

深層学習技術の価値化に向けた研究開発とそのビジネス応用

NTTではAI（人工知能）活用を重要な戦略の1つとして位置付けています。NTTソフトウェアイノベーションセンターではAIの中でも特に深層学習技術の価値化の研究開発に取り組んできました。本稿ではその中でもビジネス応用をする際に必要となるディープラーニング推論環境の最適化（高速化、軽量化、省力化）と、その技術を用いて実現される映像分析ビジネス（監視カメラ解析ビジネスなど）の検証活動について紹介します。

もりが くひろ えだ たけはる とやま まさし
森賀 邦広 /江田 毅晴 /外山 将司
みかみ けいた ひろかわ ゆたか やまだ ゆうじ
三上 啓太 /廣川 裕 /山田 佑二
むらまつ さなえ ささき たく
村松 沙那恵 /佐々木 琢
やまぐち しんや いなや かつお
山口 真弥 /稲家 克郎

NTTソフトウェアイノベーションセンター

「幻滅期」への突入をもうすぐ迎える深層学習

第三次AI（人工知能）ブームと呼ばれるこのブーム。ガートナー社が2018年8月に発表した「先進テクノロジーのハイプ・サイクル：2018年⁽¹⁾」では、2017年に続き2年連続で、「過度な期待のピーク期」に位置付けられました。それは、言い換えれば「幻滅期」の入り口にプロットされた、ともいえます。PoC（Proof of Concept）や先進的な企業が取り組んだ先行事例などの状況、ベストプラクティスが発表され、自ら深層学習技術を用いて企業の課題を解決することの難しさを企業の担当部署の方も感じているのではないのでしょうか。「過度な期待のピーク期」では、皆が想像し期待するような効果・効用は得られず、実際に担当された方はがっかりし、「幻滅期」に入っていく、それが今の状況です。しかし、その状況は、真のビジネス応用の始まりともいえます。今後は、実装や周辺技術が追い付き、徐々に現実のビジネスで採用されていきます。

そのビジネス応用の始まりを迎えている技術、それが、リアルタイムに人物に関する映像解析が可能なNTTソ

フトウェアイノベーションセンターの技術です。

複数の監視カメラの映像を高速に、そしてリアルタイムに検知

2011年より話題になり始めた深層学習は、コンピュータに「人の目と耳」を与えることに成功しました。そして、2018年の現在では、純粋な技術レベルでいえば、すでに人の目と耳の能力を超えたといっても過言ではありません。

その技術を映像分析技術としてパッケージにしたのが「リアルタイム人物トラッキング」です。施設に設置された大量の監視カメラの映像をリアルタイムに分析、ターゲット人物（不審者、VIP、要支援者、迷子等）を即時に検知、追跡することができるパッケージです。本パッケージは下記の機能の組合せで実現しています（図1）。

- ① 「人物抽出」：映像の中から人物のみを抽出する
- ② 「属性推定」：人物の性別や年齢層を推定する
- ③ 「詳細属性推定」：特定の体の部位に紐付いた属性推定。例えば、ロングヘアー、ホワイトシャツ、ブルージーンズ、サングラスなど

色や服装、アイテム（鞆の有無）などを含めた詳細な属性で人物を検索する

- ④ 「同一人物判定」：検出した人物が同一人物かを全身照合で判定する
- ⑤ 「軌跡推定」：人物が歩いた軌跡を映像より推定する
- ⑥ 「複数カメラ対応」：①～⑤を複数のカメラをまたいだ場合にも対応する
- ⑦ 「リアルタイム分析」：①～⑥をリアルタイムで分析可能とする

■他社に先駆け全身照合を実現

④の「同一人物判定」では、他社に先駆け「全身照合」での同一人物判定を実現し、人物が後ろ向きに写っていても人物を抽出できるようにしました。大量の人物画像ペアから深層学習技術を用いて特徴量を自動抽出することで、ルールベースで人間が設定した特徴（体型、服の色、髪型など）で判定するものに比べて高精度で照合することが可能となりました。

本技術の一部をサービス化したNTTコミュニケーションズの人物検索サービス「Takumi Eyes」が、第20回自動認識システム大賞を受賞⁽²⁾したことから、本技術が市場からも高い評価を

得ているといえます。

■パナソニックグループ×NTTグループ コラボレーションの成果として

全身照合も単体では完璧ではありません。全身照合は着ていたアウター（コート等）を脱ぐような外見の変化には対応することが困難です。この課題にも、パートナーとの協働で解決策を提示しました。

本プロジェクトは、2015年のパナソニック株式会社との業務提携⁽³⁾をきっかけとし、パナソニックの顔認識技術をうまく組み合わせることで精度の向上を大幅に図ることができました⁽⁴⁾。

このようにさまざまな角度、条件で

撮影されるカメラ映像の中から、高い精度で人物を照合できるようにしたものが「リアルタイム人物トラッキング」です。このように深層学習技術を用いた全身照合技術と顔認証のような技術を組み合わせるサービスはまだあまり類をみません。現時点では、全身照合+顔認証での同一人物判定ですが、お客さまのニーズの状況に応じて、追加で歩き方の特徴検出する機能を組み合わせるなど、要望に応じて最適な手法を選択することも可能な技術になっています。

2030年度には1600億円と予測される映像監視ビジネス市場

この深層学習技術を用いた「リアルタイム人物トラッキング」はどのようなビジネス活用の可能性があるのかをみてみましょう。

映像分析ビジネスは、AI市場の中でも一番高い成長率が予測されている有望市場です〔13億円（2015年実績）から、123倍の1600億円（2030年度）⁽⁵⁾〕。監視カメラで撮影された映像を分析する市場が大きくなるとの予想です。

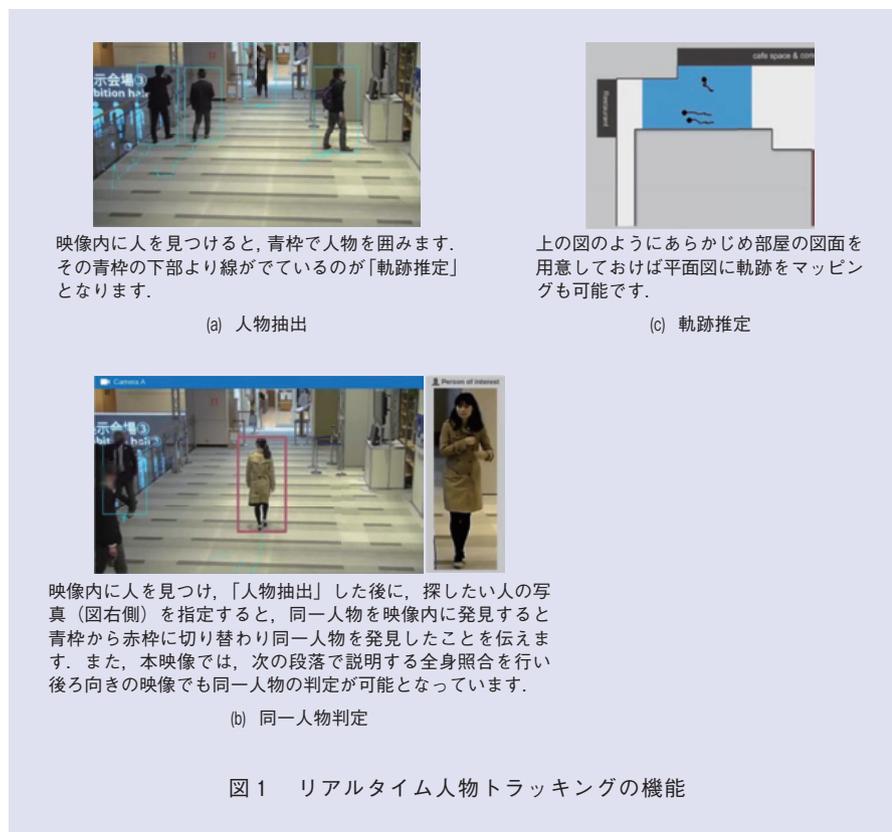
実施のユースケースをみてみましょう。

■コンビニの監視カメラ活用：同一人物判定

監視カメラの利用事例としてイメージしやすい事例がコンビニエンスストアの監視カメラ映像です。現在のコンビニでの監視カメラのユースケースは、何か事件・事故が起きた後に、実際に何が起きていたのかを“過去映像”として“人”が確認することです。こうしたときに便利な機能が「同一人物判定」です。

コンビニにて犯罪が起きた場合に、監視カメラの映像さえ残っていれば、犯行時の映像から犯人の映像を指定し、過去映像から検索することで、いつ入店したのかもすぐに見つけることができます。また、計画的な犯行か、衝動的な犯行なのか、過去に下見にきたなどの来店履歴もすぐに見つけることができます。

また、監視カメラが複数台あるよう



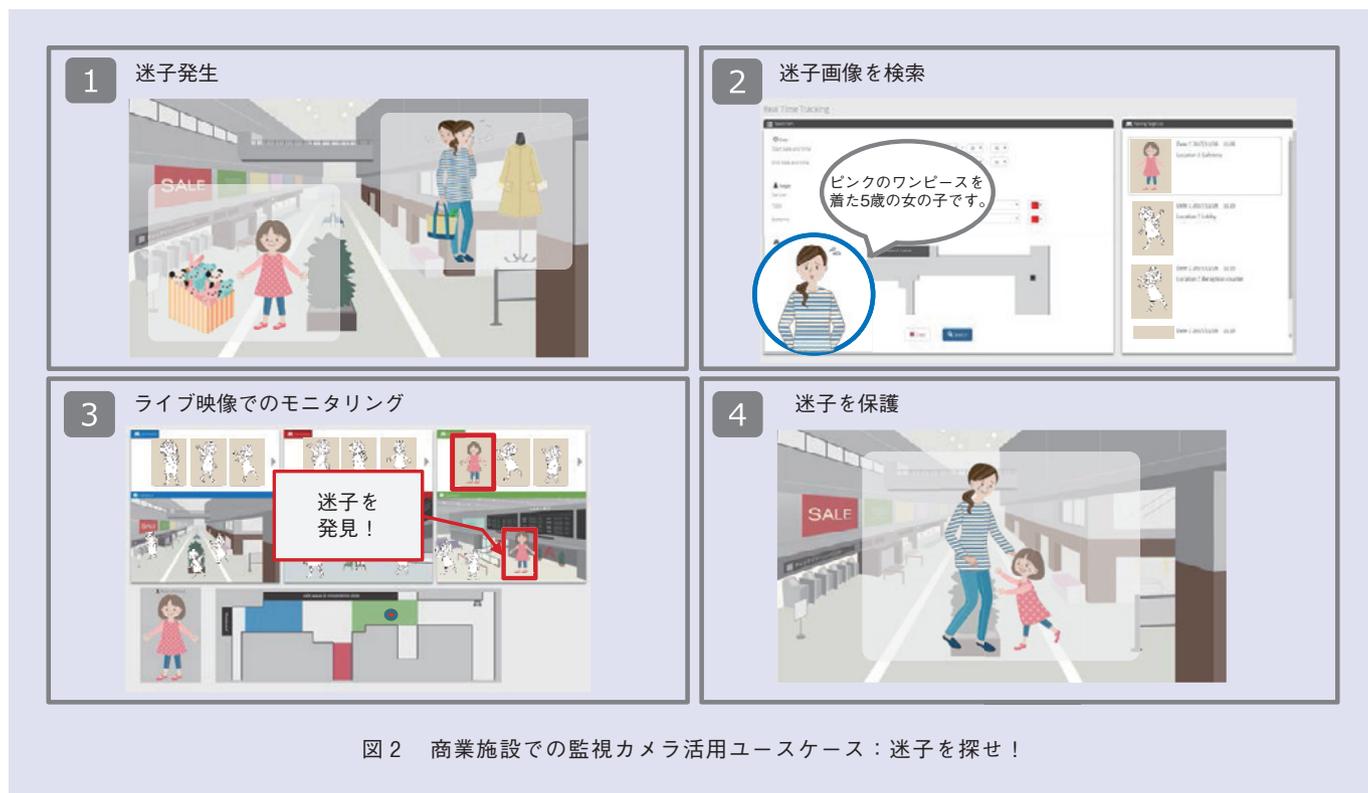


図2 商業施設での監視カメラ活用ユースケース：迷子を探せ！

なある程度の規模のマンションやショッピングモールなどの大型施設であれば、さらに効果的です。

刑事ドラマで監視カメラを徹夜で刑事がチェックするといったシーンは、本技術がサービス化された後にはもう登場することはなくなるでしょう。しかも本技術を用いれば、ヒューマンエラーによる見落としもありません。

■商業施設での監視カメラ活用：迷子を探せ！：属性推定+色検索+同一人物判定

もっと大規模な商業施設ならではの活用事例が「迷子の保護」です。百貨店やショッピングモールで買い物をしていると、迷子に関するアナウンスを

聞くことがあります。「5歳でピンクのワンピースを着た〇〇ちゃんをお母様が探しております」などのアナウンスは週末の商業施設での定番です。しかし、本技術がサービス化された後にはもう聞くことはなくなるかもしれません。「属性推定」技術で年齢を指定し、「色検索」で洋服の色を指定し検索します。100台近くある監視カメラの映像の中から、最後に映った子どもの映像を表示し、迷子を発見することができます(図2)。迷子のアナウンスを実施しない方針の大規模テーマパークなどには今後必須といえるシステムかもしれません。

■自治体向け徘徊老人の検索

先ほどの事例は商業施設×迷子のユースケースでしたが、そのベストプラクティスは徘徊老人の探索にも適用可能です。自宅から出たところの1枚の写真を指定することで短時間で自治体が設置している複数の監視カメラの映像から徘徊老人を見つけることが可能です。当然のことですが人間が目視で監視カメラ映像を確認することに比べ、短時間で見つけることができます。人間の補助として本技術を活用して、早く見つけることで、徘徊老人が事故に巻き込まれる確率も減らすことができます。

地元の警察・消防含めた自治体の人

たちが数百人で一斉にしらみつぶしに徘徊老人を探すといったニュースでの映像は、本技術がサービス化された後にはもう登場することはなくなるかもしれません。

■マーケティングに活用：同一人物判定+属性推定+軌跡推定

本技術は防犯目的以外にマーケティング目的にも利用可能です。時間帯の来店人数や属性を監視カメラ映像から分析して出すことや、導線分析にも利用できるかもしれません。限られた店舗スペースの中でどのように棚と商品を配置するかといったものもデータから分析することができるようになるでしょう。基本的な部分だけでなく、さまざまな応用も考えられます。例えば、検出するロジックのプログラムが必要になりますが、教師データを用意して学習させることで、「お客さまが手にとったが棚に戻した商品≒興味は持ったが購入されなかった商品」も抽出することができます。こうした技術の登場で、レジのPOSデータでは分からない情報も取ることができます。

「リアルタイム処理」を実現したコアな技術

本サービスを実現するうえで、もう1つ説明しておくべき重要なコア技術があります。それが、深層学習推論環境の最適化技術です。映像内の対象物を、高速に「検知」や「分類」など行い、リアルタイムで処理できるのは、この最適化技術のおかげといえます。処理に応じて、次のような世間の最新

の技術を複数組み合わせることで、10倍以上の高速化を実現しています。

- ・精度の高い人物検知や人物照合の映像解析アルゴリズム
- ・深層学習の推論処理の最適化実装技術

これらの技術をうまく組み合わせてアプリケーションを開発することで、今回紹介した「人物トラッキング」を、リアルタイムで処理を行うことが可能になったのです。

「人物トラッキング」も、「Takumi Eyes」として、最初に世間に登場したときには、コンビニの監視カメラのように「過去画像検索」が中心のサービスでした。それが時を経て、深層学習推論環境の最適化技術を研究し、組み合わせることでリアルタイム処理が可能となったのです。

今後の展開

リアルタイム処理が可能となった映像解析技術の次は、分散処理を中心とした研究を進めていく予定です。社内では「二層エッジモデル」と命名し、電話局の局舎やデータセンタ、そして、クラウドサービスも含めたシステム構築を可能にし、それぞれの場所であるべき機能を持たせる研究を進めていく予定です。また、特にエッジでは適切な機能を持たせることと同時に、さまざまな環境で活用可能なように、さまざまなデバイスで動くような取り組みも進めていきます。

NTTグループでは、AI技術「corevo[®]」により多くのお客さまの生活やビジネ

スをより良いものにしていただくために、今後もさまざまなパートナーの皆様と連携しながら、実社会への適応に向けた取り組みを進めていきます。

■参考文献

- (1) <https://www.gartner.co.jp/press/pdf/pr20180822-01.pdf>
- (2) <https://www.ntt.com/about-us/press-releases/news/article/2018/0912.html>
- (3) <http://www.ntt.co.jp/news2015/1506/150617a.html>
- (4) <http://www.ntt.co.jp/news2018/1810/181003a.html>
- (5) http://www.group.fuji-keizai.co.jp/press/pdf/161128_16095.pdf



(後列左から) 稲家 克郎/ 江田 毅晴/
廣川 裕/ 外山 将司/
森賀 邦広

(前列左から) 山田 佑二/ 村松 沙那恵/
山口 真弥/ 佐々木 琢/
三上 啓太

深層学習技術のビジネス応用はどうしても職人芸が必要な領域です。NTT研究所では深層学習の独自技術を研究し、使いやすくするために開発し、事業の優位性の確保に貢献します。

◆問い合わせ先

NTTソフトウェアイノベーションセンタ
第二推進プロジェクト
TEL 0422-59-2797
E-mail katsuo.inaya.zt@hco.ntt.co.jp