

建設機械の自動化（自動操縦）を実現する AI の開発

安本 雅 啓

近年、AIの研究が加速している。AIの中でも、模倣学習や強化学習といった技術は、熟練者の行動の模倣や、試行錯誤的な学習によって、人間と同等レベルの作業、あるいは時には人間のパフォーマンスを超えることもできると言われている。建設機械の自動化に対しても、これらのAI技術を活用することで、これまで熟練者が行ってきた操縦をAIが代替し、建設業界における人材不足等の問題を解決することが期待される。本稿では、クレーンに取り付けられた荷物の揺れを抑えるというタスクと、油圧ショベルで掘削を行うというタスクをAIにシミュレーション上で学習させ、人間と同等、あるいは時には人間を超える性能を出せることを確認したので紹介する。

キーワード：AI、模倣学習、強化学習、機械学習、自動操縦、シミュレーション

1. はじめに

2012年に、コンピュータを用いた物体認識の精度を競う国際コンテスト「ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012」で、トロント大学のチームが深層学習と呼ばれる技術を用いて、前年度と比較して大幅な精度向上を達成してから、AIの研究開発は大きく加速した。2016年には、Google傘下の企業であるDeepMindが開発した囲碁プログラム（AlphaGo）が、世界的なプロ棋士に対して勝利するという出来事も起こっている。AIの技術は画像だけでなく、自然言語、音声認識、ゲーム対局など、様々な領域に広がっており、今でも活発な研究は続いている。

これらのアカデミアでの研究の盛り上がりと並行して、産業界でもこれらのAIの研究成果を製品や業務に取り入れようという動きが急速に進んでいる。例えば、これまで人間が行ってきた製造物の検品や作業の監視といった作業を、カメラ+AIで置き換える技術は既に様々な現場で導入が進んでいる。これらのタスクは人間の認知・判断を代替するものであるが、判断した結果を実世界にまでフィードバックするというタスクとして、ロボットアームを用いたピッキング作業の自動化なども盛んに研究されており、今後、同様に産業界への適用が見込まれる。

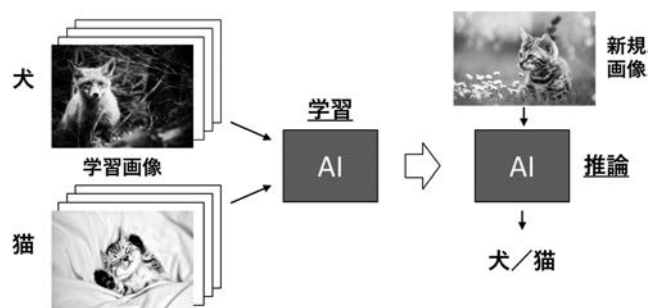
建設機械についても、ロボットアームと同じように、適切なセンサとAIをつなげることで、その操作

を自動化することができると考えられる。本稿では、シミュレータ上で建設機械の自動操縦タスクをAIに学習させたので、その内容と成果を報告する。

2. AI 技術

昨今、様々な場面でAIという言葉聞くことがあり、拡大解釈して使われている場面も多く見受けられる。ただ、最近のAIの盛り上がりは、専ら深層学習に端を発したものであり、これは機械学習の技術のカテゴリに入るものである。そこで本報では、機械学習のことをAIと呼ぶこととする。

機械学習とは、与えられたデータの中から法則性を見出すことである。例えばよくある例に、猫と犬の識別を行うというタスクがある(図—1参照)。AIには、犬と猫のラベルがついた画像が大量に与えられ、それらのデータを用いて学習を行う(学習フェーズと呼



図—1 機械学習タスクの例：犬と猫の分類

ぶ)。その後、今までに見たことがない犬または猫の画像を見せられたときに、それが犬か猫かを当てる、というものである(推論フェーズと呼ぶ)。このときAIは、大量の犬と猫の画像から、犬に共通する性質、猫に共通する性質を見出し、この性質(機械学習では特徴量と呼ばれる)に基づいて、犬か猫かを分類できるようにする。

今までに見たことがない新しい犬や猫を判別できるかどうかという性質を汎化性と呼ぶが、よい特徴量を見つけていれば、高い汎化性を得られるが、よくない特徴量を見つけてしまった場合には、汎化性は低くなる。例えば、学習用のデータに、黒い猫の画像と、白い犬の画像しか入っていないと、AIは、動物の色によって猫か犬かを判別できると「勘違い」してしまう。しかし、実際には色という特徴量は間違っており、例えば新しく白い猫を見せたときには、犬と誤認識してしまう。そのため、AIに正しい特徴量を獲得させるためには、多様性がある、大量のデータが必要となる場合が多い。

これから建設機械の自動化で用いる2つの技術「模倣学習」と「強化学習」は、いずれもこの機械学習に属する技術である。人間があらかじめ熟練者の作業データを取得し、これを用いて学習を行うのが模倣学習、エージェントが自ら環境とインタラクションをすることで学習データを生成し、これを用いて学習を行うのが強化学習である。以降、各手法の詳細について述べる。

(1) 模倣学習

建設機械を操作するとき、人間は状況を知覚し(センシング)、知覚に基づいて操作をする(アクション)。このセンシングとアクションの対をデータとして保存し、AIにも同じようなことができるように学習させるというのが模倣学習である。学習したAIは、人間と同じように、センシング情報(例えば画像やセンサの値など)を与えることで、アクションを出力する。

模倣学習の特徴としては、十分な学習データがあれば、比較的短い時間で学習ができる。ただし、学習データについては、多様性が必要となる。ここで言う多様性とは、なるべく多くの環境条件におけるデータという意味である。前節でも述べたように、データに偏りがあると、誤った特徴量を抽出してしまう可能性がある。そのため、想定しうる様々な条件でデータを取得する必要がある。

また、学習データを取得する場合は、なるべく同じ人のデータを収集した方がよい。これは、ある状況に

おけるアクションの取り方が人によって異なる場合、AIはどちらの行動を出力すべきか、うまく学習ができないためである。同じ人であれば、基本的には同じ状況に直面したときには同じアクションを取ると想定されるため、なるべく同じ人のデータを取得することが望ましいと考える。また、その人のやり方を良い部分も悪い部分も同様に真似するように学習するため、どういった人を選ぶかにも考慮が必要となる。

(2) 強化学習

強化学習とは、エージェントが環境内で探索を行い、その中で得られた報酬をもとに、学習を行っていく手法である。最初はほぼランダムに近い動作を行うが、動作する中で、どの行動に対して報酬が得られたのかという情報をもとに、より多くの報酬を得られるように自らの行動の方策を改善していく。

強化学習の特徴としては、報酬(一般的には関数として記述される)を上手く設計することができれば、人間のお手本なく、エージェントが自ら学習し、最適な行動方策を学習できる。また場合によっては、人間よりも高いパフォーマンスを実現できる可能性がある。

デメリットとしては、一般的に膨大な数の試行錯誤を必要とするため、学習に時間がかかること、試行錯誤を行うので、実世界だと危険を伴う可能性があること、学習が不安定なことなどが挙げられる。また、一般的に、適切な報酬関数を設計するのが難しい場合が多い。報酬関数が適切でない場合は、その適切でない部分を利用しようとする動作(報酬ハッキングと呼ばれる)を学習してしまうことがある。

(3) 模倣学習と強化学習の比較

基本的には、熟練者のデータをとることができるのであれば、模倣学習からスタートした方がよいと思われる。理由は、強化学習を用いる場合、膨大な数の試行錯誤が必要となる場合があり、現実世界でこれを行うのは難しいためである。また、強化学習は試行錯誤を通して学習していくため、学習最初はあまり適切でない動作を出力してしまう場合も多く、現実世界で行うのは危険であることが多い。

強化学習のこれらのデメリットを軽減するような研究も行われている。例えば、模倣学習でAIの学習を行った後、この学習済みAIを初期値としてさらに強化学習で学習を行うという方法が提案されている。この方法により、強化学習における試行錯誤の回数を減らすことができると、初期動作が模倣学習によって

作られたある程度よいものとなるため、危険な動作をするリスクが減るというメリットがある。

3. 実験設定

建設機械の自動操縦タスクにおける模倣学習・強化学習の有効性検証のため、シミュレータを用いた実験を行った。シミュレータを用いれば、強化学習については、比較的多くの試行錯誤を行うことができるのと、仮に危険な動作をしたとしても問題ないというメリットがある。また、模倣学習についても、比較的短時間で熟練者のデータを多数取得することができる。

自動化の対象とするタスクとして、クレーンによる吊り荷の揺れ止めと、油圧ショベルによる土砂の掘削の2つを選定した。これらのタスクは、土木・建設工事において基本的なタスクであるが、制御モデルの設計を行うことが比較的難しく、また人間が操作するにしてもある程度の訓練・経験が必要となる、という観点で選んだ。

これらのタスクをシミュレートするためには、クレーンのワイヤーのたわみや、土砂の動きのシミュ

レーションができる必要がある。そこで、これらの動作のシミュレーションが可能な Vortex シミュレータを使用した。Vortex シミュレータでの建設機械の動作イメージを図-2 に示す。

(1) クレーンの揺れ止め

クレーンのワイヤーの先に荷物を取り付けて吊るし、荷物を移動させようとする際、大きく動かすと荷物が揺れてしまう場合がある。その際、揺れが止まるように制御する必要がある。ここでは、静止した状態から、まずクレーンのブームを大きく左右に振り、揺れが発生した時点をスタートとし、クレーンのブームを上下左右に動かすことで、揺れを止めるというタスクを設計した。

AI への入出力を図-3 に示す。入力としては、センサ値を入力とするケースと（図中ケース1）、深度画像を入力とするケース（図中ケース2）の2つを試した。ケース1は、加速度センサなどを機体に取り付けることでデータを取得することを想定、ケース2は、運転席に取り付けたステレオカメラや LIDAR センサでデータを取得することを想定している。AI は

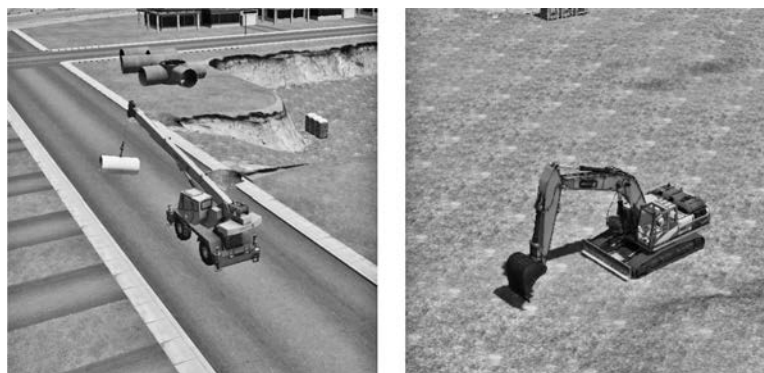


図-2 Vortex シミュレータでの建設機械の動作イメージ

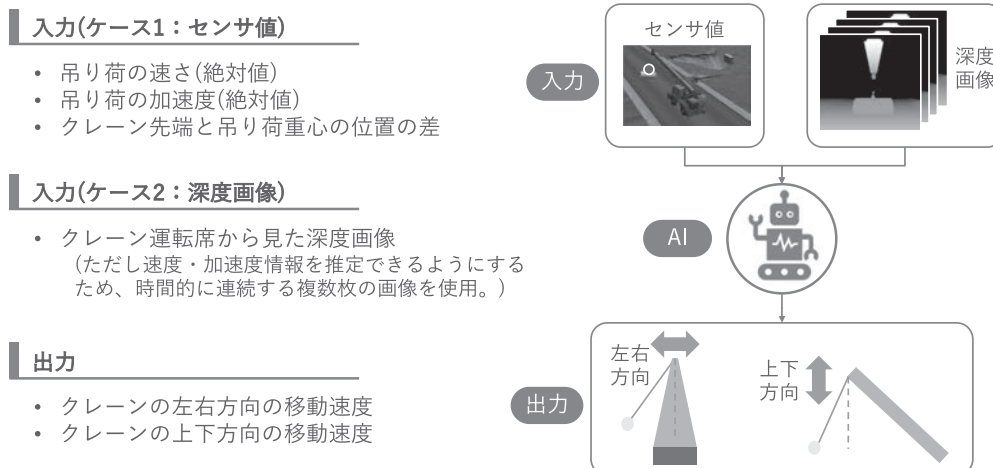
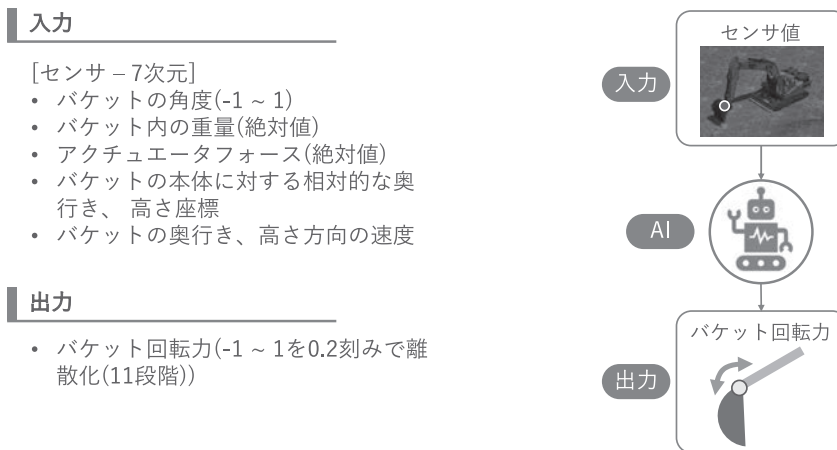


図-3 クレーンの揺れ止めタスクにおける AI の入出力情報



図一4 油圧ショベルの掘削タスクにおけるAIの入出力情報

これらを入力として受け取り、処理した結果として、クレーンの左右方向と上下方向の移動速度を出力する。

クレーン先端と吊り荷重心の位置の差が 1.5 m 以下かつ、吊り荷速度が 1.5 m/s 以下の状態が数秒間連続した時、揺れがおさまったとみなし、タスク成功として定義した。

(2) 油圧ショベルの掘削

油圧ショベルで地面の土の掘削を行う際、一度にどの程度の土を掘削できるかは、ショベルの先についたバケットの角度をどのように制御するかによって大きく変化する。このバケット角度制御を自動化し、なるべく多くの土をバケットに入れるというタスクを設計した。

AIへの入出力を図一4に示す。バケットの角度などのセンサ情報を入力とし、バケットの回転力を出力するようにAIを学習する。

最終的にバケットの中に入っている土の量がある一定値を超えた場合を、タスク成功として定義した。

(3) 学習アルゴリズム

模倣学習では、MARWIL (Monotonic Advantage Re-Weighted Imitation Learning) と呼ばれるアルゴリズムを用いた。通常まず用いられるナイーブな模倣学習の手法 (Behavior Cloning と呼ばれる) が、与えられた教師データ全てに対して同じ重みで学習を行うのに対し、MARWILは、報酬獲得のために有利な行動を重点的に模倣するという特徴を備えている。結果として、理論上は、Behavior Cloning よりも高い報酬を獲得可能と言われている。

強化学習では、PPO (Proximal Policy Optimization) と呼ばれるアルゴリズムを用いた。本アルゴリズムの

特徴としては、学習が比較的安定しているのと、学習並列化による高速化が可能である点などが挙げられる。

4. 実験結果

(1) クレーンの揺れ止め

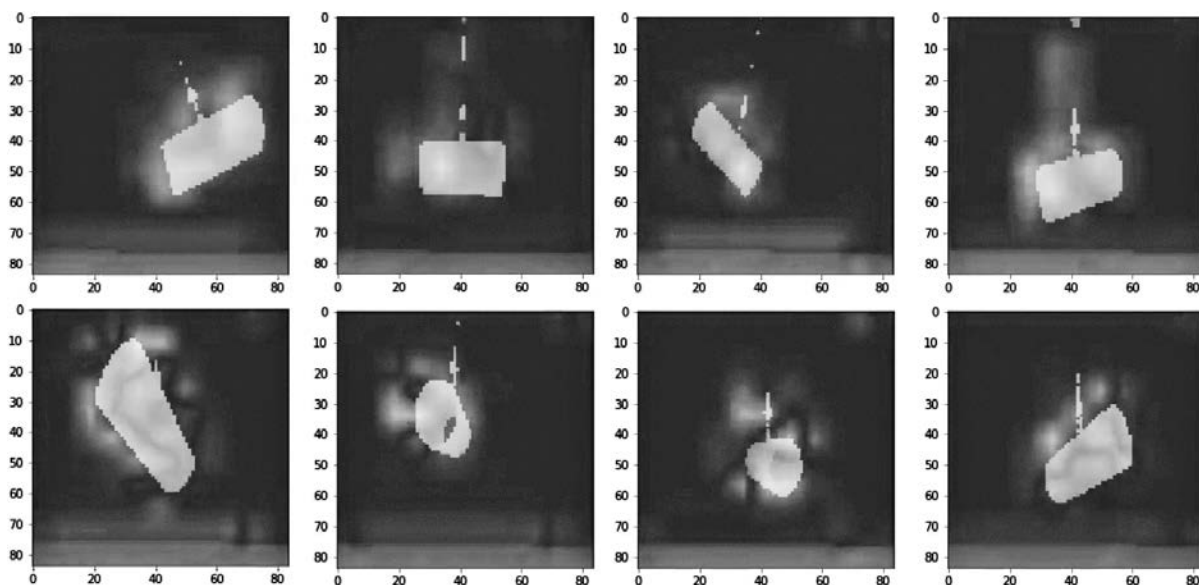
クレーンの揺れ止め実験の結果を表一1に示す。模倣学習については、入力がセンサ、深度画像のいずれの場合でも90%を超える成功率を達成できていることが分かる。強化学習については、入力がセンサの場合でも63.0%、深度画像の場合だと0%の成功率となった。

強化学習で成功率が低い原因としては、クレーンの揺れ止めタスクは、最初の数秒間、慎重に制御をしないと、加速度が逆に大きくなってしまい、または地面に衝突してしまって失敗となってしまい、なかなか成功体験を得るのが難しい、という点が想定される。強化学習は一般的に、成功した事例をもとに行動方策の改善を行うため、成功事例が全くない状態だと学習が進まない。そこで、模倣学習を用いてあらかじめ性能を向上させた方策を初期方策として強化学習を行うといった改善策が考えられる。

深度画像を入力とした模倣学習のケースについては、最終的に学習された制御方策に対して、エージェントがどこに注目しているかの可視化を行った。結果を図一5に示す。吊り荷の注目度が一貫して高く、時々、地面付近の注目度も増加していることが分か

表一1 クレーンの揺れ止めタスクの成功率

タスク	クレーン揺れ止め	
	センサ	深度画像
模倣学習での成功率	91.7%	92.0%
強化学習での成功率	63.0%	0.0%



図一5 AIの注目点の可視化結果

表一2 油圧ショベルの掘削タスクの成功率

タスク	油圧ショベル掘削
入力	センサ
模倣学習での成功率	81.0%
強化学習での成功率	96.0%

る。また、吊り荷の輪郭を抽出しているように見える。地面の注目度が高いのは、地面との接触を避けるためには地面と吊り荷との間の距離を把握しておく必要があるためと考えられる。また吊り荷の輪郭の注目度が高いのは、輪郭を抽出しその時間差分をとることで、揺れ止め制御に不可欠な、吊り荷の速度や加速度といった情報を算出するのが理由ではないかと考えられる。

なお、強化学習によって学習が進む様子は、<https://youtu.be/Ma-fFFWhWsI>で見ることができる。

(2) 油圧ショベルの掘削

油圧ショベルの掘削実験の結果を表一2に示す。模倣学習で成功率81.0%なのに対し、強化学習での成功率はこれを超える96.0%であった。人が操作しても、4回に1回は失敗（所定の土量を超えられない）してしまうことがあることを考慮すると、これは人間を超える性能を達成できていると言える。

なお、強化学習によって学習が進む様子は、<https://youtu.be/32Y4OLVCFrY>で見ることができる。

5. おわりに

模倣学習と強化学習によって、クレーンによる揺れ

止めと油圧ショベルによる土の掘削のタスクで、高い成功率を達成することができた。模倣学習については、いずれのタスクでもある程度高い成功率をコンスタントに実現できていた。強化学習については、油圧ショベルのタスクでは模倣学習を超える成功率を達成できている一方、クレーンのタスクでは低い成功率となり、学習の難しさを改めて実感した形となった。

今後は、これら技術の実機での検証を進めていく予定である。実機ではシミュレーションと異なり、データ取得のコストが高いのと、強化学習による試行錯誤に危険が伴うといった課題が存在する。これらの課題に対し、シミュレーションであらかじめ学習したものを初期値として用いるといった対策が必要になると考えている。

謝辞

Vortex シミュレータのご提供、および、建機自動操縦タスクの選定にあたっては、(株)電通国際情報サービスの飯田倫崇様、原悟様より、多大なご支援・ご協力を頂きました。誌面をお借りして厚くお礼を申し上げます。

JCMA

【筆者紹介】
安本 雅啓 (やすもと まさひろ)
(株)アラヤ
取締役 CTO

