

## ディープラーニングと著作物

会員 山田 尚史



## 要 約

ディープラーニング技術の普及と発展につれて、コンピュータを利用して著作物を解析し、著作物に類似するものを作成する研究や、ビジネスへの応用が進められている。本稿では、具体的な研究の紹介とともに、特にディープラーニングの中でも成果が目立っている CNN を中心として、その動作原理や歴史的背景について説明する。そして、コンピュータによって生成されたコンテンツが著作物として保護され得るかという点と、学習済みモデルの保護における論点を述べる。

## 目次

1. 機械学習とディープラーニング
  - (1) 機械学習とは何か
  - (2) 機械学習と人工知能の歴史
  - (3) ディープラーニングの登場
2. ディープラーニングの原理と特徴
  - (1) ニューラルネットワークの仕組み
  - (2) 代表的な二つのモデル：CNN と RNN
3. コンテンツを制作するディープラーニング
  - (1) 事例：DCGAN (Deep Convolutional Generative Adversarial Network)
  - (2) 事例：「ゴッホ風」画像作成フィルタ
  - (3) 「人工知能が作成した」という隠れ蓑
  - (4) アルゴリズムによる制作物は著作物足り得るか
4. 著作物を利用した学習と著作権
  - (1) 著作物の情報解析における原則
  - (2) 学習済みモデルは保護され得るか
  - (3) ディープラーニング特有の論点
5. まとめ

## (1) 機械学習とは何か

機械学習とは、文字通り、人間の持つ学習能力と同様の機能を、機械すなわちコンピュータ上で実現するための技術の総称である。そのアプローチとしては、過去のデータから統計的なアプローチを試みる数理モデルや、人間の脳を模倣する方法といった、様々な計算モデルが提案されており、正しい判断を下す精度を向上させるため、日夜研究が進められている。特に、知能を人の手で実現する「人工知能」の分野においては、機械学習が主たる研究領域となっている。

今日、機械学習は、アカデミックな場のみならず、ビジネスの場でも広く利用され、従来人間が行っていたタスクを代替したり、一部においてはより高度に解決したり、といった目的で利用されている。特に、ビッグデータという言葉で表されるとおり、データの転送速度や保存容量が文字通り桁違いに進歩したことや、CPU、GPU といったハードウェアの性能向上に加えて、高度なアルゴリズムの発見も相まって、機械学習の精度や成果も向上の一途を辿っている。今では、バナー広告のコンバージョンの予測といった、結果が比較的評価しやすい分野のみならず、自動運転といった、複雑かつ高度なタスクを人間に代わってこなす機械学習モデルについても、実用化に向けた社会的な実験が始まっている。

## (2) 機械学習と人工知能の歴史

ディープラーニングの詳細を紹介する前に、機械学習が今までにどのような歴史を辿ってきたかを簡単に

## 1. 機械学習とディープラーニング

昨今、ディープラーニング（深層学習）と呼ばれる技術が機械学習分野でブレイクスルーを起こし、学会誌のみならず、一般の新聞や雑誌をも賑わせている。また、内閣府の知的財産戦略本部を初めとして、多くの組織や機関が AI や AI 制作物の扱いに関する議論を行っている。本稿では、このディープラーニングと、著作物の関連について、研究事例を踏まえつつ、議論を深めたい。ここでは議論に先立ち、前提として、機械学習とは何か、そしてディープラーニングとは何かについて、簡単に説明する。

紹介したい。今でこそ、スポットライトが当たり大きな期待をされている機械学習技術であるが、このような事態は初めてではなく、歴史上、加熱した期待と、その反動の失望という二回の「冬の時代」を経験している。

機械学習は人工知能研究の一分野として発展してきたため、その歴史は人工知能の歴史と切っても切れない関係がある。人工知能の原型は1943年、McCullochとPittsが提唱したものであり<sup>1)</sup>、これは後述するニューラルネットワークの原型である。なお、後述するが、ニューラルネットワークはディープラーニングの根幹を成す思想であると言ってよい。その後、1947年にはTuringが、ロンドン数学学会の講義にて、現在の人工知能の概念を提唱している。なお、Turingは1950年、知能の定義として、かの有名なチューリングテストを提唱している。

その後の1950年後半には、FORTRANやLISPといった、今でも使われている高級なプログラミング言語が開発され、研究が加速するものの、1960年代末には、McCarthyとHayesにより人工知能研究における最大の難問の一つである「フレーム問題」が提唱された<sup>2)</sup>。

その後も人工知能の研究は進められたものの、商業的な成果に結びつかず、1974年以降、他学会からの批判と資金提供の削減をきっかけとして、業界全体が縮小を続ける「AIの冬」がやってくる。

1980年以降、エキスパートシステムと呼ばれるシステムの出現により、機械学習研究は一時的に活気を取り戻した。エキスパートシステムは、特定のタスクを解くための知識ベースから推論を行うことで、専門家の判断を擬似的に再現することを目的としたものである。しかしながら、商業的な成功も結果的には一時的なブームに過ぎず、1987年には第二の冬の時代が到来することになる。

1993年以降、計算機自体の性能向上と、解決する分野を限定することで再び徐々に活気を取り戻し、現在も種々の分野で発展を続けている。その後2012年頃のディープラーニングの発見を契機に、ハードウェアの性能向上を初めとする技術インフラの水準の向上も相まって研究が活発化し、現在に至る。

以上が、人工知能研究の歴史の概略である。このような歴史を鑑みるに、機械学習に携わる者は、「今度こそは」と思う一方で、いつか第三の冬が来るのではな

いかと危惧している部分もあるように感じる。人工知能はフィクションの世界でも数多く取り上げられるテーマであり、知名度と理解度の乖離が大きい分野であることも一因ではないだろうか。私見ながら、人工知能という言葉の乱発や、ディープラーニングが魔法であるかのように過剰に期待を煽るのではなく、技術的な成果を正確な言葉で適切に発信していくことが大切であるように思う。

### (3) ディープラーニングの登場

先にも述べたとおり、昨今機械学習が注目されているきっかけは、やはりディープラーニングの登場である。そこで、ディープラーニングとは一体何なのか、どのように登場したのか、ということについて説明する。

ディープラーニングは、層の深いニューラルネットワークを利用した機械学習モデルの学習や活用を指す言葉として使われている。ニューラルネットワークは、ニューロンをネットワーク状に結合したもので、人間の脳の神経回路を模した計算モデルである。

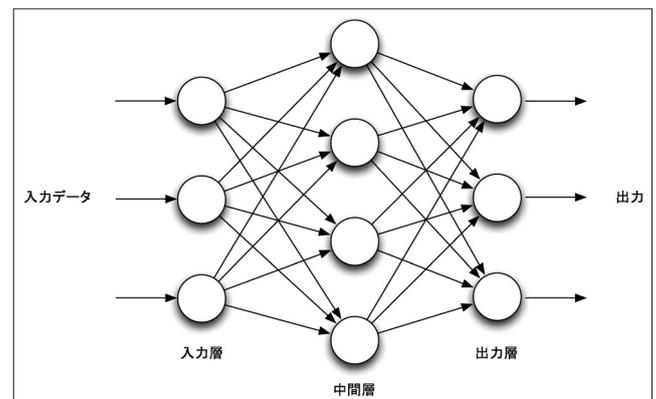


図1：標準的なフィードフォワードニューラルネットワーク

機械学習は、事前に正解データを与える「教師あり学習」と、正解データではなくデータのみから学習を行う「教師なし学習」の二つに大別できる。このうちニューラルネットワークは、教師あり学習の一手法である。ニューラルネットワークは、入力層、中間層、そして出力層によって構成される。図1は、中間層が一層の代表的なニューラルネットワークを図示している。ディープラーニングを理解する上で、ニューラルネットワークの知識は不可欠であるため、ニューラルネットワークの学習の仕組みについては後の章にて詳説する。

中間層の数が少ないモデルでは、表現できる数式に

限界があった。しかしニューラルネットワークの学習に使われる誤差逆伝播法 (Backpropagation) は、誤差が調整されて前の層に伝わるという特徴から、中間層の数が多く、すなわち層の深いモデルにおいては、収束するまでの計算量が膨大なものとなっていた。

そこで提案されたのが、オートエンコーダ (Auto-Encoder) と呼ばれる手法である。オートエンコーダは、簡単に言うと、前の層からの入力を、そのまま出力するように一層ずつ隠れ層を学習する手法である。ここで、オートエンコーダは、一層ずつの学習が行えるため、ネットワークの深さに線形にしか影響されず、従来と比して非常に層の多いニューラルネットワークの学習が可能になった。

ニューラルネットワークは、入力データの性質から様々な改良がされており、今日では、コンボリユーション (畳み込み) ニューラルネットワーク、すなわち CNN や、リカレント (再帰型) ニューラルネットワーク、すなわち RNN が、代表的なディープラーニング技術として知られている。

なお、蛇足ながら、事前に述べたハードウェア性能の向上により、今では層の深い計算モデルにおいても、誤差逆伝播法が広く使われている。オートエンコーダも学習の初期値として使われることはあるが、登場時のように、オートエンコーダがディープラーニングの代名詞であるというような文脈で使われることは少なくなっている。しかしながら、機械学習ブームの火付け役となり、その後の研究の加速をもたらしたオートエンコーダの発見・提唱は、機械学習の歴史においても非常に重要なものであると言えるだろう。

## 2. ディープラーニングの原理と特徴

本稿の論旨は著作物とディープラーニングの関連性であるが、その議論に先立ち、ディープラーニングの技術的な原理と、その特徴を述べる。ディープラーニングは多層のニューラルネットワークを指す言葉であるので、まずはニューラルネットワークの特徴と学習の仕組みについて詳説する。

### (1) ニューラルネットワークの仕組み

代表的なニューラルネットワークは、図1に示したとおりである。ニューラルネットワークは、入力層、中間層、出力層によって構成される。ここで、図中の

丸をノード、ノード同士を結ぶ線をエッジと呼ぶ。

エッジは、それぞれ独立に重みを持っている。エッジの重みは、エッジの始点となっているノードに入力された数値に掛け合わされ、終点となっているノードに受け渡される。一つのノードは、複数のエッジの終点となっているから、それらの値の和が最終的なノードへの入力として受け渡されることになる。このようにして、入力層に入力された数値は、次々ノードの重みが掛け合わされて次の層への入力となり、最終的に出力層のノードそれぞれへの入力が計算されることになる。出力層への入力は、ニューラルネットワーク全体の出力となる。

例えば、一般的なクラス分類問題を考えると、出力層のそれぞれが、クラスの一つ一つに対応した結果と考えて、出力層の数字が最も大きいクラスを答えるモデルとみなすことができる。このとき、学習という行為は、入力に対して正しい出力を得られるように、それぞれのエッジの重みを最適化する行為であると言える。

次に、学習の概要を説明する。ニューラルネットワークの学習には、誤差逆伝播法という手法が使われている。

ニューラルネットワークの学習を始めるにあたって、まず入力データと、その入力データを入力したときの理想的な出力を多数準備する。この理想的な入出力の組み合わせを教師データと呼ぶ。例えば、ある入力するときクラス A が求められるべきであるならば、出力はクラス A に対応したノードの出力が“1”で、その他のノードの出力が全て“0”となっているようなデータである。

誤差逆伝播法においては、はじめに学習対象のニューラルネットワークに対して、入力データを入力し、そのときの出力を求める。そして、それぞれの出力ノードについて、出力と理想値の差分を判定し、その差分を引き起こした原因として、一層前のエッジのうち最も重みが大きいエッジを判定し、そのエッジの重みを、理想値が出るような方向に調整する。

この調整が完了すると、その一層前についても、同様の調整を行う。そうすることで、重みの調整が、出力層から入力層に向かって逆伝播していくことになる。この調整を大量の教師データに対して繰り返すことで、各エッジの重みは理想値に近い値を出すように調整され続け、最終的に目的に適したニューラルネッ

トワークが得られる。

なお、教師データに含まれる入力と全く同じ入力について正しい出力を返すことは、機械学習の目的ではない。機械学習モデルに求められることは、未知の入力について正しい出力を得ることである。このような、未知の広範な入力について正しい出力を得ることができる性能のことを「汎化性能」と呼ぶ。一方で、教師データの入力に偏って学習され、汎化性能を發揮できない状況を「過学習」と呼ぶ。

## (2) 代表的な二つのモデル：CNN と RNN

今日、ディープラーニング分野においては、CNN すなわちコンボリュショナルニューラルネットワークが特に注目されている。CNN の基本原理は、図1に示すような一般のニューラルネットワーク（CNN、RNN と区別して、フィードフォワードニューラルネットワークと呼ばれる）と同じだが、主としてCNN は入力が多次元であることと、「畳み込み」と「プーリング」という処理をする点で異なっている。

「畳み込み」とは、重み付けされたフィルタをラスカスキャンし、繰り返し畳み込み計算を行うことで次の層への出力を得ることであり、全結合ニューラルネットワークにおける、入力に対してエッジの重みをかけることに相当する。CNN ではその後、得られた出力に対してプーリングを行う。プーリングとは、小領域から値を出力して、新たな出力に変換する処理であり、例えば、 $5 \times 5$  の領域から1つの値を得るような処理である。プーリングの一例として、特定の領域の最大値を取る Max プーリングや、合計値を取る Sum プーリングがある。

こうした特徴から、CNN は画像認識の領域で特に驚異的な性能を發揮する。二次元の入力を考えた時、従来のフィードフォワードニューラルネットワークでは、二次元の画素を一次元に並べ直すようなことになる。そうすると、画像内の物体の位置が少しずれたときでも、一次元の入力は大きくずれてしまう。また、画素を直接使うのではなく、画像を何らかの特徴量に変換して扱う試みもなされているが、この場合でも、解こうとしている問題と相関がある特徴量を技術者が適切に選ばなければ高い精度を發揮できないという問題がある。

一方で、CNN においては、画素を直接入力とするので、特徴量の抽出を事前に行う必要がない。また、

Max プーリングで考えると、その最大値は  $5 \times 5$  の領域のどこにあっても同じ結果が得られるので、二次元画像内の要素が平行移動をしても似たような入力を得られ、すなわち平行移動に強い汎化性能が期待できる。

また、リカレントニューラルネットワークすなわち RNN は、入力データ間に時系列的な関係を設定でき、前の要素の学習の結果の一部を、次ステップの学習に用いるニューラルネットワークである。ここでは技術的な詳細は割愛するが、RNN は自然言語処理や、音声の解析分野で高い成果を上げている。また、前述の CNN と組み合わせることで、動画の解析に用いられている。

## 3. コンテンツを制作するディープラーニング

以上、ディープラーニングとその技術的思想の根幹であるニューラルネットワークについて、基本的な構成と動作原理を説明した。ここで、本稿の論旨の一つである、ディープラーニングがコンテンツを制作することと、それにまつわる論点について、最新の理論や研究を紹介しつつ議論したい。

### (1) 事例：DCGAN (Deep Convolutional Generative Adversarial Network)

機械学習の基本に立ち返ると、学習の目的は、汎化性能の獲得であって、それはある種の抽象化ということができる。すなわち、多量の教師データから、汎用的に適用できるある種のルールを見出すことは、実質的にデータを抽象的に捉えて扱う能力の習得に他ならない。

ここで、学習によってデータの抽象化ができるのであれば、反対に機械学習モデルに具体化をする能力を獲得させれば、コンテンツを無尽蔵に出力できる。このような発想は以前からあったが、一定の研究成果を出し有名になったのが DCGAN (Deep Convolutional Generative Adversarial Network) である。

DCGAN は、CNN と、GAN すなわち Generative Adversarial Network<sup>3)</sup> というデータ生成の手法を組み合わせたものである。簡単に説明すると、GAN は、Generator と Discriminator という二つの部分から構成される。Generator は、教師データによく似せた、教師データではない画像を生成する部分で、Discriminator は、入力されたデータが、教師データか

生成された偽の画像かを判別する部分である。この二つのネットワークを学習し、精度を上げていくことで、結果的に教師データによく似た、教師データそのものではない画像を生成することができる。DCGANは、GANをCNNにすることで画像の生成を試みたものである。図2は、DCGANのモデルの一例を示したものである（本稿未記載の参考文献4より引用<sup>4)</sup>）。

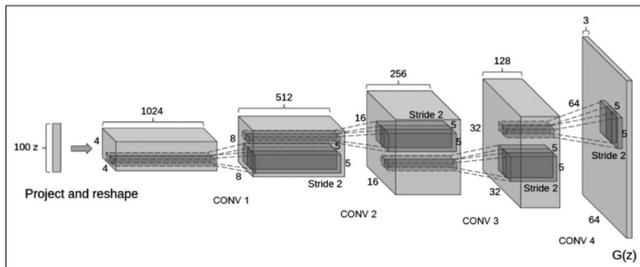


図2：DCGANのネットワーク構造の一例。参考文献4より

DCGANを用いた実例として、ベッドルームの画像で学習した学習器から、新たなベッドルームの画像を生成することや、顔画像のうち、「笑顔」や「サングラス」といった属性を抽出し、他の人の顔画像に適用することに成功している。



図3：DCGANで生成したベッドルームの画像。同じく参考文献4より

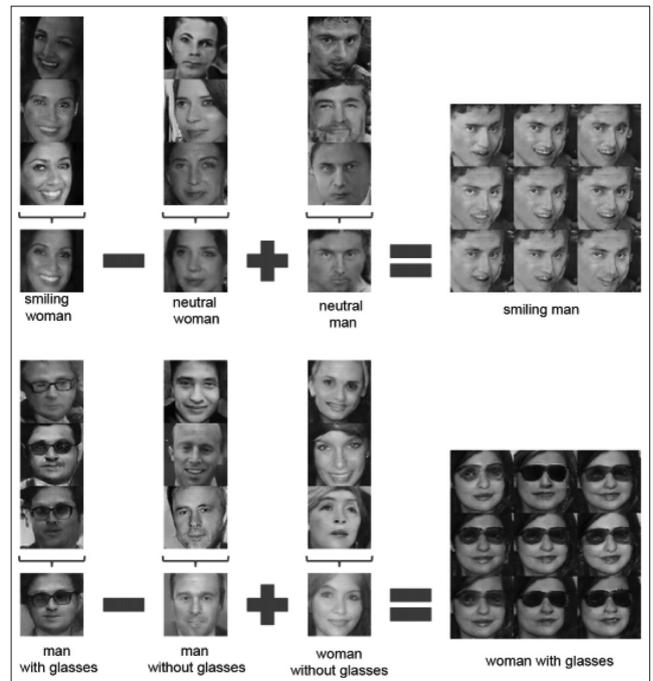


図4：DCGANで生成した属性付きの顔画像。同じく参考文献4より

これらの画像は、現実の画像ではなく、紛れもなく機械によって生成された画像である。しかしながら、画像のみを見たときに、必ずしも機械学習で作られたものとは断定できないような質の高い成果も含まれている。

同様の研究として、画像から画像を生成するpix2pixという手法が研究されている<sup>5)</sup>。

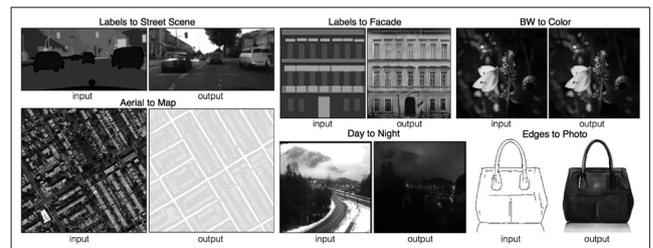


図5：pix2pixによる変換前(左)と後の画像(右)。参考文献5より

白黒では分かりにくく恐縮ながら、pix2pixでは、GANを用いて、ある画像を写真のようにしたり、逆に写真を平滑化された地図にしたり、色を付けたり、といった変換を行っている。元論文では、評価関数を変えることで、変換後の画像に様々なバリエーションを持たせられることが示されている。

## (2) 実例：「ゴッホ風」スタイルの転写

上記で紹介したGANの他にも、CNNを使って、画像に他の画像の「スタイル」を付加する技術が提唱されている<sup>6)</sup>。

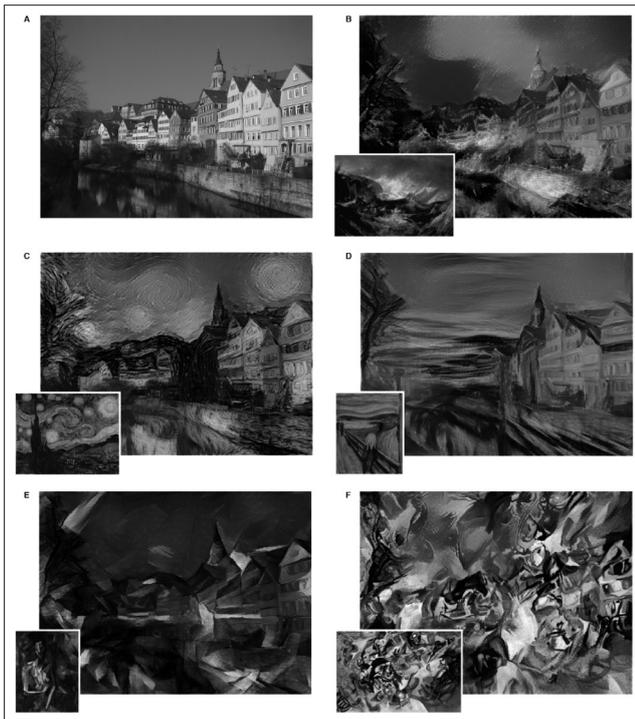


図6：様々な名画のスタイルが適用された写真。参考文献6より

こちらでも白黒画像だと分かりにくいですが、図では湖畔に佇む家々の写真（左上の画像A）に対して、有名な絵の画風を適用している。例えば、Cはゴッホの『星月夜』を、Dはムンクの『叫び』を適用したものである。

ここでは、元の画像をコンテンツ画像、画風の抽出元となった画像をスタイル画像と呼ぶことにする。生成された各画像は、コンテンツ画像と同様の意味を保持しつつ、その画風としては明らかに異なっている。適用後の画像は、元のスタイル画像とよく似ているものの、スタイル画像の一部分を切り抜いたり、拡大縮小したりして複製している部分はない。元論文によれば、スタイル画像からは、画像全体で使われている色や、線の太さ、どの色とどの色が隣り合って描かれやすいかといった情報が抽出され、元の画像が同様の特徴を満たすように元の画像に値の変化を加えている。

### (3) 「人工知能が作成した」という隠れ蓑

先のスタイルの転写においては、あくまでスタイル画像からはスタイルを示す情報のみが抽出されているに過ぎず、画像そのものが成果物たる出力画像に複製されている訳ではない。一方で、画像の一部をそのまま複製して出力するような手法が考案されたときに、「人工知能が作成した」という一言が免罪符になってしまわないだろうか。

仮に、スタイル画像自体が著作権法で保護されているとして、このような複製の事実が明らかである場合、それは著作権法第四十七条の七における『必要と認められる限度』を越えていると解釈されるだろう。著作権法第四十七条の七においては、情報解析における著作権の例外規定が示されているが、ここでいう情報解析はあくまで大量の情報から統計的な解析を行う場合であって、特定の画像の一部又は全部をそのまま複製することは、統計的な解析の範囲を越えていると言える。

では、この複製を行う際に、色調補正やガウシアンフィルタの適用、リサンプリングといった画像処理を行い、元画像と全く同一ではないが似た画像を作成し、それを複製している場合はどうだろうか。この場合も、元画像の特徴を留める限りは、多くのケースで統計的な解析の範囲を越えていると考えるのが自然と考える。

しかしながら、このような画像処理を加えた場合、成果物だけを見た限りでは、上記で説明したDCGANの成果物と見分けが付かないことが考えられる。そして、ソースコードと学習データセットが公開されるのでもない限りは、第三者から見て、どのような過程でその画像が生成されたかを明らかにすることは非常に困難である。

もちろん、DCGANの成果物においても同様に、元画像の著作権に関する議論の余地はある。一方で、DCGANにおいて適切に情報解析が行われている場合、画像処理を行った場合とは議論の前提が大きく異なるだろう。

このように、「人工知能が作成した」という言葉は、著作権法の侵害の隠れ蓑になる恐れがあり、プログラムの作成者側も、可能な範囲で情報は開示すべきであると考えられる。

### (4) アルゴリズムによる制作物は著作物足り得るか

以上、視覚的に伝えやすいため、画像の生成に着目して例示を行ったが、作曲や、自然言語の生成においても、同様にコンテンツ生成の研究が進められている。例えば、文字を人工的に読み上げるTTS（テキストトウスピーチ）分野では、DeepMind社によるWavenetが驚異的な成果を上げている<sup>7)</sup>。<sup>(1)</sup>

このように、機械学習を用いて制作されたコンテン

ツは、既に人間の作ったものと比しても遜色ないものもあり、今後さらに進歩していくものと考えられる。この時、当然に生じる疑問として、これらの成果物に著作権は発生するのであろうか。

昭和48年の著作権審議会報告書によれば、『コンピュータ創作物の著作者については、コンピュータが人間の創作行為を完全に代替するのではなく、人が思想感情を表現する「道具」として使用したと認められることが通常であり、人による創作物として著作物性が認められる』とされている。

これを鑑みれば、例えば、機械学習による学習結果を利用して、あくまで人が創作行為のみを行った場合には著作物性が認められ、反面、例えば自動稼働によって自動的にコンテンツを作り続けるような場合には、現行制度上、その制作物は権利の対象とは考えられていない。

平成28年1月に、内閣府の知的財産戦略推進事務局が公開した『AIによって生み出される創作物の取扱い（討議用）』の資料中において、同様の問題提起がなされている<sup>8)</sup>。そこでは、AI創作物の保護の要否につき、保護した場合には「権利のある創作物が爆発的に増える」との懸念を、保護しない場合には、「完全な自由利用が可能となるため、価値のあるAI創作物が生成されても保護されず、フリーライドを許してしまう」との懸念を示している。

一方で、先程の「隠れ蓑」のような議論に通じるが、制作の過程で機械学習が使われているか否かを外部から判断することは、非常に困難であるし、そのアルゴリズムがどれくらい高度かを判断することも同様に難しい。そのため、一概に「人工知能（機械学習）によって作られたものに対する保護」を規定するにしても、その線引からして非常に困難であると言わざるを得ない。

また、仮に権利が発生する場合として、権利の主体が誰かを定めることも非常に困難である。『AIによって生み出される創作物の取扱い（討議用）』においては、権利の主体の候補として、①プログラム開発者、②人工知能の学習用データの提供者、③人工知能に対し創作の意図をもって指示した者、の三種類に分けて、議論を進めている。

これらは一律に、また形式的に決定するのは難しく、個々の事例において総合的に判断されるのが適当だと考えられる。しかしながら、機械学習は、コン

ピュータで動くものであるから、24時間365日、休みなく高速に動き続けることができる。そのため、何らかの検討過程を設けるにしても、現実問題として、生産され続けるコンテンツに対して人間による審査を行うことは非常に困難である。

自律的に作られ続けるコンテンツには、同様の処理速度で自動的に著作物性を判断するシステムが必要であろう。そのためにはいっそ、著作物性の有無を判断する機械学習モデルを組むといった、大胆な施策が必要であるように思う。機械学習によって、全ての入力を100%の精度で判断することは難しくても、人間でないと判断が難しいものに人間のリソースを集中させることは可能であるから、今後は知的労働を含む様々な場面において、機械学習の活用は増えていくと思われる。

#### 4. 著作物を利用した学習と著作権

先の議論においては、ディープラーニングによる成果物が著作物として保護されるかに着目し、著作物を機械学習に利用する場合の、教師データの著作権については詳しく述べなかった。CNNの発展につれて、画像が教師データとして利用されるようになったが、これらには当然に、著作権で保護されているものもある。これらについて、対応する著作権法の条文も引きつつ、その例外や学習済みモデルの権利について、ディープラーニング特有の論点についても加えつつ説明する。

##### (1) 著作物の情報解析における原則

先に述べたような、著作物を教師データとした機械学習について、我が国の著作権法では、情報解析のための複製等（第47条の7）や、電子計算機における著作物の利用に伴う複製（第47条の8）によって、著作権者の著作権を侵害せず、著作物を情報解析に自由に利用してよい旨が定められている。第47条の7によれば、コンピュータ等を用いて情報解析を行うために、必要と認められる限度において複製・翻案を行うことができる旨が定められている。ここで、情報解析とは、『大量の情報から言語、音、映像等を抽出し、比較、分類等の統計的な解析を行うことをいう』<sup>9)</sup>。

ただし、第47条の7には例外規定があり、情報解析を行う者の用に供するために作成されたデータベースの著作物については、この限りでない、とされている。

これは例えば、情報解析のためにアレンジされた一連の教師データについては、そのアレンジメントについて著作権が発生すべきというものである。

## (2) 学習済みモデルは保護され得るか

モデルに学習を行わせるのは、プログラムのコードである。これは、創作性の要件を満たせば、著作物として保護される（著作権法第2条第1項第1号、第10条第1項第9号）。また、学習に用いるデータに関する保護については、先に述べたとおりである。それでは、学習され、出力された学習済みモデルは保護を受けられるのだろうか。

同様の議論は、平成28年5月に経済産業省の商務情報政策局が公開した「IoT時代に対応した自律分散協調に必要なアーキテクチャーと技術要素について」においても提示されている<sup>10)</sup>。現行法上、学習済みの機械学習モデルが著作物として認められるかどうかは、機械学習モデルの学習が創作行為に該当するか否かによるものと考えられる。

技術者の視点から、保護に賛成の立場に立って見ると、学習済みモデルには、下記のような創意工夫の余地があるように思える。まず、初めに行うのはモデルの構造の設計である。これは、モデルがフィードフォワードなニューラルネットワークなのか、CNNなのか、RNNなのか、それらの組み合わせなのか。また、中間層は何層あるのか、入力層と出力層は何次元になっているのか、といったことを決める。次に、各層における活性化関数や損失関数を決定する。そうして学習を開始するわけだが、データも教師データをそのまま入力するのではなく、移動、回転に対する汎化性能を得るためにオーグメンテーションをする。学習において汎化性能を下げ得るデータに関しては、何らかの前処理を行い取り除く。学習の初期値を調整する。ドロップアウトを試みる。そうして実際に学習を行ってみて、より精度を上げるために、ネットワーク構造や関数を変えて、再度の学習を行う。実際に技術者が行うこれらの行為は、ある種の創作行為と捉えても差し支えない部分があるように思う。また、上記の工夫に関して、一部はプログラムのコード中に現れるものの、一部は設定ファイルに箇条書きになっていたり、特にデータの前処理に関する部分については、ソースコード中には現れなかったりするものもある。

一方で、保護に反対の立場に立って見ると、ネット

ワーク構造はロジカルに考えればそこまでバリエーションがあるものではなく、上記の学習の手法も、採用の適否を決めているだけで、創作行為とまでは言えず、仮に誰か他の技術者の学習の手法を知ってしまった場合には、著作権により同じようなモデルを学習させることを阻害するのではないか、という懸念も考えられる。

ところで、実際に機械学習を活用したビジネスの現場を見ていると、機械学習モデルが何らかの権利によって保護されているといまいと、第三者による複製を防ぐために、セキュリティ上の手段を講じたり、契約で利用制限を設けたり、といったことは今でも通常に行われており、今後も続くように思う。一方で、JavaScriptを用いたクライアントサイドでの機械学習の活用等も進められている<sup>11)</sup>ため、権利保護が進むことで、機械学習分野における企業や技術者の投資が一層進むことは間違いないので、適切に権利が設定されることが望まれている。

## (3) ディープラーニング特有の論点

最後に、補足ながら、ディープラーニングに特有の「転移学習」と「蒸留」について説明する。これは、上記の『IoT時代に対応した自律分散協調に必要なアーキテクチャーと技術要素について』の中で、モデルの2次利用、3次利用として説明されていた部分である。

転移学習とは、ある特定のタスクを解くために学習した学習済みモデルの知識を、別ドメインのタスクを解くために活用することを指し、より具体的には、学習済みモデルの一部又は全部を、別のタスクを解くために活用することである。例えば、一般物体認識のタスクのために学習されたCNNの中間層は、画像の特徴抽出器として働くものと見なすことができる。元のCNNでは、その後の層は出力層に向けて最適化されているが、中間層で抽出された値だけを、他のタスク、例えば映っている人の姿勢を認識するための学習に活用することができる。この場合、人の姿勢判別のための教師データは、元のCNNモデルに入力され、中間層以降の出力に対するネットワークが新たに学習される。このネットワークは、ニューラルネットワークが新たに設計されることもあれば、SVMのような全く別の原理で動く機械学習モデルが採用されることもある。このように、元々画像に何が映っているかを判別する機械学習モデルは、どこかの抽象化段階で十分に

画像の特徴を抽出できているとみなせるので、人の姿勢判別といった全く別のタスクにおいても、その特徴を活用できると考えられている。このような転移学習は様々な分野で成果を上げている。

蒸留もまた、既存の学習済みモデルを活用する手法の一つである。蒸留は、学習済みのモデルとは全く別のネットワークを設計し、その教師データとして、他の学習済みモデルの入力と出力を利用する。機械学習モデルのネットワークは、大規模になることで汎化性能が上がる場合がある一方で、通常は活用時に電力や処理性能をより大きく必要とする。蒸留によって元の汎化性能が完全に再現される訳ではないものの、小型のネットワークにその性能を移し替えることで、組み込みシステムといった、より限定的な環境でも成果を上げることが可能になる。

以上の手法は、既存の学習済みモデルのいわゆるデータコピーではないので、直ちに複製ということとはできない。ただし、部分的に利用を行ったり、その結果を模したネットワークを再構築したりといった行為にあたるので、そもそも複製にあたるか否か、元の学習済みモデルの権利者（がもし存在すれば）との間でどのような利益の調整が行われるのか、承諾の要否、元のネットワークに課されていた権利上の制約（例えば、第47条の7の例外規定におけるデータベースの著作物の複製権の承諾等）、といった論点がある。これらがどのように重み付けされて考慮されるかについては、文化的な側面、及び経済的な側面から、制度設計が必要であると考えられる。

## 5. まとめ

以上に渡り、ディープラーニングと著作物の関連について、その動作原理や歴史的背景を説明し、特にディープラーニングとして成果が目されているCNNを中心として、仕組みと特徴を説明した。そして、CNNを元にして近年研究されているDCGANを初めとする生成モデルとその成果物を紹介し、著作物として保護され得るかという点における議論と、現状出始めている問題について簡単に触れた。最後に、著作物を利用した学習について、学習済みモデルは保護され得るか、ディープラーニング特有の論点について述べた。

近年、ディープラーニングは非常に注目され、今なお世間の関心は高い。アカデミックな分野において

も、数ヶ月のスパンで革新的な機械学習モデルが提案されているし、ニュースでも多くの企業が自動運転や自動応答に着目し、資金と人材の投入していることを報じている。このような流れの中で、内閣府の知的財産戦略本部は、まさに本稿を執筆している平成28年末にも、AIの作成・保護・利活用の在り方について新たな情報財検討委員会を開いている。

人工知能は、その70年に渡る歴史の中で、二回の冬の時代を乗り越えてきた。それでもなお現在、資本市場は機械学習に期待を寄せ、機械学習もまた、ディープラーニングの活用という形で成果を上げ始めている。そのようなめまぐるしい動きの中で、コンピュータやアルゴリズムの高度化に伴い、機械から人間と変わらないアウトプットが出始め、文化のあり方や著作権に求められている保護のあり方も変わり始めているように感じる。我々弁理士もまた、代理人としての責務を果たすため、旧来のフレームに捉われず、最新の技術動向を理解し、法に則って判断を下していくことが肝要であると考えられる。

## 注

- (1) Wavenet は、Dilated Causal Convolution という特殊な畳み込みの構造を用いてCNNで学習を行っているとのことだが、詳細なネットワーク構造については未だ公開されていない。続報を待ちたい。
- (2) 米国においては Google 社が特許取得済み。広告番号 WO2014105866

## (参考文献)

- 1) W.S. McCulloch, W. Pitts (1943) "A logical calculus of the ideas immanent in nervous activity". In: Bulletin of mathematical biology,
- 2) John McCarthy, Patrick J. Hayes (1969) "Some philosophical problems from the standpoint of artificial intelligence". Machine Intelligence 4: 463-502.
- 3) Goodfellow, Ian et al (2014) "Generative Adversarial Nets". Advances in Neural Information Processing Systems 27
- 4) Alec Radford, Luke Metz, Soumith Chintala (2015) "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks". arXiv: 1511.06434
- 5) Isola, Phillip and Zhu, Jun-Yan and Zhou, Tinghui and Efros, Alexei A (2016) "Image-to-Image Translation with Conditional Adversarial Networks". arXiv: 1611.07004
- 6) Leon A. Gatys, Alexander S. Ecker, Matthias Bethge

- (2015) "A Neural Algorithm of Artistic Style". arXiv: 1508.06576
- 7) Aaron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, Koray Kavukcuoglu (2016) "WaveNet: A Generative Model for Raw Audio". arXiv: 1609.03499
- 8) <http://www.kantei.go.jp/jp/singi/titeki2/tyousakai/kens>
- 9) [http://www.bunka.go.jp/seisaku/chosakuken/seidokaisetsu/gaiyo/chosakubutsu\\_jiyu.html](http://www.bunka.go.jp/seisaku/chosakuken/seidokaisetsu/gaiyo/chosakubutsu_jiyu.html)
- 10) [http://www.meti.go.jp/committee/sankoushin/shojo/johokeizai/bunsan\\_senryaku\\_wg/pdf/003\\_02\\_00.pdf](http://www.meti.go.jp/committee/sankoushin/shojo/johokeizai/bunsan_senryaku_wg/pdf/003_02_00.pdf)
- 11) <https://transcranial.github.io/keras-js/>
- (原稿受領 2016. 12. 7)

日本弁理士会の  
『特許出願等援助制度』をご活用ください  
～優れた発明・考案・意匠の創作を応援します～

**JPAA  
Information**

**特許出願等援助制度とは？**

有用な発明や考案、意匠の創作が、経済的な事情によって世の中に活用されることがなく埋もれてしまうことがないように、日本弁理士会が必要とされる費用の全部又は一部を負担する制度です。

**援助対象者は？**

発明者や教育機関、中小企業等が対象です。

**援助の費用は？**

必要となる、弁理士の報酬、費用及び特許庁の手数料の合計を超えない範囲で負担します。

**援助の条件は？**

日本弁理士会が審査を行い援助の可否を決定します。(※詳細は右の「利用の流れ」)

**利用の流れ**

申請

↓

審査

↓

審査結果の通知

↓

援助が決定したら  
弁理士の選定

↓

契約

↓

援助の開始

特許出願等援助制度の詳細、申請書様式のダウンロードは日本弁理士会のホームページで