのようなサービスを提供する新会社を容易には承認しないだろう。

また、AIの研究開発には非常にコストがかかる。十分な資金力を持ち、有能なデータサイエンティストを引き付けることができる企業は少ない。よく知られた米Google、Facebook、Microsoft、IBM、中国Baidu、Alibaba、Tencentなどの技術巨人たちに占有されることになってしまう可能性も考えられる。我が国はAIの開発競争に関しては明らかに出遅れている。わが政府がデータ利活用に向けての環境整備をオールジャパンで急ぐ理由の一つでもあろう。

7. 結 語

機械学習の紹介と、画像診断分野におけるAIの現況について概観した。

深層学習ブームは熱狂と恐怖の中にある。しかし冷静に現状を見るならば、医療分野でのAIの成果は、まだほんの一部でしかない。しかもアプリケーションの開発には、放射線科医や病理医といった専門医のドメイン知識が必須である。さらにCADの実装には解決すべきELSIその他の障壁がある。

医学は合理的でエビデンスに基づくべきというのは簡単だが、現実の医療は、あいまいで、人間関係とか、励ましとか、安心感なども同様に重視されるように、医療は単なる診断や



図 13 “AI”zen Myo'o (Ragaraja), Seated – (Tokyo National Museum)

治療のアルゴリズムではない。深層学習についても、万能の汎用人工知能となれるかについては、当のHinton^[98]も含め、AI専門家の中でも、疑問視する考えも少なからずある^[99]。医療におけるAIの立ち位置はあくまで、医師が診療を提供する方法を合理的に変革し、医師の仕事の一部を補完するものでなければならない。

しかし、画像認識は深層学習の最も得意とする分野であり、医療の中でも、画像診断分野においては、AIによる技術革命が、大きな変化をもたらすことは間違いない。

画像診断分野は数十年の間に、CTの発明からMRI、PETの導入など新しいモダリティの登場や、レントゲンフィルムからPACSへ診療形態の変革など、技術の進歩により大きく変化してきたが、放射線科医はこの変化に対応してきた。

AIは放射線科医にとってかわるのではなく、読影作業における単調な繰り返し作業などの重荷を減らし、共に働く優秀で、注意深い、忠実な助手として加わるだろう。その先では、放射線科医は画像診断の専門家として、数多くの機械学習アルゴリズムを統括管理し、AIによる読影を点検し、より直接患者に係わるかたちの診療形態におのずとシフトしていくのではないだろうか。

現時点でいえることは、近い将来、AIに対応できない放射線科医は、これをうまく活用する放射線科医にとってかわられるであろうということである。

現実はずいぶんAIが玄関口にどっかと腰を下ろしているのである（図13）。

[14] 2017年1月、FDAはベンチャー企業が開発した心臓のMRIデータの解析システムを医療機器として正式に認可した。このシステムは深層学習技術を用いており、AIがどのような理由で診断を下したか知ることができないが、現時点では解析能力の高さが、従来基準で評価されたことで認可された。

8. 引用文献リスト

- 1) J.McCarthy 教授の FAQ (質問と回答) 形式の AI の解説.WHAT IS ARTIFICIAL INTELLIGENCE ? . <http://www-formal.stanford.edu/jmc/whatisai/whatisai.html>. (Accessed 2018-09-23)
- 2) Demis Hassabis. MACHINE LEARNING; What we learned in Seoul with AlphaGo. Google Official Blog. <https://www.blog.google/technology/ai/what-we-learned-in-seoul-with-alphago/>. (Accessed 2018-07-26)
- 3) Alpaydin, Ethem. Introduction to Machine Learning (Adaptive Computation and Machine Learning series). The MIT Press. (xxxi) (Kindle 版). 2009-12-04,
- 4) Nisie A, et al. Current radiologist workload and the shortages in Japan: how many full-time radiologists are required?. Jpn J Radiol.33(5) pp 266-272 .2015.
- 5) Fenton JJ, Taplin SH, Carney PA, Abraham L, Sickles EA, D'Orsi C, et al. "Influence of computer-aided detection on performance of screening mammography". N Engl J Med, vol.356, pp 1399-1409. 2007
- 6) Lehman CD, Wellman RD, Buist DS, Kerlikowske K, Tosteson AN, Miglioretti DL, et al." Diagnostic accuracy of digital screening mammography with and without computer-aided detection". JAMA Intern Med, vol.175, pp 1828-1837., 2015
- 7) 松尾 豊. 人工知能は人間を超えるか (角川 E P U B 選書) (Kindle の位置 No.1445). KADOKAWA / 中経出版. Kindle 版.2015
- 8) Creative Destruction Lab. "Geoff Hinton: On Radiology" The Artificial Intelligence Channel. <https://www.youtube.com/watch?v=2HMPRXstSvQ&feature=youtu.be&t=29>. (Accessed 2018-07-26).
- 9) 谷田部卓. ビジネスで使う機械学習 (Kindle の位置 No.126-127). Yatabe Takashi. Kindle 版.
- 10) Chapelle O, Schölkopf B, Zien A. Semi-supervised learning. MIT Press; Cambridge, MA, USA : 2006.
- 11) 神島敏弘. 推薦システム. 2017-04-07 更新. <http://www.kamishima.net/archive/recsys.pdf>. (Accessed 2018-09-23)
- 12) Kingma DP, et. al. Semi-Supervised Learning with Deep Generative Models, Advances in Neural Information Processing Systems 27 (NIPS) , 2014.
- 13) Rasmus A., et. al. Semi-Supervised Learning with Ladder Network, 2015, <http://arxiv.org/abs/1507.02672>. (Accessed 2018-09-23)
- 14) Silver D., et al. Mastering the game of Go with deep neural networks and tree search. Nature. Volume 529, pp 484-489. (28 January 2016)
- 15) Vapnik V. and Lerner A. Pattern recognition using generalized portrait method. Automation and Remote Control. 24, pp 774-780, 1963
- 16) Rosenblatt, F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. Psychological Review.65(6) : pp 386-408. 1958
- 17) Rumelhart DE, Hinton GE, and Williams RJ: Learning representations by back-propagating errors, Nature, 323(6088) ,pp 533-536, 1986.
- 18) Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. The Journal of Machine Learning Research, Volume 15 Issue 1, January 2014 pp 1929-1958
- 19) 岡野原 大輔. 深層学習. 人工知能学会監修 神島敏弘編. 近代科学社版 (Kindle 版). pp 139-145.
- 20) Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. Neural Comput.; 18: pp 1527-1554. 2006
- 21) 松尾 豊. 人工知能は人間を超えるか (角川 E P U B 選書) (Kindle の位置 No.1794). KADOKAWA / 中経出版. Kindle 版.2015
- 22) Krizhevsky A, Sutskever I, Hinton GE, editors. ImageNet classification with deep convolutional neural networks; Proceedings of the 25th International Conference on Neural Information Processing Systems; 2012 December 3-6; Lake Tahoe, NV, USA. Curran Associates Inc.; pp 1-9. 2012.
- 23) Albelwi, S. Mahmood, A. A Framework for Designing the Architectures of Deep Convolutional Neural Networks. Entropy 2017, 19, 242. <https://www.mdpi.com/1099-4300/19/6/242> (Accessed 2018-09-23)
- 24) Rikiya Yamashita.Mizuho Nishio, Richard Kinh Gian Do, Kaori Togashi. Convolutional neural networks: an overview and application in radiology. Insights into Imaging. Volume 9, Issue 4, p611-629 August 2018, <https://link.springer.com/article/10.1007/s13244-018-0639-9>. (Accessed 2018-09-23).
- 25) Daniel LK, Yamins Ha Hong, Charles F Cadieu, Ethan A Solomon, Darren Selvert, and James DiCarlo, Performance-optimized Hierarchical models predict neural responses in higher visual cortex.Proceedings of the National Academy of Sciences of the United States of America.2014
- 26) 岡谷貴之. 深層学習. 人工知能学会監修 神島敏弘編. 近代科学社版 (Kindle 版). pp 176-177.2015

- 27) 久保陽太郎. 深層学習. 人工知能学会監修 神島敏弘編. 近代科学社版 (Kindle 版). p214-217. 2015
- 28) Gregor K, Danihelka I, Graves A, Rezende DJ, Wierstra D, editors. DRAW: a recurrent neural network for image generation; Proceedings of the 32nd International Conference on Machine Learning (ICML-2015) ; 2015 July 6-11; Lille, France. JMLR; pp 1462-1471. 2015.
- 29) Image-net.org Website. Large Scale Visual Recognition Challenge 2012(ILSVRC2012) results. <http://www.image-net.org/challenges/LSVRC/2012/results.html>. (Accessed 2018-05-03).
- 30) Image-net.org Website. Large Scale Visual Recognition Challenge 2016(ILSVRC2016) results. [Accessed April 1, 2017]. <http://image-net.org/challenges/LSVRC/2016/results>
- 31) Caffe. <http://caffe.berkeleyvision.org/>. (Accessed 2018-05-03).
- 32) Theano: theano 1.0. <http://deeplearning.net/software/theano/>. (Accessed 2018-05-03).
- 33) Chainer. <https://chainer.org/>. (Accessed 2018-05-03).
- 34) TensorFlow. An open source machine learning library for research and production. <https://www.tensorflow.org/?hl=ja>. (Accessed 2018-05-03).
- 35) Collobert R, van der Maaten L, Joulin A, editors. Torchnet: an open-source platform for (deep) learning research; Proceedings of the 33rd International Conference on Machine Learning (ICML-2016) ; 2016 June 19-24; New York, NY, USA. New York: JMLR; 2016.
- 36) scikit-learn. Machine Learning in Python. <http://scikit-learn.org/stable/>. (Accessed 2018-05-03).
- 37) AWS Marketplace: Homepage.https://aws.amazon.com/marketplace/?nc2=h_ql_mp. (Accessed 2018-09-23).
- 38) Cloud Machine Learning Engine. <https://cloud.google.com/ml-engine/?hl=ja>. (Accessed 2018-09-23).
- 39) Built with Watson. <https://www.ibm.com/watson/developer/>. (Accessed 2018-05-03).
- 40) Badrinarayanan V, Kendall A, Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. PAMI, 2017.
- 41) Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI) , Springer, LNCS, Vol.9351 : pp 234- -241, 2015
- 42) Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: CVPR. (2016)
- 43) Ren S, He K, Girshick R, Sun J, editors. Faster R-CNN: towards real-time object detection with region proposal networks; Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS-2015) ; 2015 December 7-12; Montreal, Canada. Cambridge: MIT Press; pp 91-99. 2015.
- 44) Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu CY., Berg AC. SSD: Single Shot MultiBox Detector. Computer Vision - ECCV 2016.pp 21-37.2016
- 45) Mullachery V, Motwani V. Image Captioning. ArXiv.org Web site. <https://arxiv.org/pdf/1805.09137.pdf>. (Accessed 2018-09-23).
- 46) Margaret Mitchell, C. Lawrence Zitnick, Dhruv Batra, Devi Parikh. VQA: Visual Question Answering. ArXiv.org Web site. <https://arxiv.org/pdf/1505.00468.pdf>. (Accessed 2018-09-23).
- 47) VQA とは. 原田・牛久研究室.<https://www.mi.tu-tokyo.ac.jp/research/vqa/> (Accessed 2018-09-23).
- 48) Tajbakhsh N, Shin J Y, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Liang J. Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE transactions on medical imaging, 35 (5) : pp 1299-1312, 2016.
- 49) Antropova N, Huynh BQ, Giger ML. A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Medical physics, 2017.
- 50) Wang X, Yang W, Weinreb J, Han J, Li Q, Kong X, Yan Y, Ke Z, Luo B, Liu T, Wang L. Searching for prostate cancer by fully automated magnetic resonance imaging classification: Deep learning versus non-deep learning. Scientific Reports, 7(1) : 15415, 2017. ISSN 20452322. doi: 10.1038/s41598-017-15720-y.
- 51) Xue Y, Zhang R, Deng Y, Chen K, Jiang T. A preliminary examination of the diagnostic value of deep learning in hip osteoarthritis. PLOS ONE, 12(6) : e0178992, Jun 2017. ISSN 1932-6203. doi: 10.1371/journal.pone.0178992.
- 52) Zhu Z, Albadawy E, Saha A, Zhang J, Harowicz MR, Mazurowski MA. Deep Learning for identifying radiogenomic associations in breast cancer. arXiv preprint arXiv: 1711.11097, 2017.

- 53) Simonyan K Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv : 1409.1556, 2014.
- 54) Suk HI, Lee SW, and Shen D. Deep ensemble learning of sparse regression models for brain disease diagnosis. *Medical Image Analysis*, 37 : p101–113, 2017. ISSN 13618415. doi : 10.1016/j.media.2017.01.008.
- 55) Canziani A, Paszke A, and Culurciello E. An analysis of deep neural network models for practical applications. arXiv preprint arXiv : 1605.07678, 2016.
- 56) He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp 770–778, 2016.
- 57) Szegedy C, Ioe S, Vanhoucke V, Alemi AA Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. In *AAAI*, pp 4278–4284, 2017.
- 58) Korfiatis P, Kline TL, Lachance DH, Parney IF, Buckner JC, Erickson BJ. Residual Deep Convolutional Neural Network Predicts MGMT Methylation Status. *Journal of Digital Imaging*, 30(5) : pp 622–628, Oct 2017. ISSN 1618727X. doi : 10.1007/s10278-017-0009-.
- 59) Yoo Y, Tang LYW, Brosch T, Li DK.B., KolindS, Vavasour I, Rauscher A, MacKay AL, Traboulsee A, Tam RC. Deep learning of joint myelin and T1w MRI features in normal-appearing brain tissue to distinguish between multiple sclerosis patients and healthy controls. *NeuroImage : Clinical*, 17 : pp 169–178, 2018. ISSN 22131582. doi : 10.1016/j.nicl.2017.10.015
- 60) Ciresan D, Giusti A, Gambardella LM, Schmidhuber J, eds. Deep neural networks segment neuronal membranes in electron microscopy images. *Proceedings of the 25th International Conference on Neural Information Processing Systems*; 2012 December 3–6; Lake Tahoe, NV, USA. Curran Associates Inc., 2012 : pp 2843–2851
- 61) Milletari F, Ahmadi SA, Kroll C, Plate A, Rozanski V, Maiostre J, Levin J, Dietrich O, Ertl-Wagner B, Boetzel K, Nabab N. Hough-CNN : Deep learning for segmentation of deep brain regions in MRI and ultrasound. *Computer Vision and Image Understanding*, 2017.
- 62) Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2017; 39 : pp 640–651.
- 63) Brosch T, Yoo Y, Tang LYW, Li DKB, Traboulsee A, Tam R. Deep convolutional encoder networks for multiple sclerosis lesion segmentation. In : Navab N, Hornegger J, Wells WM, Frangi AF, editors. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015*; New York. Springer; 2015. pp 3–11.
- 64) Miao He, Haibo Luo, Zheng Chang and Bin Hui. Pedestrian Detection with Semantic Regions of Interest. *Sensors* 17(12) :2699 · November 2017. DOI: 10.3390/s
- 65) Szegedy C, Reed S, Erhan D, Anguelov D, Ioe S. Scalable high-quality object detection. arXiv preprint arXiv : 1412.1441, 2014.
- 66) Tajbakhsh N, Shin JY, Gurudu SR, HurstR T, Kendall CB, Gotway MB, Liang J. Convolutional neural networks for medical image analysis : Full training or fine tuning? *IEEE transactions on medical imaging*, 35 (5) : pp 1299–1312, 2016.
- 67) Samala RK, Chan HP, Hadjiiski L, Helvie MA, Wei J, and Cha K. Mass detection in digital breast tomosynthesis : Deep convolutional neural network with transfer learning from mammography. *Medical physics*, 43(12) : pp 6654–6666, 2016.
- 68) Simonyan K Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv : 1409.1556, 2014.
- 69) Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 1–9, 2015.
- 70) Christian Szegedy, Sergey Ioe, Vincent Vanhoucke, and Alexander A Alemi. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. In *AAAI*, pp 4278–4284, 2017.
- 71) He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp 770–778, 2016.
- 72) Platania R, Shams S, Yang S, Zhang J, Lee, Park SJ. Automated Breast Cancer Diagnosis Using Deep Learning and Region of Interest Detection (BC-DROID). In *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, pp 536–543. ACM, 2017.
- 73) Cao Z, Duan L, Yang G, Yue T, Chen Q, Fu H, Xu Y. Breast Tumor Detection in Ultrasound Images Using

- Deep Learning. In International Workshop on Patch-based Techniques in Medical Imaging, pp 121–128. Springer, 2017.
- 74) Shi Y, Suk HI, Gao Y, Shen D. Joint coupled-feature representation and coupled boosting for AD diagnosis. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit.* 2014; pp 2721–2728. 2014 :
 - 75) Shin HC, Roberts K, Lu L, Fushman DD, Yao J, Summers RM, editors. Learning to read chest X-rays : recurrent neural cascade model for automated image annotation; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016 June 27–30; Las Vegas, NV, USA. IEEE; 2016.
 - 76) Chen MC, Ball RL, Yang L, Moradzadeh N, Chapman BE, Larson DB, Langlotz CP, iAmrhein TJ, Lungren MP. Deep Learning to Classify Radiology Free-Text Reports. *Radiology*, page 171115, Nov 2017. ISSN0033–8419. doi : 10.1148/radiol.2017171115.
 - 77) Shin B, Choksi FH, Lee T, Choi JD. Classification of radiology reports using neural attention models. In Proceedings of the International Joint Conference on Neural Networks, volume 2017–May, pp 4363–4370, 2017. ISBN 9781509061815. doi : 10.1109/IJCNN.2017.7966408.
 - 78) Sarvnaz Karimi, Xiang Dai, Hamedh Hassanzadeh, and Anthony Nguyen. Automatic Diagnosis Coding of Radiology Reports : A Comparison of Deep Learning and Conventional Classification Methods. *BioNLP*, pp 328–332, 2017.
 - 79) Hessam Sokooti, Bob de Vos, Floris Berendsen, Boudewijn P. F. Lelieveldt, Ivana Isgum, and Marius Staring. Non-rigid Image Registration Using Multi-scale 3D Convolutional Neural Networks. In *Medical Image Computing and Computer Assisted Intervention*, p232–239.2017. ISBN 978–3–319–66182–7. doi : 10.1007/978–3–319–66182–7 27.
 - 80) Bob D de Vos, Floris F Berendsen, Max A Viergever, Marius Staring, and Ivana I sgum. End-to-End Unsupervised Deformable Image Registration with a Convolutional Neural Network. *arXiv preprint arXiv : 1704.06065*, 2017.
 - 81) Eunhee Kang, Junhong Min, and Jong Chul Ye. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. *Medical Physics*, 44(10) : pp 360–375, Oct 2017. ISSN 00942405. doi : 10.1002/mp.12344.
 - 82) Kyong Hwan Jin, Michael T. McCann, Emmanuel Froustey, and Michael Unser. Deep Convolutional Neural Network for Inverse Problems in Imaging. *IEEE Transactions on Image Processing*, 26(9) : pp 4509–4522, 2017. ISSN 10577149. doi : 10.1109/TIP.2017.2713099.
 - 83) Dong Nie, Xiaohuan Cao, Yaozong Gao, Li Wang, and Dinggang Shen. Estimating CT Image from MRI Data Using 3D Fully Convolutional Networks BT - Deep Learning and Data Labeling for Medical Applications : First International Workshop, LABELS 2016, and Second International Workshop, DLMIA 2016, Held in Conjunction with MICC. pp 170–178. Springer International Publishing, Cham, 2016. ISBN 978–3–319–46976–8. doi : 10.1007–978–3–319–46976–8 18.
 - 84) Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Networks. 10 Jun 2014. *arXiv : 1406.2661*
 - 85) 重田高志. MRI 技術の最近のトレンド (第3回) : MRI 画像再構成技術. *MEDICAL IMAGING TECHNOLOGY* Vol. 31 No. 4 September 2013
 - 86) Guang Yang, Simiao Yu, Hao Dong, Greg Slabaugh, Pier Luigi Dragotti, Xujiong Ye, Fangde Liu, Simon Arridge, Jennifer Keegan, Yike Guo, and David Firmin. DAGAN : Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction *IEEE Transactions on Medical Imaging*, pages 1–1, 2017. ISSN 0278–0062. doi : 10.1109/TMI.2017.2785879.
 - 87) Kensuke Umehara, Junko Ota, and Takayuki Ishida. Application of Super-Resolution Convolutional Neural Network for Enhancing Image Resolution in Chest CT, Oct 2017. ISSN1618727X.
 - 88) W Yang, Y Chen, Y Liu, L Zhong, G Qin, Z Lu, Q Fenq, W Chen. Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain. *Medical Image Analysis*, 35 : pp 421–433, 2017.
 - 89) ClearRead CT. Technology Overview. <https://www.riveraintech.com/wp-content/uploads/2017/08/ClearRead-CT-Technology-Overview.pdf> (Accessed 2018–09–23).
 - 90) Colburn W, et al. Biomarkers and surrogate endpoints : Preferred definitions and conceptual framework. Biomarkers Definitions Working Group. *Clinical Pharmacol & Therapeutics*. 2001; 69 : pp 89–95. doi : 10.1067/mcp.2001.113989.

- 91) Greenland P, LaBree L, Azen SP, Doherty TM, Detrano RC. Coronary artery calcium score combined with Framingham score for risk prediction in asymptomatic individuals. *Jama*. 2004; 291 : pp 210-215. doi : 10.1001/jama.291.2.210.
- 92) Luke Oakden-Rayner, Gustavo Carneiro, Taryn Bessen, Jacinto C. Nascimento, Andrew P. Bradley, and Lyle J. Palmer. Precision Radiology : Predicting longevity using feature engineering and deep learning methods in a radiomics framework 2017 May 10. doi : 10.1038/s41598-017-01931-
- 93) William Gale, Luke Oakden-Rayner, Gustavo Carneiro, Andrew P Bradley, and Lyle J Palmer. Detecting hip fractures with radiologist-level performance using deep neural networks. *arXiv preprint arXiv : 1711.06504*, 2017.
- 94) Monika Grewal, Muktabh Mayank Srivastava, Pulkit Kumar, and Srikrishna Varadarajan. RADNET : Radiologist Level Accuracy using Deep Learning for HEMORRHAGE detection in CT Scans. *arXiv preprint arXiv : 1710.04934*, 2017.
- 95) Jameson Merkow, Robert Luftkin, Kim Nguyen, Stefano Soatto, Zhuowen Tu, and Andrea Vedaldi. DeepRadiologyNet : Radiologist Level Pathology Detection in CT Head Images. *arXiv preprint arXiv : 1711.09313*, 2017.
- 96) 厚生労働省 . 保健医療分野における AI 活用推進懇談会 報告書 . 2017 年 .[http : //www.mhlw.go.jp/stf/shingi/other-kousei.html?tid=408914](http://www.mhlw.go.jp/stf/shingi/other-kousei.html?tid=408914). (Accessed 2018-09-23).
- 97) 日本医師会 . 第 IX 次学術推進会議報告書「人工知能 (AI) と医療」[http : //dl.med.or.jp/dl-med/teireikaiken/20180620_3.pdf](http://dl.med.or.jp/dl-med/teireikaiken/20180620_3.pdf). (Accessed 2018-09-23).
- 98) Steve LeVine. Artificial intelligence pioneer says we need to start over. Sep 15, 2017. [https : //www.axios.com/artificial-intelligence-pioneer-says-we-need-to-start-over-1513305524-f619efbd-9db0-4947-a9b2-7a4c310a28fe.html](https://www.axios.com/artificial-intelligence-pioneer-says-we-need-to-start-over-1513305524-f619efbd-9db0-4947-a9b2-7a4c310a28fe.html). (Accessed 2018-09-23).
- 99) Gary Marcus. Deep Learning : A Critical Appraisal. *arXiv preprint arXiv : 1801.00631*.2018. [https : //arxiv.org/ftp/arxiv/papers/1801/1801.00631.pdf](https://arxiv.org/ftp/arxiv/papers/1801/1801.00631.pdf). (Accessed 2018-09-23).

