

iCOIL: 制約付き最適化と模倣学習の統合によるシナリオを考慮した自律駐車

Lexiong Huang^{1,2,*}, Ruihua Han^{1,3,*}, Guoliang Li¹, He Li⁴, Shuai Wang^{1,†}, Yang Wang^{1,†}, and Chengzhong Xu⁴

概要 自律駐車(AP)は、人間の介入なしにインテリジェントな車両を駐車スペースにナビゲートする新しい技術である。数学的最適化や機械学習に基づく既存のAP手法は、実行時間が長すぎたり、汎化が不十分であったりするために、潜在的な失敗につながる可能性がある。このギャップを埋めるために、本論文では、効率的で信頼性の高いAPを実現するための統合制約付き最適化・模倣学習(iCOIL)アプローチを提案する。iCOIL法は、COとILの2つの作業モード候補を持ち、ハイブリッドシナリオ解析(HSA)モデルを採用して、様々なシナリオの下でより良いモードを決定する。マカオ・カーレース・メタバース(MoCAM)プラットフォーム上でiCOILを実装し、検証する。結果は、iCOILがAP手順全体を通して異なるシナリオに適切に適応し、他のベンチマークよりも有意に大きな成功率を達成することを示している。

I. INTRODUCTION

自律駐車(AP)は、インテリジェント運転の中核タスクであり、人間が介在することなく、車両を現在位置から駐車スペースに移動させるための一連の制御コマンドを決定する[1]–[4]。次世代APは、APの範囲を数メートルから数十メートルに拡大することで、より多くの人間のドライバーの駐車時間を節約することを目的としている[5], [6]。しかし、AP範囲が長いと、動的なシナリオでの高い適応と、衝突のない軌道の高速生成が必要となる。最適化ベースのアルゴリズム[1], [6]は計算複雑度が高く、実行時間が長すぎるため、失敗の可能性がある。一方、深層学習アルゴリズム[2]–[5]は、フィードフォワード演算により、入力画像から駆動信号をミリ秒単位で生成することができる。しかし、ターゲットシナリオに学習データセットの分布外のデータが含まれていれば、破綻する可能性がある。その結果、現在のところ、優れた汎化と高い計算効率を同時に達成するAPアルゴリズムは存在しない。

本論文では、最適化手法と学習手法の統合を検討し、異なるシナリオにおける2つの手法間のインテリジェントな切り替えを実現することを目的とする。しかし、このような統合には、以下のような技術的な課題がある：

本研究は、マカオ科学技術開発基金(FDCT) (No. 0081/2022/A2)、広東省基礎応用基礎研究プロジェクト (No. 2021B1515120067)、深圳科学技術プログラム (No. RCB20200714114956153)、深圳先端技術研究所と直接駆動技術との協働プロジェクトの支援を受けた。

* Equal contribution

† Corresponding author: Shuai Wang (s.wang@siat.ac.cn) and Yang Wang (yang.wang1@siat.ac.cn).

1 Center for Cloud Computing, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences

2 University of Chinese Academy of Sciences, Beijing 100049, China

3 Department of Computer Science, University of Hong Kong

4 IOTSC, University of Macau



(a) Virtual simulation

(b) Real-world testbed

図1: MoCAM上でのiCOILの実装。

- (1) 学習アルゴリズムのシナリオの不確実性シナリオは、複雑(静的オブジェクトと動的オブジェクトを含む)、異種(照明や天候の異なる非IIDデータを含む)、マルチモーダル(画像、点群、その他の感覚データを含む)[7]である。このため、与えられたシナリオが学習アルゴリズムに脅威を与えるかどうかを判断することは困難である。
- (2) **Scenario complexity for optimization algorithms.** The computational time of optimization is proportional to the size of state-action space and the number of surrounding objects [8]. This makes it challenging to determine whether a given scenario poses a threat to optimization.
- (3) **Real-time implementation of the AP system.** This requires highly efficient implementation of the learning and optimization algorithms, and a fast switching mechanism between two algorithms.

上記の課題を解決するために、本論文では、効率的で信頼性の高いAPソリューションとして機能する、統合制約付き最適化・模倣学習(iCOIL)を提案する。iCOILは、COを活用して、数学的な状態進化モデルによって理論的な保証のある衝突のない軌道を生成し、ILを活用して、ディープニューラルネットワーク(DNN)によって人間の専門家の論理と行動を模倣する。COモードとILモード間の移行を可能にするため、iCOILはハイブリッドシナリオ解析(HSA)モデルを採用し、情報エントロピーに基づくシナリオの不確実性と、自車両と障害物の形状に基づくシナリオの複雑性を計算する。iCOILソリューションはロボットオペレーティングシステム(ROS)パッケージとして実装され、CARLA-ROSブリッジを介してCar Learning to Act(CARLA)[9]に接続される。このように、提案されたiCOILベースのAPシステムは、Macao Car Racing Metaverse (MoCAM)プラットフォーム上で検証された。MoCAMは、高忠実度の仮想シミュレーション(すなわち、図1a)を通じてAPの設計と検証を可能にし、その結果をROSインターフェースを通じて実世界のテストベッド(すなわち、図1b)に転送するデジタルツインシステムである。また、iCOILとILの成功率と駐車時間を比較する。実験結果は、iCOILの優位性を確認するものである。ソースコードはオープンソースのROSパッケージとして公開される。

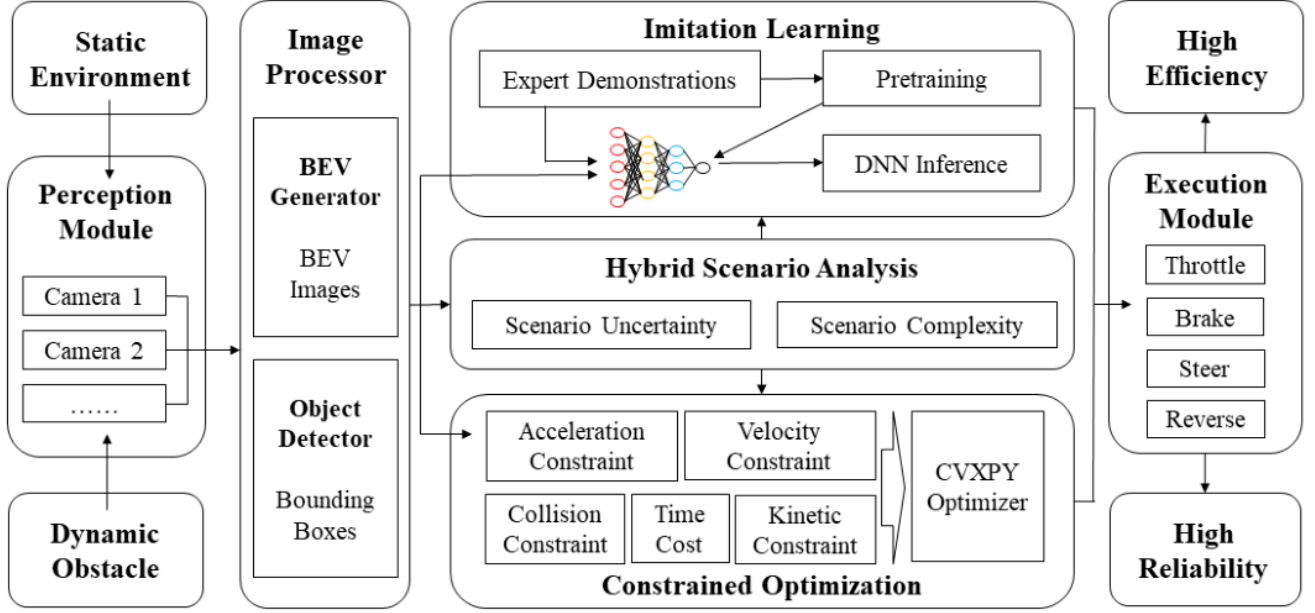


図2:提案するiC0ILベースAPシステムの全体アーキテクチャ。

II. 関連研究

学習ベースのアルゴリズムは、入力シナリオ(すなわち、一連の画像または点群)から運転行動(すなわち、スロットル、ブレーキ、ステア、リバース)へのブラックボックス推論マッピングと見なすことができる[2]–[5]、[10]、[11]。感覚と行動のデータセットを仮定せず、逆伝播によってデータセットから直接学習する脳のようなDNNを構築している[10]。このような方法は、効率的な特徴抽出と表現を提供し、また、非反復的なフィードフォワード操作による高速な推論を提供する。学習ベースのアルゴリズムは、強化学習(RL)[3]、[4]、[10]、[12]とIL[2]、[5]、[11]、[13]、[14]に分けられる。RLアルゴリズムは、環境や他のエージェントと相互作用することで、インテリジェントな車両が経験から学習することを可能にする[10]。それにもかかわらず、RLベースのアプローチは、長いトレーニングに悩まされ、限られたシナリオでのみ機能し、センサーノイズに敏感である。ILベースのアルゴリズムは教師あり学習に属し、デモデータセットから学習することで専門家の運転方針を模倣する[2]、[5]、[11]。しかし、模倣学習の有効性は、データセットの質とDNNの汎化能力に依存する。データの品質を向上させるために、HG-DAGGERが[15]で提案され、実世界システムにおける人間の専門家からの対話的なデータセット収集をサポートしている。一方、DNNの汎化能力を向上させるために、条件付きIL[11]や連合IL[16]がDNNの選択と集約のために提案されている。

最適化ベースのアルゴリズムは、目標ウェイポイントと実際のウェイポイント間の距離を最小化するために、手作業で設計された数学的操作(主に反復)である [1]、[6]、[17]、[18]。これは、有限水平マルコフ決定過程を用いて未来と現在の状態の関係を構築するモデル予測制御(MPC)[1]によって実現することができる。学習ベースのアルゴリズムと比較して、最適化ベースのアルゴリズムは、車両が安全領域内で移動することを強制する、境界のあるアクションを生成することができる[18]。

グリッド探索やランダムサンプリングなどの従来のヒューリスティックと比較して、最適化ベースのアルゴリズムは、その解が局所最適または大域最適であることが保証されているため、より高い運転性能を達成することができる[17]。C0ベースのアルゴリズムは、障害物のある乱雑な環境に対する最適化アルゴリズムの制約付きバージョンであり、衝突回避制約が非凸であるため、制約のない対応するものよりも導出が困難である[1]。さらに、障害物の数が増えるにつれて、これらのアルゴリズムは学習や発見的アルゴリズムよりも長い実行時間を必要とする[18]。そのため、C0ベースのアルゴリズムに関する最近の研究は、主に非凸性を解決する方法 [1]、[6] や計算を高速化する方法 [18] に焦点を当てている。

III. コイルベースのAPシステム

APシステムのアーキテクチャを図2に示す。ここで、 $X = \{x_1, x_2, \dots\}$ 、 x_i は*i*番目のフレームのエゴビュー画像、 $A = \{a_1, a_2, \dots\}$ 、 a_i はスロットル、ブレーキ、ステア、リバース要素を含む*i*番目のフレームの行動ベクトルである。*i*番目のフレームでは、知覚モジュール(すなわち、図2の左側)が、搭載カメラを介して環境を画像 x_i にマッピングする。画像処理装置(すなわち知覚モジュールの隣)には、鳥瞰図(BEV)変換器 $y_i = g(x_i)$ と物体検出器 $z_i = h(y_i)$ が含まれる。ここで、 y_i と z_i はそれぞれ*i*番目のフレームにおけるBEV画像とバウンディングボックスを表すベクトルである。図3に示すように、関数 g はエゴビュー画像(すなわち図3a)をBEV画像(すなわち図3b)に変換し、関数 h は障害物のバウンディングボックスを生成する(すなわち図3c)。提案するiC0IL(すなわち、図2の中央部)は、IL $f_{IL}(-)$ 、C0 $f_{CO}(-)$ 、HSA $f_{HSA}(-)$ モジュールを含み、以下に詳細に示す。

ILモジュール(すなわち、図2の中央上部)は、BEV画像をアクションにマッピングする、すなわち、 $a_i = f_{IL}(y_i | \theta^*)$ 、ここで θ^* は事前学習されたDNNパラメータである。



図3: iCOILベースAPシステムの入力データ、生成されたBEV画像、検出されたボックス。

この θ^* は、専門家のデモを収集し、そのデータセットでDNNを学習することで得られる。一方、C0モジュール(すなわち、図2の中央下段)は、検出されたバウンディングボックス z_i を利用して、衝突のない行動、すなわち、 $a_i = f_{CO}(z_i)$ を生成する。この f_{CO} は、加速度、速度、運動量、衝突回避の制約条件下で時間効率コストを最小化するAP最適化問題を解くことによって得られる。

HSAモジュール(すなわち、図2の中央)は、過去期間(すなわち、 $t = i-T$ から $t = i$ までの動的時間窓)のフレーム i における平均シナリオ不確実性 U_i と平均シナリオ複雑度 C_i を計算し、それらを運転モードを決定する指標として使用する関数 $f_{HSA}(\{x_t\}_{t=i-T}^i)$ である。平均不確かさ(複雑さ)がある閾値を超えると、現在のシナリオはIL(C0)アルゴリズムに脅威を与え、我々のAPシステムはより高い信頼性(効率)を得るために他のモードに切り替えることになる。その結果、 $f_{HSA}(\{x_t\}_{t=i-T}^i) = U_i C_i^{-1}$ とし、モード切替に以下の条件を採用することができる：

$$f(x_i) = \begin{cases} f_{IL}(g(x_i)|\theta^*), & \text{if } f_{HSA}(\{x_t\}_{t=i-T}^i) \leq \lambda \\ f_{CO}(h(g(x_i))), & \text{if } f_{HSA}(\{x_t\}_{t=i-T}^i) > \lambda \end{cases}, \quad (1)$$

ここで、 λ は経験的に微調整される予め定義された閾値である。

IV. THE iCOIL ALGORITHM DESIGN

A. IL

提案するILモジュールでは、ILを多カテゴリ分類問題として定式化するために、連続運転行動を離散値に変換する[2], [5], [11]。これにより、特徴抽出ネットワークと状態アクションネットワークからなるDNNを活用することが可能となり、前者は画像から高次元の特徴を抽出し、後者は特徴をアクションにマッピングする。特徴抽出ネットワークは3つの層を採用し、各層は畳み込み、ReLU活性化、最大プーリングの3つのユニットを持つ。状態-行動ネットワークは4つの完全接続層と、確率的行動を出力するソフトマックス層を持つ。最も確率の高いアクションが、自車両での実際の実行のために選択される。DNNの全パラメータをまとめて θ と表記する。

上記の構造に基づき、DNN学習は θ に対する最適化問題であり、DNN出力とエキスパートアクションの差を最小化することを目的とする：

$$\theta^* = \arg \min_{\theta} \sum_{(x'_i, a'_i) \in D} L(f_{IL}(g(x'_i)|\theta), a'_i), \quad (2)$$

ここで、 D は学習データセット、 $L(-)$ はクロスエントロピー関数を表す。

$$L = -\frac{1}{|D|} \sum_j A'_{i,j} \log(f_{IL,j}^{\text{Prob}}(g(x'_i)|\theta)) \quad (3)$$

ここで、 $|D|$ は D のカーディナリティ(すなわちサンプル数)、 M は行動離散化後のクラス数、 $A'_{i,j}$ は a'_i の j 番目の要素、 $f_{IL,j}^{\text{Prob}}(-)$ は j 番目のDNN出力の確率値を表す。

B. CO

C0モジュールは、アッカーマン運動モデル $s_{i+1} = u(s_i, a_i)$ を活用して、軌跡 s_{i+1}, \dots, s_{i+H} を推定する。ここで、 s_i は $h(g(x_i))$ から得られるフレーム i における自車両の位置と向きを表し、 H は予測水平線の長さである。以上の動力学モデルにより、すべての $h \in [0, H-1]$ に対して関数 $s_{i+h+1} = v_{h+1}(s_i, \{a_t\}_{t=i}^{i+h})$ を導出できる。APの目標は、最短時間を使って車両を駐車スペースに移動させることである。これは距離コストを最小化することで実現できます[1]。

$$C(\{a_t\}_{t=i}^{i+H-1}) = \sum_{h=0}^{H-1} \|v_{h+1}(s_i, \{a_t\}_{t=i}^{i+h}) - s_{i+h+1}^*\|^2, \quad (4)$$

ここで、 $\{s_i^*\}$ は目標ウェイポイントのリストを表し、例えば、現在位置から目標駐車スペースまでの最短経路を表す。一方、自車両は周囲の物体と衝突することはない。

$$\|v_{h+1}(s_i, \{a_t\}_{t=i}^{i+h}) - o_{h+1,k}\| \geq d_{\text{safe},k}, \forall h, k \quad (5)$$

ここで、 $o_{h,k}$ はフレーム $i+h$ における k 番目の障害物の位置、 $d_{\text{safe},k}$ は障害物 k の大きさに関係する安全距離である。以上の解析から、無衝突APは以下のC0問題[18]として定式化される：

$$\{a_t^*\}_{t=i}^{i+H-1} = \arg \min_{a_t \in A, \forall t} \{C(\{a_t\}_{t=i}^{i+H-1}) : (5)\}, \quad (6)$$

ここで、 A は運転行動の境界集合である。上記の問題は、目的と制約の両方において関数 $\{v_{h+1}(-)\}$ により非凸である。原始問題をいくつかの凸最適化問題に変換し、オープンソースの最適化ソフトウェア(CVXPYなど)を用いて解く。フレーム i の出力アクションは $f_{CO}(h(g(x_i))) = a_i^*$ とする。

表1: シナリオの複雑さに関するパラメータ

Type	Symbol	Parameters
Ego-vehicle	H	行動空間の予測水平線の長さ
Ego-vehicle	N_a	の次元
Obstacle	D_0	フレームjにおける障害物kの最も
Obstacle	$D_{j,k}$	危険な障害物距離距離

C. HSA

HSAモジュールは、過去Tフレームにわたるシナリオの不確実性 U_i とシナリオの複雑さ C_i を計算する必要がある。特に、シナリオの不確実性はILの「信頼度」を測定し、DNN出力の確率的分布に関する[19], [20]。例えば、偶数分布はより大きな不確実性を示し、偶数でない出力分布はより小さな不確実性を示す[19], [20]。その結果、フレームiにおけるインスタントシナリオの不確実性 ω_i は、DNN出力の確率分布 $\omega_i = -\frac{PM}{j=1} p_{i,j} \log p_{i,j}$ のエントロピーとして定義される。過去Tフレームの ω_i を合計すると、次のようになる。

$$U_i = \frac{1}{T} \sum_{h=0}^{T-1} \omega_{i-h} = -\frac{1}{T} \sum_{h=0}^{T-1} \sum_{j=1}^M p_{i-h,j} \log p_{i-h,j}. \quad (7)$$

一方、シナリオの複雑さはC0の計算遅延を測定するものであり、(6)を解くための計算複雑度に比例する。過去Tフレームの計算量を平均すると、シナリオの平均複雑度は[18]となる。

$$C_i = \frac{1}{T} \sum_{h=0}^{T-1} \left[H \left(N_a + \sum_{k=1}^K e^{-\|D_0 - D_{i-h,k}\|} \right) \right]^{3.5}. \quad (8)$$

ここで、パラメータ H , N_a , D_0 , $D_{j,k}$ は表Iで定義される。 C_i の解釈: まず、式(8)の3.5のべき乗は、計算遅延が予測地平の長さや障害物の数に対して超線形に増加することを示す。第二に、式(8)の項 $e^{-D_0 - D_{j,k}}$ は、障害物の一部のみが計算量に影響を与える可能性があることを意味し、 D_0 は最も危険な障害物位置である。これは、障害物が非常に近い場合、計画空間が縮小され、障害物が遠い場合、衝突確率がゼロに近くなるためである[8]。

V. EXPERIMENTS

A. 実装の詳細

提案するiCOILbased APの性能を検証するために、Macao Car Racing Metaverse (MoCAM)プラットフォームを採用した。MoCAMは、実世界のサンドボックスとCARLA[9]をベースに開発された高忠実度仮想シミュレータからなるデジタルツインシステムである。iCOILは3つのPython ROSノード(すなわち、IL、CO、HSAノード)として実装され、CARLA-ROSブリッジを介してMoCAMに接続される。MoCAMノードとiCOILノードの他に、既製のオープンソースソフトウェアであるBEV変換ノードと物体検出ノードも配置する。これらの登録ノード間のデータ共有は、ROS通信によって実現され、ノードはエゴビュアの情報を運ぶROSトピックを公開または購読する。

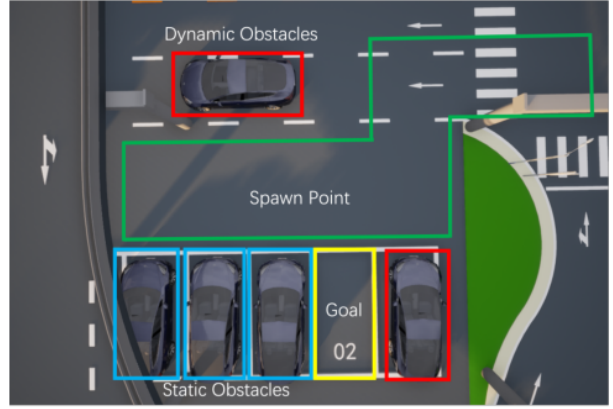


図4: CARLAにおけるAPシナリオのシミュレーション。

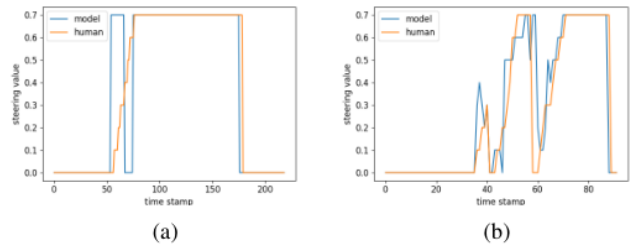


図5: ILと人間のドライバーのステアリング値。

画像、BEV画像、バウンディングボックス、制御コマンド。提案するiCOILの他に、APに多層DNNを直接採用した従来のIL方式[2]も実現した。すべての実験は、Intel i9 CPUとNVIDIA 3080 GPUを搭載したデスクトップ上で実施した。

B. Settings

図4に示すマップを考える。ここでは、自車両の始点は産卵地点領域(すなわち、緑色の領域)内にランダムに生成され、駐車スペースはゴール領域(すなわち、黄色のボックス)内にある。APタスクは、イージーレベル(青でマークされた3つの静的障害物のみ)、ノーマルレベル(赤でマークされた3つの静的障害物と2つの動的障害物)、ハードレベル(すべての障害物が存在し、入力画像とバウンディングボックスにノイズを追加する)に分類できる。敵対的データは実世界の不確実性をシミュレートし、提案するiCOILの頑健性を検証することができる。

C. 提案するiCOILベースAPの検証

IL DNNを学習するためのドライビングデータセットを生成するためにMoCAMを採用する。5171個のデータサンプル(前進2624個、逆転2547個)を収集し、300エポック後に学習プロセスを終了した。ILと人間のドライバーが生成したステアリング動作を図5で比較する。学習されたILは、人間のドライバーと同様の動作を生成していることがわかる。しかし、行動空間に離散化が適用されるため、ILの行動曲線は人間のドライバーの行動曲線よりも段差があり、滑らかでない。iCOILベースのAPの有効性を検証するために、始点から終点までの完全な駐車プロセスを図6aに示す。

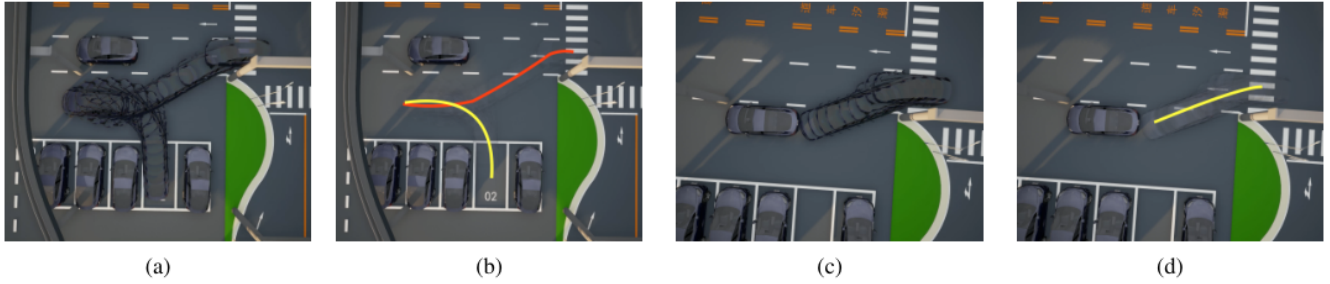


図6:iCOILとILの駐車プロセスと軌跡の比較。(a) iCOILベースAPの駐車プロセス。(b) iCOILベースAPの軌跡。(c) ILベースAPの駐車プロセス。(d) ILベースAPの軌跡。赤い曲線はC0モード、黄色い曲線はILモードを表す。

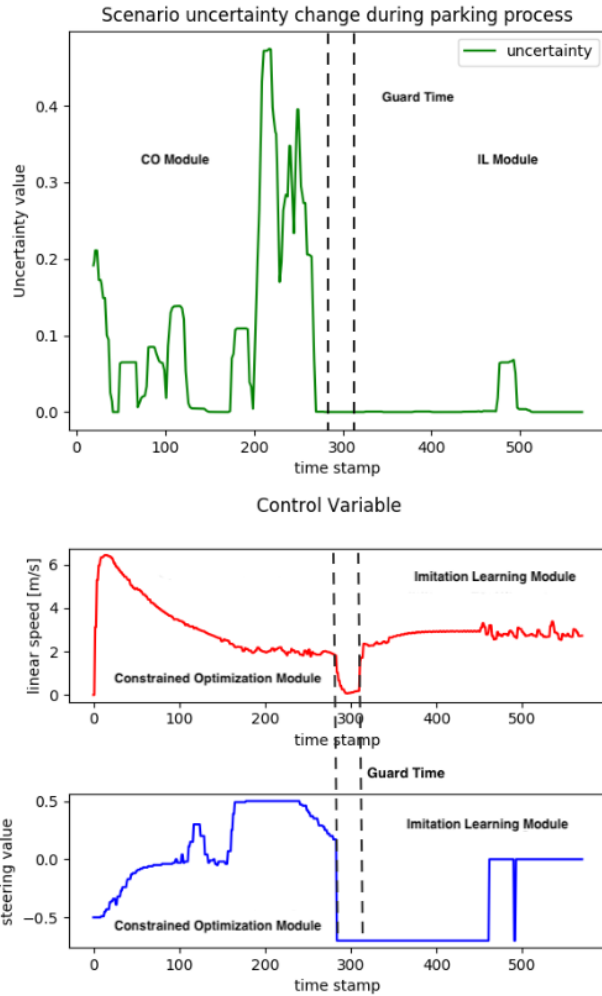


図7:HSAに基づくモード切替。

提案するiCOILbased APは、衝突することなく、自車両を目標駐車空間までナビゲートすることに成功していることがわかる。上記の手順でHSAモジュールが生成したシナリオの不確実性を図7に示す。不確実性は、初期には0から0.5の間で変動するが、280前後のタイムスタンプの後、0.1以下に低下し、安定した状態を保つ。これは、提案するHSAの優れたシーン理解能力を示している。

HSAに基づき、iCOILベースのAPシステムは、図6bに示すように、AP効率を向上させるためにC0からILへのモード切替を行う。関連する制御コマンドの経時変化を図7に示す。モード切り替え後はリバースボタンがオンになり、470タイムスタンプ前後でステアリング値がゼロに近くなり、自車両が駐車スペースに入り、APタスクを完了するためにゆっくりと後方に移動していることがわかる。我々の実装では、異なるモード間の遷移を滑らかにするために、20個のタイムスタンプを持つガードタイムが追加されていることに注意してください。以上の実験により、提案するiCOILベースAPの有効性が実証された。

D. 性能比較

この小節では、iCOIL方式とIL方式の駐車時間と成功率を比較する。駐車場時間は、出発点から駐車場までの総時間として定義される。自車両がある時間内にゴールに到達できないか、他の障害物と衝突した場合、タスクは失敗とみなされ、そうでない場合は成功とみなされる。

異なる難易度における駐車時間と成功率を表IIに示す。簡単なレベルでは、iCOILは94%の成功率を達成し、ILは72%の成功率を達成した。しかし、ILの駐車時間は若干短くなっている。通常のレベルでは、動的な障害物を考慮するため、どちらの方法も成功率は低下した。しかし、汎化が不十分なため、ILの成功率は著しく低下する。ILによる駐車障害のスナップショットを図6cと図6dに示す。一方、iCOILの成功率はあまり影響を受けていない。これは、iCOIL方式が衝突回避にC0を採用しているため、IL方式よりも正確で安全な行動を生成しているためである。ハードレベルに関しては、センシングの不確かさが大きくなるため、ILの性能はさらに低下する。すべての難易度において、提案するiCOILは成功率メトリクスの点で常にILを上回った。なお、ILの成功率は、より多くのデータとより大きなDNNによって向上させることができる[2]。この場合、ILとiCOILの差は小さくなるが、コーナーケースシナリオではiCOILがILを上回る。

E. 感度分析

最後に、iCOILの駐車性能を異なる開始点で検証する。特に、図4では、近い出発点、遠い出発点、ランダムな出発点を考慮し、異なる障害物数の下での関連結果を以下に示す。

表 II: 駐車場時間と成功率の比較

Easy Task				
Method	Average	Max	Min	Success Ratio
iCOIL	26.02	27.21	24.89	94%
IL [2]	23.65	25.16	22.52	72%
Normal Task				
Method	Average	Max	Min	Success Ratio
iCOIL	25.40	26.29	24.01	91%
IL [2]	25.81	26.54	23.77	36%
Hard Task				
Method	Average	Max	Min	Success Ratio
iCOIL	25.72	26.70	24.58	92%
IL [2]	24.12	26.44	23.31	33%

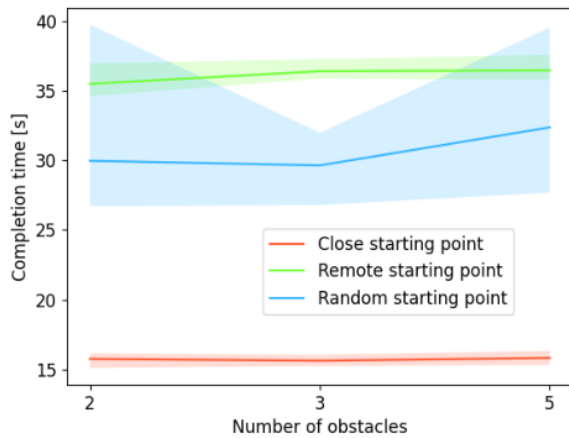


図8: 出発点と障害物の数を変えた場合のiCOILの駐車時間。

図9: 駐車時間の比較。

図8. クローズ・スタート・ポイントの場合、障害物の数は駐車時間にほとんど影響を与えない。しかし、遠隔地やランダムな場合と同様に、障害物の数が増えるにつれて駐車時間が長くなる。これは、この場合、始点が障害物に近い可能性があり、駐車時間が大幅に増加するためである。ランダムな始点を持つ場合、移動軌跡が異なると駐車距離が大きく変化する可能性があるため、駐車時間の変動が大きいことに注意してください。

最後に、iCOILの実行頻度をテストする。ILとC0の平均周波数はそれぞれ75Hzと18Hzである。これは我々の理論的分析を裏付け、モードスイッチングの必要性を示している。両周波数とも、低速駐車シナリオの要件を満たすことができることに注意してください。

VI. CONCLUSION

本論文では、C0とILを統合することで、高い効率と信頼性を実現するシナリオを考慮したAPソリューションであるiCOILを紹介した。iCOILはROSパッケージとして実装され、リアルタイムAPシステムにデプロイされた。iCOILベースのAPシステムをMoCAMで検証し、

様々な構成でその有効性を実証した。iCOIL方式は、成功率の点で従来のILを上回ることが示された。

REFERENCES

- [1] X. Zhang, A. Liniger, and F. Borrelli, "Optimization-based collision avoidance," *IEEE Transactions on Control Systems Technology*, vol. 29, no. 3, pp. 972–983, May. 2020.
- [2] R. Chai, A. Tsourdos, A. Savvaris, S. Chai, Y. Xia, and C. P. Chen, "Design and implementation of deep neural network-based control for automatic parking maneuver process," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 4, pp. 1400–1413, April. 2022.
- [3] P. Zhang, L. Xiong, Z. Yu, P. Fang, S. Yan, J. Yao, and Y. Zhou, "Reinforcement learning-based end-to-end parking for automatic parking system," *Sensors*, vol. 19, no. 18, p. 3996, September. 2019.
- [4] Z. Du, Q. Miao, and C. Zong, "Trajectory planning for automated parking systems using deep reinforcement learning," *International Journal of Automotive Technology*, vol. 21, pp. 881–887, July. 2020.
- [5] J. Ahn, M. Kim, and J. Park, "Autonomous driving using imitation learning with look ahead point for semi structured environments," *Scientific Reports*, vol. 12, no. 1, p. 21285, December. 2022.
- [6] Z. Han, Y. Wu, T. Li, L. Zhang, L. Pei, L. Xu, C. Li, C. Ma, C. Xu, S. Shen *et al.*, "Differential flatness-based trajectory planning for autonomous vehicles," *arXiv preprint arXiv:2208.13160*, 2022.
- [7] S. Wang, C. Li, D. W. K. Ng, Y. C. Eldar, H. V. Poor, Q. Hao, and C. Xu, "Federated deep learning meets autonomous vehicle perception: Design and verification," *IEEE Network*, pp. 1–10, December. 2022.
- [8] J. Canny, *The complexity of robot motion planning*. MIT press, 1988.
- [9] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," in *Proc. CoRL*, 2017, pp. 1–16.
- [10] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. A. Sallab, S. Yogamani, and P. Pérez, "Deep reinforcement learning for autonomous driving: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 4909–4926, June. 2022.
- [11] J. Hawke, R. Shen, C. Gurau, S. Sharma, D. Reda, N. Nikolov, P. Mazur, S. Micklethwaite, N. Griffiths, A. Shah *et al.*, "Urban driving with conditional imitation learning," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France, 2020, pp. 251–257.
- [12] R. Han, S. Chen, and Q. Hao, "Cooperative multi-robot navigation in dynamic environment with deep reinforcement learning," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France, 2020, pp. 448–454.
- [13] L. Tai, S. Li, and M. Liu, "A deep-network solution towards model-less obstacle avoidance," in *2016 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, Daejeon, Korea (South), 2016, pp. 2759–2764.
- [14] Y. Pan, C.-A. Cheng, K. Saigol, K. Lee, X. Yan, E. A. Theodorou, and B. Boots, "Imitation learning for agile autonomous driving," *The International Journal of Robotics Research*, vol. 39, no. 2-3, pp. 286–302, October. 2019.
- [15] M. Kelly, C. Sidrane, K. Driggs-Campbell, and M. J. Kochenderfer, "Hg-dagger: Interactive imitation learning with human experts," in *2019 International Conference on Robotics and Automation (ICRA)*, Montreal, QC, Canada, 2019, pp. 8077–8083.
- [16] B. Liu, L. Wang, X. Chen, L. Huang, D. Han, and C.-Z. Xu, "Peer-assisted robotic learning: a data-driven collaborative learning approach for cloud robotic systems," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, Xi'an, China, 2021, pp. 4062–4070.
- [17] D. González, J. Pérez, V. Milanés, and F. Nashashibi, "A review of motion planning techniques for automated vehicles," *IEEE Transactions on intelligent transportation systems*, vol. 17, no. 4, pp. 1135–1145, April. 2016.
- [18] R. Han, S. Wang, S. Wang, Z. Zhang, Q. Zhang, Y. C. Eldar, Q. Hao, and J. Pan, "Rda: An accelerated collision free motion planner for autonomous navigation in cluttered environments," *IEEE Robotics and Automation Letters*, vol. 8, no. 3, pp. 1715–1722, March. 2023.
- [19] A. Kendall and Y. Gal, "What uncertainties do we need in bayesian deep learning for computer vision?" in *Advances in Neural Information Processing Systems*, vol. 30, Long Beach, CA, USA, 2017.
- [20] A. Kendall, Y. Gal, and R. Cipolla, "Multi-task learning using uncertainty to weigh losses for scene geometry and semantics," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, Salt Lake City, UT, USA, 2018, pp. 7482–7491.