

AI 特許紹介(29)
AI 特許を学ぶ！究める！
～Grasp2Vec 特許～

2021年6月10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 Google

出願日 2019年6月17日

登録日 2020年4月28日

登録番号 US10635944

発明の名称 自己教師あり (Self-Supervised) ロボットオブジェクトの相互作用

944 特許は、教師データを与えることなく、ロボットにオブジェクトを認識させる Grasp2Vec 技術に関する。

2.特許内容の説明

図1は、ロボット制御システム100の例を示している。

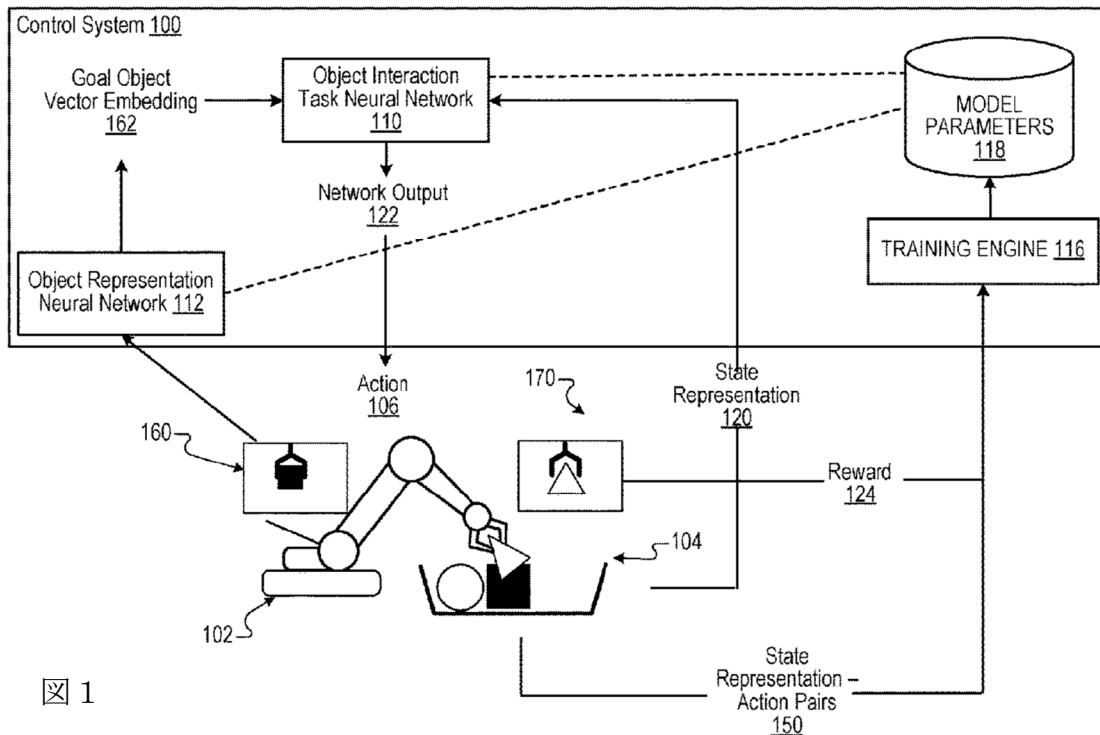


図 1

システム 100 は、ロボット 102 によって実行されるアクション 106 を選択し、次にロボット 102 に選択されたアクション 106 を実行させることによって、環境 104 と相互作用するロボット 102 を制御する。

システム 100 は、オブジェクト相互作用タスクニューラルネットワーク 110 およびオブジェクト表現ニューラルネットワーク 112 を含む。システム 100 はまた、訓練エンジン 116 を含み、オブジェクト相互作用タスクニューラルネットワーク 110 のパラメータ（オブジェクト相互作用パラメータ）、および、オブジェクト表現ニューラルネットワーク 112 のパラメータ（オブジェクト表現パラメータ）の現在の値を含むモデルパラメータのセット 118 を維持する。

複数の時間ステップのそれぞれで、オブジェクト相互作用タスクニューラルネットワーク 110 は、ネットワーク出力 122(アクション選択ポリシー)を生成するために、オブジェクト相互作用パラメータの現在の値に従って、環境 104 の現在の状態を特徴付ける現在の状態表現 120 を含む入力を処理するように構成される。

ネットワーク入力はまだ、環境内のゴールオブジェクト 160 の埋め込み 162 を含む。このゴールオブジェクトには、タスクを完了するためにロボット 102 が指定された相互作用を有するべきオブジェクトを含む。

システム 100 は、ネットワーク出力 122 を使用して、現在のタイムステップでロボットによって実行されるアクション 106 を選択する。ロボットは、各タイムステップでそれぞれの報酬 124 を受け取ることができ、報酬 124 は、スカラー数値によって指定され、例えば、タスクを完了するためのロボットの進行を特徴付ける。

オブジェクト表現ニューラルネットワーク 112 は、環境内でロボットが遭遇するオブジェクトのベクトル埋め込みを生成するニューラルネットワークである。ベクトル埋め込みは、オブジェクトの数値表現として機能する、順序付けられた数値のコレクションであり、たとえば、浮動小数点値のベクトルである。

トレーニングエンジンは、教師なしの方法でオブジェクト表現ニューラルネットワーク 112 をトレーニングして、環境内でロボットが遭遇するオブジェクト間の関係を反映するベクトル埋め込みを生成することを学習する。

システム 100 は、オブジェクト表現ニューラルネットワーク 112 を使用して、オブジェクト相互作用タスクニューラルネットワーク 110 をトレーニングし、ニューラルネットワーク 110 が、環境内の異なるオブジェクトとの相互作用を含むタスクを実行できるようにロボットを効果的に制御する。

特に、オブジェクト相互作用タスクニューラルネットワーク 110 によって受信された各ネットワーク入力、環境内のゴールオブジェクト 160 のオブジェクト表現ニューラルネットワーク 112 によって生成されたベクトル埋め込み 162 を含む。言い換えれば、各ネットワーク入力は、状態表現 120 と目標オブジェクトベクトル埋め込み 162 の両方を含み、オブジェクト相互作用タスクネットワーク 110 は、ロボットが相互作用するべきオブジェクトの埋め込みを条件とするネットワーク出力を生成する。

オブジェクト表現ニューラルネットワーク 112 のトレーニング中、システム 100 は、オブジェクト相互作用タスクニューラルネットワーク 110 を使用してロボット 102 を制御することによってタスクのエピソードを実行し、ニューラルネットワーク 110 は、オブジェクト表現ニューラルネットワーク 112 によって生成された埋め込み 162 を条件とする。

システム 100 は、ロボットを制御して1つまたは複数のカメラの視野内に物体を配置し、次に物体が視野内にある間に画像を捕捉することによって、物体の画像を捕捉する。各タスクエピソードについて、システム 100 は、ロボット 102 が環境内の任意のオブジェクトのタスクを正常に完了したかどうかを判断し(ゴールオブジェクト 160 と

は異なっている)、もしそうであれば、オブジェクト表現ニューラルネットワーク 112 を使用してタスクエピソードの報酬値 124 を決定する。

ロボット 102 が任意のオブジェクトのタスクを正常に完了すると、システム 100 は、タスクが完了した結果オブジェクト 170 の画像を取得する。図 1 の例では、タスクは、ゴールオブジェクト 160 (矩形) をつかむことであったが、ロボット 102 は、実際には、ゴールオブジェクト 160 とは異なる結果オブジェクト 170 (三角形) をつかんでいる。

次に、システムは、オブジェクト表現ニューラルネットワーク 112 を使用して結果オブジェクト 160 の画像からエピソードの報酬値 124 を生成する。オブジェクト表現ニューラルネットワーク 112 を使用して報酬を生成することにより、システム 100 は、正確に報酬値 124 を生成することができる。結果オブジェクト 170 がゴールオブジェクト 160 と異なる場合でも、タスクニューラルネットワーク 110 のトレーニングをガイドする。

次に、トレーニングエンジン 116 は、報酬値 124 および状態表現のセット-アクションペア 150 についてニューラルネットワーク 110 をトレーニングして、オブジェクト相互作用パラメータの現在の値を更新する。

オブジェクト表現ニューラルネットワーク 112 を使用して報酬値を生成することにより、システムは、完全に自己教師ありデータ **self-supervised data** 上でロボットの効果的な制御ポリシーを学習することができる。特に、システムは、ニューラルネットワークを効果的に訓練するために、結果オブジェクト 170 とゴールオブジェクト 160 が同じであるかどうかを明示的に識別する必要はない。

図 2A は、オブジェクト表現ニューラルネットワーク 112 のアーキテクチャを示す。

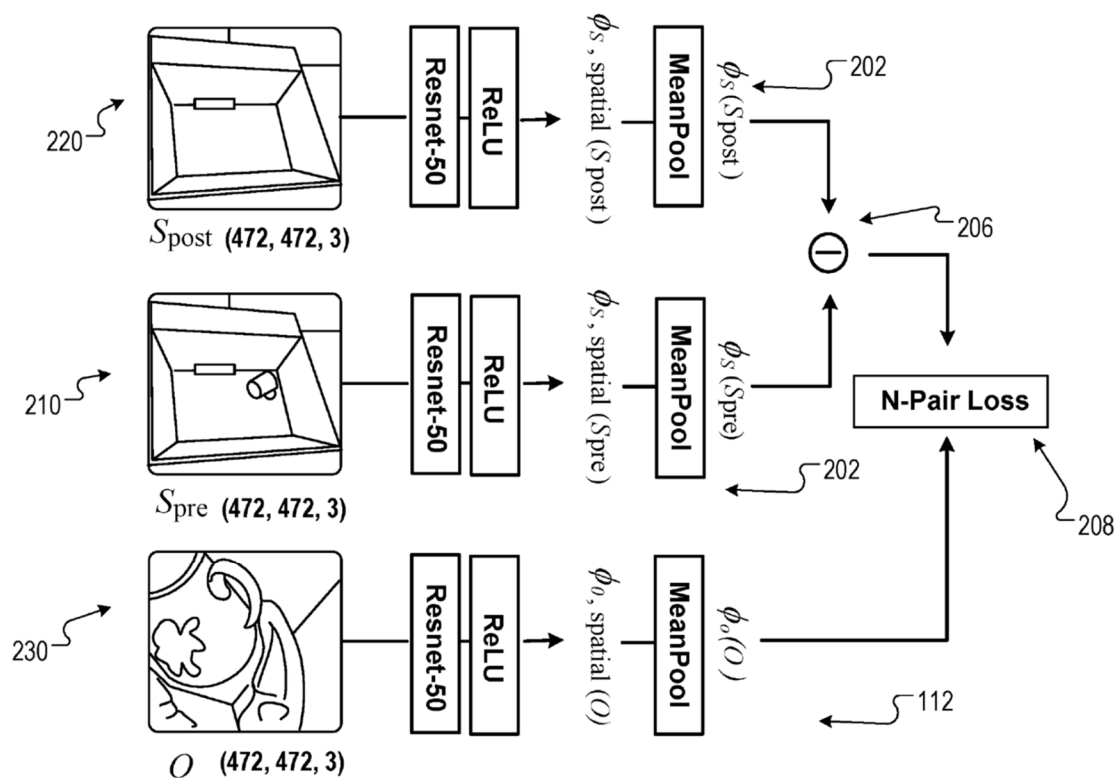


図 2 A

シーン表現ニューラルネットワーク 202 とオブジェクト表現ニューラルネットワーク 112 は両方とも、画像（この場合は $472 \times 472 \times 3$ 画像）を受信し、画像を処理してベクトル埋め込みを生成する畳み込みニューラルネットワークである。

シーン表現ニューラルネットワーク 202 は、シーンの画像を受信し、シーンの画像を処理して、シーンのベクトル埋め込みを生成するように構成される。シーン表現ニューラルネットワーク 202 とオブジェクト表現ニューラルネットワーク 112 は共に畳み込みニューラルネットワークである。

図 2A に示されるように、双方のニューラルネットワークは、畳み込みニューラルネットワーク層のセットと、それに続く正規化線形ユニット (ReLU) 活性化関数を含む。畳み込み層は、図の例では、「resnet-50」とラベル付けされている。図 2 に示されるように、層は、ResNet50 アーキテクチャによって指定された接続性を有するニューラルネットワークの畳み込み層である。次に、各ニューラルネットワークは、ReLU の 3D 出力にグローバル平均プーリング (「MeanPool」) 操作を適用して、ベクトル埋め込みを生成する。

ニューラルネットワークを訓練するために、システムは、シーン表現ニューラルネッ

トワーク 202 を使用して1つまたは複数のカメラの視野内に1つまたは複数のオブジェクトを有する環境の前シーンの前画像 210(spre)を処理して、前シーン Φ_s (spre) のベクトル埋め込みを生成する。

システムはまた、ロボットがシーン表現ニューラルネットワーク 202 を使用して1つまたは複数のカメラの視野から1つまたは複数のオブジェクトの特定のオブジェクトを除去した後、環境の後のシーン 220 (spost) を処理して、後のシーン Φ_s (spost) のベクトル埋め込みを生成する。

システムはまた、オブジェクト表現ニューラルネットワーク 112 を使用して視野から除去された特定のオブジェクトのオブジェクト画像 230(O)を処理して、特定のオブジェクト $\Phi_o(O)$ のベクトル埋め込みを生成する。

次に、システムは、後のシーン Φ_s (spost) のベクトル埋め込みと前のシーン Φ_s (spre) のベクトル埋め込みとの間の差 206 を決定し、オブジェクト表現パラメータおよびシーン表現パラメータの更新を決定することによってニューラルネットワークを訓練する。これは、特定のオブジェクト Φ_o (O) のベクトル埋め込みが差 206 に近づくことを促進する。

特定のオブジェクト Φ_o (O) のベクトル埋め込みを差 206 に「近づける」ように「促進」することにより、特定のオブジェクト Φ_o (O) のベクトル埋め込みと差 206 との間の距離（または差）を減少させることができる。

システムは、差 206 と特定のオブジェクト $\Phi_o(O)$ のベクトル埋め込みに依存する目的関数を最適化することによってニューラルネットワークを訓練して、特定のオブジェクトのベクトル埋め込み (NO) と差 206 との間の距離を減らす。目的関数はN対損失 (N-pairs loss) 208 である。

図 2B は、オブジェクト相互作用タスクニューラルネットワーク 110 の例示的なアーキテクチャを示す。

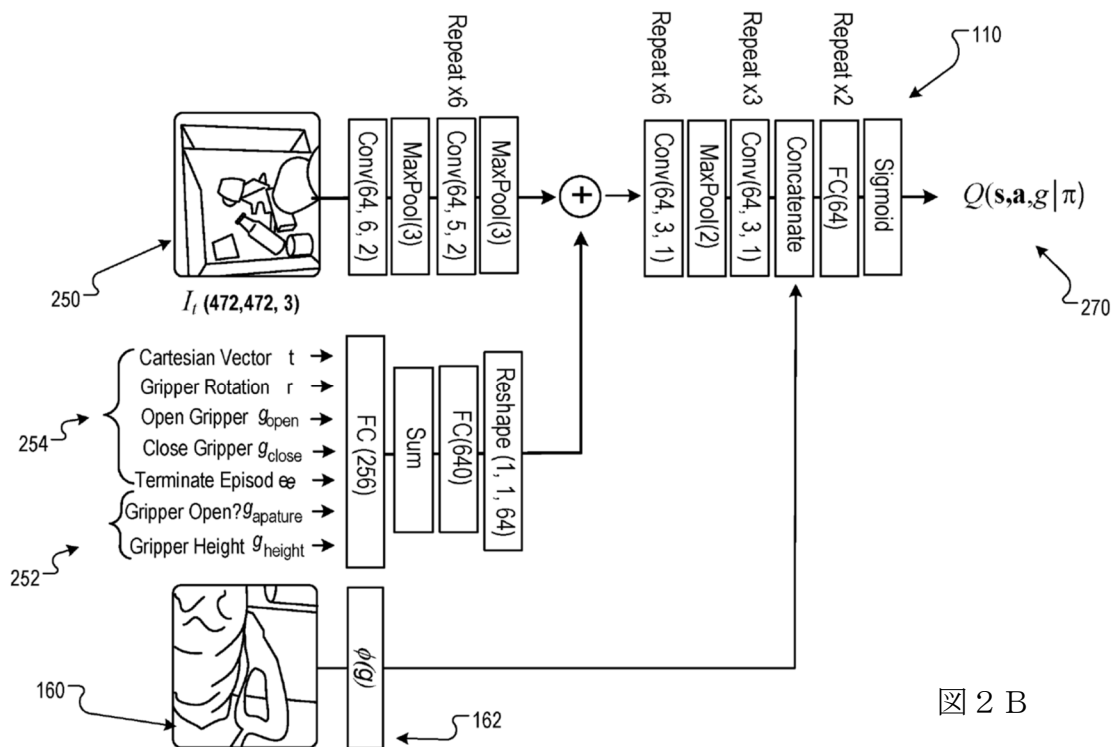


図 2 B

図 2B に示されるように、オブジェクト相互作用タスクニューラルネットワーク 110 は、ロボットの現在の状態を特徴付ける環境 250 の画像およびデータ 252 を含む状態表現を入力として受け取る。

データ 252 は、ロボットのグリッパーが現在開いているかどうか、およびグリッパーの現在の高さを示すデータである。ニューラルネットワーク 110 への入力はまだ、アクションのアクションデータ 254 を含む。

アクションデータには、デカルトベクトル、グリッパー回転値、オープングリッパー値、クローズグリッパー値、およびエピソード終了値が含まれる。オブジェクト相互作用タスクニューラルネットワーク 110 はまた、オブジェクト表現ニューラルネットワーク 112 によって生成されるゴールオブジェクト 160 のベクトル埋め込み 162 を条件とする。

オブジェクト相互作用タスクニューラルネットワーク 110 の出力は、環境が状態表現によって特徴付けられる状態にあり、かつ、タスクエピソードのゴールオブジェクトがゴールオブジェクト 160 である場合、アクション 254 を実行することによって受け取られるリターン値の推定値を表す Q 値 270 である。

図 2B に示されるように、アクションデータ 254 および状態データ 252 は、256 ノ

ードを有する完全結合層(“FC(256)”)、続いて合計層(“sum”)、64 ノードを有する完全結合層(“FC(64)”)、および、FC(64)層の出力を $1 \times 1 \times 64$ テンソルにリシェイプするリシェイプ層(“Reshape(1,1,64)”)を含む完全結合サブネットワークを介して処理される。

画像 250 は、複数の畳み込み層を含む畳み込みサブネットワークを介して処理される。特に、畳み込みサブネットワークには、畳み込み層(“Conv (64,6,2)”)層と“Conv (64,5,1)”)層が 6 回繰り返される) と最大プーリング層(“maxpool”)が含まれる。

畳み込みサブネットワークと完全結合サブネットワークの出力が追加され、追加の畳み込み層(“Conv(64,3,1)”)を介して処理されてから、ゴール埋め込み 162 が連結層によって連結される。次に、この連結された出力は、全結合層(“FC(64)”)を介して処理され、最後に出力層(“concatenate”)を介して処理され、Q 値 270 が生成される。

3.クレーム

944 特許のクレーム 1 は以下の通りである。

1. (i) 複数のオブジェクト相互作用パラメータを有し、(ii) ロボットが実行するアクションを選択して、特定の対象オブジェクトの画像を条件とする環境で、特定の対象オブジェクトとの特定のインタラクションを実行することを含むタスクをロボットに実行させるために使用される、オブジェクト相互作用タスクニューラルネットワークを訓練する方法において、

環境内に現在配置されている複数のオブジェクトから選択されたゴールオブジェクトのゴールオブジェクト画像を取得し、

複数のオブジェクト表現パラメータを有するオブジェクト表現ニューラルネットワークを使用してゴールオブジェクト画像を処理し、前記オブジェクト表現ニューラルネットワークは、複数のオブジェクト表現パラメータの現在の値に従ってゴールオブジェクト画像を処理して、ゴールオブジェクトのベクトル埋め込みを生成するように構成され、

オブジェクト相互作用タスクニューラルネットワークがゴールオブジェクトのベクトル埋め込みを条件とし、複数のオブジェクト相互作用パラメータの現在の値に従って、オブジェクト相互作用タスクニューラルネットワークを使用してロボットによって実行されるアクションを選択することにより、タスクのエピソードを実行するようにロボットを制御し、

エピソード中に実行されたアクションから、一連の状態表現(アクションのペア、各状態表現の状態表現)を生成し、アクションペアは、状態表現のアクションがエピソード中にロボットによって実行されたときの環境の状態を特徴付けており、

エピソード中に、ロボットが環境内の複数のオブジェクトのいずれかに対してタスクを正常に実行したかどうかを判断し、

ロボットが環境内の複数のオブジェクトのいずれかに対してタスクを正常に実行した場合：

環境内の複数のオブジェクトのうちの1つに対してタスクを成功裏に実行するロボットに基づいて1つまたは複数の報酬値を決定し、

タスクが成功裏に実行された複数のオブジェクトのうちの1つの成功したオブジェクト画像を取得し、

成功したオブジェクトのベクトル埋め込みを生成するために、オブジェクト表現パラメータの現在の値に従ってオブジェクト表現ニューラルネットワークを使用して成功したオブジェクト画像を処理し、

成功したオブジェクトのベクトル埋め込みとゴールオブジェクトのベクトル埋め込みの間の類似性測度を決定し、

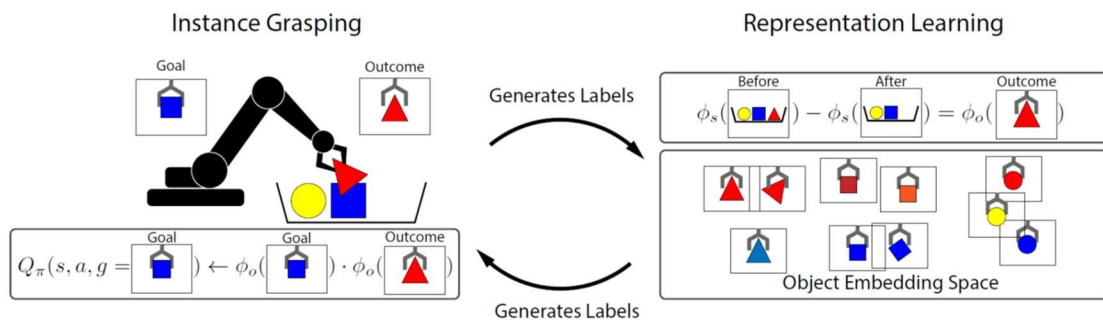
成功したオブジェクトのベクトル埋め込みとゴールオブジェクトのベクトル埋め込みとの間の類似性尺度に基づいて最初の報酬値を決定し、

1つ以上の報酬値の報酬値ごとに、状態表現のシーケンス（アクションペアと1つ以上の報酬値の報酬値）を使用してオブジェクト相互作用タスクニューラルネットワークをトレーニングする。

4. Grasp2Vec に関する論文

本特許に関連する論文¹「Grasp2Vec: Learning Object Representations from Self-Supervised Grasping」が Eric Jang 氏らにより発表されている。

下記図は、Grasp2Vec の学習プロセスを示す説明図である。



インスタンスの把持と表現学習プロセスは、完全に自己教師あり (self-supervised) 法

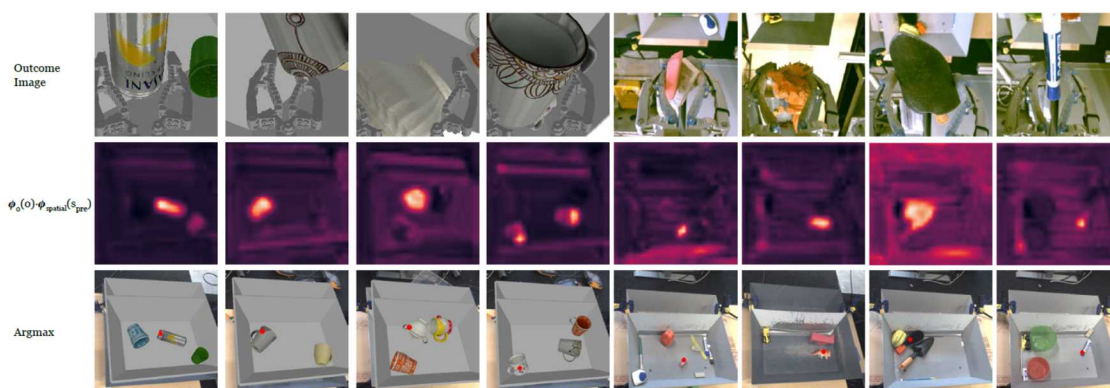
¹Eric Jang, Coline Devin, Vincent Vanhoucke, Sergey Levine “Grasp2Vec: Learning Object Representations from Self-Supervised Grasping”, arXiv:1811.06964v2 [cs.RO] 19 Nov 2018

により互いのラベルを生成する。

把持からの表現学習においては、ロボットアームがシーンからオブジェクトを取り除き、結果のシーンとグリッパー内のオブジェクトを観察する。そして、シーンの埋め込みの相違がオブジェクトの埋め込みと一致するように学習する。

学習した表現による把持の監視においては、オブジェクトの埋め込み間の類似性メトリックを、インスタンの把持の報酬として使用し、把持結果に手動でラベルを付ける必要をなくしている。

下記図は把持結果を示す説明図である。



4 番目の列は、2 つの異なるマグカップが同様の特徴を持ち、**argmax** が間違ったマグカップにあるという失敗のケースを示している。右半分は実際の画像の結果を示している。実際の把持結果で訓練された表現は、高い精度でオブジェクトをローカライズすることができる。

ロボットの動作については Youtube にて公開されている。

https://www.youtube.com/watch?v=QzII_ny4l8s

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 2.0](#)」、「[ブロックチェーン 3.0](#)(共著)」がある。