

AI 特許紹介(70)
AI 特許を学ぶ！究める！
～VPT 特許～

2024 年 11 月 8 日
河野特許事務所
所長弁理士 河野英仁

「AI 特許紹介」シリーズは、注目すべき AI 特許のポイントを紹介します。熾烈な競争となっている第 4 次産業革命下では AI 技術がキーとなり、この AI 技術・ソリューションを特許として適切に権利化しておくことが重要であることは言うまでもありません。

AI 技術は Google, Microsoft, Amazon を始めとした IT プラットフォーマ、研究機関及び大学から毎週のように新たな手法が提案されており、また AI 技術を活用した新たなソリューションも次々とリリースされています。

本稿では米国先進 IT 企業を中心に、これらの企業から出願された AI 特許に記載された AI テクノロジー・ソリューションのポイントをわかりやすく解説致します。

1.概要

特許権者 OpenAI

出願日 2023 年 4 月 19 日

登録日 2024 年 1 月 30 日

登録番号 US11887367

発明の名称 機械学習を使用してモデルをトレーニングし、ビデオと入力データセットに基づいて自動インターフェースアクションを実行する

367 特許は、少量のラベル付けされた動画により学習された逆動力学モデル (IDM : inverse dynamics model) を用いて、ラベル付けされていない動画に対する疑似ラベルを生成し、生成した疑似ラベル及び動画により、動画から行動を予測する基盤モデルをトレーニングする VPT(Video PreTraining)技術に関する。

2.特許内容の説明

図 1 は、自動化されたアクションを実行するために機械学習モデルをトレーニングする方法 100 を示すフローチャートである。

100

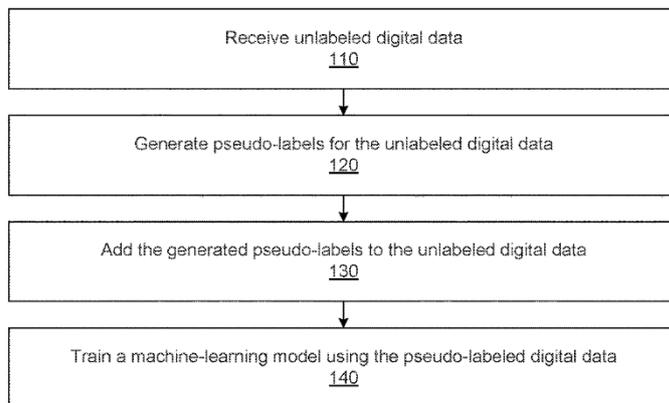


FIG. 1

方法 100 は、ラベルなしデジタルデータを受信するステップ 110 を含む。方法 100 は、ラベル付けされていないデジタルビデオデータに対して疑似ラベルを生成するステップ 120 を含む。ここで使用される疑似ラベルとは、デジタルビデオデータ（たとえば、1 つ以上のデジタルビデオフレーム）の属性を示す。

方法 100 は、疑似ラベルをラベルなしデジタルデータに追加し、それによってラベルなしデジタルデータを疑似ラベル付きデジタルデータに変換するステップ 130 を含む。疑似ラベル付きデジタルデータには、予測されたユーザアクションデータ（たとえば、上記の例のようなユーザアクションを表すまたは示すデータ）と関連するビデオタイムステップデータが含まれる。

方法 100 は、疑似ラベル付きデジタルビデオデータを使用して基盤モデルとなる機械学習モデルをトレーニングするステップ 140 を含む。疑似ラベルの結果として、利用可能なラベル付きデータの量が増加する。したがって、因果機械学習モデルをトレーニングするために利用可能なトレーニングデータが増加し、その結果、因果機械学習モデルはより正確かつ効率的になる。因果機械学習モデルは、行動クローニング（例えば、模倣学習）モデル、または、強化学習モデルである。

図 2 は疑似ラベルの生成処理を示すフローチャートである。

200

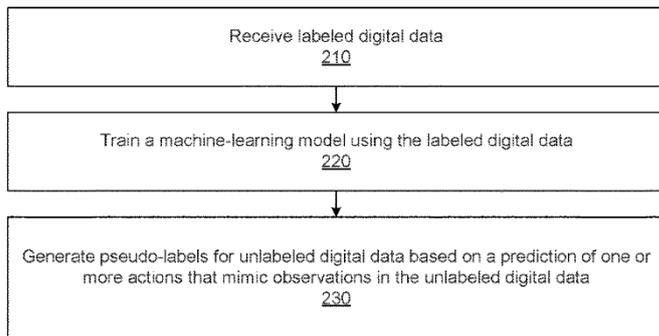


図 2 に示すように、疑似ラベルを生成する方法 200 は、ラベル付けされたデジタルビデオデータ（たとえば、アクションと観察のペアを含むデータ）を受信するステップ 210 を含む。受信したラベル付けされたデジタルビデオデータには、実際のユーザアクションデータと関連されたビデオデータが含まれ、これにより、デジタルビデオデータ内にアクションと観察のペアが形成される。たとえば、ラベル付けされたデジタルビデオデータは、人間のエージェントによって提供される、デジタルビデオデータの 1 つ以上のタイムステップに対応するアノテーションを含む。アノテーションは、ユーザインターフェースを介して実行可能な 1 つ以上のユーザアクションへの参照を含み、キーの押下、ボタンの押下、タッチスクリーンへの入力、ジョイスティックの動き、マウスのクリック、スクロール ホイールの動き、またはマウスの動きの少なくとも 1 つを表す。

疑似ラベルを生成する方法 200 は、ラベル付けされたデジタルビデオデータを使用して、機械学習モデル（たとえば、逆ダイナミクスモデル (IDM)）をトレーニングするステップ 220 をさらに含む。

たとえば、IDM は、少量のラベル付き請負業者 **contractor** データ（人間の請負業者 **contractor** によって作成された観察とアクションのペアを含む 100 ~ 2,000 時間のデータなど）からの非因果データを使用して最初にトレーニングされる。このような IDM は、観察の軌跡が与えられた場合に、特定のタイムステップでのアクションの負の対数尤度を最小化する。たとえば、模倣学習モデルよりも IDM を使用する利点の 1 つは、IDM が非因果的である可能性があることである。つまり、IDM によるユーザアクションの予測は、ラベルなしデータからの過去と将来の両方のイベント（たとえば、参照フレームに対する過去と将来のフレーム）の関数である。

疑似ラベルを生成する方法 200 は、ラベル付けされていないデジタルデータ（たとえば、最初はアクションと観察のペアを含まないデータ）に対して少なくとも 1 つの疑似ラベルを生成するステップ 230 をさらに含む。少なくとも 1 つの疑似ラベルは、ラ

ベル付けされていないデジタルデータの少なくとも 1 つのタイムステップを模倣するためにユーザインターフェース（たとえば、ネイティブのヒューマンユーザインターフェース）を介して実装できる 1 つ以上のアクションの機械学習モデルによる予測に基づいて生成される。

3.クレーム

367 特許のクレーム 1 は以下の通りである。

1. 機械学習モデルを訓練して自動化されたアクションを実行する方法であって、ラベルのないデジタルビデオデータを受信し、ラベル付けされていないデジタルビデオデータに疑似ラベルを生成し、該生成は以下を含み、
 - ラベル付けされたデジタルビデオデータを受信し、
 - ラベル付けされたデジタルビデオデータを使用して、逆動力学モデル（IDM : inverse dynamics model）を含む第 1 の機械学習モデルをトレーニングし、
 - ラベル付けされていないデジタルビデオデータに対して少なくとも 1 つの疑似ラベルを生成し、
 - 少なくとも 1 つの疑似ラベルは、IDM によって生成された、ラベルなしデジタルビデオデータの少なくとも 1 つのタイムステップを模倣する 1 つ以上のアクションの予測に基づいており、
 - 1 つ以上のアクションの予測は、ラベルなしデジタルビデオデータ内の過去の情報と将来の情報の非因果的組み合わせに基づいて生成され、過去の情報と将来の情報は、ラベルなしデジタルビデオデータ内の 1 つ以上の参照フレームを基準とし、
 - 少なくとも 1 つの疑似ラベルをラベルなしデジタルビデオデータに追加して疑似ラベル付きデジタルビデオデータを形成し、
 - ラベルなしデジタルビデオに対して少なくとも 1 つの追加の疑似ラベルを生成するために、疑似ラベル付きデジタルビデオデータを使用して第 1 の機械学習モデルまたは第 2 の機械学習モデルをさらにトレーニングする。

4. 本特許に関連する論文

本特許に関する論文 “Video PreTraining (VPT): Learning to Act by Watching Unlabeled Online Videos”¹が、OpenAI の Bowen Baker 氏らにより公表されている。

¹ Bowen Baker, et al. “Video PreTraining (VPT): Learning to Act by Watching Unlabeled Online Videos” arXiv:2206.11795v1 [cs.LG] 23 Jun 2022

下記図は、Minecraft のクラフト GUI の例である。



エージェントはマウスとキーボードを使用してメニューを移動し、アイテムをドラッグアンドドロップする。

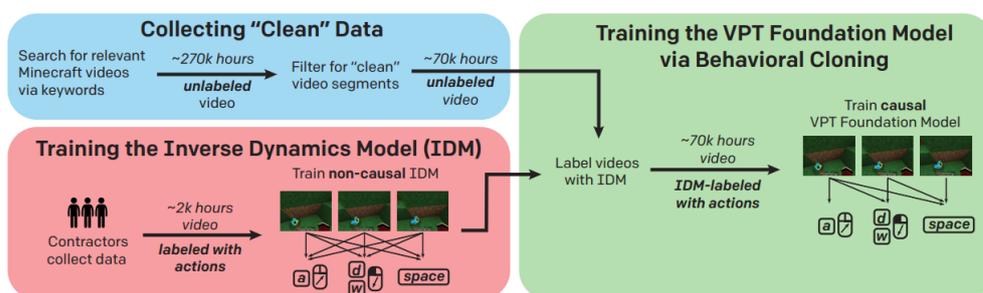
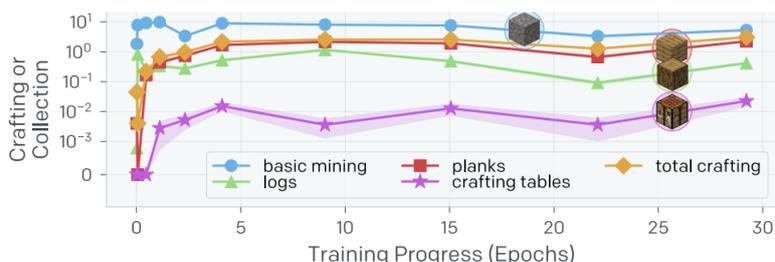


Figure 2: Video Pretraining (VPT) Method Overview.

図 2 に示すように、VPT は、まずラベル付けされた請負業者のデータを少量収集し、それを使って逆ダイナミクスモデル $p_{IDM}(a_t | o_{1..T})$ をトレーニングする (ピンク部分)。このモデルは、 T 回の観測値の軌跡 $o_t : t \in [1..T]$ が与えられた場合に、タイムステップ t でのアクションの負の対数尤度を最小化することを目指す。模倣学習ポリシーとは対照的に、IDM は非因果的であり、これは a_t の予測が過去と未来の両方のイベントの関数になる可能性があることを意味する (つまり、 $o_t \rightarrow o_{t+1}$)。

次に、ラベル付けされていないクリーンなビデオデータ (水色部分) に対し、学習済みの IDM を用いて疑似的なラベル付け (アクション) を行う。疑似ラベルが付されたビデオデータを用いて VPT 基盤モデルをトレーニングする (緑色部分)。この基盤モデルは重要なゼロショット動作を示し、模倣学習と RL の両方を使用してファインチューニングすることで、さらに複雑なスキルを実行できる。



上記グラフはトレーニングエポックの関数として、60 分間のサバイバルエピソード

2500 回の平均で、エピソードごとに特定のアイテムが収集された量を示す。平均の標準誤差で網掛けされている。基本的な採掘とは、土、砂利、砂（ツールなしで収集できるすべての材料)の収集を指す。多くのタスクで人間レベルのパフォーマンスを示し、熟練した人間でも 20 分以上 (24,000 の環境アクション) のゲームプレイで達成できるダイヤモンドツールを作成できるコンピュータエージェントを論文にて報告している。

OpenAI の HP に VPT 技術が紹介されている。

<https://openai.com/index/vpt/>

以上

著者紹介

河野英仁

河野特許事務所、所長弁理士。立命館大学情報システム学博士前期課程修了、米国フランクリンピアースローセンター知的財産権法修士修了、中国清華大学法学院知的財産夏季セミナー修了、MIT(マサチューセッツ工科大学)コンピュータ科学・AI 研究所 AI コース、生成 AI ビジネスコース修了。

[AI 特許コンサルティング](#)、[医療 AI 特許コンサルティング](#)の他、米国・中国特許の権利化・侵害訴訟を専門としている。著書に「世界のソフトウェア特許(共著)」、「FinTech 特許入門」、「[AI/IoT 特許入門 3](#)」、「[ブロックチェーン 3.0](#)(共著)」がある。