

三竹レポート：AI活用の現場から(2)

AI画像解析の大幅な精度向上の社会的インパクト

NTTコミュニケーションズ株式会社
アプリケーション&コンテンツサービス部

AI推進室長 三竹 保宏



本レポートでは、ビジネスの現場で実際にAIが活用されている事例を、そのビジネスインパクトや課題も含めてご紹介していく。今回は、Deep Learning（深層学習）によるAI画像解析を活用した実ビジネスの動向についてお伝えする。

Deep Learningによる AI画像解析精度の飛躍的な向上

現在は世界的に第3次AIブームの最中にあると言われているが、近年のAI技術のブレークスルーは画像解析の分野から起こっている。2012年、世界的な画像認識コンテストで、ヒントン教授率いるトロント大学が初参加で圧勝し、世界中のAI技術者に驚きを与えた。トロント大学は、コンピュータに大量のデータを学習させる機械学習の一つであるDeep Learningを活用することで高い精度の画像解析が可能となることをコンテストで証明した。その後、Deep Learningによる画像解析は、広く利用可能な基礎技術として、様々なユースケースで応用と実用が進んでいる。

簡単にDeep Learningによる画像解析の仕組みについて説明する。Deep Learningでは、人間の脳の信号伝達構造に似たニューラルネットワークで情報に重みづけをしながら信号伝達を行い、その過程で画像データの特徴となる成分を抽出する(図1)。

簡単なイメージ例で言うと、アルファベットの「A」の画像であれば、

上端の尖った部分と、中央部付近にある2カ所のT字型の結合部に特徴があるといったことを、AIがDeep Learningのアルゴリズムで自律的に捉えていく(図2)。

ここでは、人間が教えることなく、AIプログラムが自ら画像内の特徴を獲得するということが重要な点で、言い方を変えれば、大量の画像データをDeep Learningで学習させると、AIプログラムは人間が気付かない画像の特徴点まで獲得する可能性がある。こうして、AIが人間の眼を超える画像認識力を取得し、高精度な画像解析を行っていく。

AI画像解析の実ビジネスへの 適用事例の拡大

Deep Learningを活用した画像解析は、既に多くのビジネスに活用されている。

NTTグループの事例で言えば、自動車のドライブレコーダーの録画映像から、一時停止の標識で正しく停止しているかAIで判定する取り組みや、車載カメラで撮影した道路上のマンホールの鉄蓋の劣化度合いをAIで判定する取り組み、人の表情や身振りをカメラを通してAIが分析してコミュニケーション力を判定する取り組み等の多くの事例がこ

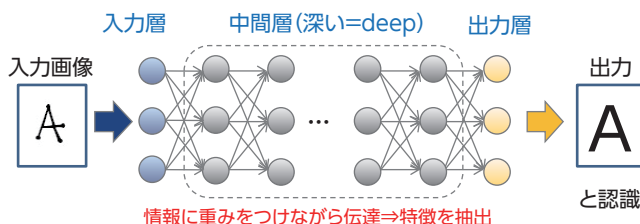


図1 ディープニューラルネットワーク (DNN) イメージ



図2 DNNによる画像の特徴点の抽出イメージ

れまでに発表されている。これらの事例は全て、Deep Learning で画像の特徴点を捉える仕組みを活用して実現されている。

NTT Com の最近の取り組みでは、人物の顔と全身の特徴を Deep Learning による画像解析で捉えて、リアルタイムで特定の人物を検知する実証実験を東京都内の大規模商業施設で行っている (図3)。

監視カメラの映像は人手 (目視) で確認されることが一般的であるが、AI を活用することで、目視をする人的稼働の削減と注意すべき映像の見逃しの削減に寄与することができる。NTT Com の AI 画像解析サービス『Takumi Eyes』は、あらかじめ同一人物を特定するための要素を学習済であるため、一枚の全身画像をもとにして、カメラに映る角度が違っていても高い精度で同一人物を検知することが可能である。

また、顔の画像が撮れるカメラで顔認証による人物検出を行い、顔が鮮明には映らない位置のカメラでは、全身の特徴で人物検出をする形で顔認証と全身照合を組合せて利用することによって、更に検出精度を上げることができる。

顔認証と全身照合を組み合わせたリアルタイムの AI 人物照合サービスは現在のところ他に例が無く、Takumi Eyes は、ショッピングモールなどの大規模施設内での不審者の追跡や、道に迷って戻れなくなった高齢者の探索、といったユースケースで実利用の検証が進んでいる。

AI 映像解析の実ビジネスでの活用における課題として、教師データ画像の収集や分析モデルの生成・

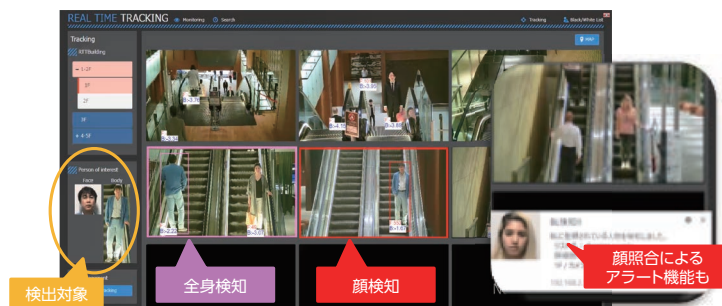


図3 AI人物解析「Takumi Eyes」リアルタイムトラッキング画面

チューニングにかかるコストの検証と、作成した分析モデルが実用に耐えうる精度が出るか、という導入効果の見極めが必要である。

Takumi Eyes の場合であれば、同じ人物を様々な角度や違う明るさで撮影した画像を教師データとして、同一人物を検出するモデルの生成とチューニングを行っている。教師データを収集/作成する手間はかかるが、一度学習済みの人物照合モデルが生成できれば、それを横展開できる (複数の案件で同じ分析モデルが活用できる) 利点がある。

教師データを新たに作成しなくてもよい事例もある。例えば、医療機関等で保存されているレントゲン画像データは、そのまま教師データとして活用できる可能性があり、AI 画像解析の医療診断への適用は、医師の診察をサポートする機能として実用化が進む分野であると期待されている。

一方で、教師データを揃えるのが難しい事例もある。製造業の工場で、傷がついた不良品を取り除く検品処理に AI 映像解析を活用しようとするような場合、傷のつき方には無限のバリエーションがあるため、教師データを揃えづらい。こういった場合には、傷のない正常な製品の画像を学習させて、正常な製品のみを検

知することで、検知から外れたものを「正常ではないもの=不良品」としてピックアップするような逆転の発想で AI 画像解析ソリューションが提供されている例もある。

新たに教師データを揃えることがビジネス展開上の課題となることも多いため、少ない教師データ画像から、多数のバリエーションを持つ教師データ画像を自動で生成する取り組みの研究開発も行われている。また、作成した分析モデルを実利用する中で、新たな教師データが増えていき、追加学習による分析モデルのチューニングが進む、という正のスパイラルが形成できれば、継続的な精度向上という面で理想的である。

AI 画像解析の適用領域は多岐に渡っており、ロボットや自動運転車両の眼となって周囲の情報の判別を行ったり、ドローン等に搭載することで、インフラの劣化検知や、農作物の生育状況の分析等へ活用する取り組みもある。これらの取り組みでは、音声情報やセンサー情報等の様々な情報と画像解析をマルチモーダル (複合的) に組み合わせることで、AI を活用したより高度な分析が実現されていくことが期待される。

< AI 活用のことなら下記へ >
ai-strategy-ac@ntt.com