SE Report

連載 第3回

AIと安全工学

―実社会でのAIの活用―

横浜国立大学　准教授

杉本　千佳 Chika Sugimoto

1. はじめに

近年AIの活用が急速に進み、身近なところにAI技術を搭載した製品やサービスが増えている。例えば、スマホのロック解除のために用いられる顔認証システムがある。機械学習を導入したAppleのFace IDは、TrueDepthカメラにより何千以上もの目に見えない赤外線ドットを顔に投射し、それを赤外線カメラで撮影、顔の凸凹などの深度情報を取得して顔の3Dモデルを構築し、登録済みの顔のデータと照合する[1)]。スマホにはニューラルエンジンが搭載され、今年になりマスク着用での顔認証にも対応している。スマホのセキュリティ対策は個人情報を保護する上で非常に重要である。Face IDの誤認証率はおよそ100万分の1といわれており、暗号化やプライバシー保護のための対策が取られている。こうした顔認証技術の研究開発は長い間行われてきた。顔認証では、カメラで取得される画像の中から人間の顔がどこにあるかを検出し、瞳中心、鼻翼、口端など目、鼻、口の特徴点の位置や顔領域の位置や大きさを求め、それらをもとに事前に登録した画像データと照合して本人かどうかを識別する。近年では高度な深層学習の導入により、認証精度と検索速度を大幅に向上させている。顔認証は入退室管理や決済など様々な用途で用いられており、セキュリティ対策においてAIは重要な技術となっている。その他の例としては、コロナ禍において店舗の入り口などに設置されるようになった非接触式の体温測定装置がある。こうした装置は、AI技術を活用して画像に映る顔を認識して額部を検出し、得られた表面温度分布データの中から額の皮膚温度を表示している（**図1**）。その他にも、AIには音声認識技術、音声合成技術、画像処理技術、文字認識技術、自然言語処理技術、情報検索技術などの様々な技術があり（**表1**）、機能に応じたAIが活用されている。このように、近年では意識せずにAI技術が搭載されたシステムを利用していることが多い。

一方で、日本の企業におけるAI導入状況は、中国・米国・欧州主要国を下回っている。情報処理推進機構（IPA）が公開した「DX白書2021」[2)]によると、日本企業のAI技術の活用状況は20.5％で、米国企業の44.2％と比較し半分以下となっている。前年の調査結果

図1　非接触式体温測定装置

表 1　AI 技術の活用例

AI 技術	活用例
音声認識	AI スピーカー、Pepper、ポケトーク（通訳機）
音声合成	AI スピーカー、Pepper、ボーカロイド
物体検知	自動運転車、ドローン、Pepper、監視モニタ
顔認識	Face ID、AI カメラ、顔認証カメラ、Pepper
文字認識	AI-OCR、Google 翻訳
自然言語処理	AI スピーカー、Google 翻訳、Pepper、Watson
情報検索	チャットボット、全文検索システム、AI スピーカー、Pepper

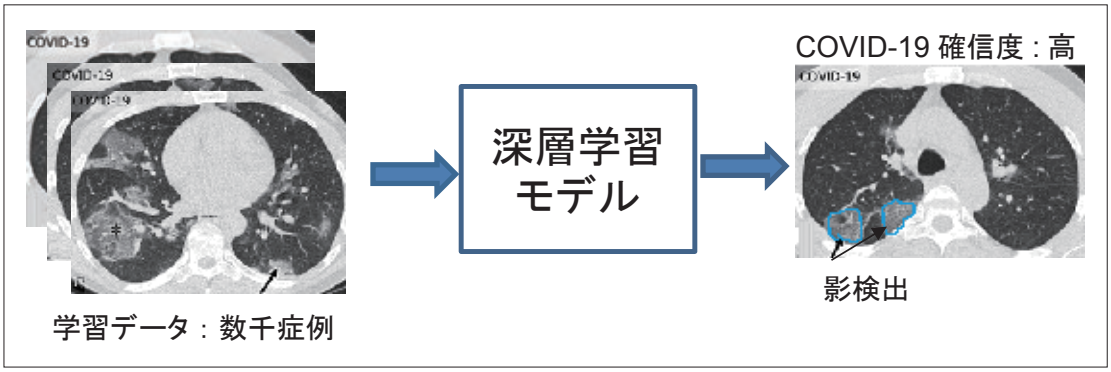
4.2％と比較すると約5倍に増加し、AI の利用は急速に拡大しているといえるが、AI 人材の不足が深刻化し導入課題となっている。また、全くの AI 未導入企業の割合が米国では7％であるのに対し、日本では33％と高くなっている。米国は AI を積極的に活用し、データドリブンな経営により意思決定の改善やより良い顧客体験の創出を行い、製品やサービスの革新や業務効率の向上を進め、AI 投資からリターンを得ている。これに対し、日本では金融・保険業、製造業などの一部の産業分野で AI 導入が進められ、業務の効率化や省人化、生産性の向上、新たな製品やサービスの創出の一躍を担っているが、いまだ限定的である。DX（デジタルトランスフォーメーション）推進の旗が振られる昨今の状況の中で、それを支える技術として AI が期待される。

そこで、第3回目の本稿では、様々な分野での AI 技術の活用について述べ、その中での安全工学の役割に言及する。

2. 医療分野での AI 活用

高齢化や医療費増大が進む日本の医療において、業務の効率化や医師への支援は急務であり、AI 活用が期待される分野である。画像診断支援、手術支援ロボット、診断・治療支援、ゲノム医療、医薬品開発、見守り・介護支援など、保健医療分野では政府が中心となり AI の本格導入に向けた検討が進められている。とりわけ、AI を活用した医用画像解析・診断支援の研究開発は早くから行われており、内視鏡画像解析 AI[3] や COVID-19肺炎画像解析プログラム[4] などの AI 技術を利用した医療機器プログラムが実用化されている。COVID-19肺炎患者の胸部 CT 画像には、淡いすりガラス影・網状影・浸潤影などが特徴的に見られるとの報告がある[5]。コロナ禍では医療の逼迫が深刻化し、感染有無を確認するために効率的な検査や画像診断を支援するソリューションが求められた。このため、世界中で短期間に COVID-19肺炎画像解析 AI が開発された（**図2**）。画像認識は AI が得意とするところであり、検査や治療において取得された CT や MRI などの大量の画像データを学習させることで、専門医と同等以上の判定精度をもつ AI が開発されている。これらの AI が搭載されたコンピュータ支援診断システムが医師の読影作業の効率化や診断の質の向上をサポートする。このように、蓄積されてきた診断画像や生体信号データ、健康診断の数値、学術論文や症例データ報告などの医療データを AI 技術を活用して解析し、高い性能を持つ診断・治療支援システムを実用化しようと

図 2　COVID-19 肺炎画像解析 AI



する取り組みが進められている。一方で、医療は命を扱うものであり、AI による診断ミスが命の危険に直結することがあることから、導入には十分な検証と正しい理解が求められる。また、医療データはプライバシー性が高いため慎重な取り扱いが必要であり、AI 診断結果の利用には倫理的問題をはらむケースがあることにも留意する必要がある。

3. 産業分野での AI 活用

生産現場では自動化に伴いロボットの導入が進められているが、AI を搭載したロボットの導入も進んでいる。近年は産業用ロボットが部品等の対象物を認識し掴んで移動させるピッキング作業や不良品等の選別作業を担うようになっているが、そのロボットの眼として AI の画像認識能力が活用されている。また、移動を伴うピッキング作業においてロボットが工場内で最適な移動動線をリアルタイム制御するのにも強化学習による AI が活用される。計算機性能が向上し非常に高速な処理能力を持つことから、人間を上回る作業効率を達成できる。人手不足により選別や診断に熟練が必要な作業の人材確保が難しくなる中、知能化したロボットは有用な代替手段となる。こうした現場では同じような作業を反復して行うため、成功と失敗の評価を与えて強化学習により繰り返し学習させることで精度を向上させることができることから、AI 活用に適している。一方で、人とロボットが共存して働く場合には人の安全性確保が重要であり、AI の振る舞いを人間がコントロールできるようにし、人間と協調した作業環境を構築することが必要である。

また、交通分野では自動運転技術で AI が大きな役割を担っている。自動運転では、車自体が道路上の車線や信号、周辺車両や歩行者等の移動体、運転に影響を及ぼす障害物などの周辺環境を正確かつ瞬時に認識することが必要である。この重要な認知プロセスにおいて AI が利用されている。車が走行する環境は、道路環境、人間環境、自然環境の組み合わせにおいて無限のパターンがある。車や人間だけではなく犬や猫が急に飛び出すこともあれば、信号機や車の故障、大きな地震が来て道路に影響が出ることもあり得る。また、霧や雪で視界や道路状況が悪くなることもある。医療分野と同様に、自動運転でのエラーは事故につながり人命を危険にさらしかねない。よって、自動運転車の導入は自動運転レベルを設定して段階的に進められており、導入により起こり得るリスクや課題への十分な対策が必要とされている。

また、情報化社会においてはセキュリティ対策が非常に重要になっている。いずれのシステムでも、不正アクセスにより情報が危険にさらされたり、誤作動を招いたりする可能

性がある。様々なシステムがインターネットに接続される中、AI 導入により人手による作業を無くし自動化した場合でも、システムやサーバーの平常状態を AI に学習させておき、常に状態を分析して平常状態とは異なる異常をリアルタイムに検知し、即座に対応できるようにしておくことが必要である。不正アクセスやサイバー攻撃は後を絶たないことから、フィッシング詐欺などの不正な手口を早期に発見する不正検知のための AI 導入も進んでいる。

4. 安全工学の役割

実社会への AI の導入が進むにつれ、AI 技術を用いたシステムやサービスの安全性検証の必要性が高まり、今後安全工学でも AI のアルゴリズムで動作する機器などの安全性を評価するケースが増えるだろう。また、近年はバーチャルの世界も広がっており、大きな注目を集めるメタバースにも多くの AI 技術が活用されている。

大量のデータを学習して構築される AI 搭載システムは、過去から現在までに蓄積されたノウハウを反映し高い性能を持つと期待される。一方で、そのデータに偏りがあり、使用する環境においてタスクについての十分な学習ができていなければ、適切な結果を得ることはできない可能性がある。また、安全検査や性能評価は従来そのサービスやシステムが使用される想定条件下で行われる。実際の製品開発では安全率に対してある程度の余裕をもって設計されるが、近年の自然現象による災害や人為的な要因を含む事故例を見ると、想定をはるかに超える事象が起こっており、不確実性が高まっている。こうした中で、AI の評価においてどこまで安全率を見るかは難しい問題である。人命に関わるものに対して安全率は大きめに取られ、十分に余裕を持った設計がなされるべきであるが、実用化の観点からは効率の低下やコストの上昇は出来るだけ抑えることが求められる。自動運転車のようにシステムエラーが起きた時のリスクが高いものは、常に安全性を確保するためにフェールセーフ機能が設けられるだろう。AI の活用を進めていくためには、AI のリスクや課題を整理し社会実装における基準を設けることが必要であり、こうした基準作りに安全工学の果たす役割は大きい。

5. まとめ

近年の AI 技術の進歩は著しいが、実用的には AI はデータが豊富な限定された分野での機能にとどまっている。しかし、データ共有やデータのオープン化が進む中で、これまで入手できなかったデータを活用できるようになり、さらに幅広い分野において AI 導入が進むことが期待される。こうした動きを後押しするためにも、安全工学が AI に関する安全の指針や基準を策定し、評価できるような仕組みを作っていくことが望まれる。

参考文献

1) Apple：先進の Face ID テクノロジーについて，2021. https://support.apple.com/ja-jp/HT208108
2) （独）情報処理推進機構：DX 白書2021エグゼクティブサマリー，14，2021. https://www.ipa.go.jp/ikc/publish/dx_hakusho.html
3) 国立研究開発法人科学技術振興機構：内視鏡 AI 診断支援医療機器ソフトウェア，2021. https://www.jst.go.jp/pr/announce/20210112/pdf/20210112.pdf
4) （独）医薬品医療機器総合機構：COVID-19肺炎画像解析プログラム FS-AI693型，承認番号：30300BZX00145000，2021.
5) S. Simpson et al.,“Radiological Society of North America Expert Consensus Document on Reporting Chest CT Findings Related to COVID-19：Endorsed by the Society of Thoracic Radiology，the American College of Radiology，and RSNA”，Radiology：Cardiothoracic Imaging，Vol.2，Number 2，2020.

すぎもと●ちか

生体計測工学、知覚情報処理、アフェクティブ・コンピューティングなどの研究分野において、生体・行動・環境情報の認識とその応用に関する研究に従事。東京大学大学院新領域創成科学研究科助教を経て、2010年から横浜国立大学大学院工学研究院准教授。

表 2　用語説明

	説明
1. 機械学習	AI を支える技術の1つ。コンピューターがデータから反復的に学習し、そこに潜むパターンを見つけ出すこと。
2. TrueDepth カメラ	フロントカメラとその近くに配置されている赤外線カメラ、近接センサ、環境センサ、ドットプロジェクターなどのセンサ群の総称。機械学習をする超高性能なニューラルエンジンの働きにより、高度なトラッキング機能を持つ。
3. セキュリティ対策	不正アクセス、サイバー攻撃、ウイルス感染、個人情報の漏えいなどを防ぐためのセキュリティ脅威に対する対策。
4. データドリブン	収集・蓄積されたデータを分析し、その結果に基づいて様々な課題に対して判断・意思決定を行うこと。
5. デジタルトランスフォーメーション（DX）	進化したデジタル技術を浸透させることで人々の生活をより良いものへと変革することで、既存の価値観や枠組みを根底から覆すような革新的なイノベーションをもたらすもの。
6. 手術支援ロボット	低侵襲技術を用いて複雑な手術を可能とするため、コンピューター機能を備えたインタフェースを介在させ、精密な操作やアクセスが困難な場所でのアプローチを可能とする安全な手技を支援するロボット。代表例が「da Vinci（ダビンチ）」。
7. ゲノム医療	タンパク質の設計図である遺伝子を網羅的に調べ、遺伝子変異を明らかにすることにより、一人一人の体質や病状に合わせてより効率的・効果的に治療などを行う医療。
8. 内視鏡	口または鼻、肛門から先端にカメラのついた細い柔軟なチューブ（ビデオスコープ）を挿入し、先端に組み込まれた小型撮像素子からの映像を直接見ながら、食道・胃・十二指腸などの消化器官内部の検査や治療・処置を行うための医療機器。
9. CT（Computed Tomography）	体の周囲から X 線をあてて、体の中の吸収率の違いをコンピューターで処理し体の断面を画像化する技術。
10. MRI（Magnetic Resonance Imaging）	強い磁石と電磁波を使って強力な磁場が発生しているトンネル状の装置の中で、FM ラジオで用いられている周波数の電波を体にあて、体の内部の断面をさまざまな方向から画像化する技術。
11. 読影	レントゲンや CT、MRI、超音波などの検査によって得られた画像から病変や異常所見を読み、診断を下すこと。
12. ピッキング作業	伝票やリストなどに従って、保管されている必要な品物を集める（ピックアップする）作業のこと。
13. サイバー攻撃	PC やスマホなどの情報端末、サーバーや Web サイトなどのコンピューターシステムに対し、ネットワーク経由で情報の改ざんや漏えい、データの搾取、システムの破壊などを行うこと。
14. アルゴリズム	ある特定の問題を解いたり課題を解決したりするための処理手順や計算方法。
15. フィッシング詐欺	インターネットのユーザから経済的価値がある情報を奪うために行われる詐欺行為。
16. メタバース	コンピュータやコンピュータネットワークの中に構築された、現実世界とは異なる3次元の仮想空間やそこでのサービスのこと。