

中山英樹

Abstract

現在の機械学習の成功を支える中心的なアプローチである教師あり学習では、対象とするタスクの入出力関係を網羅した大規模なラベル付き訓練データセットの存在を前提とする。しかしながら、実際の応用場面では大規模なデータを用意することは必ずしも容易ではない。また、ある一時点において固定されたデータセットによる学習では、ダイナミックに変化する実世界をカバーすることは困難である。本稿では、限られたデータからの学習や、オープンワールドへの対応を念頭に、機械学習の先進的な研究分野について解説を行う。

キーワード：限られたデータ、オープンワールド、弱教師あり学習、転移学習、継続学習

1. はじめに

機械学習が様々な知能システムを実現するための中心的な技術として活躍するようになってから既に久しい。特に、深層学習（ディープラーニング）のブレイクスルー以降の発展は目覚ましく、従来は非常に困難であると考えられてきた技術が次々と実現されている。

機械学習には前提とするデータに応じた幾つかの枠組みが存在するが、現在の成功を支えているのは、事例に対する正解を教示する教師あり学習と呼ばれる枠組みである。ある入力 $\mathbf{x} \in \mathcal{X}$ に対し、適切な出力 $\mathbf{y} \in \mathcal{Y}$ を与える関数、すなわち予測モデル $f: \mathcal{X} \rightarrow \mathcal{Y}$ を得ることを考えよう。ここで、 \mathcal{X} , \mathcal{Y} はそれぞれ入力、出力の空間を表す。これを実現するために、教師あり学習では、入出力のペアからなる事例を多数集めた訓練データセット $(\mathbf{x}_i, \mathbf{y}_i)_{i=1}^N$ (N はデータ数) を、いわば「練習問題」として利用してシステムを訓練する。 \mathbf{y}_i は i 番目の入力事例 \mathbf{x}_i に対する答えであり、教師あるいはラベルと呼ばれる。信頼性の高いラベルをどのように得るかが教師あり学習における重要なポイントであり、多くの場合、人手によりデータをアノテーションして作成する必要がある。このような良質なラベル付きデータセットが得られ

れば、深層学習を基盤とする教師あり学習により、様々なタスクで非常に良い性能が得られるようになっている^{(1)~(3)}。一方で、実際の応用を考えた場合、単純な教師あり機械学習には課題も多い。本稿では、以下に述べる二つの問題に焦点を当てる。

第1に、教師あり学習ではラベル付きの訓練データが大量に存在することが必要である。例えば、深層学習を用いた画像識別では百万枚以上のラベル付き画像を用いることが一般的になっている⁽⁴⁾が、このような大規模データセットをユーザが自力で構築することは現実的でない場合も多い。例えば、医療などプライバシーが重要なアプリケーションの場合、データの収集自体を慎重に行う必要があるだろう。また、教師あり機械学習には生データだけがあればよいわけではなく、対となるラベルの付与は多くの場合人手に依存し、膨大な人的コストが必要となることを忘れてはならない。

第2に、システムの訓練時と実際の運用時（テスト時）が同じ状況であること、すなわちクロズドワールドを想定していることも大きな制約である。具体的には、入出力の空間 \mathcal{X} , \mathcal{Y} や、確率分布 $p(\mathbf{x}, \mathbf{y})$ が訓練の前後で変わらないことが前提である。これは、画像識別を例にとると、訓練データに含まれるクラス（既知クラス）のみがテスト時にも現れ、その見え方も訓練時と変化しないことを意味する。しかしながら、より実世界を対象とするアプリケーションでは必ずしもこの前提は成り立たない。例えば、家庭用ロボットを例にとると、時々刻々と変化する家庭環境の全ての物事をあらかじめ

中山英樹 正員 東京大学大学院情報理工学系研究科創造情報学専攻
E-mail nakayama@ci.i.u-tokyo.ac.jp
Hideki NAKAYAMA, Member (Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, 113-8657 Japan).
電子情報通信学会誌 Vol.104 No.12 pp.1263-1270 2021年12月
©電子情報通信学会 2021

教えておくことは不可能であり、訓練データに含まれないクラスも必然的に現れる。また、一度教えたクラスであっても、その分布は絶えず変化していくであろう。このように、ダイナミックに変化するオープンワールドに対応するためには、新奇な物事を発見し、逐次的に知識をアップデートするための仕組みが必要である。

以上述べたとおり、本稿では、(1)限られたデータからの学習、(2)オープンワールドへの対応、の2点を、機械学習が実世界で更に活躍するために重要なポイントであると考え、それぞれについて画像認識分野を中心に近年の研究動向を紹介する。

2. 限られたデータからの学習

有限個のサンプルデータから一般的な規則を推定することはそもそも機械学習の目的そのものであり、基本的には、経験誤差の最小化において何らかの正則化を行い、汎化性能を上げることが王道となる考え方である。しかしながら、与えられたラベル付きデータの数が極端に少ない場合には限界があり、何らかの別のリソースから知識を得るための工夫が必要となる。

2.1 弱教師あり学習

弱教師あり学習とは、教師あり学習と比べて不完全な教師情報が与えられる場合の学習手法の総称である⁽⁵⁾。

古くから研究されてきた半教師あり学習では、ラベル付きデータは少量しか与えられないが、多数のラベルなしデータを活用することで性能向上を図る⁽⁶⁾。実際のアプリケーションでは、生データ自体は容易にとれる場合が比較的多いため、ラベルなしデータを活用することは重要な方向性の一つであるといえる。

半教師あり学習には非常に多くのアプローチがあるが、比較的単純な手法である擬似ラベリング⁽⁷⁾では、まず与えられたラベル付きデータのみで予測モデルの学習を行い、この予測モデルでラベルなしデータへラベルを付与する。これを元のラベル付きデータセットへ加えて

学習を再実行する、という処理を繰り返す。このアプローチは、モデル自体は一般的な教師あり学習のものをそのまま利用できるため扱いやすく、深層学習時代において広く用いられるようになっていく。

このように、半教師あり学習手法の多くはラベルなしデータのラベルをアルゴリズムの中で自動的に推定するが、関連する研究領域である能動学習⁽⁸⁾と呼ばれる枠組みでは、特に重要なラベルなしデータを選択的にユーザ(人間の教示者)へ提示し問い合わせることで、対話的にラベル付けを行う。

以上のように、ラベルを与えるデータの量を減らすアプローチのほか、ラベル自体を簡素化することもアノテーションコストを減らす有効なアプローチである。例えば、正例のみラベルを付与するPU学習⁽⁹⁾や、あるデータ集合が正例を含むか否かを示す集合レベルのラベルのみ付与する Multiple Instance Learning (MIL)⁽¹⁰⁾などが汎用的な枠組みとして挙げられる。また、各タスクに特化したラベルの簡素化も多く研究されている。例えば、画像認識のタスクの一つであるセマンティックセグメンテーションでは、一般的には教師ラベルとして、画像中の各物体をピクセルごとに詳細に区分けしたラベルが必要となり、アノテーションに非常に大きな手間を要する。これを例えば、画像全体に対してタグ付けした物体ラベルのみから学習できるようになればアノテーションコストは大きく下がるであろう(図1)^{(11),(12)}。

2.2 転移学習

転移学習^{(13),(14)}とはかなり広範な概念や技術を指す言葉であるが、「新規タスクの効果的な仮説を効率的に見つけ出すために、一つ以上の別のタスクで学習された知識を得て、それを適用する問題」というものが広く受け入れられている定義の一つである⁽¹⁵⁾。

まず、文献(13)、(14)に従い、2.2で必要となる幾つかの言葉を定義する。

- ・ ドメイン $D = \{\mathcal{X}, p(\mathbf{x})\}$: ある入力の空間 \mathcal{X} と、

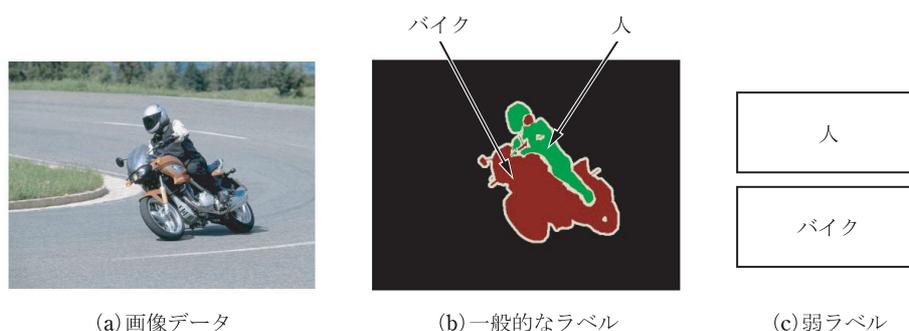


図1 セマンティックセグメンテーションにおける弱ラベルの例(画像、ラベルはPascal VOCデータセット⁽¹²⁾より引用)

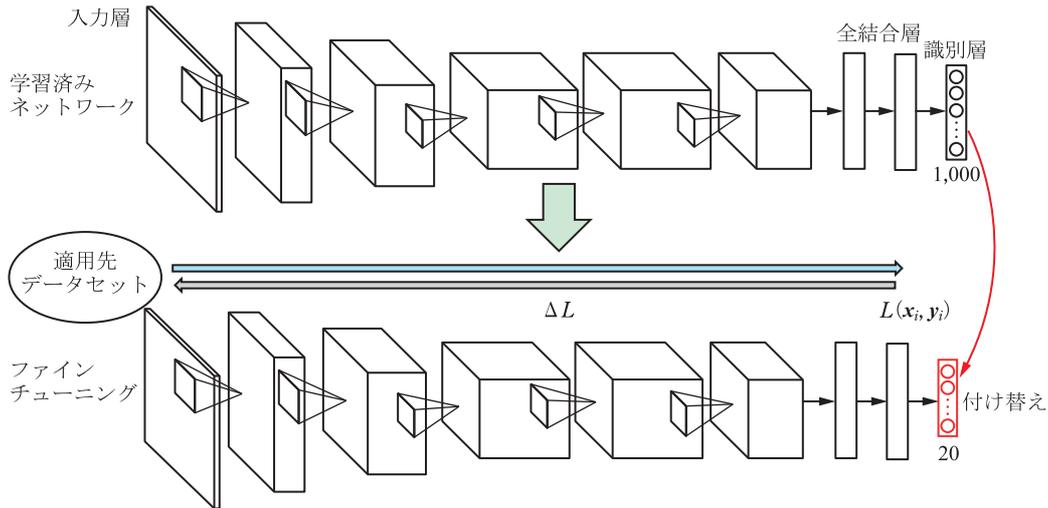


図2 深層学習モデルを用いたファインチューニング (文献(16)から引用)

入力データ $\mathbf{x} \in \mathcal{X}$ の分布.

- ・ タスク $\mathcal{T} = (\mathcal{Y}, p(\mathbf{y}|\mathbf{x}))$: ある出力の空間 \mathcal{Y} で, $\mathbf{x} \in \mathcal{X}$ に対する望ましい出力 $\mathbf{y} \in \mathcal{Y}$ を与える規則.

ここで, 我々が最終的に実現したいシステムを構成するドメイン及びタスクをそれぞれターゲットドメイン $\mathcal{D}^t = \{\mathcal{X}^t, p(\mathbf{x}^t)\} (\mathbf{x}^t \in \mathcal{X}^t)$, ターゲットタスク $\mathcal{T}^t = (\mathcal{Y}^t, p(\mathbf{y}^t|\mathbf{x}^t)) (\mathbf{y}^t \in \mathcal{Y}^t)$ と呼ぶ. 我々が考えるのは, ターゲットドメイン及びタスクに関する十分大きなラベル付きデータセットが存在しない場合である. このようなときに, ターゲットに何らかの形で関連した別ドメイン・タスクのデータから得られる知識を活用することを考える. これらをそれぞれソースドメイン $\mathcal{D}^s = \{\mathcal{X}^s, p(\mathbf{x}^s)\} (\mathbf{x}^s \in \mathcal{X}^s)$, ソースタスク $\mathcal{T}^s = (\mathcal{Y}^s, p(\mathbf{y}^s|\mathbf{x}^s)) (\mathbf{y}^s \in \mathcal{Y}^s)$ と呼ぶ.

以上が問題の一般的な定義となるが, 本稿では $\mathcal{X}^s = \mathcal{X}^t$ の場合を前提とするため, 以下のように簡略化した表記を用いることにする.

- ・ $\mathcal{D}^s = \{\mathcal{X}, p(\mathbf{x})\}$
- ・ $\mathcal{D}^t = \{\mathcal{X}, p(\mathbf{x})\}$
- ・ $\mathcal{T}^s = (\mathcal{Y}^s, p(\mathbf{y}^s|\mathbf{x}))$
- ・ $\mathcal{T}^t = (\mathcal{Y}^t, p(\mathbf{y}^t|\mathbf{x}))$

ただし, $\mathbf{x} \in \mathcal{X} (= \mathcal{X}^s = \mathcal{X}^t)$ である.

2.2.1 ファインチューニング

深層学習時代において最も一般的に用いられる転移学習法は, ソースタスクで事前学習されたニューラルネットワークを, ターゲットタスクで更に学習させるファインチューニングと呼ばれるアプローチである (図2)⁽¹⁶⁾. これは, ソースタスクで得られた知識をニューラルネッ

トワークのパラメータとして表現し, これをターゲットタスクにおける学習の初期値として用いることで知識転移を行っていると解釈できる. ファインチューニングは汎用性の高い枠組みであり, $\mathcal{D}^s \neq \mathcal{D}^t$ の場合はもちろん, 出力層の一部を付け替えることで $\mathcal{T}^s \neq \mathcal{T}^t$ の場合でも適用可能である. 元々画像認識で成功したアプローチであるが⁽¹⁷⁾, 近年は自然言語処理においても大規模なテキストデータから事前学習した言語モデルが極めて強力に働くことが分かり, 注目されている^{(18), (19)}.

なお, この枠組みでは, ソースタスクにおける事前学習時には大規模な訓練データが存在することが前提である. 画像認識における事前学習は基本的に教師付きデータによって行われてきたが, 自然言語処理における言語モデルは自己教師あり学習と呼ばれるアプローチが成功してきた⁽¹⁸⁾. これは, 自分自身の構造を一部隠して予測することを通して表現学習を行う手法であり, 教師なし学習の一種である. 近年では, 自己教師あり学習が画像認識モデルの事前学習でも高い性能を得るようになったことから注目を集めている⁽²⁰⁾.

2.2.2 マルチタスク学習

マルチタスク学習は, 関連する複数のタスクを一つのモデルで同時にこなすことで, それらの共通因子をモデル内部に獲得し, 各タスクの予測精度向上に役立てるものである⁽²¹⁾. 典型的なモデル構成では, 入力からの特徴抽出部を一本化し, この上にタスク固有の予測レイヤを接続する (図3). ファインチューニングでは最終的にターゲットタスクの性能を上げることが目的であったが, マルチタスク学習では全てのタスクの性能を並立させることが目的である. ソースタスク, ターゲットタスクの明示的な区別はないものの, 各タスクは共有する特徴表現を介して互いに有用な情報を交換していることか

ら、転移学習の一種であると解釈することができる。

一見、複数のタスクを同時にこなすことはモデルにとってより複雑な作業であるため、各タスクのデータでそれぞれ独立にモデルの学習を行う方が合理的に思えるかもしれない。しかしながら、特にそれぞれのタスクのデータが少ない場合、マルチタスク学習は他のタスクのデータを間接的に利用できるため、結果として各タスクで独立にモデルを学習した場合より高い汎化性能を得られることが多い。

2.2.3 ドメイン適応

ドメイン適応では主に、 $\mathbf{y}^s = \mathbf{y}^t$ であるが、 $p(\mathbf{x}^s) \neq p(\mathbf{x}^t)$ となる場合を扱う。つまり、ソースとターゲットでタスクの目的は同じであるが、入力データの周辺分布が異なる場合である。例えば、CG 画像をソースとして訓練に用い、ターゲットとして実画像中の人検出を行う場合が挙げられる⁽²²⁾。ドメイン適応は、ターゲットのデータに一切ラベルが与えられない教師なしドメイン適応と、多少のラベルは得られる半教師ありドメイン適応

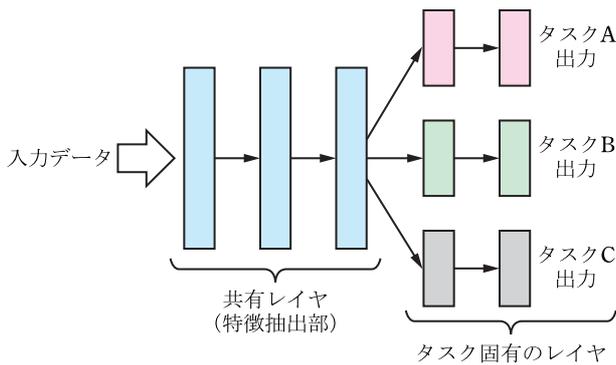


図3 ニューラルネットワークによるマルチタスク学習モデル

に分けられるが、単にドメイン適応という場合には教師なしを指す場合が多い^(注1)。

ドメイン適応では、 $p(\mathbf{x}^s)$ と $p(\mathbf{x}^t)$ のずれ及び $p(\mathbf{y}^s|\mathbf{x})$ と $p(\mathbf{y}^t|\mathbf{x})$ のずれ、すなわちドメインシフトを吸収することが必要になる。しかしながら、 $p(\mathbf{y}^t|\mathbf{x})$ に関する情報はほとんど得られないため、ドメインシフトを正確に推定することは非常に難しい。特に、教師なしドメイン適応は一般的に解くことはほぼ不可能であるが、経験的には、周辺分布 $p(\mathbf{x}^s)$ と $p(\mathbf{x}^t)$ を大域的にアラインメントするように特徴空間の学習を行うだけでもある程度うまくいくことが知られている (図4)^{(23), (24)}。ここでは、分布の内側での相対的なクラスの配置は大きく変化しないことが暗に仮定されているといえる。

ドメイン適応の進化形の一つに、ドメイン一般化 (Domain Generalization)⁽²⁵⁾ が挙げられる。これは、ターゲットタスクでは教師なしデータを含め一切のデータが得られない状況を想定した更に挑戦的なものである。多くの手法では、教師あり訓練データセットが存在する複数のソースドメインを利用した学習により、ドメインの差異に対して汎化した汎用的な特徴表現を得ることで、未知のドメインに対しても自然にアラインメントを行うことを狙う。

2.3 外部知識の利用

ここまでは、ターゲットタスクの教師ありデータのほかに何らかの別データを活用するアプローチを見てきたが、より抽象化されたりリソースを用いる研究も盛んに行われている。代表的な例として、言語的知識リソースを

(注1) そもそも、ターゲットタスクである程度のラベル付きデータが得られるならばファインチューニングやマルチタスク学習ができるので、あえてドメイン適応を行うメリットは薄いだろう。

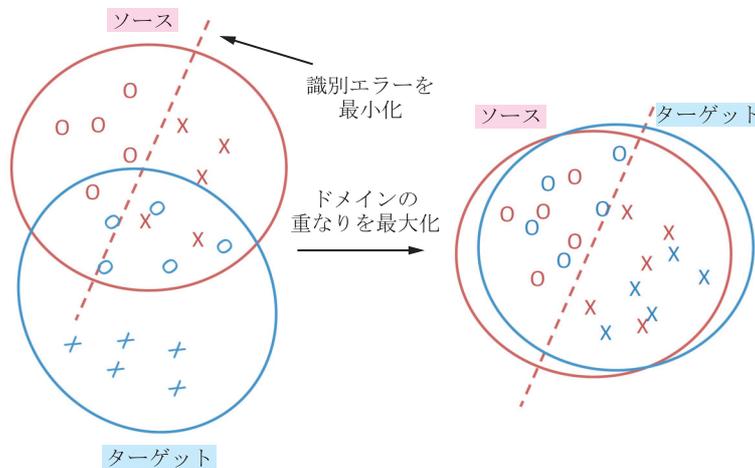


図4 ドメイン適応における周辺分布のアラインメント (文献(23)から引用して和訳)

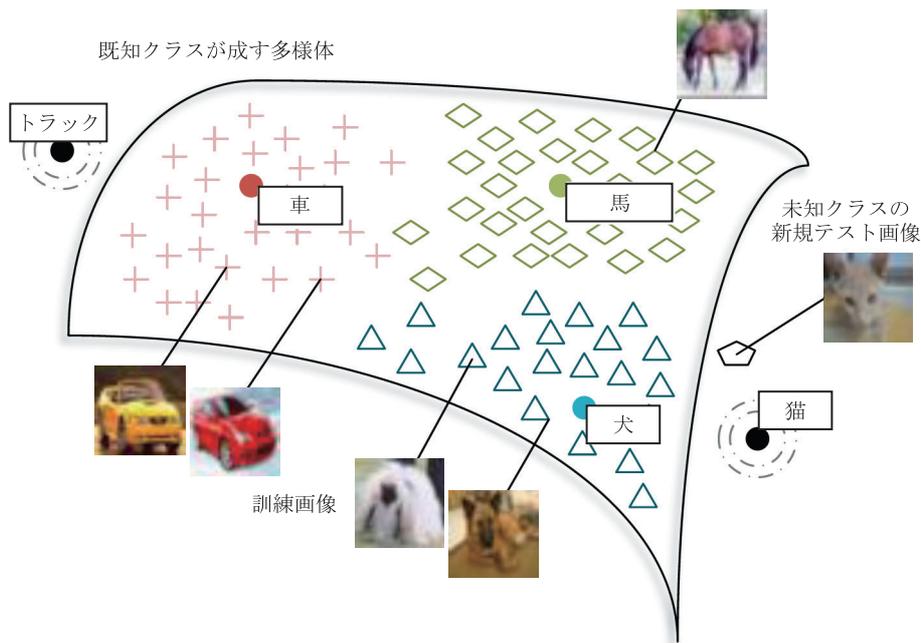


図5 単語ベクトルとのクロスモーダルな埋込空間を活用したゼロショット画像認識 (文献(27)から引用して和訳)

活用したゼロショット画像認識を紹介する。

ゼロショット画像認識とは、対象のラベル付きデータが一切与えられない状況で認識を行うものであり、何らかの知識を外部から与えない限り原理的に実現不可能なタスクである。最初期には、属性情報 (attribute) を活用するアプローチから研究が始まった⁽²⁶⁾。これは、認識の手掛かりとなるプリミティブな視覚的属性の識別器と、対象クラスの辞書的な知識を与えることで未知クラスの認識を可能とするものである。例えば我々人間は、もし「シマウマ」という動物を生まれて初めて見たとしても、シマウマとは馬とほぼ同じ形状で全身しま模様の動物である、という知識を持っていれば、それがシマウマということは推測できるであろう。これと同じように、形状や縞模様の属性識別器の出力と辞書情報を照合することにより認識を行うのである。

現在のゼロショット認識も基本的に同じアプローチの延長にあるが、深層学習時代に入ると知識リソースがニューラルネットワークの中にシームレスに接続されるようになっていく。例えば、自然言語処理の分野で広く用いられる単語ベクトルを利用するのが典型例として挙げられる⁽²⁷⁾。単語ベクトルは、テキストコーパスを用いた自己教師あり学習によって単語をベクトル空間上の点として埋め込んだものであり、単語の意味的な関係性が空間上での距離や位置関係として構造化されることが知られている。したがって、教師付きデータセットを用いて画像特徴と単語ベクトルのマッチングができれば、未知の入力に対しても最も近い単語ベクトルを検索

することで、そのクラスを推定することができる (図5)。

より近年では、知識リソースとして、Wikipediaなどの一般的なテキストや、知識グラフなどの構造化されたデータベースを用いるものも多い^{(28), (29)}。

3. オープンワールドへの挑戦

ここまで紹介した技術は、学習の前提や手順の違いはあるものの、最終的にシステムが運用されるターゲットタスクにおいては基本的にクローズドワールドが前提となっていた。しかしながら、実世界はダイナミックに変動するものであり、訓練データセット中に存在しない未知の物体は次々に現れるため、システムはまだ学習していない事物を正しく未知であると判断し、新しいクラスとして追加的に学習しなければならない。ここでは、そのようなオープンワールドを想定した機械学習の研究分野について紹介する。

3.1 オープンセット認識

オープンセット認識^{(30), (31)}とは、既知クラスを正しく認識すると同時に、既知のどのクラスとも類似していない入力データを未知クラスとして判定できる識別器の構築を目的とするものである。しかしながら、未知事象の棄却は、そもそもの機械学習の目的である汎化とは背反するため、両者の両立は非常に難しい課題である。

直感的な例を図6⁽³²⁾に示す。ここでは、二つの既知ク

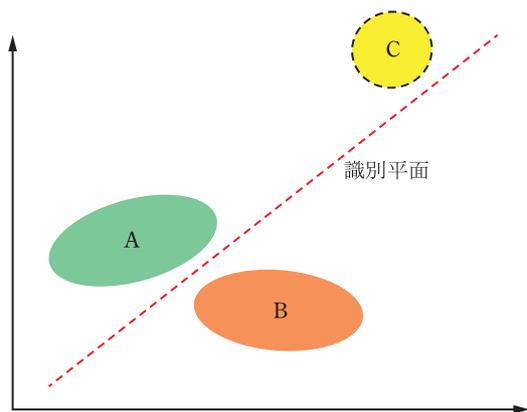


図6 識別的なモデルにおいて、未知事象が誤認される例 A と B が既知クラス、C が未知クラスを示す (文献(32)から引用)。

ラス A と B のラベル付き訓練データが与えられて、図の位置に分布をなしているとする。代表的な識別的手法の一つである線形 SVM では、二つのクラスの訓練データを最も合理的に分離するように識別平面を定める。識別平面は、訓練データが与えられていない領域（オープンスペース）までカバーしているため、多少訓練データから外れた位置に入力があったとしても、既知クラスのデータであれば高い汎化性をもって識別できることが期待される。しかしながら、もし訓練データに存在しない未知のクラス C が図の領域に存在し、そのデータが入力された場合何が起るであろうか。人間の常識的な感覚では、A から B からも離れているため、どちらにも属さないデータであるという推測がある程度可能であろう。しかしながら、システムはあくまで A と B から学習された識別平面を基準とするため、高い確信度で A と識別してしまうことになる。このような現象は、実世界で稼動するシステムでは予期しない誤動作を引き起こす可能性があり、深刻なぜい弱性となるだろう。

このように、オープンセット認識は、オープンスペースのどこまでを既知事象へ含め、どこから先を未知と判断するかというトレードオフを扱う問題となる。Bendale ら⁽³⁰⁾は、画像識別タスクを題材とし、既知クラスのスコア関数のしきい値処理に基づくオープンセットの識別ルールを提案した。具体的には、既知クラスの中で最大スコアを得たクラスがあるしきい値 τ を超えていればそのクラスへ識別し、どのクラスのスコアも τ を超えていない場合には未知クラスと判定する。これは一見妥当に見える戦略であるが、スコア関数は何を用いてもよいわけではなく、各クラスの訓練データ集合からの距離に応じて単調に減少することが求められる。これは図6の例で見たように、SVM や畳込みニューラルネットワークのような識別的な手法では一般には満たされない条件である。

そこで Bendale らは更に、ニューラルネットワーク

のソフトマックス層をオープンセット認識のために改良した OpenMax と呼ばれる手法を提案した⁽³³⁾。これは、ソフトマックス層の手前の層が成す特徴空間において、入力と既知クラス中心との距離を基準として各クラスの重みを計算し、未知クラスを含めて正規化したソフトマックスを計算するものである。最終的な識別スコアだけでなく、特徴空間上での距離を考慮することで、識別モデルにおいてもオープンスペースの評価をある程度可能にしている手法であるといえる^(注2)。

3.2 継続学習

継続学習とは、同じモデルで新しいタスクを追加的かつ継続的に学習させる枠組みであり、逐次学習やライフロング学習などとも呼ばれる⁽³⁴⁾。継続学習では、新規に学習するタスクのみならず、過去に学習したタスクの性能も保つ必要がある。複数のタスクをこなす点においてマルチタスク学習と似ているが、継続学習ではあるタスクについて学習が終わると、そのタスクの訓練データには二度とアクセスできなくなる点が大きく異なる。イメージとしては、人間が人生において蓄える様々な経験を活用しながら新しい物事を効率良く学ぶように、機械学習においても過去のタスクから得られる抽象的な知識を新規タスクの学習に役立てることを狙うものである。

このように、継続学習は過去タスクからの知識転移を行う点において転移学習の一種であると言えるが、ファインチューニング等の一般的な枠組みを単純に適用することはできない。なぜならば、継続学習ではある一時点においてモデルに与えられるのはある一つの新規タスクの訓練データのみであり、単純にこのデータのみを用いて学習を進めると、その新規タスクに対してのみモデルがフィッティングし、過去のタスクに対する学習結果が消えてしまう破滅的忘却 (Catastrophic Forgetting) という現象が起きるためである⁽³⁵⁾。このため、継続学習では、過去のデータ自体に依存せず過去のタスクに関する知識を保持する手段が必要である。例として、新規タスクの学習時にパラメータの勾配降下において、過去タスクへの影響を少なくする制約を加えるもの⁽³⁵⁾や、過去タスクのデータ分布を生成モデルとして保持するアプローチ⁽³⁶⁾などが挙げられる。また、過去タスクのデータを少しだけ保持することを許すように条件を緩和し、新規タスクの学習時に過去タスクのデータも再利用するメモリベースと呼ばれるアプローチも有力である⁽³⁷⁾。一般に、メモリベースのアプローチはシンプルであり、安定的に良い性能を得られることが知られているが、実質的には少量の教師データを用いたバッチ学習

(注2) もちろん、畳込みニューラルネットワークでは特徴表現そのものが識別的に学習されるため、完全に問題が解決されているわけではない。

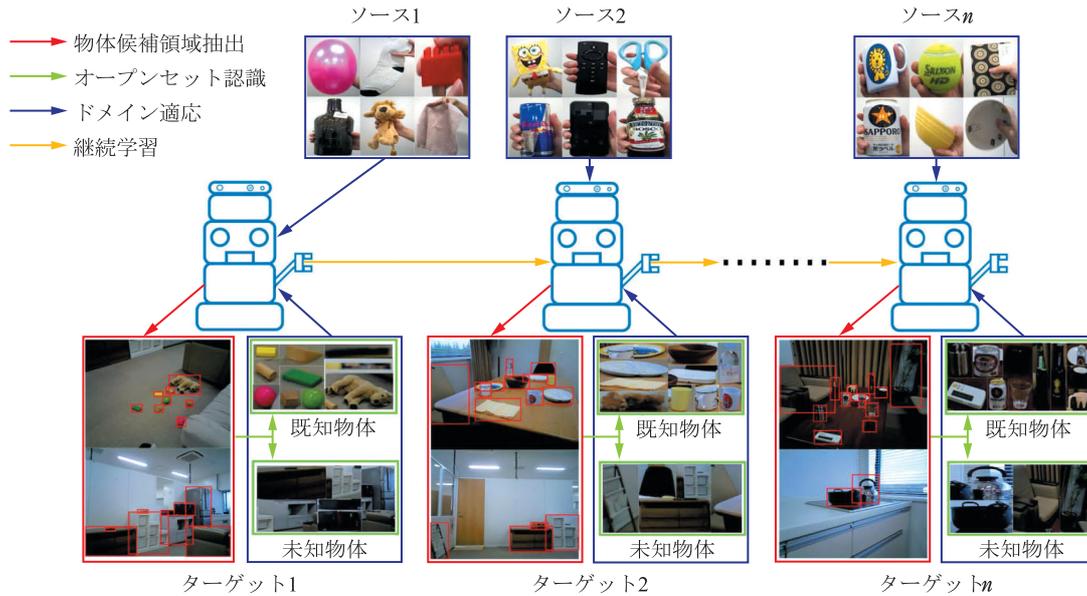


図7 家庭用ロボットにおける継続的オープンセットドメイン適応 (文献(38)から引用して和訳)

に近いものになっているといえる。

4. 実世界への応用へ向けて

本稿で見てきたように、機械学習は質的な進化が続いており、様々な観点から伝統的な枠組みを超えた試みがなされている。一方で、実世界のアプリケーションは、これまで紹介してきた技術や枠組みのどれかへ排他的に分類されるものではなく、実際には様々な要素が複雑に絡み合うことが多い。最初に述べた、家庭環境で活動するロボットの例をもう少し考えてみよう (図7)。既に議論したとおり、家庭環境に現れる全ての物事を教えることは現実的ではないためオープンセット認識は必須の機能であるし、新しい商品など次々に現れる新規物体を認識するために継続学習の機能も必要であろう。また、各家庭では照明等の条件や、物体の配置のされ方や背景等も異なることから、必然的に訓練時とはドメインの違いが発生するであろう。

このように、実際のアプリケーションにおいては複数の要素を同時に考慮する必要があるが、これは各要素技術の単純な組合せでは不十分であることが多い⁽³⁸⁾。例えば、オープンセット認識とドメイン適応の両立は非常に困難であることが知られている。これは、ターゲットに未知クラスのデータが存在する場合、一般的なドメイン適応手法で核となっている分布アラインメントの前提が崩れるためである。この例のように、複数の要素を並立させるためには本質的に新たな研究が必要であり、オープンセットドメイン適応⁽³⁹⁾、オープンセットドメイン一般化⁽⁴⁰⁾、継続的ドメイン適応⁽⁴¹⁾など様々な試み

が始まっている。

機械学習はまだまだ大きな可能性を秘めており、新しい課題や問題設定が次々に生まれている。本稿で紹介したような挑戦的な研究がより一層進展し、新たな価値創出へとつながっていくことを期待したい。

文 献

- (1) K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on imagenet classification," Proc. ICCV, pp. 1026-1034, 2015.
- (2) W. Xiong, L. Wu, F. Alleva, J. Droppo, X. Huang, and A. Stolcke, "The microsoft 2017 conversational speech recognition system," Technical Report MSR-TR-2017-39, 2017.
- (3) M. Popel, M. Tomkova, J. Tomek, L. Kaiser, J. Uszkoreit, O. Bojar, and Z. Žabokrtský, "Transforming machine translation: a deep learning system reaches news translation quality comparable to human professionals," Nature Communications, vol. 11, no. 1, pp. 1-15, 2020.
- (4) O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A.C. Berg, and L. Fei-Fei, "Imagenet large scale visual recognition challenge," Int. J. Comput. Vis., vol. 115, no. 3, pp. 211-252, 2015.
- (5) Z.H. Zhou, "A brief introduction to weakly supervised learning," National Science Review, vol. 5, no. 1, pp. 44-53, 2018.
- (6) Semi-Supervised Learning, O. Chapelle, B. Schölkopf, and A. Zien, eds., Cambridge: MIT Press, 2008.
- (7) D.-H. Lee, "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks," ICML 2013 Workshop: Challenges in Representation Learning, pp. 1-6, 2013.
- (8) B. Settles, "Active learning literature survey," Computer Sciences Technical Report 1648, University of Wisconsin-Madison, 2009.
- (9) C. Elkan and K. Noto, "Learning classifiers from only positive and unlabeled data," Proc. KDD, pp. 213-220, 2008.
- (10) F. Herrera, S. Ventura, R. Bello, C. Cornelis, A. Zafra, D. Sánchez-Tarragó, and S. Vluymans, Multiple instance learning: foundations and algorithms, Springer, 2016.
- (11) L. Chan, M.S. Hosseini, and K.N. Plataniotis, "A comprehensive analysis of weakly-supervised semantic segmentation in different

- image domains,” *Int. J. Comput. Vis.*, vol. 129, no. 2, pp. 361-384, 2021.
- (12) M. Everingham, S.M.A. Eslami, L. Van Gool, C.K.I. Williams, J. Winn, and A. Zisserman, “The pascal visual object classes challenge : A retrospective,” *Int. J. Comput. Vis.*, vol. 111, pp. 98-136, 2014.
- (13) S.J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345-1359, 2010.
- (14) F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, “A comprehensive survey on transfer learning,” *Proc. IEEE*, vol. 109, no. 1, pp. 43-76, 2021.
- (15) 神島敏弘, “転移学習,” *人工知能誌*, vol. 25, no. 4, pp. 572-580, 2010.
- (16) 中山英樹, “深層畳み込みニューラルネットワークによる画像特徴抽出と転移学習,” *信学技報*, SP2015-45, pp. 55-59, July 2015.
- (17) P. Agrawal, R. Girshick, and J. Malik, “Analyzing the performance of multilayer neural networks for object recognition,” *Proc. ECCV, Lect. Note Comput. Sci.*, vol. 8695, pp. 329-344, 2014.
- (18) J. Devlin, M.W. Chang, K. Lee, and K. Toutanova, “BERT : Pre-training of deep bidirectional transformers for language understanding,” *Proc. NAACL*, pp. 4171-4186, 2019.
- (19) T.B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D.M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, “Language models are few-shot learners,” *Proc. NeurIPS*, 2020.
- (20) T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, “A simple framework for contrastive learning of visual representations,” *Proc. ICML*, no. PMLR 119, pp. 1-11, 2020.
- (21) S. Ruder, “An overview of multi-task learning in deep neural networks,” *arXiv preprint arXiv : 1706.05098*, 2017.
- (22) S.R. Richter, V. Vineet, S. Roth, and V. Koltun, “Playing for data : Ground truth from computer games,” *Proc. ECCV, Lect. Note Comput. Sci.*, vol. 9906, pp. 102-118, 2016.
- (23) E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, and T. Darrell, “Deep domain confusion : Maximizing for domain invariance,” *arXiv preprint arXiv : 2003.04382*, 2014.
- (24) Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, “Domain-adversarial training of neural networks,” *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 1-35, 2016.
- (25) K. Muandet, D. Balduzzi, and B. Schölkopf, “Domain generalization via invariant feature representation,” *Proc. ICML*, vol. 28, pp. 10-18, 2013.
- (26) C.H. Lampert, H. Nickisch, and S. Harmeling, “Attribute-based classification for zero-shot visual object categorization,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 3, pp. 453-465, 2014.
- (27) R. Socher, M. Ganjoo, C.D. Manning, and A.Y. Ng, “Zero-shot learning through cross-modal transfer,” *Proc. NIPS*, pp. 935-943, 2013.
- (28) S. Bujwid and J. Sullivan, “Large-scale zero-shot image classification from rich and diverse textual descriptions,” *Proc. the Third Workshop on Beyond Vision and LANGUAGE : inTEgrating Real-world kNowledge (LANTERN)*, pp. 38-52, 2021.
- (29) N.V. Nayak and S.H. Bach, “Zero-shot learning with common sense knowledge graphs,” *arXiv preprint arXiv : 2006.10713*, 2020.
- (30) A. Bendale and T. Boulton, “Towards open world recognition related work,” *Proc. CVPR*, pp. 1893-1902, 2015.
- (31) C. Geng, S.J. Huang, and S. Chen, “Recent advances in open set recognition : A survey,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 10, pp. 3614-3631, 2021.
- (32) 中山英樹, “オープンワールド認識理解のための機械学習と評価,” *日本ロボット学会誌*, vol. 34, no. 6, pp. 382-385, 2016.
- (33) A. Bendale and T.E. Boulton, “Towards open set deep networks,” *Proc. CVPR*, pp. 1563-1572, 2016.
- (34) M. Delange, R. Aljundi, M. Masana, S. Parisot, X. Jia, A. Leonardis, G. Slabaugh, and T. Tuytelaars, “A continual learning survey : Defying forgetting in classification tasks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, no. c, pp. 1-29, 2021.
- (35) J. Kirkpatrick, R. Pascanu, N. Rabinowitz, J. Veness, G. Desjardins, A.A. Rusu, K. Milan, J. Quan, T. Ramalho, A.G. Barwinska, D. Hassabis, C. Clopath, D. Kumaran, and R. Hadsell, “Overcoming catastrophic forgetting in neural networks,” *Proc. Natl. Acad. Sci. USA*, vol. 114, no. 13, pp. 3521-3526, 2017.
- (36) H. Shin, J.K. Lee, J. Kim, and J. Kim, “Continual learning with deep generative replay,” *Proc. NIPS*, pp. 1-10, 2017.
- (37) S.-A. Rebuff, A. Kolesnikov, G. Sperl, and C. Lampert, “iCaRL : Incremental classifier and representation learning,” *Proc. CVPR*, pp. 2001-2010, 2017.
- (38) I. Kishida, H. Chen, M. Baba, J. Jin, A. Amma, and H. Nakayama, “Object recognition with continual open set domain adaptation for home robot,” *Proc. WACV*, pp. 1517-1526, 2021.
- (39) P.P. Busto and J. Gall, “Open set domain adaptation,” *Proc. ICCV*, pp. 754-763, 2017.
- (40) K. Katsumata, I. Kishida, A. Amma, and H. Nakayama, “Open-set domain generalization via metric learning,” *Proc. ICIP*, 2021.
- (41) Q. Lao, X. Jiang, M. Havaei, and Y. Bengio, “Continuous domain adaptation with variational domain-agnostic feature replay,” *arXiv preprint arXiv : 2003.04382*, 2020.

(2021年5月11日受付)



なかやま ひでき
中山 英樹 (正員)

2011 東大大学院情報理工学系研究科知能機械情報学専攻博士課程了。博士 (情報理工学)。2012 から同研究科創造情報学専攻講師。2018 から同准教授, 現在に至る。画像認識, 自然言語処理, 機械学習の研究に従事。情報処理学会, IEEE, ACM 各会員。