

学習済みプログラムのパラメータを物の発明として把握できるのか

弁理士 松下 正

要 約

膨大なデータの収集や管理が可能になったことから、画像認識や予測ができる AI 学習用プログラムの提供が実用的になりつつある。かかるプログラムも自然法則を利用する限り、従来のコンピュータプログラムと同様に特許の保護対象となる。

ここで、AI 学習用プログラムは、通常のプログラムと異なり、本質は、そのパラメータにある。にもかかわらず、かかるパラメータ（データの集合物）の特許法による保護について、特許制度小委員会報告書案では、前記パラメータは法上の「物」に該当するのか疑義があるため、当該パラメータをネット配信等する行為を侵害とできない問題点について指摘がなされているものの保護すべきか否かについては言及がなされていない。

本稿では、前記パラメータを、特定の処理を行うプログラムを構築するための専用部品、すなわち、「プログラムの部品」として、「物」の発明に該当すると解釈できないかについて検討する。

目 次

1. はじめに
 - 1) 現在の AI ブーム
 - 2) AI 適用発明とは
 - 3) 特許制度小委員会報告書について
2. 条文および審査基準の考え方
 - 1) 「物の発明」について
 - 2) 「物の生産方法の発明」について
 - 3) 「プログラム等」
 - 4) プログラムについての保護の変遷
3. AI 適用発明における学習済みプログラムのパラメータとは
 - 1) ディープラーニングと機械学習の関係
 - 2) ディープラーニングとは
 - 3) ディープラーニングの演算におけるパラメータの意義
 - 4) パラメータ以外の部分の重要度
4. 「学習済みプログラムのパラメータのみを学習済みプログラムの部品の発明と解釈できないか？」
 についての検討
 - 1) 学習済みプログラムのパラメータの性質からみた検討
 - 2) 実施行為との関係
5. クレームにおける特定について
 - 1) パラメータは最適値である
 - 2) 範囲指定について
 - 3) 一部のパラメータの特定について
 - 4) パラメータを一部特定した形での権利化のメリット
 - 5) 用途を特定する必要があるのか
 - 6) データ構造クレームについて

6. 最後に

- 1) 社会的要請の増大について
- 2) 自然法則の利用性をもともとある場合のパラメータについて

1. はじめに

1) 現在の AI ブーム

現在は、第三次 AI ブームといわれている。過去 2 回のブームと異なり、注目されているのは、コンピュータの演算処理が向上したこと、膨大なデータの収集や管理が可能になったことから、画像認識や予測など実用的な判断ができるようになったからである。

(総務省白書平成 28 年版「人工知能 (AI) 研究の歴史」⁽¹⁾より引用)

	人工知能の置かれた状況	主な技術等	人工知能に関する出来事
1950年代			チューリングテストの提唱 (1950年)
1960年代	第一次人工知能ブーム (探索と推論)	<ul style="list-style-type: none"> ・探索、推論 ・自然言語処理 ・ニューラルネットワーク ・遺伝的アルゴリズム 	ダートマス会議にて「人工知能」という言葉が登場 (1956年) ニューラルネットワークのパーセプトロン開発 (1958年) 人工対話システムELIZA開発 (1964年)
1970年代	冬の時代	<ul style="list-style-type: none"> ・エキスパートシステム 	初のエキスパートシステムMYCIN開発 (1972年) MYCINの知識表現と推論を一般化したEMYCIN開発 (1979年)
1980年代	第二次人工知能ブーム (知識表現)	<ul style="list-style-type: none"> ・知識ベース ・音声認識 	第五世代コンピュータプロジェクト (1982~92年) 知識記述のサイクプロジェクト開始 (1984年) 誤差逆伝播法の発表 (1986年)
1990年代	冬の時代	<ul style="list-style-type: none"> ・データマイニング ・オントロジー 	
2000年代	第三次人工知能ブーム (機械学習)	<ul style="list-style-type: none"> ・統計的自然言語処理 	ディープラーニングの提唱 (2006年)
2010年代			ディープラーニング技術を画像認識コンテストに適用 (2012年)

この種の学習済みプログラムは当然、特許の対象となる。この AI 学習済みプログラムの処理実行部分については OSS (オープンソースソフトウェア) として提供されており、学習済みプログラムの重要な部分は、後述するように、学習済みプログラムのパラメータそれ自体であるともいわれる。

前記パラメータを特許法の保護対象とできるのかについて以下検討する。

2) AI 適用発明とは

まず、問題となる AI 適用発明について整理する。特許庁は、AI 関連発明を以下の 2 つに分類している。

① AI コア発明

ニューラルネットワーク、ディープラーニング (深層学習)、サポートベクタマシン、強化学習等を含む各種機械学習技術のほか、知識ベースモデルやファジィ論理など、AI の基礎となる数学的又は統計的な情報処理技術に特徴を有する発明

② AI 適用発明

画像処理、音声処理、自然言語処理、機器制御・ロボティクス、診断・検知・予測・最適化システム等の各種技術に、AI の基礎となる数学的または統計的な情報処理技術を適用したことに特徴を有する発明

すなわち、前者は AI 技術そのものに関する発明であり、後者は AI 技術を用いた応用発明である。本稿では、後者について論ずる。

(1) <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h28/html/nc142120.html>

3) 特許制度小委員会報告書について

このような AI 適用発明について、産業構造審議会知的財産分科会の特許制度小委員会報告書案「AI・IoT 技術の時代にふさわしい特許制度の在り方—中間とりまとめ—」では、以下の3点が指摘されている⁽²⁾。

① 特許方法で生産された学習用データの保護について

物の生産方法の発明については、その生産方法で生産した結果物にも効力が及ぶ（特2条3項3号）。しかし、AI 学習用データが同号の規定する「物」か否かについては疑義があるため、前記生産方法について特許を取得しても、その方法で生産された学習用データを、インターネット等を通じて提供するような行為については、これを侵害とすることはできない。海外で同一の方法で生成された学習用データを国内に提供する行為についても、同様である。

② 学習済みプログラム⁽³⁾のパラメータの保護について

AI 学習済みプログラムを利用した推定装置の特許を取得しても、当該装置の生産に用いる学習用データや学習済みパラメータを、インターネット等を通じて提供するような行為については、これを侵害とすることはできない。

③ 特許製品に関連する 3D プリンタ用データ

特許製品の生産に用いる 3D プリンタ用データを生産・提供する行為については、これを侵害とすることはできない。

以下、上記①②に関係する、前記パラメータを、あるプログラムを構築するための専用部品、すなわち、「プログラムの部品」として、「物」の発明として取り扱うことができないかについて検討する⁽⁴⁾。

2. 条文および審査基準の考え方

1) 「物の発明」について

特許法上、「発明」の定義は存在するが、「物の発明」の定義は存在せず、学説（通説）は「物の発明」とは、時間的（経時的）な要素を含まない形での発明の実施形態であると説明されている。

しかしながら、かかる定義は、方法発明との対比であり、平嶋⁽⁵⁾は、「物の発明」とは、「技術的思想が“物”の形で体现されている発明」と説明している。

ここで、伝統的には「物」とは民法 85 条の「有体物」と同義であると解されていた⁽⁶⁾。しかし、後述するように、現行法では、特許法 2 条 3 項 1 号の「実施」の定義として、「物」には「プログラム等」が含まれることが明文化されている。

2) 「物の生産方法の発明」について

平嶋は、「物の生産方法の発明」を「方法の発明」をさらに細分化した発明であり、その方法を用いた結果として一定の成果物が得られるという性質を有する発明と分類している⁽⁷⁾。「結果物」に「プログラム等」が含まれることは当然である。

3) 「プログラム等」

「プログラム」とは、「コンピュータ（電子計算機）に対する指令であって、一の結果を得ることができる

(2) 正確には、「『プログラム等』などとして特許法上の『物』に該当する場合を除き」と指摘されている。

(3) 中間とりまとめでは、「学習済みモデル」の語を使用しているが、本稿では、「学習済みプログラム」を統一して使用している。

(4) 上記③「特許製品に関連する 3D プリンタ用データ」については、特許法 101 条で保護すべきとの見解がある。

(5) 中山信弘＝小泉直樹編著「新・注解特許法 [第2版] 上巻」(青林書院 2017 年) P33

(6) 中山ほか・前掲 (5) P33

(7) 中山ほか・前掲 (5) P35

ように組み合わせられたものをいう」と定義されている⁽⁸⁾。そして、「プログラム等」とは、「その他電子計算機による処理の用に供する情報であって、プログラムに準ずるものをいう」と定義されている。

現行の特許・実用新案審査基準「特許・実用新案審査ハンドブック附属書B第1章」には、「プログラム等」について、以下の説明が存在する。

〔(vii) プログラムに準ずるもの

コンピュータに対する直接の指令ではないためプログラムとは呼べないが、コンピュータの処理を規定するものという点でプログラムに類似する性質を有するものをいう。例えば、データ構造が『プログラムに準ずるもの』に該当することがある。

(viii) データ構造

データ要素間の相互関係で表される、データの有する論理的構造をいう。〕

かかる定義からすると、プログラムに類似する性質を有するものは、特許法2条3項1号の「プログラム等」に含めるという解釈は可能といえよう。では、プログラムに類似する性質を有するとはいかなることかについて検討する。

なお、この問題については、プログラムが、法上の「物」の発明として保護されるようになった意義および経緯を検討すべきである。

4) プログラムについての保護の変遷

平成5年(1993年)の「特許・実用新案審査基準」[第Ⅷ部特定技術分野の審査基準第1章コンピュータ・ソフトウェア関連発明]では、「ソフトウェアによる情報処理が、対象の物理的性質又は技術的性質(構造上の性質を含む)に基づいてなされている場合は、ソフトウェアによる情報処理に自然法則が利用されるとみられるから、その発明は、自然法則を利用したものといえる」と発明として成立することが規定された⁽⁹⁾。しかしながら、「～装置。」という発明の特定は問題ないが、「～プログラム。」という特定は発明ではないとして登録を認めなかった。

その後、海外、特に米国における登録実務との関係で、「～プログラムを記録した記録媒体。」という表記については、平成9年(1997年)4月以降は、認められるようになった。これは、侵害形態としてプログラムを記録したCD-ROMのような形式での取引が増えてきたためである。

その後、平成14年(2002年)4月の特許法改正(平成14年9月施行)により、特許法における「物」に「プログラム等」が含まれることが規定された。これにより、媒体を介さずネットワーク上を伝送されるプログラムについても、直接侵害を追及できるようになった。

なお、かかる規定は、従来は、「物」に含まれていなかったプログラム等を、新たに「物の発明」に含まれるように改正されたわけではなく、確認規定であると説明されている⁽¹⁰⁾。また、工業所有権法(産業財産権法)逐条解説[第21版]では、「平成14年の一部改正では、2条3項1号において、「物(プログラム等を含む。以下同じ。)」と定義することにより、以降の条文において『物』に『プログラム等』が含まれることが明確化された」と説明されている⁽¹¹⁾。

(8) 特許法2条4項

(9) 「コンピュータ・ソフトウェア関連およびビジネス分野等における保護の在り方に関する調査研究報告書」平成22年3月社団法人 日本国際知的財産保護協会 AIPPI・JAPAN のP19～24に詳しい。

(10) 広実郁郎「特許法等の一部改正の概要」NBL No.739 P37～38, 同氏説明SOFTICセミナー「特許法の一部改正と実務のインパクト」(平成14年7月1日開催)

(11) 特許庁「工業所有権法(産業財産権法)逐条解説[第21版]」P15・14～17行

<https://www.jpo.go.jp/system/laws/rule/kaisetu/kogyoshoyu/document/chikujokaisetsu21/tokkyo.pdf>

3. AI 適用発明における学習済みプログラムのパラメータとは

一般にプログラムにおける演算処理では、特定の項目については重み付けされることも多い。この種の重み付けは、当業者にとっては設計変更レベルであることが多いであろう。ディープラーニングによる前記パラメータが、もし、単なる重み付け程度なのであれば、特許の対象とするのは困難といえる。これに対して単なる重み付けを越える何かがある場合、特許の対象とできるかもしれない。

以下、AI 適用発明におけるパラメータの意義について検討する。

1) ディープラーニングと機械学習の関係

まず、本稿で説明する用語「AI」、「機械学習」、「ディープラーニング」について整理しておく。

AI (Artificial Intelligence) とは、いわゆる人工知能であり、コンピュータに知識を与えて、目的となる答えを探し出す手法である。機械学習とは、与えられたデータ群からコンピュータが学習し、それに基づいて予測、判断する学習手法である。ディープラーニング (深層学習) とは、この機械学習の一種であり、後述する CNN のような複数の中間層を用いて、データに含まれる特徴を段階的により深く学習する手法である。3 手法の関係は、以下のような関係となる。

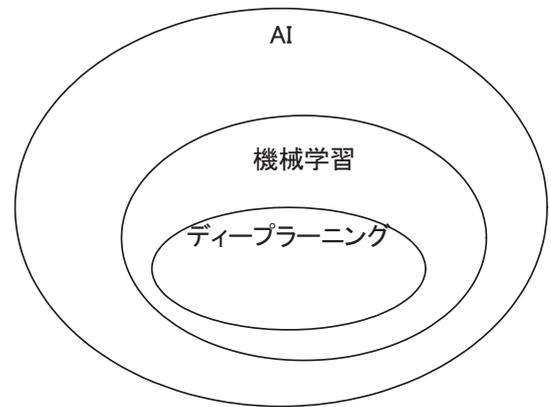
2) ディープラーニングとは

1) 概要

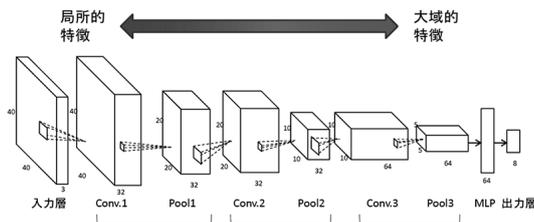
ニューラルネットワークはニューロンの組み合わせであり、今日、期待されているのは、中間層 (隠れ層) を有するディープラーニングである。中間層を有する演算手法の典型例として、画像処理などに用いられる CNN (Convolutional Neural Network)、言語処理などでよく使われる RNN (Recurrent Neural Network) がある⁽¹²⁾。

CNN は、もともとの画像にフィルタリングを行い、複数の特徴を抽出する処理を行う。かかる特徴は、たとえば、画素の並びや色の配置傾向などである。かかるフィルタはランダムに生成され、学習によって取捨選択される。

CNN の説明として、よく下記のような図が用いられる。



典型例：畳込みニューラルネットワーク (CNN)

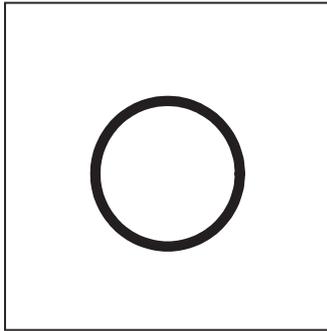


パテント 2020 Vol.73 No.13 (別冊 No.24) P115 より引用

以下、CNN における処理を具体的に説明する。CNN 演算の典型としては、画像処理の場合、隣接する画素との配置関係の特徴抽出する。たとえば、手書きの○か×を判定する場合に、元の画像が 9*9 画素 (ピク

(12) RNN の改良バージョンとして、LSTM (Long Short Term Memory) がある。単純な RNN では、長いデータを処理すると計算量が爆発的に増えてしまうという問題がある。そこで、再帰的に全ての情報を与えるのではなく、一部については無視するという計算手法である。

セル) の白黒画像化し、白を 0, 黒を 1 とすると下記のようなデータとなる。



0	0	0	0	0	0	0	0	0
0	0	0	1	1	1	0	0	0
0	0	1	0	0	0	1	0	0
0	1	0	0	0	0	0	1	0
0	1	0	0	0	0	0	1	0
0	0	1	0	0	0	1	0	0
0	0	0	1	1	1	0	0	0
0	0	0	0	0	0	0	0	0

これに対して、下記のような 3*3 のフィルタ⁽¹³⁾を複数用いて、畳み込み演算を行う。

0	0	1
0	1	0
1	0	0

1	0	0
0	1	0
0	0	1

.....

畳み込み演算とは以下のような処理をいう。

0	0	1
0	1	0
1	0	0

0	0	0	0	0	0	0	0	0
0	0	0	1	1	1	0	0	0
0	0	1	0	0	0	1	0	0
0	1	0	0	0	0	0	1	0
0	1	0	0	0	0	0	1	0
0	1	0	0	0	0	0	1	0
0	0	1	0	0	0	1	0	0
0	0	0	1	1	1	0	0	0
0	0	0	0	0	0	0	0	0

太線で囲んだ 3*3 の領域を見ると、左上が 0*0, その右側が 0*0, さらにその右側が 1*0, 真ん中の左端は、0*0, その右側は 1*0, その右側は 0*0・・・が得られる。これを 3*3 の領域全てに求めて、合計すると、

$$0*0+0*0+1*0+0*0+1*0+0*0+1*0+0*0+0*1=0$$

これを右に、1 画素ずらして演算する⁽¹⁴⁾と、同様に、「0」となる。これを次々繰り返すと、下記のような 7*7 の特徴マップが得られる。

0	0	1	1	1	0	0
0	3	1	1	0	1	0
2	1	0	0	1	0	1
1	1	0	0	0	1	1
1	0	1	0	0	1	1
0	1	0	1	1	3	0
1	0	1	1	2	0	0

(13) フィルタはカーネル (Kernel) とも呼ばれる。ここではイメージをわかりやすくするためにフィルタという。

(14) ずらす量は 1 以上である。これをストライドという。

得られた特徴マップでは、左上の領域と右下の領域に「3」の部分がある。

これは前記フィルタが示す画像とマッチする領域を持つ部分を検出していることとなる。数学的には、フィルタとの内積を計算している。すなわち、内積とは同じ方向を有するベクトルであれば大きく、方向が異なれば小さくなる⁽¹⁵⁾。

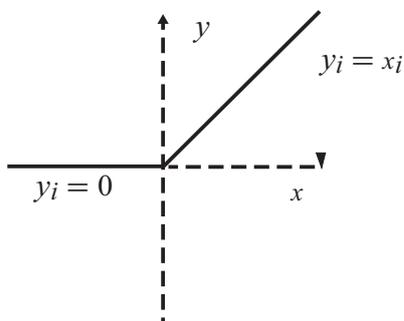
上記特徴マップをプーリング処理（粗め処理）する。プーリング処理の代表が最大値プーリング（Max Pooling）⁽¹⁶⁾であり、以下のように複数の領域の最大値をとる（この場合、4つの領域の内最大値が「3」）。プーリング処理では、領域を重ねず、同じ大きさですらす。すなわち、次の「1, 1, 1, 1」の最大値の「1」となる。これを繰り返す。これにより演算量を減らしつつ、どのような特徴を有するかを抽出することができる。

0	0	1	1	1	0	0	3	1	1	1
0	3	1	1	0	1	0	2	0	1	1
2	1	0	0	1	0	1	1	1	3	3
1	1	0	0	0	1	1	1	1	3	3
1	0	1	0	0	1	1				
0	1	0	1	1	3	0				
1	0	1	1	2	0	0				

つぎに、活性化関数と呼ばれる非線形関数による計算を行う。今日、CNNによく利用されているのは、ReLU と呼ばれる非線形モデルである。

$ReLU(x) = \{0(x < 0), x(x \geq 0)\}$ で表される。

すなわち、0未満は0として出力し、0以上はその値をそのまま出力する。グラフで表すと以下のような非線形関数⁽¹⁷⁾である。



上記の畳み込み、プーリング⁽¹⁸⁾および活性化関数による演算処理が、CNNの1つの隠れ層の演算に該当する。

このような畳み込み演算で用いるフィルタの重みを学習させることで、当該領域にそのような特徴を有する画像か否かを判断できる。

上記畳み込みに使用するフィルタについて、設定しなければならないパラメータは、フィルタの数 (K)、フィルタの大きさ (F)、フィルタの移動幅 (S) である⁽¹⁹⁾。

また、活性化関数としてどれを用いるのか？、何層で構成するのか？などが推論の精度などに影響する。

(15) CNNの説明は、下記がわかりやすい。「【深層学習】畳み込み層の本当の意味、あなたは説明できますか？【ディープラーニングの世界 vol. 5】」 <https://www.youtube.com/watch?v=vU-JfZNBdYU&t=325s>

(16) 平均値プーリング（Average Pooling）という、各領域の平均値をとる演算手法もある。

(17) 正規化線形関数とも呼ばれる。

(18) プーリングは任意である。

(19) 実際には、パディング、すなわち、画像の端の領域を0で埋める処理が指定されるが、説明を簡略化するために、ここでは無視して説明する。

学習させる際には、CNN で用いるフィルタの構成（配置，大きさ，ストライド），層の深さを指定すれば，学習するためのプログラムが試験データのマッチング処理を行い，最終的に，これらをパラメータとして記憶する。

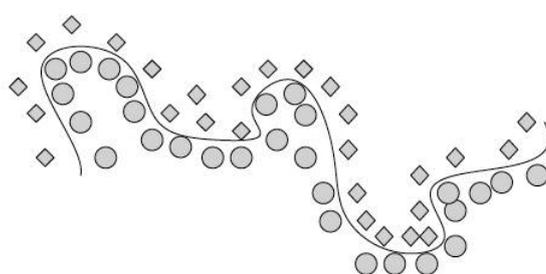
なお，中間層の層数，ニューロン数，活性化関数の種類については，学習済みプログラムが推論するデータの種類，データ量に応じて試行錯誤して設計さえすれば⁽²⁰⁾，上記複雑な関数をコンピュータで扱えるパラメータ数で特定した疑似関数を得ることができる。すなわち，学習済みプログラムについては，結果的にこれらのパラメータが指定されていることとなる。

3) ディープラーニングの演算におけるパラメータの意義

ディープラーニングでは，上記の隠れ層が複数あり，各フィルタの構成および重み付けが学習されて，学習済みプログラムが完成する。

学習済みプログラムにおいては，入力された値から一定の値が出力されるため，一種の関数演算が行われている。ディープラーニングとは，適切な関数を見つけ出すことである。

たとえば，下記のようなデータ分布があり，これを表す関数を見いだすことである。



ここで，適切な関数は，理論的には次元数の制限がなければ，これを定義することは可能である。

しかし，現実問題として上記のような複雑な関数を見いだすのは，困難である。これは与えられる入力データ数が限定されること，さらに計算能力の問題である。

そこで，上記のような畳み込み演算，プーリング，単純な非線形関数を用いることにより，パラメータ数を飛躍的に下げて演算することで現実的な関数を見いだしているのである。

このように，ディープラーニングにおいて得られる関数は，多数次元の関数を疑似的に定義したものである。単純な非線形関数を複数層で合成した関数（以下，合成関数という）とすることにより，複雑な関数の疑似関数を生成することができる⁽²¹⁾。

前記非線形関数の合成関数によって，複雑な関数を生成できるのは以下のような理由による。たとえば，シンプルな2次元の関数 $y=2x^2-1$ をグラフで表すと，下記のようなになる。



(20) これらについても，学習過程において動的に変化させて，最適な結果を得られる構造をシミュレーションすることも行われている。

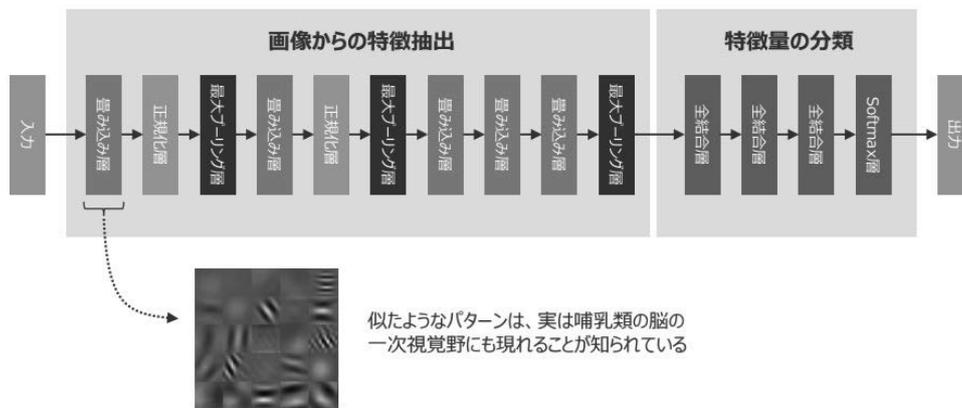
(21) <https://thinkit.co.jp/article/14103>

さらに、かかる左辺の y を同じ関数に入力とする $z=2y^2-1$ を x と z の関係でみると、4次関数となる。さらに、再度同じ関数に z を入力すると、8次関数となる。このように関数の結果を同じ関数に入力することを繰り返す演算を、10回やると1024次関数となる。このように非線形関数の演算結果を、さらに同じ非線形関数によって繰り返し演算させることで、複雑な関数と、同等な関数として表すことができる。

また、一般的に1024次関数だと、 $f(x)=P1x^{1024}+P2x^{1023}+\dots+P1024x+P1025$ で表されるため、パラメータ数は1025となるが、上記合成関数であれば、処理を10回繰り返すだけなので、パラメータ数が $30^{(22)}$ と極端に少なくなる。

このように、ディープラーニングは多次元の関数を、簡易な形で表すことができるので、現在のパソコンレベルでも画像判定などが可能となるのである。

画像認識で有名な AlexNet⁽²³⁾ が画像認識に効果がある理由は、似たような特徴抽出が、哺乳類の脳の一次視野に現れているからだといわれている。



「基礎から始める機械学習・深層学習」(MathWorks Japan)⁽²⁴⁾ P73 より引用。

4) パラメータ以外の部分の重要度

得られた学習済みプログラムにおけるパラメータとそれ以外のプログラムの関係について説明する。

学習前は、CNN の中間層における演算手法を特定する層構造特定パラメータおよび各演算における重みは、乱数設定された学習前パラメータである。一方、学習済みプログラムでは、各演算における重みが設定される。

学習前のプログラムのコンポーネントは、OSSなどで、学習済みプログラムのライブラリとして提供されているものもある。これを読み込んでデータを与えれば、学習済みプログラムを生成することもできる。

AI プログラム言語としては Python⁽²⁵⁾ が有名である。これは、AI・機械学習関連のライブラリ、フレームワークとして以下のように豊富に提供されているためである。

NumPy (行列演算パッケージ)

SciPy (プログラミング数学, 科学, 工学のための数値解析)

(22) 2次関数のパラメータ数は $f(x)=ax^2+bx+c$ で表されるので、 a, b, c の3つであり、これが10であれば $3 \times 10 = 30$ となる。

(23) 2012年に、ImageNetコンペティションで優勝したトロント大学 SuperVision チームの開発したネットワークである。

(24) https://www.cc.kyushu-u.ac.jp/scp/doc/users/lecture/2018/matlab_handout1.pdf

(25) Pythonは、プログラミング言語の命令を一つずつ、機械語に解釈しながら実行するインタプリタ言語であるので、処理速度に限界がある。そこで、処理を高速化するために JIT コンパイラという技術を利用したライブラリであるコンパイラ Numba が提供されている。

Pandas (データ解析)

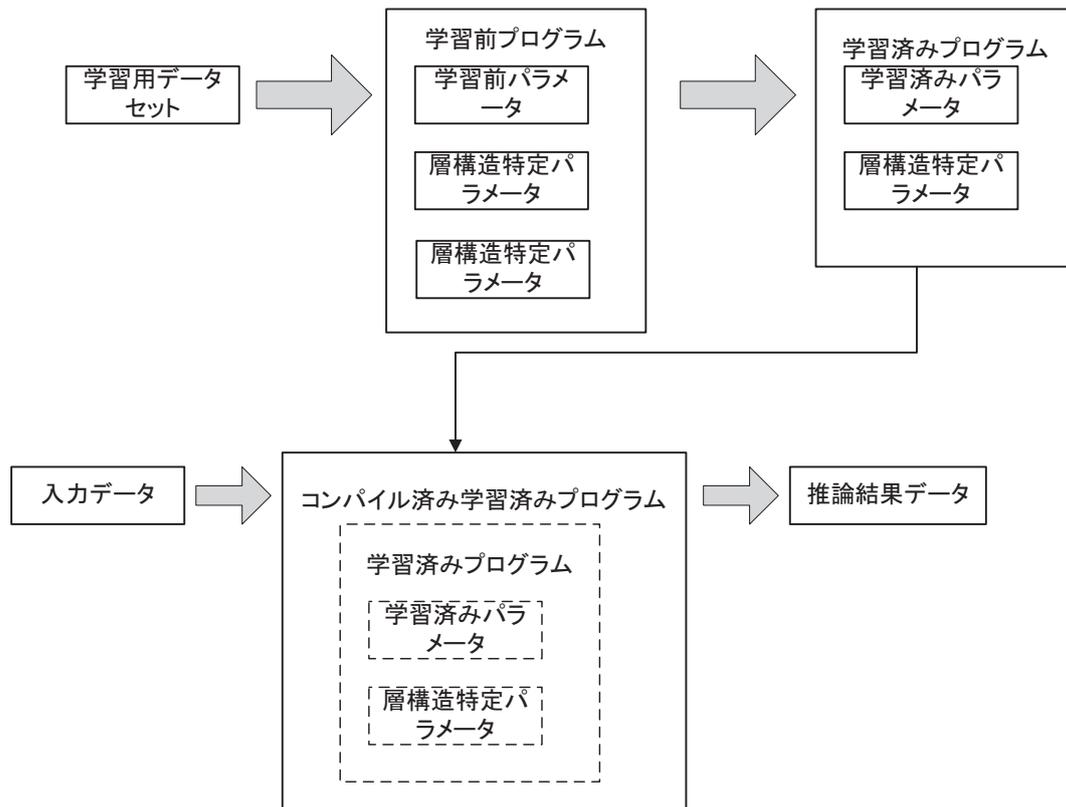
TensorFlow (機械学習に用いるためのソフトウェアライブラリ)

Chainer (ニューラルネットワークの計算および学習を行うためのライブラリ)

Scikit-Learn (機械学習ライブラリ)

Keras (ニューラルネットワークライブラリ)

このように、AI 適用発明における学習済みプログラムのコアの部分は、パラメータそれ自体といえる。



4. 「学習済みプログラムのパラメータのみを学習済みプログラムの部品の発明と解釈できないか？」 についての検討

1) 学習済みプログラムのパラメータの性質からみた検討

① 学習済みプログラムのパラメータの実体

上記3. で説明したように、学習済みプログラムのパラメータ（以下、単に、前記パラメータという）は、一種の関数を特定するためのデータの塊であり、複雑次元の関数を深さ方向で表したものである。すなわち、形式的にはデータの集合体にすぎず、それだけでは、単体では価値を見いだせないものである。

しかし、現在発明として扱われているプログラムも、単独では価値を見いだすことができない。プログラムが単独で発明として保護されるようになったのは、ハードウェアと一体化されて、「～装置」を構成することで、価値を見いだし得るからである。すなわちプログラムは、情報の塊にすぎないが、発明である「～装置」を構成する中核部品として、発明と認識されたものといえる。

② 前記パラメータも、このような考え方を採用できないかについて検討する。

前記パラメータは学習されていないプログラムに読み込まれて、一体不可分となり、学習済みプログラムを構成する。そしてかかる学習済みプログラムは、ハードウェアと一体化されて何らかの演算処理を行う。ということは同じ理屈で前記パラメータを学習済みプログラムの部品として把握できるはずである。

このように形式的にみると、前記パラメータは単独では単なる情報にすぎないし、コンピュータに対する指令とまではいえないが、未学習のプログラムと組み合わせされると、特許法上価値のあるプログラムとなる。すなわち、コンピュータプログラムの部品と考えることができる。

なお、プログラムについては、歴史的には、当初は、「装置」のみから「記録媒体との一体化」さらに、プログラム単体で発明として成立させている。前記パラメータについては、これを一気に飛び越す形となるが、これは特に問題にはならないといえよう。

なお、下記2つについては、プログラムの場合と同様に要件を具備できる。

ア) 無体物である点

特許法においても、伝統的には「物」とは民法85条の「有体物」と同義と解されていた。しかし、平成14年(2002年)4月の特許法改正(平成14年9月施行)により、特許法における「物」には、無体物である「プログラム等」が含まれることが確認された。改正の解説としては、「プログラム等を、新たに「物の発明」として含まれるように改正されたわけではない」とされている。

イ) 単独では技術的效果がない点

前記パラメータは、関数を特定するためのデータの塊であり、単体ではなんら価値が見いだせない。しかしながら、その点は、プログラムと同様である。すなわち、プログラムもプログラム単体では実質的な価値はなく、ハードウェアと一体化してはじめて技術的效果を奏する。

すなわち、プログラムはシステムの中核部品であるので、単体では技術的效果を奏しなくても発明性を認められたといえる。

このように、プログラムについては、発明性があるのは、当初は「システム全体」だけ(又はサブコンピネーション)であったが、そのうち、「～のプログラムが記録された記録媒体」が認められ、現在ではプログラムそれ自体も物の発明として認められている。これは、そのようなものが単独取引対象となっていったという社会的要請があるからであろう。

パラメータは学習されていないプログラムに読み込まれて、一体不可分となり、学習済みプログラムを構成する。したがって、同じ考え方が可能である。

また、学習済みプログラムのパラメータを0から作るよりも、これを用いてすこし学習を進めることで、より短時間で所望の結果を得られるプログラムを得ることもできる。その意味で、途中まで学習させたパラメータは社会的価値を有するようになるかも知れない。

③ ブラックボックス化との関係

前記パラメータは、技術思想を内包しているとしても、ブラックボックス化されているので、数字の羅列に見えてしまう。しかし、技術思想が見えないからといって発明でないとはいえない。たとえば、化学の発明は、なぜそれが起こるかわからないが、現象を実験結果で示すことで発明として認識され得る。

ただ、化学の発明は、もともと自然法則に則って現象が起きている。この点については、コンピュータというハードウェア資源との協同要件を満足すれば発明として認識するという現在の実務であれば、ブラックボックス化されていることは特に発明として保護しないことを決定づけるものとはならない。

④ パラメータの本質

ディープラーニングの代表ともいえるCNNの場合、一般的には畳み込み演算の数が多いほどよいが、そのためには、膨大な学習用のデータが必要であり、また、生成過程で多くの演算リソースが必要となり、その分コストもかかる。そこで、現実には、できる限りシンプルでかつ、精度の高い組み合わせを特定した結果、得られたのが学習済みプログラムである。すなわち、AIのパラメータの価値は、理論上のベストとまではいえないが、入力されたデータを前提とすると現実にはほぼ問題がないという値を見いだした点にある。その意味で、化学分野の発明と似たものとなるといえる。

当然、パラメータとして最適値のみを特定すれば、点の特許とならざるをえない。しかし、化学分野の発

明でも、数値範囲をある程度広げて実験結果を示して、幅のある特許を取得する。前記パラメータについても同様の考え方をすればよい。この点については後述する。

⑤ 数学的なブロックの組み合わせが発明となり得るのか

似た構成としては、電気回路に関する発明がある。電気回路を構成する個々の部品については公知であるので、その接続の仕方に発明の進歩性を見いだしている。

前記パラメータもその意味では同じである。すなわち、どのようなフィルタ構成とするのか？、何層で構成するのか？、活性化関数として何を用いるのか？などの個々の構成については、公知であるが、その接続手法が特殊であり、ハードウェア資源と一体化すれば、発明として認められるべきである。

2) 実施行為との関係

歴史的には、特許法2条3項1号の「物」とは有体物と理解されていた。これは、特許も物権的権利であるので、民法の考え方がベースとなったからである。また、発明が具現化された「物」を生産して、販売行為などを専有するというインセンティブが特許権の中核をなしていたので、有体物に限定して解釈しても、事実上は問題なかった。

しかし、特許法で、「物」と「方法（生産方法を含む）」に発明を区分けしたのは、そもそも、侵害行為を明確に定義するためである。であれば、実施行為に該当するような実施が可能なのであればこれを保護対象から外す必要はない。

前記パラメータを物の発明として把握した場合、その生産、使用等はなんら問題がない。したがって、実施行為の定義からみた場合に、特許法2条4項の「プログラムに準ずるもの」に含めて解釈することに何ら問題はない。

5. クレームにおける特定について

1) パラメータは最適値である

酒井は、「学習済みモデルの特定方法は、製造方法以外はない」と説明する⁽²⁶⁾。これは、一般的に学習済みプログラムはパラメータが複雑で特定ができないためであろう。

確かに、前記パラメータは、すでに述べたように、その最適値を見いだした発明であるので、これらはパラメータとして数値で限定されているため、これらを特定すれば発明として成立し得る。しかし、パラメータの全てを特定すると、点の特許となり、侵害回避が簡単となってしまう。

2) 範囲指定について

これをサポートするために、前記パラメータを変えた場合の、正答率の変化をサポートする開示をして、これを含むような特定ができれば点の権利とはならない。したがって、最適値で予測精度⁽²⁷⁾が、98.2%であった場合、パラメータを調整して、予測精度が95%以上となる組み合わせをシミュレーションして、その組み合わせを特定することも可能であろう。かかる手法は、化学発明における数値範囲の発明特定手法と同様である。

層構造をどうすれば予測精度が上がるのか？などは、学習過程でプログラムが自動的に変更して、予測精度が高くなる並べ替えなどを自動的に行うサービスも存在する⁽²⁸⁾ので、複雑化するとしても不可能ではない。

(26) 酒井將行「AI・IOT技術によるビジネスモデルに対する知的財産権」パテント2018 Vol. 71 No.11 (別冊 No.20) P230

(27) 正確には、正解率（全予測正答率）、適合率（正予測の正答率）、再現率（正に対する正答率）、特異率（負に対する正答率）、F値（適合率と再現率の調和平均）などの表し方がある。

(28) <https://dl.sony.com/>

3) 一部のパラメータの特定について

また、パラメータの特定手法としては必ずしも、全てを特定する必要はないと解する。たとえば、前出の AlexNet が見いだしたように、哺乳類の脳の一次視野に現れる特徴を抽出するフィルタとして機能していることがわかれば、そのようなフィルタを用いるとの限定が可能であろう。

また、ある層構造である場合にだけ、予測精度が高い場合には、当該層構造を特定することも可能かもしれない。なお、このような場合に、全パラメータでなく、クレームにおいて一部の特定でよいのかについては、実施可能要件との関係で問題となる。この場合も、実施形態には全てのパラメータを記載すれば何ら問題とはならない。サポート要件については、シミュレーション結果として、それを示す必要はあるがこれも可能であろう。

特に、入力データが技術的な性質を有する場合、上記のような一部の特定で発明として認定できるであろう。たとえば、地球上に存在する数十億の花の画像から花の種類を特定する学習済みプログラムがあったとして、それは、自然界に存在する花の画像が有する何らかの特徴から花の種類を判別するプログラムであり、自然法則に則った発明である。このように、画像認識、音声認識など何らかの自然法則に基づいているので、その一部であっても遺伝子配列と同様に考えるべきである。

4) パラメータを一部特定した形での権利化のメリット

現在の審査基準における AI 発明は、入力と出力パラメータだけで、演算部分は全く特定しない形で行われる。確かに権利範囲は広いが、どこかに願望として記載されている先行文献が存在すれば、簡単に無効となってしまうおそれがある。これに対して、上記のようなフィルタ特性又は層構造などを具体的に特定した従属クレームがあれば、それだけでは無効にはできない場合がほとんどであろう。その意味で、より深みのある特許が取得できる。

5) 用途を特定する必要があるのか

たしかに、前記パラメータは、ディープラーニングの学習済みプログラムに用いるという用途以外の用途は想定できない。しかし、上述したように、前記パラメータを物の発明として保護する前提として、学習済みプログラムとの一体性が問題となる。したがって、クレームにて一体性担保のための特定は何らかの形で必要であろう。

6) データ構造クレームについて

既に述べたように、審査基準には「データ構造」を特定したクレーム記載形式が認められている。これにより、特殊なデータ圧縮技術の結果物である特殊なデータ構造を有するデータそれ自体をクレームできるというメリットがある。ただし、現在の実務では、単にデータ構造だけではなく、それにさらにデータ処理の特定がないと登録できない。その意味で、上記のパラメータには使いにくいであろう。

6. 最後に

1) 社会的要請の増大について

今後は、前記パラメータの位置づけが、より重要視されるであろう。

妹尾⁽²⁹⁾は、機械世代論において、19世紀には、人間の動力が外在化され、20世紀は計算系と記憶系が外在化された。そして、21世紀は感覚系、五感が外在化され、IoT時代が一般化されると複数のロボットがネッ

(29) 妹尾堅一郎 (抄録執筆：加藤)「モノのサービス武装，サービスのモノ武装 ～IoT時代の産業生態系を見通し，ビジネスモデルを仕掛ける～」 <https://www.i-ise.com/jp/column/hiroba/2016/20160525.html>

トワークされた。これは神経系の外在化を意味し、頭脳の外在化を意味すると、指摘している。

小川⁽³⁰⁾は、18～19世紀は、フィジカル空間（現実の世界）で価値形成がなされたが、20～21世紀は、サイバー空間での価値形成がなされる、すなわち、現実世界のデータがサイバー空間で価値を持つというようにデータの位置づけが変わると指摘している。

このようなデータの重要性が叫ばれる中、IoT 機器からのデータに基づくデータ収集はますます加速していく。

その意味で、ビッグデータは従来のプログラムにおける数値データとは質的なレベルが異なるといえ、膨大なデータ（データ群）から得られた知見であり、知財そのものともいえる。

また、データ量が膨大だと、学習処理にはかなりの日数がかかる。これに対して、一般的なデータ群で一定の学習が済んでいると、これをベースとして後は、特有のデータを追加学習することで、新たな学習済みプログラムを作成するというやり方も提案されている。その場合、一定の学習が済んでいるパラメータそれ自体が単独で取引対象となることも考えられる。

AI 関連発明のパラメータは、発明としてどのような取り扱いが可能なのかを今一度検討すべきである。

なお、現実の AI プログラムのパラメータについては、演算能力の向上等とともに、複雑化していくと考えられる。そうすると、遺伝子配列のように、別の形式でデータ提出が必要となるかもしれない。

2) 自然法則の利用性がもともとある場合のパラメータについて

現在の審査基準では、株価の予測システムのように、もともと、自然法則の利用性がない場合でも、ハードウェアとの協働要件を満たせば発明該当性ありと認定される。これに対して、画像のエッジ検出のように自然法則の利用性がもともとあるコンピュータシステムについては、ハードウェアとの協働要件を特定しなくても発明として認定される。

その意味では、学習済みプログラムのパラメータも、画像認識のように自然法則をもともと用いている場合と、販売データのような自然法則には則っていない場合とを、分けて議論することも検討が必要であろう。

以上

(30) PwC コンサルティング合同会社主催 ウェビナー（特許庁委託事業）「オープンイノベーションを活用した新事業創造に資する知財戦略の実践へ向けて」（2020年12月21日開催）の小川紘一講演資料