

映像分析と生成AIによるリアル世界の行動理解

神南 吉宏 梶木 善裕

要 旨

NECは、安全・安心・公平・効率のためにリアル世界の行動理解が重要だと考えています。しかし、従来の映像分析技術は複雑な行動や想定外の行動を理解するのが困難です。そこで、最近の生成AIを活用することで、行動の文脈を把握し、行動の意図や未来を予測する、リアル世界の行動理解技術を提案します。本稿では、映像分析技術と生成AIを活用したリアル世界の行動理解を実現する具体的なアーキテクチャを提案するとともに、提案したアーキテクチャによりオフィスビルにおける不審行動の理解が可能であることを実証した実験の結果について紹介します。



リアル世界／行動理解／映像分析技術／生成AI／大規模言語モデル(LLM)

1. リアル世界の行動理解の重要性

NECは、安全・安心・公平・効率という社会価値を創造し、誰もが人間性を十分に発揮できる持続可能な社会の実現を目指しています。その具体的な取り組みとして犯罪抑止・治安向上、街の危険や混雑の解消、高齢者や子どもの見守りなどに注力していますが、このような人々の安全・安心のための取り組みを効果的に推進するためには、リアル世界の人々の行動を深く理解することが極めて重要です。

2. 従来の映像分析の課題

リアル世界の人々の行動理解へ向け、これまでNECでは顔認証、入退場・滞留検知、行動検知などの映像分析技術群により、カメラ画像からの顔識別・人物属性分析や、人流マーケティング、作業現場の安全管理など、さまざまなソリューションを提供してきました。

しかし、現在の映像分析技術は、リアル世界を事前に想定したモデルに当てはめて、単純な動作や振る舞いを認識するものであるため、多様で変化の大きいリアル世界における人々の複雑な行動、特に想定外の行動を十分に理解

することができません。加えて、単純な動作や振る舞いを認識するだけでは、行動の意図理解や予測につなげにくいことが課題でした。

3. 生成AIの活用とその効果

近年発展の著しい大規模言語モデル(Large Language Model、以下、LLM)は、自然言語で記述された文章の複雑な文脈を理解し、適切な応答文を生成することができます。LLMは、文章だけではなく静止

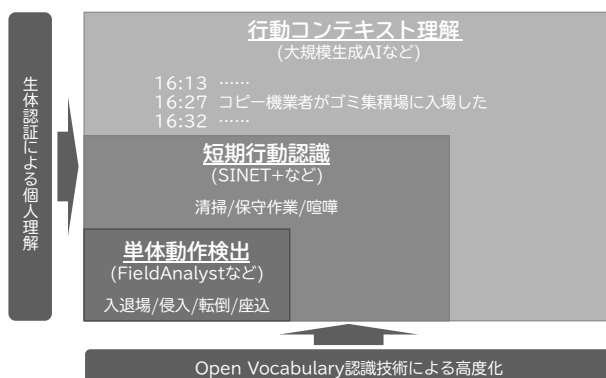


図1 リアル世界行動理解の価値

画像、動画画像、点群、音声、構造化データ、時系列データなどを理解し、生成する大規模生成AIモデル (Large Generative AI Model、以下、LGAIM) へと進化しようとしています。

NECは、LLM・LGAIMをリアル世界の行動理解に活用することにより、現在の映像分析の課題が解決可能になると考えています。なぜなら、これらの技術は画像などのリアル世界の情報から人々の行動の文脈を把握して、行動の理由・意図の理解や、次の行動の予測につなげることができるためです。そして、行動の理由・意図の理解により、その行動を支援することや、次の行動の予測により、リスクを未然に防止することなど、価値の高いサービスを提供することができます(図1)。

4. リアル世界行動理解のアーキテクチャ

4.1 リアル世界の事象の認識と記録

リアル世界行動理解を実現するためには、まずリアル世界の個別事象を認識し、記録する必要があります。将来的にはLGAIMが個別事象の認識も担えるようになると予想していますが、現時点では従来の映像分析技術を活用します。

また、人々の個別の行動を一連の行動履歴として記録するため、生体認証技術による個人の特定や、複数カメラ情報の識別技術を活用します。個人の特定は、工場やオフィスなど被撮影者が個人情報の取得を認識できるような空間で、適切な利用目的のもとで実施し、そうでない場合は服装などの外観特徴による識別技術など個人を特定しない技術を活用し、匿名性を確保します。

また、映像分析だけでなく、音声認識技術や音響分析技術により、更に詳細な個別事象認識を行うことも可能です。

このようにして認識した事象は、「誰が・いつ・どこで・何をした」など、個々人の識別子や日時や場所と動作や行

表 行動事象データの例

人物ID	人物属性	場所	開始時刻	終了時刻	行動
鬼海秀里	保守業者	エントランス	08-04 13:00:05	08-04 13:00:05	入場
鬼海秀里	保守業者	エレベーター	08-04 13:00:31	08-04 13:00:31	乗車
日電太郎	一般社員	エントランス	08-04 13:00:45	08-04 13:00:45	退場
五味綺麗	清掃業者	エントランス	08-04 13:00:50	08-04 13:00:50	入場
鬼海秀里	保守業者	エレベーター	08-04 13:00:59	08-04 13:00:59	降車
鬼海秀里	保守業者	機械室	08-04 13:01:23	08-04 13:09:12	滞留

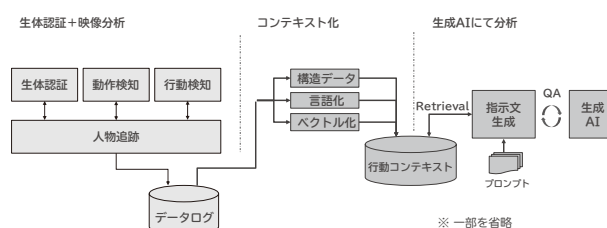


図2 リアル世界行動理解のアーキテクチャ

動をセットにして、構造データとしてデータベースへ記録されます(表)。

更に、文章や映像などから動作や行動の意味的な特徴を表す特徴ベクトル (embedding) を抽出して記録します。

4.2 生成AIによる行動理解

このようにしてリアル世界の事象を記録したデータベースから、個人ごとの一定期間の事象群を抽出します。

次に、抽出した事象群に対し、文章の形式での記録を時系列に並べ、「以下の行動履歴に対し、不審な行動はないか確認せよ」などの指示文とともに、LLMに提示して解析を行わせます。一般に、生成AIには処理できるトークン数に制限があり、長期間のすべての事象群を入力して分析させることは不可能です。このためNECは、分析に必要な事象のみを抽出したり、事象群を要約したりすることにより、長時間の事象群のLLMによる分析を可能にしています。更に、この言語化の際に一般用語で事象を表すことにより、工場内などの独自性の高いリアル世界であっても、一般世界の知識を活用した分析を可能にしています。

なお、独自性の高いリアル世界の分析においては、特徴ベクトルが類似する一般的な他の行動事象などを抽出して参考情報として付加すると、In-context Learningにより生成AIの分析精度を高めることができます。LGAIMを活用できるようになれば、事象群の映像や音声なども入力することで、より高い精度での分析ができるでしょう(図2)。

このような手法により、LLMに一連の行動を理解させ、行動の意図や未来の行動を推測することができます。その結果をもとに、意図に沿った支援、意図に反するリスクの予測・防止などが可能になることが期待されます。

5. 実証実験と結果

5.1 ユースケースと事象の記録

前述したコンセプトの有効性を検証するために、オフィスビルにおける不審行動の理解を対象とした実証実験を行いました。

本実証実験では、オフィスビルのエントランス、エレベーターホールなど8箇所にカメラを設置し、NECの状況推定システムFieldAnalyst¹⁾、多様な行動の高精度認識技術²⁾などを利用して入退場や滞留などの行動・振る舞いを検出し、行動事象データを作成しました。また、エントランスでの顔認証による個人の特定を行っています(図3)。

このようにして、個々人の動作や行動を検知・認識したうえで、事象ごとに「誰が、いつ、どこで、何をした」などの文章に変換するとともに、汎用のSentence Similarityモデルを使用して文章の意味に相当する特徴ベクトルを抽出してデータベースに記録しました。

5.2 行動事象の抽出と分析

次に、個々の人物がオフィスビルから退場したなどの適切なタイミングで、当該人物が入場してからの一連の事象群をデータベースから抽出し、自然言語の文章に変換し、「以下の行動履歴のうち、業者の業務内容からみて不必要な疑わしい行動をその理由とともに指摘してください」などの指示文とともに、LLMに入力して推論を行いました(図4)。

通常、映像からなんらかの行動を理解するには、何が理解すべき行動にあたるかを定義し、定義した行動を表現するルールや学習データを作成し、これを映像分析システムに教示する必要があります。

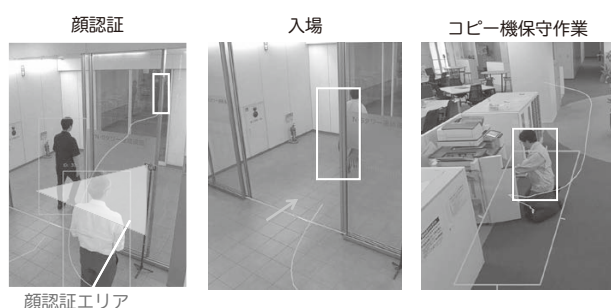


図3 検証映像からの事象の記録例

映像分析+大規模生成AIデモ



図4 行動コンテキストのLLMによる推論

しかし、「不審な行動」のように厳密に定義することが困難な行動は、すべてのケースを網羅したルールや学習データを作成することは事実上不可能です。

LLM・LGAIMが持つ一般的な知見を活用することにより、この問題を解決し、現実的なコストで行動理解を実現できると期待されます。NECは、本実証実験においてその可否を検証しました。

5.3 LLMによる推論結果

第5章3節では、この実証実験により得られた特徴的な結果を記します。

まず、清掃業者がゴミ集積場へ入場してゴミ収集作業を行った行動事象に対しては、LLMは指摘する事項がない旨を回答しました。しかし、コピー機業者がゴミ集積場へ入場して物品を物色した行動事象に対しては、LLMは「ゴミ集積場は業者の業務内容とは関係がない場所であり、滞留する必要はないため」という理由とともに、不審な行動が検出された旨を指摘しました。

次に、一般社員がロッカールームでロッカーを開閉した行動事象に対しては、LLMは「不審な行動は見当たらない」旨を回答しましたが、前述の清掃業者が社員のロッカールームでロッカーを開閉した行動コンテキストに対しては、不審な行動として検出しました。

このように、オフィスビルという非公開のリアル世界に対しても、映像分析とLLMを組み合わせ、ルール設定や学習なしに不審な行動を検出できることを実証しました。

一方、一般的な知見だけでは、不審ではない行動を過検出したり、不審な行動を見逃したりするケースがあります。例えば、清掃業務に使用する器具がロッカーに保管さ

れているのであれば、清掃業者がロッカーを開閉するのは業務に必要な行動ですが、そうでなければ不審な行動です。この問題に対して、NECはプロンプトに業務固有のルールを記載するなど、LLMへの指示の方法を改善することにより対処できることを確認しています。

6. むすび

本稿では、映像分析と生成AIによるリアル世界の行動理解の現状と可能性について紹介しました。NECはこの技術を活用し、リアル世界における人々の安全・安心に貢献する、次世代フィジカルセキュリティの実現を目指していきます。またこの技術は、工場・倉庫などでの作業を支援したり、オンライン・オフラインでの顧客の行動からニーズや不満を発見したりといった用途にも活用可能であり、次世代フィジカルセキュリティにとどまらず、さまざまな領域での活用を推進していきます。

本技術には大きな可能性がある一方、いくつかの課題も残されています。例えば、現在広く活用されているLLMが学習している情報はインターネットや書籍などから収集した一般公開情報が中心であるため、企業や官公庁の特殊施設などの独自性・機密性の高い場における行動を理解できない可能性があります。本技術をこのようなリアル世界にまで適用するには、機密保持を担保しながら独自性・機密性の高いリアル世界の情報を学習する必要があります。また、AIが人々の行動理解に踏み込むと、人権侵害やプライバシー侵害を起こす懸念があり、技術上で保護を行ったうえで、法令や社会ルールに適合する形で運用することが極めて重要になります。

NECは、機密性の高いオンプレミス環境において独自データによる個別チューニングが可能なNEC開発の生成AI「cotomi」³⁾の活用などにより、このような課題の解決にも積極的に取り組み、人々が安心して人間性を発揮できる社会の実現を目指します。

参考文献

- 1) NECソリューションイノベータ: FieldAnalyst
<https://www.nec-solutioninnovators.co.jp/sl/fieldanalyst/index.html>
- 2) NEC プレスリリース: 映像から複数の人物の多種多様な作業内容を高精度に認識する技術を開発, 2022.11
https://jpn.nec.com/press/202211/20221128_01.html
- 3) NEC プレスリリース: NEC、生成AI「cotomi (コトミ)」の強化・拡充と共に生成AI事業戦略を発表, 2023.12
https://jpn.nec.com/press/202312/20231215_02.html

執筆者プロフィール

神南 吉宏

生体認証・映像分析統括部
シニアプロフェッショナル

梶木 善裕

生体認証・映像分析統括部
プロフェッショナル

NEC 技報のご案内

NEC 技報の論文をご覧くださいありがとうございます。
ご興味がありましたら、関連する他の論文もご一読ください。

NEC技報WEBサイトはこちら

NEC技報（日本語）

NEC Technical Journal（英語）

Vol.75 No.2 ビジネスの常識を変える生成AI特集 ～社会実装に向けた取り組みと、それを支える生成AI技術～

ビジネスの常識を変える生成AI特集によせて
生成AI技術への取り組み ～基盤から応用、ルール作りまで～

◇ 特集論文

急速に広まる生成AIの市場適用

NECにおける生成AIの取り組みについて
電子カルテと医療文書の作成支援による医師業務効率化
映像×LLMによる報告書作成業務の自動化
映像分析と生成AIによるリアル世界の行動理解
サイバー脅威インテリジェンス生成自動化
生成AIの社内活用を進めるNEC Generative AI Service (NGS)
ソフトウェア・システム開発への生成AIの活用
LLMとMIで革新する素材開発プラットフォーム
LLMと画像分析を活用した被災状況の把握

生成AIの可能性を高める基盤技術

日本語性能に優れたNECの大規模言語モデル (LLM)
NECの生成AIを支える国内企業で最大規模のAIスーパーコンピュータ
より安全な大規模言語モデル (LLM) を目指して
データを秘匿したまま連携できる連合学習技術とLLMへの適用可能性
大規模言語モデル (LLM) によるFew-shotクラスタリングの可能性
オープンドメイン常識推論のための知識拡張型プロンプト学習
AI連携とオーケストレーションのための基盤ビジョンLLM
クエリを考慮した新規手法により関連する企業データを減らしLLM APIの使用コストを最適化

AI技術が社会へ浸透するために

AI標準化・ルールメイクに関する動向とNECの取り組み
人権尊重に向けたNECのAIガバナンスの取り組み
RCModelを用いたAIリスクマネジメントのための人材育成事例

◇ NEC Information

2023年度C&C賞表彰式典開催



Vol.75 No.2
(2024年3月)

特集TOP