

安心・安全な社会構築を支える 映像解析AI技術

公共施設の安心・安全のため、防犯カメラと映像解析AI技術の活用が広がっている。日立は、大規模カメラから迷子や不審人物などを瞬時に探索するソリューションを製品化して、公共空間の安心・安全に貢献してきた。COVID-19の流行や働き方改革を背景に、映像解析AIのニーズは安心・安全をつかさどる人々の業務をさまざまな形でサポートする方向に拡大している。本稿では、AIを公共空間における感染症対策や人流解析に展開した例を紹介するとともに、さらなる支援に向けて、人とその周辺環境を深く理解する映像解析技術の取り組みと展望について述べる。

吉永 智明 | Yoshinaga Tomoaki

福田 安宏 | Fukuda Yasuhiro

渡邊 裕樹 | Watanabe Yuki

廣池 敦 | Hiroike Atsushi

1. はじめに

都市の持続的な成長、QoL（Quality of Life）の高い暮らしの実現のためには、安心・安全を高い水準で維持していくことが必要不可欠である。治安維持を目的に、駅・空港・商業施設・街頭などでは防犯カメラが広く普及している。そして、膨大な防犯カメラの映像をAI(Artificial Intelligence)によって解析し、侵入などの問題事象を迅速に発見することで、安全確保をサポートしている。

日立グループは、顔検知機能を搭載した映像監視レコーダーの製品化（2006年）や、100台超のカメラ映像から高速に顔を検索する技術（2009年）など、都市の安心・安全をAIで支援する技術の開発に先行的に取り組んできた。2019年には、深層学習技術を導入した高速人物発見・追跡ソリューションを製品化し¹⁾、大量の防犯カ

メラ映像の中から特定人物を高速かつ高精度に検索できるソリューションによって、社会インフラならびに公共エリアの大規模警備・安全サポートを行ってきた。

近年、映像解析AIの主目的は、危険を発見することから、大規模施設を少人数で運営管理するスタッフや、施設の利用者をサポートすることにシフトしつつある。本稿では、安心・安全な施設運営を支援する最新のソリューションや技術を紹介する。また、AI技術をより高度化していくための取り組みについても紹介する。

2. 映像解析AIを用いた監視・警備業務の高度化支援ソリューション

防犯カメラや映像解析AIを、犯罪の抑止や発見のみでなく施設運営業務のサポートにも役立てるソリューションを新たに開発した。ここでは2020年初頭より世界的に流行しているCOVID-19の影響下において密を避けて施

設運営を行うためのソリューションと、施設の経営判断や避難計画に活用が可能な人流可視化ソリューションを紹介する。

2.1

感染症対策支援ソリューション

世界各国の行政機関が、新型コロナウイルスワクチンの接種が完了するまではマスク着用や社会的距離の確保といった基本的な感染防止措置が続く可能性があるとの見解を示しており、日常生活に戻るまでには数年はかかるとの予想もある。このような状況に対して、日立が日本国内で販売している「高速人物発見・追跡ソリューションのアドオン機能」では、感染症対策への貢献のため密集した人だかりを検知する機能を強化した新バージョン v1.3²⁾、^{※1)}を提供している。

この機能は、映像内で検知された人物の位置を二次元の座標情報に変換して人物間の距離を分析するソーシャルディスタンス表示と、防犯カメラに映っているすべての人数でなく、人物間が一定の距離以下に密集している人数を測定することで、人だかりを検知・発報するアラート通知が可能である（図1参照）。

また、COVID-19の収束後もインシデントが発生した際にできる人だかりを検知することに利用できるため、警備業務において異常事態の発生をいち早く把握する用途でも活用が期待できる。

さらに、新バージョンではサーマルカメラシステムと連携し発熱者を検知した後の追跡を可能とするインタフェースを拡充するとともに、車やバイクなどの乗り物の検索が行える機能を追加しており、POI（Person of Interest）やVOI（Vehicle of Interest）を発見・追跡す

る機能を拡充している。

本製品は人々が多く集まる公共施設、市街地などの監視業務において施設内で発生した迷子や、不審者の発見・追跡を中心とした活用事例以外に、例えば現金やチケットを利用する交通機関での乗車と降車の起終点（OD：Origin Destination）情報生成にも活用の可能性がある。その他にも、トイレなどで着替えを行った人物の検知などカメラ映像をAI解析してIoT（Internet of Things）データとして利用するケースで効果を発揮しやすい。本製品を活用することで、これまで大量のカメラ映像を目視で確認していた業務を軽減し、IoTデータとしての自動処理や対象の人物に類似する人物画像を高速かつ自動的に絞り込み、人が判断するHuman-in-the-loop（人間参加）型の処理を確立して効率を向上させることで、働き方改革の推進、さらには社会課題の解決のカギとなるデジタルトランスフォーメーション加速に寄与できる。

2.2

広域人流解析ソリューション

ここでは、人物追跡技術を用いた広域人流解析ソリューションを紹介する。

広域人流解析のダッシュボードの実装例を図2に示す。

本ダッシュボードは、(1)広域マップ上での人流、(2)混雑度の推移、(3)流入元の内訳、(4)人物画像の一覧、(5)カメラ内の人流、(6)人物ごとのタイムラインを可視化する機能を備える。(2)、(4)、(5)は単一カメラ内の人検出・追跡結果を可視化したものであり、多くの映像解析製品で利用可能な人流解析機能である。日立の開発した複数カメラ間の人物追跡技術を人流解析に活用することで、カメラ間をまたいだ広域での人流解析が可能となる。(1)、(3)では、指定したカメラに現れた群衆が、どのエリアから流入してきたかを広域マップと内訳

※1) 海外ではHitachi Vantara社よりHitachi Multifeature Video Searchとして販売。

図1|密集した人だかりを検知する機能

映像内で検知された人物から人物間の距離を分析し、人だかりを検知する。COVID-19収束後も、異常事態発生時の把握など活用可能である。

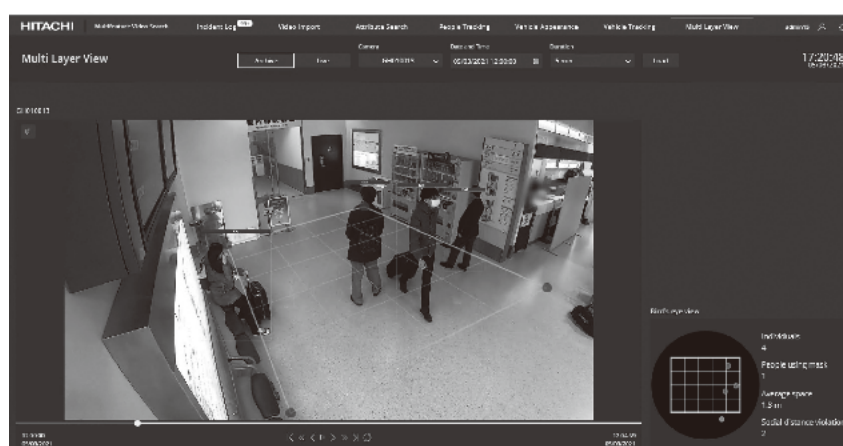




図2| 広域人流解析ソリューションのダッシュボードの例

商業施設での人流解析や、事件・事故発生時の状況把握が支援できる。

チャートで可視化している。これにより、混雑が発生したときに、その要因を分析することが可能になる。また、(6)では人物ごとにカメラ間の移動情報を可視化することで、群衆をグループ化して分析することや、エリア間の移動時間を調査することが可能になる。

これらの広域人流解析の結果は、ショッピングモールやテーマパークなどで混雑の少ない快適な空間を実現するためのビッグデータ分析に利用することができる。また、突発的な事件・事故が発生した際にいち早く状況を把握して対応するためには、直感的な人流可視化が有効である。日立は、広域人流解析によって快適で安心・安全な空間の実現に貢献していく。

3. 施設運営をサポートする次世代の映像解析AI技術

前章で述べたように、映像解析技術は防犯目的だけでなく、大規模な公共施設を少人数で効率よく運営していくために必須の機能となってきた。施設で働く従業員が、少ない人数ながらも、より質の高いきめ細かなサービスを提供していくためには、公共空間の一人ひとりの状況をAIが深く理解することが重要である。AIが人を深く理解できて初めて、スタッフや施設利用者への正確で最適な情報提供と、一部業務の代替が可能になる。

ここでは、人を深く理解する次世代の映像解析技術として、行動と荷物を深く理解することで人物を検索する技術を紹介する。さらに、人を深く理解する映像解析AIの開発を支援する3D (Three Dimensions) シミュレーション技術を紹介する。

3.1

行動・荷物の検索技術

警備や施設管理をより効率的に行うためには、ひたすらくりや不正侵入などの犯罪、転倒者など困っている人物をより迅速に見つけることが求められる。また、公共施設においては、多くの人はバックパックやスーツケースといった荷物を携行しており、荷物の扱いも重要な業務となっている。そこで日立は、防犯カメラに映る人物の行動や荷物を検索する技術を開発した。

行動検索技術は、防犯カメラ中の人物の行動を分類して、分類結果をデータベースに保存することで、特定行動を行った人物を検索する技術である(図3参照)。深層学習で構築したネットワークモデルによって、現在9種の行動(歩く、走る、転ぶ、直立、屈む、蹴る、投げる、指差し、見回す)を分類できる。服装や背格好だけでは困難だった人物検索の効率化や、施設内の統計取得(例えば、転倒した人物が1日に何人いたかなど)への活用が期待できる。

従来技術では、人物の向きや位置、個人差などで骨格形状が大きく変化するため、防犯カメラ用途では、精度面で課題があった。日立は、骨格の形状のみでなく、骨

図3| 行動検索技術

公共空間でニーズが高い9種の行動を認識する技術によって、特定の行動があった映像シーンを検索できる。深層学習によって認識対象の行動は追加可能である。



図4| 荷物の所有関係認識技術

人と荷物を結ぶ直線が、所有関係の推定結果を示している。混雑した状況で、所有者以外の人が荷物の近くにも正しく所有関係を判別している。



格を形成する各関節点の動きの変化を考慮した深層学習モデルを開発した³⁾。図3に示すとおり、人物の向きや形状が多種多様に変わるシーンにおいても、特定行動の検索が可能になった。

荷物検索技術の特徴は、荷物と所有者をひも付けて検索可能な点である。本技術では、映像から人と荷物（スーツケースやバックパックなど）を検出すると同時に、どの荷物を誰が持っているか、あるいは背負っているかの判別が可能である（図4参照）。所有関係を判別できることで、荷物の置き引きや受け渡しなどを高精度に発見できるほか、特定の荷物を複数人が持っていた場合にそれぞれの人物を検索できる。従来技術では、人と荷物の位置関係だけで所有関係を判定していたため、荷物の近くに複数の人物がいる場合に誤認識が多く発生していた。本技術では、人・荷物を検知すると同時に、「掴んでいる」、「背負っている」といった両者の関係性を推定する技術⁴⁾によって、高精度に所有関係が推定可能となり、誤認識を3割以上削減できることを確認している。

これらの技術はまだ研究段階であるものの、その精度

はすでに実用に耐えるレベルまで向上してきている。人物の行動、荷物の所有関係を認識する技術によって、施設内で起きている状況を深く理解することが可能となるため、転倒した人物を迅速に発見してサポートしたり、置き忘れた荷物を迅速に発見したりといったサービスが提供でき、利用者は安全で快適な施設の利用が可能になると期待される。さらに公共空間のみでなく、工場や作業現場においても、機器や道具を正しく使った安全な作業が成されているかを認識し、労務災害防止を支援したり、作業ログを記録したりするといった応用が見込める。

3.2

AI開発を支援する3Dシミュレーション技術

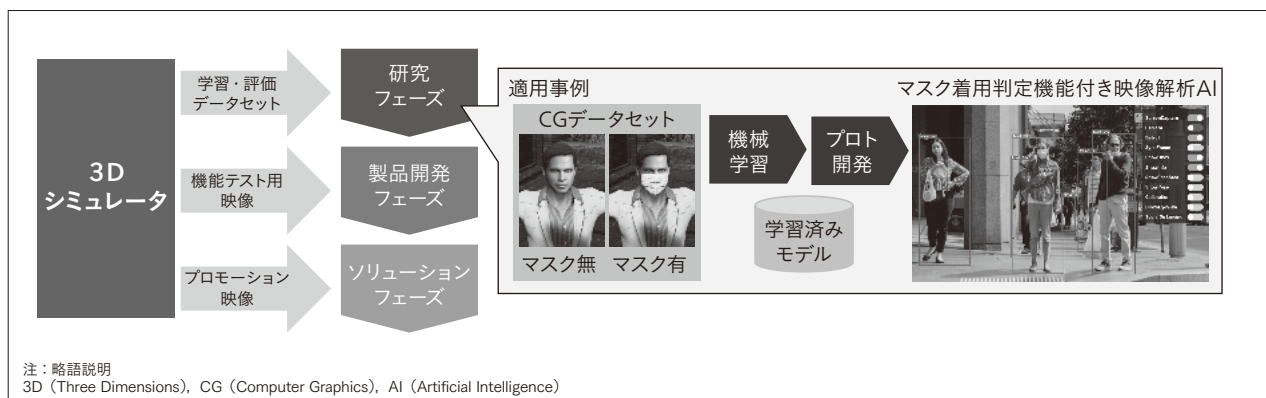
こうした人を深く認識するAI技術を高精度化するためにはデータが欠かせない。最新の深層学習技術はオープンソース化されているため、データセットを用意することができれば、多くの問題を解決することができる。一方で、解きたい問題に対して適切なデータセットを用意するためのコストが増大している。特に前節で述べたような人物認識のタスクにおいては、映像の権利やプライバシーの問題により、データセットが入手できない場合が多い。この場合、自前で映像撮影・アノテーション作業を行う必要があるが、これには膨大な時間と費用が必要になる。

そこで日立では、3Dシミュレータを活用したAI開発の導入を進めている（図5参照）。

人物や背景オブジェクトをシミュレータ上で描画することで、さまざまなシーンの映像を再現することができる。また、シミュレータが管理するオブジェクトの状態を抽出することで、正確なアノテーションが自動で行える。例えば、自前で映像撮影・アノテーションした場合に比べて、費用は1/20に、期間は1/16に短縮することが

図5| 3Dシミュレータを用いた映像AI開発支援

研究開発段階のみならず、各フェーズでCG映像データの活用が有効である。



できる。

本シミュレータは、マスク着用判定を行うAI開発に活用された。公共空間でのマスク着用が義務付けられたことを受け、AIによる自動判定のニーズが急速に高まった。一方で、パンデミック初期においては、マスク着用の習慣のない地域では実映像が少なく、また、人を集めて映像撮影・アノテーションすることが難しい状況であった。これに対して、シミュレータで生成したデータセットを用いることで、短期間でプロトタイプを開発することが可能になった。

ここでは、主に学習フェーズでのシミュレータ活用について述べたが、機能テストやプロモーションなど多くのフェーズで3Dシミュレータは有効と考える。今後、映像AI開発の各フェーズにおいてシミュレータを導入し、社会ニーズの急激な変化に対して迅速にソリューションを提供していく。

4. 映像解析AIの研鑽に向けた対外活動

本章では、安心・安全を支援する映像解析AIの技術強化に向けた日立の取り組みの一端を紹介する。

映像解析AI技術は日進月歩の進化を遂げており、グローバルで開発競争が加速している。中でも、中国と米国が大きく先行して技術開発を進めている。本稿で紹介したソリューションや技術開発を推進し、社会や顧客の課題を解決し続けるには、社外の顧客やパートナー・大学とのオープンイノベーションを進めるとともに、自社の技術力強化を進めることが必須である。ここでは、その中でも国際コンペティション（以下、「コンペ」と記す。）への参画を通じた取り組みを紹介する。

日立は、TRECVIDという米国NIST(National Institute of Standards and Technology)主催の映像解析のコンペに2015年から継続的に参加している。近年では、多くのコンペが存在し、そのプラットフォームとしてKaggle^{※2)}が注目を集めており、技術的な興味を持った個人や、研究開発に取り組む大学や企業が多く参加している。日立は、技術力強化や外部PR (Public Relations)のための活動ではなく、社会課題の解決というゴールを見据えた活動としてこうしたコンペを活用している。そのため、(1) 日立のビジネス領域に近いタスクを選定し、(2) 順位を競うだけでなくワークショップなどを通じて手法を比較

図6| 空撮映像からの災害状況の認識

災害発生時に迅速な状況把握や人による目視作業の支援が可能である。



するスタイルのコンペに積極的に参加している。上位入賞を真剣にねらって技術力強化をしつつ、目先の順位にとらわれず社会実装を見据えた開発を行ってきた。3.1節で紹介した行動認識技術も、コンペを通じて自社技術の研鑽と技術動向の把握をしながら生み出した成果の一つである。そして、2020年には、DSDI (Disaster Scene Description and Indexing) という空撮映像から災害状況を認識するタスクにおいて、トップレベルの精度を達成した⁵⁾ (図6参照)。

本技術についても、災害状況の迅速把握という重大な社会課題を解決したいという強いモチベーションから生まれたものである。

コンペへの参加のみならず、2021年には、CVPR (Conference on Computer Vision and Pattern Recognition) 併設のワークショップであるActivity Netの中で「MMAct Challenge」というコンペを主催した。防犯カメラを模した映像に対する行動分類と検知というタスクを設定したMMAct Challengeには、大学や企業など20を超えるチームが参加した。コンペ主催を通じて、社外とオープンなネットワークを構築することができた。今後もさまざまなアイデアを共有し、技術開発のスピードをより高めていきたいと考える。

今後、日立がグローバルなリーダー企業をめざしていくにあたっては、自社の技術・ビジネスの発展とともに、現状のままでは解決できない社会課題を解決可能な次世代技術の牽引が必要である。日立は国際コンペへの参加と主催を通じて業界をリードしていくことで、ときに競争しつつ、ときにオープンに協力して、次々世代のAI技術を創成していく。

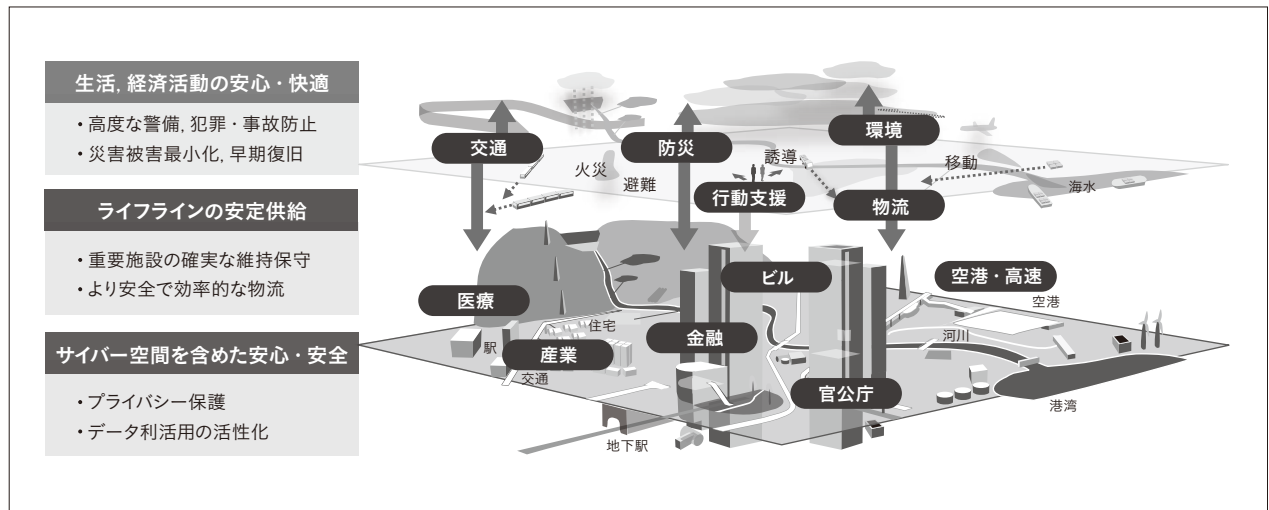
5. おわりに

本稿では、警備をはじめとした施設運営のさまざまな

※2) Kaggleは、Google LLCの登録商標または商標である。

図7| 人々の生命・生活・経済活動を支える組織の安心・安全を支援する日立の映像解析AI

公共空間・街中の人物の行動を深く理解することで、日々の移動や生活、企業における都市運営をサポートする。さらに、人を取り巻く環境を広く解析して、防災や環境保全など広く支援する。



業務を映像解析AIで支援するソリューションや技術を紹介した。

映像解析AIによって人の行動や荷物の所有者といった見逃しがちな小さな変化を人に代わって見つけることで、人に寄り添ったサービス提供が可能になると考える。また、安心・安全をより広域で考えた場合には、災害などの状況を迅速に把握する必要がある。これに対しては、映像解析対象をより広域に展開していくことで、人々の生活や都市全体の安心・安全をサポートできるようになると考えている（図7参照）。

本稿で紹介したAI技術やソリューションは、個人のプライバシーを侵害しないように配慮して利用されるとともに、人々が不利益を被らないよう、日立の「AI倫理原則」に則った管理の下で社会実装される。日立は、高精度でさまざまな環境に適用可能な信頼性の高いAIソリューションを広く都市全体に提供することで、安全で安心な社会の構築に貢献していく。

参考文献など

- 1) 日立ニュースリリース, AI画像解析技術を活用し、駅・空港・商業施設・公共機関などの監視・警備の業務効率化と高度化を支援する「高速人物発見・追跡ソリューション」を販売開始 (2019.9), <https://www.hitachi.co.jp/New/cnews/month/2019/10/1004a.html>
- 2) 日立ニュースリリース, AI画像解析技術を活用し、監視・警備業務の高度化を支援する「高速人物発見・追跡ソリューション」を強化 (2021.3), <https://www.hitachi.co.jp/New/cnews/month/2021/03/0325.html>
- 3) Y. Ito et al.: Multi-Stream Adaptive Graph Convolutional Network Using Inter- and Intra-Body Graphs for Two-Person Interaction Recognition, IEEE Access, vol. 9, pp. 110670-110682 (2021.8)

- 4) M. Tamura et al.: QPIC: Query-Based Pairwise Human-Object Interaction Detection With Image-Wide Contextual Information, Proceedings of CVPR, pp. 10410-10419 (2021.3)
- 5) 日立ニュースリリース, 災害状況をAIで把握する映像解析の基礎技術を開発 (2021.2), <https://www.hitachi.co.jp/New/cnews/month/2021/02/0219.html>

執筆者紹介



吉永 智明

日立製作所 研究開発グループ Lumada Data Science Lab.
映像解析ソリューションプロジェクト 所属
現在、Public Safety向け映像解析AI技術の研究開発に従事
電子情報通信学会会員



福田 安宏

日立製作所 社会ビジネスユニット 公共システム事業部
パブリックセーフティ推進本部 パブリックセーフティ第一部 所属
現在、株式会社日立システムズに出向し、生産技術本部業務に従事
技術士(情報工学)



渡邊 裕樹

日立製作所 研究開発グループ Lumada Data Science Lab.
映像解析ソリューションプロジェクト 所属
現在、Public Safety向け映像解析AI技術の研究開発に従事
博士(情報科学)



廣池 敦

日立製作所 研究開発グループ Lumada Data Science Lab.
映像解析ソリューションプロジェクト 所属
現在、Public Safety向け映像解析AI技術の研究開発に従事
工学博士