



Technical  
Survey

技術解説

# ディープラーニング

岡谷 貴之<sup>†</sup>

多層ニューラルネットを用いる機械学習の方法（ディープラーニング）が最近注目を集めており、画像や音声の認識などの難問を解決し得る有力な方法として期待されている。本解説では、ニューラルネットの基礎から始め、過学習の克服を可能にし今の研究ブームのきっかけを作ったプレトレーニングと、さらには画像認識で本命視される畳込みネットワークについて述べる。

キーワード：ニューラルネットワーク、画像認識、機械学習

## 1. まえがき

ディープラーニング（深層学習）とは、多層（＝ディープ）の人工ニューラルネットワーク（以下ニューラルネットあるいは単にネットと呼ぶ）を用いる機械学習の方法論を指す。近年、人工知能の諸問題で、このディープネットを用いた方法が従来法を大きく上回る性能を示すことが確認され、その研究開発はブームと言ってよいような活況を呈している。音声認識、画像認識、自然言語処理や製薬（化合物反応予測）など、さまざまな問題に応用されて成果を挙げている。

本稿では、ディープネットの画像認識への応用を主にとりあげる。ディープネットが画像認識にも高い性能を示すことが誰の目にも明らかになったきっかけは、2012年9月開催の一般物体認識のコンテストILSVRC (Image Net Large Scale Visual Recognition Challenge) であった。一般物体認識とは、1枚の画像に写る物体のカテゴリー名を答えるタスクで、画像認識の中でも難しいタスクの一つである。同コンテストは毎年開催されているが、この年、トロント大のチームが持ち込んだディープネットが、群を抜いた性能を示したのだった<sup>1)</sup>。彼らの方法は、それなりに長い研究の歴史を持つ従来手法に圧倒的な性能の差を示し、このことは、新たな時代の到来を印象づけた。

画像認識以外でもさまざまな問題で成功しつつあるディープネットであるが、その成功の理由は、特徴量そのものを学習する能力にあると考えられる。音声や画像を対象とした多くの認識問題は、入力となる音声信号や画像から、まず特徴を取り出し、取り出した特徴を分類するという2段階に分けて考えることができる（図1）。後半の分類は、サポートベクターマシンに象徴される1990年代の機械

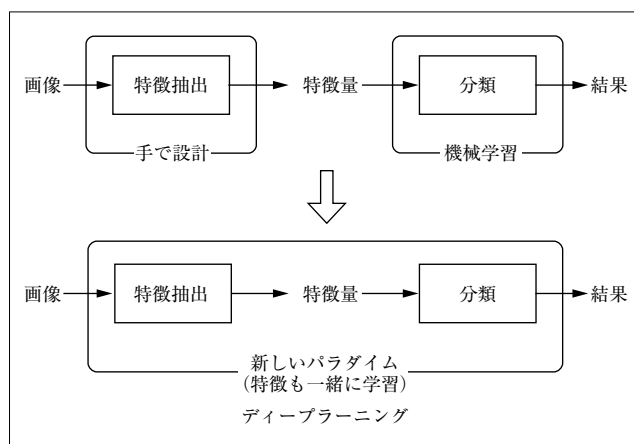


図1 画像認識研究のパラダイム変化

学習の進展により、それほど難しい問題ではなくなった。

今でも難しいことは、入力からいかなる特徴を抽出するかである。画像認識では、過去長期にわたる研究の結果、手書き数字の認識や顔の検出・認識など、いくつかの問題が解決されてきた一方で、上述の一般物体認識や動画を使った人の行動認識など、未解決の問題も多く残されている。これらが未解決である最大の理由は、何を特徴として取り出すべきかがはっきりしないことにある。

ここにディープラーニングへの期待がある。これまで、特徴抽出の処理は研究者が手で設計していた。これに対しディープラーニングは、特徴そのものの学習を可能にしてくれる（図1）。特徴学習や表現学習と呼ばれるこの新しい方法は、手で設計するのが難しい特徴量の獲得を可能にし、未解決の問題を解決し得る可能性を持つ。この意味で、画像認識を始めとする人工知能の研究は、一つのパラダイムの変化を迎えようとしている。本稿では、このあたりの状況を踏まえつつ、現在に至る研究の経緯と現況について解説する。

<sup>†</sup> 東北大学 大学院情報科学研究科

"Deep Learning" by Takayuki Okatani (Graduate School of Information Sciences, Tohoku University, Sendai)

## 2. ニューラルネットワーク

最初に、ニューラルネットの構造、および学習の仕方について簡単に述べておく。ここでは、入力から出力まで一方向のみに信号が伝播するフィードフォワード型の最も基本的なもののみを考える。これは、画像認識をはじめとする多くの問題で、もっともよく使われている。

最初に、生体のニューロンを模したユニットを考える(図2(a))。これは一つの入力 $x$ をとり、一つの出力 $y$ を返す働きを持ち、この入出力間の関係は、活性化関数と呼ばれる非線形関数

$$y = f(x) \quad (1)$$

によって決まる。活性化関数 $f$ には、生体のニューロンの入出力応答をシミュレートしたシグモイド関数 $f(x) = 1 / (1 + e^{-x})$ が昔から使われているが、最近は $f(x) = \max(0, x)$ という関数(Rectified Linearと呼ぶ)がより一般的である<sup>2)</sup>(図2(b))。

このようなユニットを層状に並べ、各層間でのみユニット間に結合を持たせたネットワークを考える(図2(c))。一つのユニット(図中の $j$ )は、直前の層の複数のユニットからの出力、 $y_i (i = 1, \dots, m)$ の重み付き和に、バイアスと呼ぶ値を加算した

$$x_j = b_j + \sum_{i=1}^m y_i w_{ij} \quad (2)$$

が入力されると考える。このユニットは、活性化関数を経由した出力、 $y_j = f(x_j)$ を、図2(c)のように上の層に等しく伝える。

このような層を任意の数だけ並べ、多層のネットワークを作る(図3)。入力信号は順に層を伝播してゆき、最後の層から最終的な結果を出力する。なお、解きたい問題がクラス分類の場合には、最後の出力層のみ特別に、目的のクラス数 $n$ と同数のユニットを配置する。その上でこれらのユニットの出力を、それらへの入力 $x_j (j = 1, \dots, n)$ を元にしたソフトマックス、 $p_j = e^{x_j} / \sum_{k=1}^n e^{x_k}$ とする。これらの $p_1, \dots, p_n$ は、 $n$ 個のクラスに対する確からしさとなり、入力を分類するときは、 $p_j$ が最大値をとるクラス $j$ に分類する。

ニューラルネットは全体で一つの関数を表現する。この関数が望みのものになるように、多数の学習サンプルを使って、ネットワークの重み $w_{ij}$ およびバイアス $b_i$ をうまく調整することがニューラルネットの学習である。画像認識などの複雑なタスクでもこれは変わらない。入力された画像に対し、そこに写る物体のクラスを表す出力( $p_1, \dots, p_n$ )を与える関数を求めたい。この場合、画像と物体のクラスラベルがペアになった学習サンプルを多数(数万以上のオーダー)用意し、学習サンプルの画像を実際にネットに入力したときの出力と、望みの出力のずれを誤差 $C$ として定義し、これを小さくすることを考える。具体的には、目

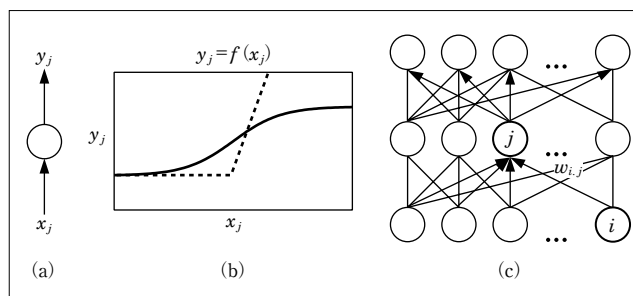


図2 (a)のユニット一つは、(b)の入出力関数(活性化関数)を持つ。これを(c)のように層状に並べ、層間でユニットを結合したものを多層重ねてネットワークを作る。

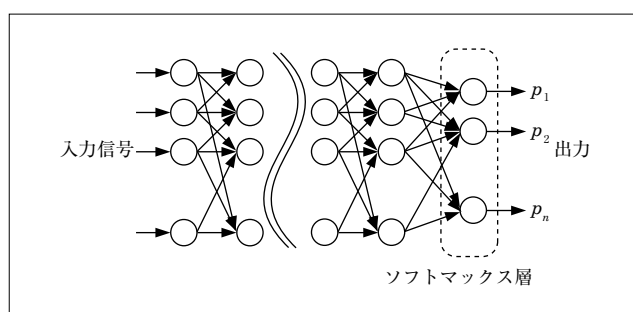


図3 クラス分類の問題を解くニューラルネット  
出力層の $n$ 個のユニットから、入力が $n$ 個のカテゴリのどれくらいかが数値で出力される。

標とする出力 $d_1, \dots, d_n$ と、実際の出力 $p_1, \dots, p_n$ の乖離を、交差エントロピー

$$C = - \sum_{j=1}^n d_j \log p_j \quad (3)$$

によって測る。目標出力 $d_1, \dots, d_n$ は、正解クラス $j$ のみ、 $d_j = 1$ となり、それ以外の全 $k (k \neq j)$ では、 $d_k = 0$ となるようにとる。

このように定義した $C$ の最小化には、勾配降下法が一般的に用いられる<sup>3)</sup>。規模の小さなニューラルネットでは、より使い勝手のよい準ニュートン法を用いることが一般的だが、大規模なディープネットでは今でも勾配降下法がよく使われる。

勾配降下法では、重みやバイアスに関する誤差勾配( $\partial C / \partial w_{ij}$ など)を計算する必要がある。これは、活性化関数層の数分だけ入れ子になった合成関数の微分を計算することになり、出力層に近い層では簡単だが、出力層から離れた深い層では煩雑になる。これを系統的に行えるようにしたものが誤差逆伝播法(Back Propagation)である。活性化関数の合成関数の微分で生じる連鎖規則を、数値計算の内部に組み入れることで、出力層での各ユニットの誤差を、出力層に近い側から逆向きに伝播していくアルゴリズムと

なる。ここでは詳細は割愛する。

なお、誤差  $C$  は全学習サンプルに対して評価する（バッチ計算）のではなく、数個～数百程度のサンプル集合（ミニバッチ）に対して評価する方法がとられる。サンプルをランダムに選んでミニバッチとすることで、大域的な収束性を得る（確率的勾配降下法と呼ばれる）。重みとバイアスの修正は、このミニバッチごとに一度ずつ行う。最も一般的には

$$\Delta w_{ij}^{(t)} = -\varepsilon \frac{\partial C}{\partial w_{ij}^{(t)}} + \alpha \Delta w_{ij}^{(t-1)} - e \lambda w_{ij}^{(t)} \quad (4)$$

のように修正する（バイアスも同様）。ここで  $\Delta w_{ij}^{(t)}$  と  $\Delta w_{ij}^{(t-1)}$  は、それぞれ今回と前回の重み更新時の修正量を表す。右辺第1項は誤差勾配の項で、 $\varepsilon$  は学習係数と呼ばれ、勾配降下のステップ幅を決定する。第2項はモメンタムと呼ばれ、前回修正量の  $\alpha$  ( $\sim 0.9$ ) 倍を今回の修正量に加算することで、ミニバッチの選択に伴う修正量のランダムネスを減じる。第3項は重み減衰項 (Weight Decay) と呼ばれ、重みが過大に大きくならないようにする働きがある。

これらのパラメータ  $\varepsilon$ 、 $\alpha$ 、 $\lambda$  はうまく選ぶ必要がある。実際、学習の行方を左右する。特に学習係数  $\varepsilon$  については、ある程度を大きな値（例えば、 $\varepsilon \sim 0.01$ ）から開始し、学習の進展にともなって  $1/10$  にすることを何度か行うことが一般的である。

ここで述べたフィードフォワード型のニューラルネットの他、出力を入力にフィードバックし、主に時系列データの処理を行うリカレントニューラルネットが有名である。また、確率的にその挙動を記述するボルツマンマシン、特に上で述べたネット単層と同じ構造を持つリストリクティッドボルツマンマシン (RBM) や、それを重ねたディープボルツマンマシンやディープビリーフネットも、ディープラーニングの研究ではよく登場する。これらは、紙面の都合で省略する。

### 3. ディープラーニングの中核技術

#### 3.1 ディープネットの過学習

前章で述べたニューラルネットの構成および学習方法は、1980年代にはすでに知られていた。当時ニューラルネットは活発に研究されていたが、1990年代半ばころから、研究は下火になる。その最大の原因は、ディープネットが学習時にいわゆる過学習を起こし、これを克服する手だてが当時は見つからなかったことによる。

過学習とは、学習サンプルに対する誤差（訓練誤差）は小さくできる一方、汎化誤差が小さくできないことである。当時のディープネットは、未知のサンプルに対する汎化性能が低く、したがって役に立たないと見なされた。誤差逆伝播時によく見られる現象として、出力層から遠ざかるにつれて情報が拡散し、重みがほとんど修正されなくなるこ

とが知られていた。

今のディープネットのブームは、この過学習の問題を克服する方法が見つかったことに端を発する。その最初の方法として登場したのが、次に述べる Hinton らの研究である<sup>4)</sup>。

#### 3.2 過学習の回避：プレトレーニング

ディープネットを誤差逆伝播法で学習する際、ネットワークの重みは、ランダムに初期化することが最も一般的である。Hinton らが発見したことは、その初期値の決め方が適切でないから上述の過学習が起こるのであって、ある方法でよい初期値を得ておけば、過学習を回避できるということであった。

彼らの方法は、図4のように、目的とするディープネットを2層ずつ分離し、各2層ネットを別々に（ただし入力層側から順番に一つずつ）訓練する。各層の訓練は、それぞれをオートエンコーダと考えることで行う。

オートエンコーダとは、入力に対する出力が、なるべく入力に近くなることを理想とするシステムである。今考えている層が、入力  $x$  から出力  $y$  を計算するとき（図5左）、これと逆向きの計算、つまり出力  $y$  を入力側  $x'$  に戻すような計算を新たに考える。この一連の計算つまり  $x \rightarrow y \rightarrow x'$  の計算は、元のネットを出力で折り返した、図5右のような構造の2層ネットによって表現できる。この2層ネット

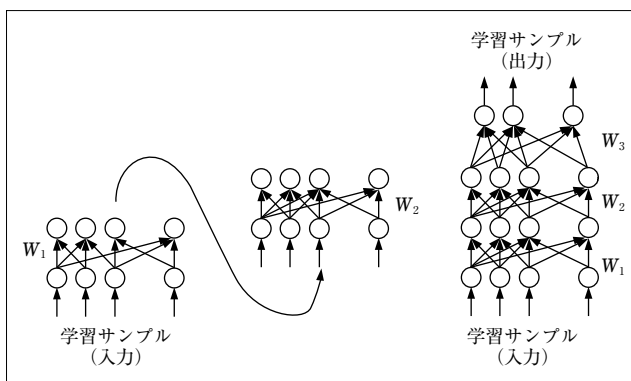


図4 3層ネットのプレトレーニング  
一番左のネットワークの第1層の重み  $W_1$  と第2層の重み  $W_2$  の初期値を、各層のオートエンコーダの学習によって得る。

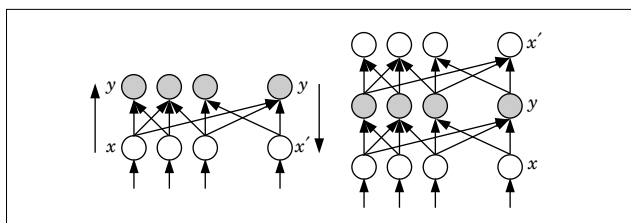


図5 オートエンコーダの概念図  
左の単層のネットワークを折り返してつないだ右のネットワークを考え、入力になるべく近いものを出力するように学習を行う。

が、個々の入力になるべく近い出力を返すように2層の重みを決定する。なお第2層の重みとバイアスは第1層のそれとは一般に無関係でよい。

学習で用いる入力には、最終的にディープネットでは扱おうとしている学習サンプル(の入力)を使う。多数の学習サンプルを対象に、個々のサンプルを入力したときの出力と入力の差が小さくなるよう、通常の誤差逆伝播法を実行し、2層の重みを決定する。自然画像のように、自然界で発生するデータには一般に偏りがある。オートエンコーダの狙いは、そんなデータそのものに内在する偏りをうまくとらえようということである。どんな入力に対しても近い出力を返す必要はないことに注意する(それを目的とすると恒等写像が答えになり、意味がなくなる)。

なおオートエンコーダは、中間層のユニット数が多いほど自由度が高まり、表現能力が増す。その一方で、自由度があまりに大きいと、入力がそのまま出力される恒等写像を学習してしまう。そこで、オートエンコーダが各入力を再現する際、中間層のユニットのうち少数しか使用されないように制約を与えて学習させることを行う。中間層のユニットがごくまばら(=スパース)にしか活性化しないことから、スパース正則化と呼ばれる。

ディープネットのプレトレーニングに話を戻すと、以上のオートエンコーダによる各層の学習を、データが入力される第1層から始め、順に第2層、第3層と上方へと繰り返す。第1層の入力には、上述のように、ディープネットが対象とする学習サンプルと同じものをそのまま使う。第2層の入力には、第1層の学習後決定された重みを使って得られる学習サンプルに対する出力(図4)を用いる。以降はこの繰り返しである。こうして定めた各層の重みが、目的とするディープネットの対応する層の重みの初期値となる。

その後、本来の教師あり学習を誤差逆伝播法で行う。なお、最後の出力層だけは例外で、(この層だけは望みの出力=ディープネットの教師信号があるので)プレトレーニングの対象とせず、重みは従来通りランダムに初期化する。

その後の研究で、この他にも、効果の差こそあれ、同様に過学習を回避する方法がいくつか見つかった。その中で最も有名なものは、ディープネットの誤差逆伝播時に、複数の中間層のユニットを半数程度、確率的に無効化し、残りのネットを訓練するドロップアウト(Dropout)<sup>5)</sup>である。この他にも、ドロップアウトを一般化したドロップコネクト(Dropconnect)、マックスアウト(Maxout)や、識別的プレトレーニング<sup>6)</sup>などが知られている。これらの研究の結果、過学習を回避する方法としてのプレトレーニングの重要性は低下しつつある。ただし、このプレトレーニングの考え方で得られる重みは、入力サンプルの特徴量を教師なし学習でできる方法であり、重みの初期化以外にもさまざまな使い道があり、その意味で今も重要である。

### 3.3 畳込みネットワーク

畳込みネットワーク(Convolutional Neural NetworkあるいはConvolutional Network)とは、畳込み層とプーリング層と呼ばれる2種類の層を、交互に積み重ねた構造を持つディープネットである。そのルーツはFukushimaらのネオコグニトロン<sup>7)</sup>にさかのぼり、さらにその基礎に神経科学の分野でのHubel-Wieselの発見<sup>8)</sup>がある。

畳込みネットの最大の特徴は、それが多層のネットワークであるにも関わらず、前節のようなプレトレーニングを行わなくても、重みをランダムに初期化した状態から直接、誤差逆伝播法で学習を行えてしまうことである。実際、LeCunらは、1980年代末にはすでに多層ネットワークの学習に成功し、文字認識のタスクに適用して高い性能を得ていた<sup>9) 10)</sup>。なぜ畳込みネットが過学習を起こさない(起こしにくい)かは後述する。

畳込みネットは、図6のように畳込み層とプーリング層のペアを複数回積み重ね、最後に通常構造の層を1~3層程度積み重ねて出力につなげる構造をとる(カテゴリー認識の場合、最後の出力は、2章に述べたソフトマックス層とする)。

畳込み層は、多チャネルの画像に同じチャネル数の小サイズの2次元フィルタ(例えば、7×7画素程度)を畳込む演算を行う(図7)。これは、一般的な画像処理でのフィルタの畳込み、すなわち小サイズの画像を入力画像に2次元的に畳込んで、画像をぼかしたり、あるいはエッジを強調するものと基本的に同じである。このフィルタの畳込みによって、局所的なパタンの特徴抽出が行われる。畳込むフィルタが抽出する特徴を表す。

一方プーリング層は、畳込み層の出力を入力とし、入力された多チャネル画像の小領域(例えば、5×5画素程度)での値を、一つの値に集約(プーリング)する(図7)。簡単

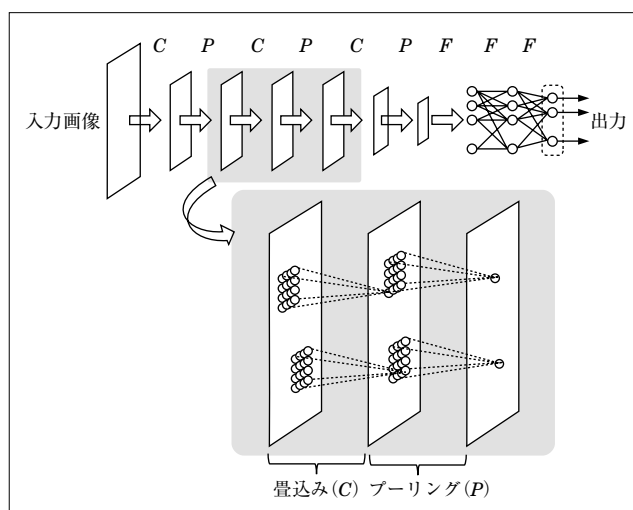


図6 畳込みネットの全体像  
C, P, Fはそれぞれ畳込み層、プーリング層、全結合層を表す。

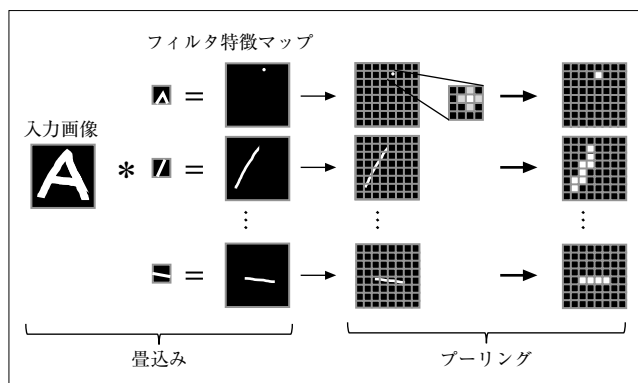


図7 畳込み層およびプーリング層の概要

に言えば、入力された多チャネルの画像をダウンサンプリング（解像度を減らす）するような演算である。これによって、位置ずれにある程度不変な特徴（画像内での特徴の微小な位置変化を無視することで、多少位置がずれても変わらない特徴）が得られる。小領域内の複数の値を一つの値に集約する方法にはいくつかあるが、その中の最大値を選ぶマックスプーリングが、画像認識では最も一般的である。

畳込みネットは、以上の畳込み層とプーリング層を結合し、このペアを何度か繰り返す構造を持つ。畳込み層とプーリング層のペア一つで、局所特徴の抽出と位置の微小変位に対する不変性を実現するが、さらにこれを多層化することで、より複雑な幾何学的変形に対する不変性を実現していると考えられる。

畳込み層およびプーリング層はそれぞれ、その計算内容に対応した結線を持つ2層構造のネットワークとして表せる。どちらも、出力層のユニット一つにつながった入力層のユニットは少数で、画像上で局在しており（神経科学の用語で受容野が局所的であるという）、この点で層間で全ユニットが結合しているようなニューラルネットとは異なっている。ただし、この構造を除けば、畳込みネットも通常のニューラルネットと同じで、誤差逆伝播による学習が可能である。学習の対象となるものは、畳込み層の重み、すなわち上述のフィルタと、出力層付近の全結合層の重みである。プーリング層の結線は固定されており、学習で変化することはない。

先述の通り、畳込みネットでは、特にプレトレーニング等を行わずとも、過学習を起こすことなく、学習を行うことができる<sup>\*</sup>。過学習を起こしにくい理由は、結線が局所的に限定されている畳込み層の構造にあると考えられている。層間で全ユニットが結合したネットの場合、先述のように誤差逆伝播時に、層を経るにしたがって誤差が拡散し小さくなる傾向にあるが、畳込みネットではその局在性ゆえ、同じような誤差の拡散が起こりにくい。

<sup>\*</sup> ただし、出力に近い全結合層は、ドロップアウトなどの過学習をさける方法を用いることが一般的である。

この畳込みネットは現在、ディープラーニングの画像認識応用では最も重要な技術となっている。上述のコンテストILSVRCでのKrizhevskyらのネットをはじめ、今のところ画像認識のいろいろなタスクで良い性能を示しているものはほぼすべて畳込みネットである。興味深いことに、これら最近のネットは、文字認識に適用された1980年代のLeCunらのネットと、その構造および学習方法はほとんど変わっていない。違いは、適用ドメインが広がった（文字認識から画像認識全般へ）ことと、学習サンプル数およびネットワークサイズが大きくなったことである。

#### 4. むすび

最後に、ディープラーニングの画像関係の代表的な応用例について紹介し、本稿のまとめとする。

最初に述べたように、ディープネット（畳込みネット）が従来法よりも特に高い性能を示すタスクが一般物体認識である。この問題の難しさは、同一カテゴリ内の見えの変動がきわめて大きいことから生じる（図8）。大きな見えの変動を乗り越えて同じカテゴリと認識できるためには、変動に不変な特徴を画像から抽出する必要がある。同時に、類似カテゴリと区別できるためには、弁別力（違いに対する敏感さ）が必要である。この二つの相反する目標をいかに両立できるかが課題となる。多層の畳込みネットが持つ変形への柔軟な不変性により、従来法よりも高い性能が実現されていると考えられている。

なお、神経科学のこれまでの研究で、脳の視覚皮質では、階層的に特徴が取り出されているらしいことがわかっている<sup>11)</sup>。畳込みネットの構造自体が神経科学の知見に基づいているが、畳込みネットが学習する特徴にも、同様の階層性が現れる<sup>12)</sup>。この点で、畳込みネット（ディープネット）と視覚皮質（の神経回路網）との類似性が強く意識されている。関連研究に、YouTubeの動画からほぼ無作為に切出

図8 ImageNet (<http://image-net.org/>) から取り出した "Television" の画像例

した画像1,000万枚を学習サンプルとして、9層のニューラルネットの教師なし学習（基本的に3.2節で説明したものと同様のもの）を実行すると、「おばあさん細胞」（＝特定の物体（画像）に選択的に反応するニューロン）の再現に成功したというLeらの研究がある<sup>13)</sup>。

一般物体認識以外の画像認識のタスクでも、畳込みネットはよい成果を挙げている。Ciresanらの畳込みネットは、数多くの画像認識のコンテスト（文字認識、交通標識の認識や医用画像応用など）でトップの成績を収めている<sup>14) 15)</sup>。認識だけでなく、物体の「形」を学習・生成するシェイプボルツマンマシン<sup>16)</sup>や、画像のノイズ除去や欠損復元への適用例<sup>17)</sup>もある。これらは、自然画像や物体の形といった複雑な構造を持つ高次元データ（それ自身が極めて高い自由度をもっているものの、データの空間では偏って分布するようなもの）を、ディープネットがそれなりに正確に表現できることによると考えられる。

このように応用範囲が拡大しつつある一方で、ディープラーニングには問題が山積している。従来方法では望めないほど高い性能を示す反面、なぜそのような高い性能が得られるのか、きちんとした答えは今のところない。付随して、チューニングすべきパラメータが多数あり、そのチューニングいかんで性能が大きく左右されるなど、使い勝手はお世辞にも良いとは言えない。また、特徴そのものの学習を可能にする対価として、従来法よりも（その柔軟性・自由度に見合った）大量の学習サンプルを必要とする傾向がある。今まさにこれらの問題の解決を目指し、盛んに研究が行われているところである。（2014年3月2日受付）

## 【文 献】

- 1) A. Krizhevsky, I. Sutskever and G. E. Hinton: "ImageNet Classification with Deep Convolutional Neural Networks", In Proc., NIPS (2012)
- 2) V. Nair and G.E. Hinton: "Rectified linear units improve restricted boltzmann machines", InProc., ICML (2010)
- 3) Q.V. Le, A. Coates, B. Prochnow, and A.Y. Ng: "On Optimization Methods for Deep Learning", In Proc., ICML (2011)

- 4) G. Hinton, S. Osindero and Y.-W. Teh: "A fast learning algorithm for deep belief nets", Neural Computation, 18, pp.1527-1544 (2006)
- 5) G.E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever and R. Salakhutdinov: "Improving neural networks by preventing co-adaptation of feature detectors", CoRR, abs/1207.0580 (2012)
- 6) F. Seide, G. Li, X. Chen and D. Yu: "Feature engineering in context-dependent deep neural networks for conversational speech transcription", InProc., IEEE ASRU (2011)
- 7) K. Fukushima and S. Miyake: "Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position", Pattern Recognition, 15, pp.455-469 (1982)
- 8) D.H. Hubel and T.N. Wiesel: "Receptive fields and functional architecture of monkey striate cortex", The Journal of physiology, 195, 1, pp.215-43 (Mar. 1968)
- 9) Y. Lecun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard and L.D. Jackel: "Backpropagation applied to handwritten zip code recognition", Neural Computation, 1, 4, pp.541-551 (1989)
- 10) Y.L. Cun, L. Bottou, Y. Bengio and P. Haffner: "Gradient-based learning applied to document recognition", In Proc., IEEE (1998)
- 11) N. Kruger, P. Janssen, S. Kalkan, M. Lappe, A. Leonardi, J. Piater, A.J. Rodriguez-Sanchez and L. Wiskott: "Deep Hierarchies in the Primate Visual Cortex: What Can We Learn for Computer Vision?", IEEE Trans. PAMI, 35, 8, pp.1847-1871 (2013)
- 12) M.D. Zeiler and R. Fergus: "Stochastic pooling for regularization of deep convolutional neural networks", In Proc., ICLR (2013)
- 13) Q.V. Le, M.A. Ranzato, R. Monga, M. Devin, K. Chen, G.S. Corrado, J. Dean and A.Y. Ng: "Building high-level features using large scale unsupervised learning", In Proc., ICML (2012)
- 14) D.C. Ciresan, U. Meier, J. Masci, L.M. Gambardella and J. Schmidhuber: "Flexible, high performance convolutional neural networks for image classification", In Proc., IJCAI (2011)
- 15) D.C. Ciresan, U. Meier and J. Schmidhuber: "Multicolumn deep neural networks for image classification", In Proc., CVPR (2012)
- 16) S.M.A. Eslami, N. Heess and J. Winn: "The Shape Boltzmann Machine: A Strong Model of Object Shape", In Proc., CVPR (2012)
- 17) J. Xie, L. Xu and E. Chen: "Image Denoising and Inpainting with Deep Neural Networks", In Proc., NIPS (2012)



おかたに たかゆき  
**岡谷 貴之** 1999年、東京大学大学院工学系研究科博士課程修了（計数工学）。同年、東北大学大学院情報科学研究科助手、その後、同大学講師、助教授（准教授へ名称変更）を経て、2013年、同大学教授となり、現在に至る。コンピュータビジョンの研究に従事。数理統計学および数値最適化を理論的基礎に、その多視点幾何から物体認識までの幅広い応用に関心を持つ。