

機械学習と AI 技術

Machine Learning and Artificial Intelligence

米 倉 一 男 技術開発本部基盤技術研究所数理工学研究部 博士（情報理工学）
齊 藤 弘 樹 技術開発本部基盤技術研究所数理工学研究部
服 部 均 技術開発本部基盤技術研究所数理工学研究部

近年、深層学習に代表される AI 技術が多くの技術的ブレークスルーを果たしている。最初は画像認識の精度向上から始まり、囲碁のトップ棋士を打ち負かす AI も登場し、現在では AI 技術を基に多くのビジネスが生まれている。本稿ではまず、近年の AI 技術について深層学習を中心に概観するとともに、社会的な要請が高まっている AI によるサービスを提供する者が負うべき説明責任について紹介する。当社では AI 技術を性能予測や形状生成などに活用する研究を行っており、その活用事例についても紹介する。

An emerging AI technology is spreading its capability. The deep learning, which is the most promising AI technique, has shown a surprising success in image processing. Also, a top Go player has been beaten by AlphaGo. Since then, the AI technologies have been applied to in industrial fields, and succeed in business. The present article provides a brief overview of AI technologies, and introduces accountability for AI which is based on high social demands. We apply AI technologies to various fields including performance prediction and shape generations in IHI. We briefly describe the applications as well in the present article.

1. 緒 言

2010年代前半から、人工知能（AI：Artificial Intelligence）の話題がメディアをにぎわすようになった。IBM社（アメリカ）の Watson がクイズ番組で優勝し、Google傘下の DeepMind社（イギリス）が開発した AlphaGo がプロの囲碁棋士に勝つなど、特定の領域で人間をしのぐ AI が出現し、話題に拍車をかけている。このような AI の基礎になっているのが機械学習と呼ばれる技術である。機械学習とは、ガウス過程回帰（GP：Gaussian Process）やサポートベクターマシン（SVM：Support Vector Machine）に加え、近年、急激に注目されている深層学習や強化学習（RL：Reinforcement Learning）などを含んだ技術であり、データから何かを学習することを指している。その目的は、一般に重要なパターンやトレンドを見つけ出し、データが何を言っているのかを理解すること⁽¹⁾とされる。この機械学習で見つけ出したパターンを用いることで、AIによる画像認識や文章理解などの動作が可能になる。

近年の AI 技術の成功は、深層学習の成功がきっかけになっているといつてよい。特に一般画像認識と呼ばれる、画像に何が映っているかを判断するタスクにおいて、深層学習は驚くべき精度を実現し一躍注目を浴びた。その後研究が加速度的に進み、現在では、多様なアプリケーション

に使用されている。松尾⁽²⁾はこのような不連続な技術的変化を、太古の昔に生物が眼を持ったことがきっかけになって多様な生物が生まれたカンブリア大爆発になぞらえ、機械が眼を持ったと表現している。

このように AI は、一連の機械学習技術のブレークスルーに伴って、急に注目を集めて社会実装が進んでいるため、それに対する法整備や AI の社会的な影響に対するコンセンサスの形成が遅れていた。しかし 2018 年 5 月から、ヨーロッパで EU 一般データ保護規則（GDPR：General Data Protection Regulation）の適用が開始され、徐々にデータの利活用と保護および AI に関する国際的なルール作成が進んでいる。

これらの AI 技術は幅広い分野に適用できるものの、学術界では画像認識を中心に研究が進んでおり、機械メーカーのニーズにすぐさま適用できないものも多いため、当社は AI 技術を社内適用するための研究を進めている。本稿では AI 技術の変遷と、説明責任に関する情勢を紹介したうえで、これらの技術を社内活用している事例を紹介する。

2. AI 技術の変遷

2.1 画像認識

画像に何が映っているかを判断するタスクは深層学習以前から取り組まれていたが、間違い率が 25～30%であつ

た。これに対して 2012 年に、Hinton らのグループ⁽³⁾は深層学習を用いて間違い率を一挙に 16%にまで低減させ、さらに 2017 年現在は 2.3%⁽⁴⁾にまで低下している。第 1 図に画像認識の一例を示す。当初は深層学習がほかの機械学習よりも優れている理由が分からない面もあったが、近年はその理論的な裏付けと理解が徐々に進んでいる⁽⁵⁾。

画像認識では、多くの場合に畳込みニューラルネットワーク (CNN : Convolutional Neural Network) が用いられる。深層学習が登場する前は、物体のエッジの検出など (特徴量抽出と呼ぶ) のプログラムを作成する必要があり、特徴量抽出の巧拙が精度を左右していた。これに対し、深層学習を用いることで最適な特徴量を自動で抽出できるようになったことが、大幅な精度向上の一因である。

2.2 画像生成

現実にはない画像を自動で生成する技術も登場している。例えば、ゴッホやモネの画風を学習させて家の画像と合成すれば、現実に存在しないゴッホ風の家の絵を描くことができる⁽⁶⁾。さらに最近では、敵対的生成ネットワーク (GAN : Generative Adversarial Network) と呼ばれる手法を用いて、写真のように鮮明な偽の画像を生成することが可能であり、人間が一目見ただけでは真贋が区別できないほどの精度が得られている (第 2 図)⁽⁷⁾。これらの技術は白黒写真や線画の自動彩色に応用されたり、低解像度画像を高解像度化する超解像度技術 (Hyper-Resolution) として活用され始めたりしている。また一方で、偽の動画と音声を生成し、あたかも政府要人の発言のような動画が作成されたこともあり、悪質なフェイクニュースとして問題になった。

2.3 強化学習

強化学習は、少なくとも 1990 年以前から存在する手法で、以前から制御などに使用されてきた。物理的なルール



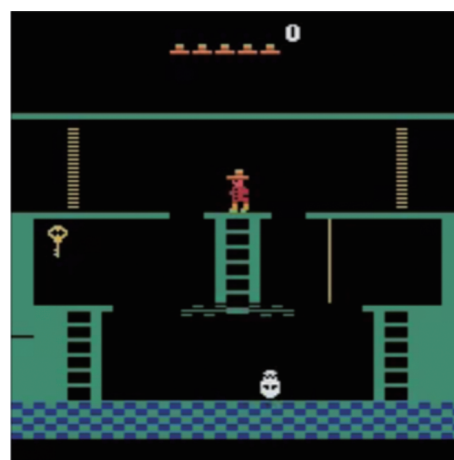
第 1 図 画像認識の一例⁽⁴⁾
Fig. 1 Example of image recognition⁽⁴⁾



第 2 図 GAN を用いて生成された実在しない人の顔⁽⁷⁾
Fig. 2 AI-generated face images⁽⁷⁾

や制御ロジックを明示的に記述しなくても、妥当な制御則が得られることが特徴である。この強化学習を前述の深層学習と組み合わせることで、従来は不可能だったタスクが可能になった。例えば、AI にテレビゲームをプレイさせる場合 (第 3 図)⁽⁸⁾、従来は画面内の情報を適切に処理することが困難だったが、深層学習を用いれば動画情報の処理が容易になるため、これを強化学習と組み合わせることで人間をはるかにしのぐ点数を獲得する AI が開発された⁽⁹⁾。

その後、この研究は AlphaGo へとつながり、囲碁棋士を打ち負かす AI が誕生することになる。AlphaGo はトップ棋士であるイ・セドル九段との第二局で、それまでの常識を覆す一手、黒 37 手目を指して棋界を驚かせた⁽¹⁰⁾。AlphaGo の棋譜には常識を覆すものも多いが、それらは同時に理にかなったものでもあり、新しい定石が幾つも生まれ始めている。また将棋の世界でも AI の性能は向上しており、AI と対局する電王戦が開催され、AI を将棋の研究に用いる棋士が増えるなど、AI の登場で将棋界が急激に変化している⁽¹¹⁾。



第 3 図 強化学習でプレイされた Montezuma's Revenge というレトロゲーム⁽⁸⁾
Fig. 3 Snapshot of "Montezuma's Revenge" by RL⁽⁸⁾

2.4 言語の分析と自然言語処理

文章の理解と処理は、昔から自然言語処理という分野で研究が続けられてきた。これは 2011 年 2 月に IBM 社が開発した Watson が、アメリカのクイズ番組である Jeopardy! に参加し、前年と前々年の全米王者に大差で勝利したことが有名である⁽¹²⁾。膨大な文献データベースにアクセスして情報を処理し、巨大な知識データベースを作成できるようになったことがブレイクスルーにつながった。Google 社も同様に大量のデータに基づいた翻訳技術を開発し、Google 翻訳として使用している。2016 年秋ごろに実装された際、翻訳の精度が劇的に向上して話題になった。また、質問者が何を聞いているかを分析する技術は、SNS (Social Networking Service) やホームページなどで自動応答するチャットボットなどとして応用されている。

文章理解に関しては、大学共同利用機関法人 情報・システム研究機構 国立情報学研究所 (NII) が中心となって AI による東京大学への合格を目指す東ロボくんプロジェクト⁽¹³⁾があった。このプロジェクトは科目ごとに異なる AI を集めたもので、数学や世界史などでは偏差値 60 を超える性能が得られたものの、読解力に関しては技術的なハードルがあるとされ、2016 年に開発は凍結された。一方で 2018 年に Google 社が開発した BERT⁽¹⁴⁾は、読解力テストで人間のスコアを上回る結果を残している。このように読解力についても長足の進歩が見られており、今後さらに発展していくと考えられる。

3. AI の説明責任

AI はコンピュータプログラムであるため、理屈のうえでは動作を明示的に記述することが可能である。しかし特に深層学習は、複雑な記述が必要で人が理解しにくいという特徴がある。また、常に学習し続けるシステム (オンライン学習) では、AI の動作が常に変化するため、常に正常な動作を保証することは単純ではない。これが AI が

ブラックボックスであるといわれるゆえんである。この性質のために、人の想定外の動作をする場合がある。例えば第 4 図は、画像に特定のノイズを加えると別の画像と判定する例⁽¹⁵⁾である。また、第 5 図のように道路標識に特定の模様を加えることで AI が誤認識してしまう例⁽¹⁶⁾も報告されている。これらは AI の動作を理解するための研究の途上で得られた特殊な例であり、すぐさま AI の暴走や悪用につながるものではない。しかし、AI の動作を正確に理解することの重要性を示唆している。

また、言語を処理する AI を二つ用意して互いに会話させたところ、人に理解できない独自の言語を発明して会話し始めたことが報道された。この実験では目的に合わせて AI が言語を変更させることを意図しており、技術的にはその意図どおりに動作したものだが (そして Facebook 社 (アメリカ) は実験が失敗だったとしているが)、その報道はある種の不気味さを感じさせ、社会が AI を許容できるかどうかを考えさせるものである。また、2.2 節で紹介したように政府要人の会見動画を捏造する例も報道され、新たなフェイクニュースの形として問題視されている。

このように、AI 技術の進歩に伴って新たな課題がもたげられている。悪意の有無や事の大小はさまざまだが、AI が社会に影響を与え始めていることは論を俟たないだろう。これに対し、GDPR では AI の決定に対して説明

(a) 実際のいたずら書きされた標識 (b) AI が誤認識する標識
改変の例



第 5 図 画像認識をだます標識改変の例⁽¹⁶⁾
Fig. 5 Adversarial image modification⁽¹⁶⁾

(a) 正しく車と認識される画像 (b) 車と認識されない画像 (c) 両者の画素の差を表示した画像



第 4 図 画像にノイズを加えると正しく認識できなくなる例⁽¹⁵⁾
Fig. 4 Image recognition with perturbation⁽¹⁵⁾

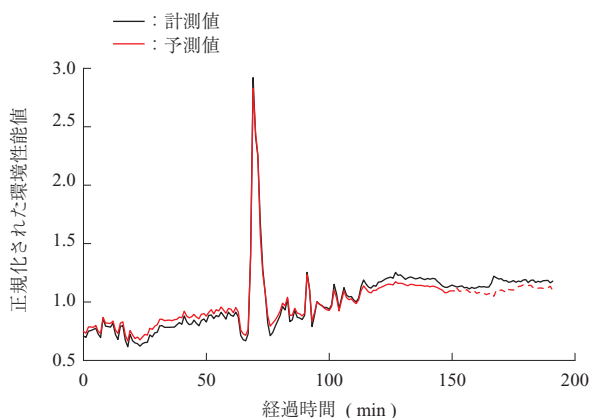
を受ける権利が明示された。これは AI を使用したサービスや AI が意思決定を行ったものに対し、サービスなどを受けた側が説明を求めることができるという権利である。例えば、ある社員が AI を使用した人事査定を受けた場合、社員は会社に査定の根拠の説明を求めることができる。同様に日本政府も、AI を用いた意思決定に対して公平性、説明責任及び透明性の原則を求める案⁽¹⁷⁾を提示しており、説明責任は社会的な要請となりつつある。このような社会的要請に応える形で、説明可能な AI (XAI: Explainable AI) の研究はさらに盛んになっている。

4. IHI での活用事例

4.1 性能予測への活用

機械学習技術の活用法の一つが、性能予測である。これまでに蓄積された機器の運転データを分析することで、異常検知や状態量の予測を行っている。例えば、深層学習を利用して車両用過給機の性能を、実験データベースを基に予測することができる。これにより試験をする前に性能値を予想することができ、効率的な試験計画が策定できる。

また第 6 図に、あるプラントの環境性能値を時々刻々と予測した例を示す。時間遅れなく、精度良く予測できていることが分かる。深層学習以外の機械学習手法との比較も実施したが、深層学習が最も精度良く予測することができた。プラントの運転データはセンサ数が数百～数千個と膨大で、さらに時系列に沿って数か月～数年のデータが蓄積されたビッグデータである。また、付帯設備の ON/OFF によって内部状態が大きく変わる場合がある。深層学習はこのような大規模データや、急激な状態変化の処理⁽⁵⁾に適していると考えられる。



第 6 図 プラントの環境性能の予測結果
Fig. 6 Performance prediction of a plant

4.2 設計への活用

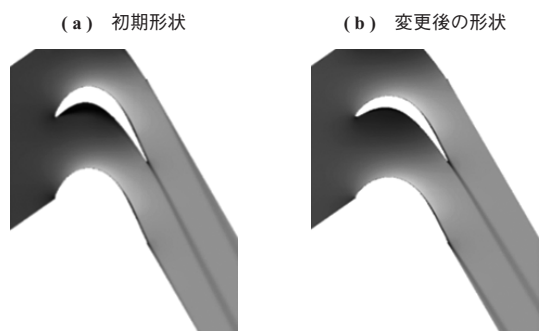
設計データから設計の特徴量を抽出して分析することも可能である⁽¹⁸⁾⁽¹⁹⁾。単なる分析にとどまらず、2.2 節で紹介した画像生成技術を設計データに援用し、新しい形状を得ることも可能になった。簡単な例として翼型とその揚力係数を学習させる。すると逆に所望の揚力係数を指定すれば、それに見合う形状が得られる。元のデータベースを分析しているため、突拍子もなく高性能な形状を得ることは難しいが、第 7 図に示すようにもっともらしい形状が得られる。当社でこれまでに設計した形状をデータベースとして学習させることで、IHI らしい形状を生成する AI が訓練できる。それに加え、文献などで公表されているデータも加えれば、IHI らしさに大学などの知見を加えた形状を創出できる可能性がある。本手法は検討のベースとなる設計解を得るために活用できると考えている。

設計された形状は、微妙な修正を加えてさらに性能を向上させる必要がある。例えば、翼形状の微妙な曲率を変えることで流れを滑らかにするなどである。この作業には強化学習を活用できる。第 8 図に強化学習を使用した形状修正の例を示す。強化学習は制御に使用されることが多いが、パラメータの最適化にも使用されることがある。強化学習の特徴は汎化性能と呼ばれ、一度学習すれば条件が変わっても、似た条件であれば簡単に最適化できることが特徴である。設計作業では似た条件での最適化を繰り返し行うことが多く、その場合に強化学習が有用と考えられる。



第 7 図 画像生成法を援用して得られた、指定した揚力係数を満たす翼形状

Fig. 7 Shapes generated by image generating methods which satisfies specified lift coefficient



第 8 図 強化学習を用いた翼形状の変更例
Fig. 8 Shape modification using RL

4.3 天気予報への活用

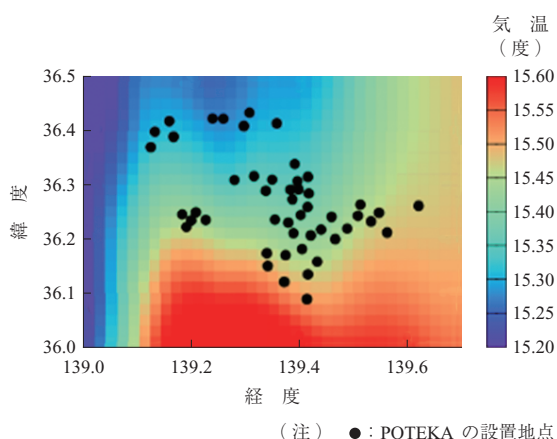
明星電気株式会社では超高密度気象観測情報提供サービスシステム POTEKA® を販売している。POTEKA は日本全国に設置され、場所によっては 2 ~ 3 km に 1 基設置されている。これだけの高密度な観測網はほかに例がなく、独特なデータといえる。この高密度データに対し、テンソル学習と呼ばれる深層学習の一手法を用いることで、任意の地点における天気を予測することができる⁽²⁰⁾。この手法を利用して気温を予測したものを第 9 図に示す。POTEKA の大規模データとテンソル学習の組み合わせにより、地理的にも時間的にも精度の良い予測が可能になった。

4.4 説明可能性への貢献

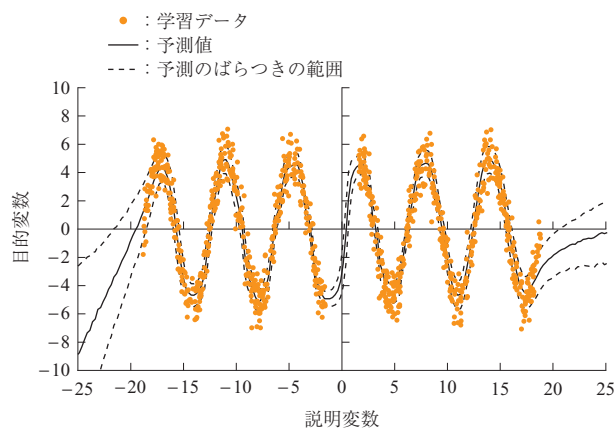
説明可能性は AI 活用における大きなトレンドである。前述した性能予測や天気予報への活用では、予測結果の不確実性を定量化することが有効と考えられる。深層学習においては Bayesian Neural Network (BNN)⁽²¹⁾ と呼ばれる手法を使用して不確実性を定量評価している。BNN では予測値を期待値と分散に分けて評価でき、分散が大きい場合は不確実性が高く、使用する場合には注意が必要と判断できる。また分散を使用すれば、予想が外れた場合の最悪ケースの値も推測できるため、意思決定の材料として活用できる。第 10 図に例を示す。データの少ない領域ではばらつきが大きくなる傾向が分かる。

5. 結 言

AI 技術の変遷を概観し、その技術が多様な発展を遂げていることを紹介し、それらの技術を社内でどのように活用しているかを紹介した。AI 技術は広い分野での活用が期待されているが、学術研究は画像認識に特化した研究が多い。当社では AI 技術を機械メーカーに適した形で取り入



第 9 図 テンソル学習による緯度経度ごとの気温の予測結果⁽²⁰⁾
Fig. 9 Prediction of temperature using Tensor learning⁽²⁰⁾



第 10 図 BNN による予測の不確実性の評価
Fig. 10 Present uncertainty of prediction by BNN

れる研究を行っている。今後、AI 技術が社会に浸透するにつれて、AI 技術を社会がどのように受け入れるべきかに関するコンセンサス形成が進むだろう。技術が社会的要請に応える形で進歩すると同時に、囲碁や将棋の世界が AI 技術によって変化したように、社会の側も技術に合わせた変化を余儀なくされるかもしれない。AI 技術については、技術的側面と社会的側面の両面を見据えた研究開発が必要である。

参 考 文 献

- (1) T. Hastie, R. Tibshirani and J. Friedman : 統計的学習の基礎 — データマイニング・推論・予測 —, 共立出版, 2014 年 6 月
- (2) 松尾 豊 : 人工知能は人間を超えるか ディープラーニングの先にあるもの, 角川 EPUB 選書, 2015 年 3 月
- (3) A. Krizhevsky, I. Sutskever and G. E. Hinton : ImageNet Classification with Deep Convolutional Neural Networks, Advances in Neural Information Processing Systems 25, (2012), pp. 1 097 - 1 105
- (4) IMAGENET : Large scale visual recognition challenge (ILSVRC) : http://www.image-net.org/challenges/talks_2017/ILSVRC2017_overview.pdf, (参照 2018-12-7)
- (5) M. Imaizumi and K. Fukumizu : Deep Neural Networks Learn Non-Smooth Functions Effectively, arXiv Preprint, (2018. 7), arXiv:1802.04474
- (6) L. A. Gatys, A. S. Ecker and M. Bethge : A Neural Algorithm of Artistic Style, arXiv Preprint, (2015. 9), arXiv:1508.06576

- (7) T. Karras, T. Aila, S. Laine and J. Lehtinen : Progressive Growing of GANs for Improved Quality, Stability, and Variation, arXiv Preprint, (2018. 2), arXiv:1710.10196.
- (8) Y. Aytar, T. Pfaff, D. Budden, T. L. Paine, Z. Wang and N. de Freitas : Playing hard exploration games by watching YouTube, arXiv Preprint, (2018. 11), arXiv:1805.11592
- (9) V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra and M. Riedmiller : Playing Atari with Deep Reinforcement Learning, NIPS Deep Learning Workshop, (2013)
- (10) 黒 37 手と白 78 手 AlphaGo とイ・セドルが再定義した「未来」: WIRED, 2016 年 9 月 13 日, <https://wired.jp/special/2016/alphago-vs-sedol/>, (参照 2018-12-25)
- (11) 山中伸弥, 羽生善治 : 人間の未来 AI の未来, 講談社, 2018 年 2 月
- (12) Computer wins on 'Jeopardy!': Trivial, It's Not, The New York Times, (2011. 2. 16)
- (13) 新井紀子, 東中竜一郎 : 人工知能プロジェクト「ロボットは東大に入れるか」第三次 AI ブームの到達点と限界, 東京大学出版会, 2018 年 9 月
- (14) J. Devlin, M-W. Chang, K. Lee and K. Toutanova : BERT : Pre-training of Deep Bidirectional Transformers for Language Understanding, arXiv Preprint, (2018), arXiv:1810.04805
- (15) C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow and R. Fergus : Intriguing properties of neural networks, arXiv Preprint, arXiv:1312.6199
- (16) K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno and D. Song : Robust Physical-World Attacks on Deep Learning Visual Classification, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2018), pp. 1 625 – 1 634
- (17) 人間中心の AI 社会原則 (案) : 内閣府, https://www8.cao.go.jp/cstp/tyousakai/humanai/ai_gensoku.pdf, (参照 2018-12-27)
- (18) 渡邊 修, 米倉一男, 久保世志 : 主成分分析を用いた形状表現法と形状最適化に関する研究, IHI 技報, Vol. 53, No. 3, 2013 年 10 月, pp. 76 – 81
- (19) K. Yonekura and O. Watanabe : A shape parameterization method using principal component analysis in applications to parametric shape optimization, ASME Journal of Mechanical Design, Vol. 136, Iss. 12, (2014. 10)
- (20) K. Yonekura, H. Hattori and T. Suzuki : Short-term local weather forecast using dense weather station by deep neural network, 2018 IEEE International Conference on Big Data (Big Data), (2018), pp. 1 683 – 1 690
- (21) Y. Gal : Uncertainty in Deep Learning, PhD Thesis, (2016)