

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/324562456>

自律走行車のための視覚と推測航法に基づくエンドツーエンドの駐車

Conference Paper · June 2018

DOI: 10.1109/IVS.2018.8500558

CITATIONS

12

READS

515

2 authors:



Vijay John

RIKEN Keihanna

77 PUBLICATIONS 1,514 CITATIONS

[SEE PROFILE](#)



Swarn singh Rathour

Toyota Technological Institute

14 PUBLICATIONS 52 CITATIONS

[SEE PROFILE](#)

自律走行車のための視覚と推測航法に基づくエンドツーエンドの駐車

概要- 本論文では、エンドツーエンドの運転のための視覚と推測航法を組み合わせた駐車システムを提案する。標準的な自律駐車フレームワークは複数のモジュールを含み、各モジュールにはそれぞれ制限がある。一方、提案する駐車フレームワークは、単一のエンドツーエンドモジュールで構成されており、これらの固有の制限を軽減している。提案するディープラーニングベースの駐車システムでは、前方および後方に取り付けた単眼カメラを用いて、ステアリング角度とギアの状態を予測するために、新しい反復2段階学習フレームワークを利用する。提案するフレームワークの第1段階では、エンコーダ・デコーダアーキテクチャを用いて、前方または背面の単眼カメラの複数のフレームから、操舵角の軌跡の初期推定値を予測する。ギアの状態推定値を用いて、ステアリング推定に使用するカメラを選択する。ギアの状態は、初期化中にあらかじめ定義され、その後、提案するフレームワークの第2段階で推定される。提案フレームワークの第2段階では、最適なステアリング角度とギアの状態を推定するために、車両のヘディング角度と絶対位置とともに、ステアリング角度の軌跡の初期推定値が長期短期記憶ネットワークへの入力として与えられる。提案するフレームワークは、取得したデータセットで検証される。ベースラインアルゴリズムとの比較分析、詳細なパラメトリック分析を行う。実験結果は、提案フレームワークがベースラインのエンドツーエンドアルゴリズムよりも優れていることを示している。

I. INTRODUCTION

自律走行と先進運転支援システム(ADAS)の分野では、大きな成果を上げている。主要な自動車メーカーの多くは、インテリジェント・パーキング・アシスト・システム(IPAS)を採用しており、パーキング・アシストと半自律型パーキングが市場における主流の自動車技術となっている。初期の駐車支援技術は、1台の屋台(前後クリアランス $\pm 70\text{cm}$)を使って、空の平行駐車スペースに逆転させる機能を備えている。また、これらのシステムは、 $0.5\sim 1.5\text{m}$ の距離から駐車空間を検出する能力を有している。前バージョンのアップグレードである現在のIPAS技術は、複数人(前後クリアランス $\pm 40\text{cm}$)の駐車が可能である。市販されているにもかかわらず、駐車支援システムには固有の限界があり、駐車問題を完全に解決していない。さらに、構造化された駐車場だけでなく、あるレベルでの人的介入 [2], [3] が必要である。したがって、本稿では、完全自律型セルフパーキングシステムを開発するために、視覚と推測航法を組み合わせたパーキングシステムを紹介する。本論文の主な目的は、現在の自律駐車システムの限界(構造化された駐車場、隣接車両などへの依存)を克服するための学習ベースのフレームワークを開発することである。

本論文の主な貢献は以下の通りである:

- 新しい2段階の深層学習ベースのエンドツーエンド駐車システム(図2)。構造化された駐車場と構造化されていない駐車場で駐車できる。
- エンドツーエンドの運転におけるエンコーダ・デコーダのアーキテクチャの使用。ベースラインのエンドツーエンドドライビングフレームワーク[セクションIV]では、画像から抽出された特徴が回帰ネットワークへの入力として与えられる。一方、エンコーダ・デコーダのアーキテクチャでは、デコーダ出力マップの顕著な特徴が回帰ネットワークへの入力として与えられる。これは推定精度を向上させることが示されている(セクションV)。
- 提案する深層学習フレームワークの予測精度を向上させるために、DRから得られる車両の方位と移動距離を利用する。(第III章)

本稿の残りの部分は以下のように分割される。セクションIIでは、レビュー作業の概要と現在の駐車技術の限界について述べる。セクションIIIでは、エンドツーエンドの駐車のために開発された提案アルゴリズムを定義する。セクションIVでは、提案するディープラーニングベースのエンドツーエンド駐車アルゴリズムの訓練と検証のためのデータセットの準備について説明する。セクションVでは、提案する学習フレームワークと他のベースラインであるエンドツーエンド学習フレームワークの比較検討と、提案するフレームワークのパラメトリックなバリエーションを示す。最後に、セクションVIにおいて、本論文の主な貢献を列挙することで、本論文の結論を述べる。

II. 文献調査

ほとんどの主要自動車メーカー(トヨタ、BMW、フォード、フォルクスワーゲン、メルセデス・ベンツなど)は、半自律型駐車やIPASシステムを搭載した自動車をロールアウトしている。様々な半自律型またはIPAS技術の基本は、依然として類似している。まず、システムは適切な駐車スペースの検出から始まる。次に、システムは、車両周囲の障害物の安全な距離をドライバーに知らせるために、最適なアプローチを定式化する。上記の駐車手順を実行するために、半自律駐車またはIPASシステムは、複数の距離ベースのセンサーとカメラで構成され、関係する変数の大部分を検出するために、車両の前部と後部のバンパーに取り付けられている。したがって、駐車システムは、環境認識、経路生成、制御、衝突回避など、複数のサブモジュールに分けることができる。これらのモジュールはそれぞれ、それ自体が困難な研究課題であり、研究コミュニティによる個別の注意が必要である。半自律型または駐車支援システムは、使用されるセンサーによって、アクティブセンサーベースの駐車(超音波またはレーザーベース)、ビジョンベースのシステム、またはビジョンとアクティブレンジセンサーの組み合わせに分けることができます[4]。

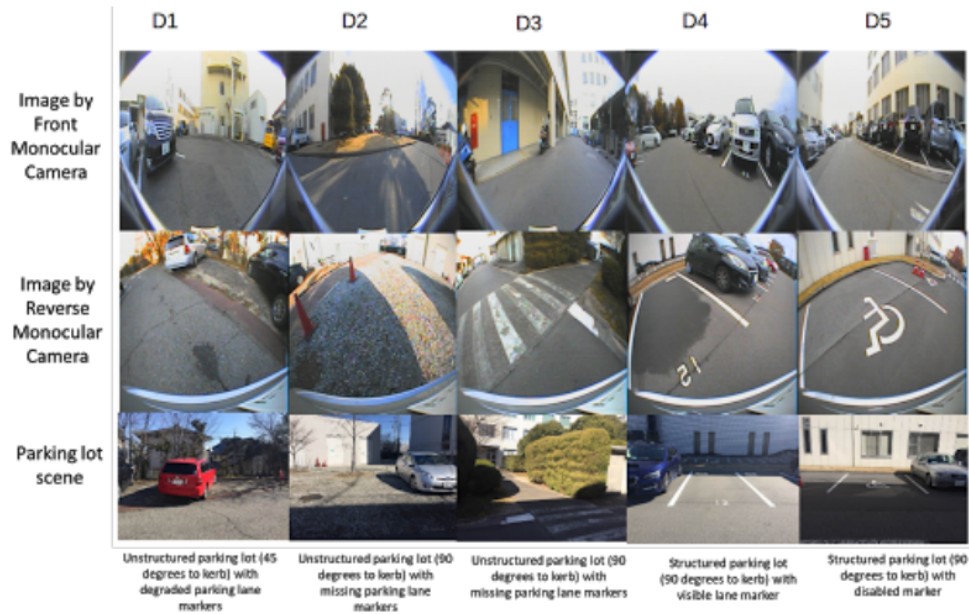


図1: エンドツーエンドの駐車を学習するために、フロントカメラ(上段)とバックカメラ(中段)で撮影された画像のコレクション。(最後の行) 手持ちカメラから撮影されたそれぞれの駐車場の画像。

超音波またはレーザーベースの駐車支援は、駐車支援に使用されるセンサーとして最も一般的であるが [5], [6]、独自の制限 [7]、例えば、超音波センサーベースのシステムは、短距離に制限され、環境内の特定の物体を知覚することが困難である。また、レーザーを用いたレンジセンサーは高精度であるが、高コストで寿命が短いため、使用が制限される。一方、ビジョンベースの駐車支援システムは、駐車レーンマーカーや隣接する駐車車両の存在のような構造化された環境を必要とする [8]、[9]、[10]。しかし、図1)に見られるように、駐車レーンマーカーが欠落しているか、隠されていることが多く、場合によっては隣接する駐車車両が存在しないこともある。さらに、視覚システムは照明の変動や環境ノイズの影響を受けやすい [11]。

本論文では、現在の半自律型駐車支援システムの上記の限界に対処するために、自律走行車のためのディープラーニングを用いた視覚と推測航法に基づく駐車システムを提案する。近年、ディープラーニングフレームワークへの関心が高まっており、人間レベルの精度で画像認識とセグメンテーションの分野で最先端の結果が得られている [11], [12], [13]、自律走行 [14], [15], [16], [17]。従来の自律走行フレームワークと比較して、深層学習ベースのフレームワークは、画像から直接操舵角を予測する単一モジュールまたはエンドツーエンドの学習フレームワークで構成されている。このようなシステムでは、環境認識、経路計画、障害物回避、制御など、複数のモジュールが不要になる。エンドツーエンド学習ベースのアルゴリズムは、エンドツーエンドで完全に学習可能であるため、最小限の人的労力で済む。

出力としてmand。したがって、エンドツーエンドの学習は、異なるモジュールを明示的にモデル化する必要がないため、魅力的である。しかし、ディープラーニングに基づくエンドツーエンドの運転に関する既存の文献は、高速道路や公共道路での運転に限られている [14]、[15]、[16]、[17]。本研究では、エンドツーエンドの深層学習フレームワークを自律駐車に拡張する。さらに、従来の駐車支援システムの問題点、特に構造化された環境を必要とする視覚ベースの駐車支援の問題点にも取り組む。

III. ALGORITHM

本論文の主な目的は、前後に取り付けた魚眼カメラを用いて得られた時間同期された連続画像観測から、専門家の操舵角のシーケンスとギアの状態を予測することである。さらに、移動距離と車両の方位(推測航法)もステアリング角度の予測に使用される。提案するネットワークは、エンコーダ・デコーダ段階とLSTM段階からなる2段階のエンドツーエンド学習フレームワークである。図2は、提案するフレームワークの各ステージの詳細なアーキテクチャを示す。

第一段階では、セマンティックセグメンテーション [18] に使用されるエンコーダ・デコーダアーキテクチャに基づく深層学習フレームワークを、前後に取り付けられた魚眼カメラから得られた画像を用いて、操舵角を予測するように修正する(図2)。前方カメラ画像は前進運動中のステアリング角を予測するために使用され、後方カメラ画像は後進運動中のステアリング角を予測するために使用される。カメラの選択はギアの状態に依存する。初期フレームでは、車が前進していると仮定し、フロントカメラを使用する。この後のフレームワークでは、歯車の状態を

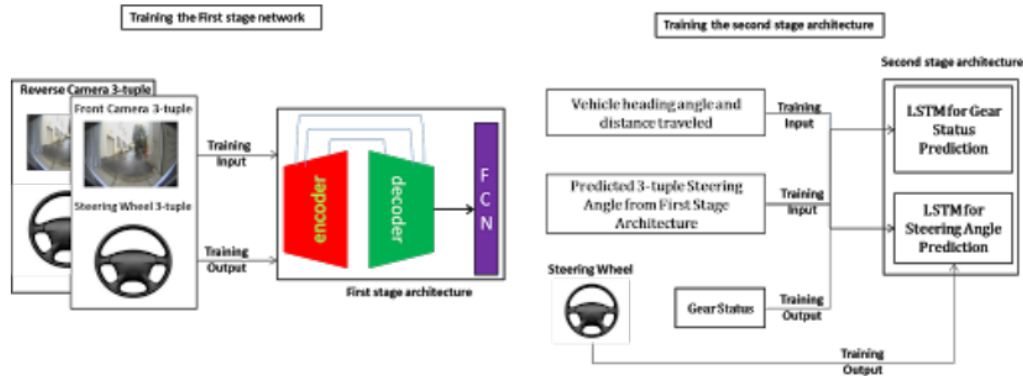


図2:学習ステップの概要。

提案するフレームワークの第2段階

第一段階の操舵角を予測するために、[18]で提案されたようなディープセグメンテーションネットワークフレームワーク、すなわちUNETが、その単純さと他のベースラインセグメンテーションネットワーク[12]、[13]に対する優位性から選択された。図2に示すように、ディープラーニングフレームワークの第1段階は、UNETと洗練ネットワークを用いて定式化される。より具体的には、2つのネットワークを結合するために、U-Netデコーダブロックの出力が洗練ネットワークへの入力として与えられる。精密化ネットワークは3つの完全接続層から構成される。最初の2つの完全接続層は、それぞれ512ユニットと最後の256ユニットを含む。最初の2つの完全接続層はReLU非線形を使用し、最後の層は線形活性化関数を使用する。損失モデルとしての平均二乗誤差は、モデルの重みを学習するためにAdam optimizer (1e-5)と共に使用される。第2段階では、LSTMネットワークを用いて第1段階の操舵角予測を精緻化し、ギヤの状態を推定する。LSTMネットワークへの入力は、車両の推測航法測定値とともに、ステージ1の操舵角予測値のシーケンスである。次に、アルゴリズムの学習とテストのステップを詳しく説明する。

1) Training Step:

a) エンコーダ・デコーダの学習: エンコーダ・デコーダベースの完全接続ネットワークは、3タプルの画像観測、すなわち (I_{i-1}, I_i, I_{i+1}) を学習入力とし、3タプルの操舵角、すなわち (s_{i-1}, s_i, s_{i+1}) を出力として学習する。典型的な駐車操作は、前方走行と後方走行の両方から構成される。ステアリング角予測の精度を高めるため、画像取得に2台のカメラを別々に利用する。前方マヌーバの間、3タプルの画像観測は、フロントマウントカメラを使用して取得される。一方、リバースマヌーバでは、バックマウントカメラを用いて3タプルの画像観測を取得する。3タプルの画像観測は、常に車両CANBUSから得られる3タプルの操舵角と同期していることに注意。両カメラから取得した画像は、第1段階のネットワークの学習に使用される。

学習データが与えられると、第1段階のネットワークにおける画像からステアリングへのマッピング関数は、教師ありの方法で学習される。

より具体的には、提案するエンコーダ・デコーダに基づく完全接続フレームワークで近似された回帰関数のパラメータ θ を最適化する。

b) LSTMの学習: b) LSTMの学習: ネットワークの第2段階では、デッドレコニング(DR)測定値を用いて、第1段階のステアリング角度推定値を精緻化することで、提案するディープラーニングフレームワークの性能を向上させる。より具体的には、第一段階の操舵角推定値とDR測定値が、操舵角の最適推定値を生成するLSTMネットワークへの入力として与えられる。さらに、LSTMネットワークは、第一段のカメラ選択のためのテスト段階で使用されるギヤの状態も予測する。

DR測定値は、車両の移動距離(すなわち d)に対応し、4輪回転速度、車速、ギヤ状態 g 、ヨーレート $\dot{\phi}$ からなる時間同期された缶バスデータを用いて導出される。提案するLSTMネットワークは、第1段階のディープネットワークから予測された s^* を、移動距離 d と絶対方位角 ϕ とともに取る。LSTMネットワークは2つの出力からなり、最初の出力は g^* を予測するLSTMベースの分類器である。 g^* は2つの状態のみの分類器から構成されるため、LSTMモデルがこのタスクに適している。次に、LSTMモデルは線形活性化ユニットと平均二乗誤差損失モデルで $s \sim_f$ を予測する。

2) テスト段階: 2段階の提案ディープネットワークフレームワークを学習した後に得られたパラメータを用いて、時間同期した I, d と ϕ を $s \sim_f$ (最終予測制御操舵角)と g^* (予測ギヤ状態)に対応付ける。テストの概要を図3に示す。

初期化中、提案アルゴリズムは、車両がギヤ状態 $g = 1$ で前進していると仮定して初期化される。その結果、初期化中、第1段階のネットワークはフロントカメラを使用して s^* を推定する。

ステアリング角の初期推定値 s^* と車両DRの測定値が第2段階のLSTMネットワークの入力として与えられる。LSTMは、ギヤの状態 g^* とともに最適なステアリング角 $s \sim_f$ を推定する。次に、ギヤの状態は、入力用の背面カメラまたはフロントカメラを選択するために、第1ステージのネットワークによって再帰的に使用される。 $g = 1$ はフロントカメラ、 $g = 0$ はバックカメラを意味する。

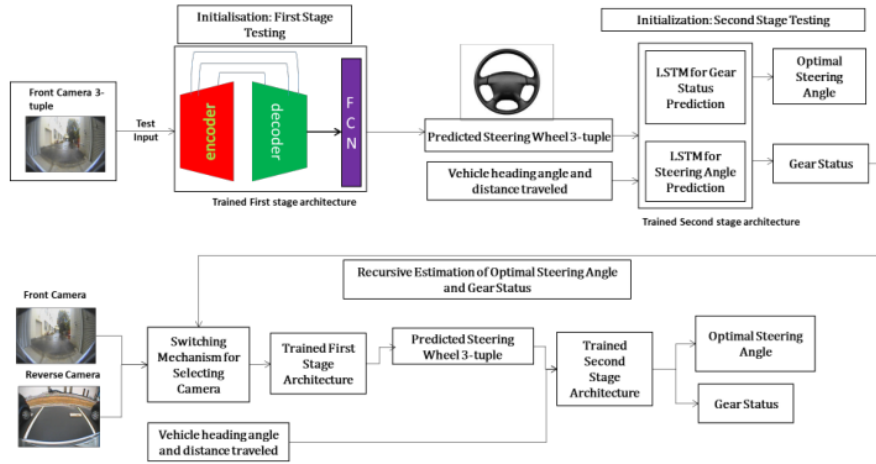


図3:テストステップの概要。

IV. 実験結果

まず始めに、提案する深層学習フレームワークの学習とテストに使用したデータセットを紹介する。最後に、実験セットアップと結果の考察を述べる。ベースラインモデルとの比較分析を行う。さらに、2段階ネットワークのバリエーションによるパラメータ分析も行う。すべてのトレーニングおよび検証は、以下の仕様のシステムで実行された: 64ビットIntel Core i7-6850K CPU @ 3.60GHz×12, GeForce GTX 1080, RAM 64GB、テンソルフローバックエンド付きkerasを使用。

A. 実験データセット

自律的な逆駐車のためのエンドツーエンドの視覚認識ベースのディープラーニングフレームワークを学習するために、エキスパートドライバーが実証した駐車は、トレーニングとテストのためのデータセットを準備するために使用された。エキスパートドライバーは、様々なデータセットを作成するために、多数の駐車場を実行するよう求められた。D1、D2、D3、D4、D5の5つのデータセット(図1)をそれぞれ2つのシーケンスから構成し、トレーニング用とテスト用に用意した。各データセットは、時間同期された前後カメラ画像観測値 l, ϕ, s, d から構成される。各データセットは別々のトレーニングシーケンスとテストシーケンスに分割された。図1は、5つのデータセットすべてについて、フロントカメラとバックカメラで撮影した画像である。データセットD1、D2、D3は、非構造化駐車場、すなわち、劣化した白線と欠落した白線のある駐車場を用いて作成した(図1、最終行)。したがって、D1-D3は非構造化駐車場を表し、D4-D5は構造化駐車場データセット(よく定義された白線(図1、最後の行))を表す。

提案アルゴリズムの性能を検証するために、各データセットの学習シーケンス(すなわちD1、...、D5)に対して提案モデルを学習させ、最後に学習後に得られた最適化パラメータ θ を用いて、対応するデータセットのテストシーケンスに対するテストを行った。例えば、D1の訓練シーケンスで提案モデルを訓練することによって得られる最適化されたパラメータ θ は、D1のテストシーケンスでテストされた。

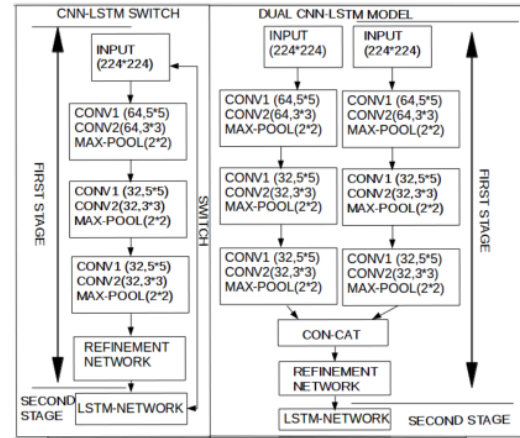


図4:パラメータ分析に使用した深層学習ベースのフレームワークのCNNアーキテクチャ。

B. 比較分析

提案モデルとの比較のために、4つのベースライン深層学習ベースのエンドツーエンドネットワークを選択した。画像ピクセルをステアリング角にマッピングするために[11]で使用されたCNNアーキテクチャ(すなわちVGG16)は、最初のベースラインを表す。残差ベースのCNNアーキテクチャ(RESNET50)は、[20]で提案されているように、2番目のベースラインを表す。3番目と4番目のベースラインでは、VGG-16とRESNET50の最終畳み込み層から抽出した特徴マップを抽出し、木ベースの回帰器(すなわち、VGG16 + ET & RESNET50 + ET)[21]を追加で学習させた。

提案するフレームワークで2台のカメラを利用することの利点を実証するために、ベースラインモデルはフロントカメラ「単独」を使用して学習された。さらに、これは先行研究[15]、[21]で使用された自動運転ネットワークを模倣するためにも行われる。比較のために、5つのデータセットすべてについて、4つのベースラインモデルの性能を調べた。

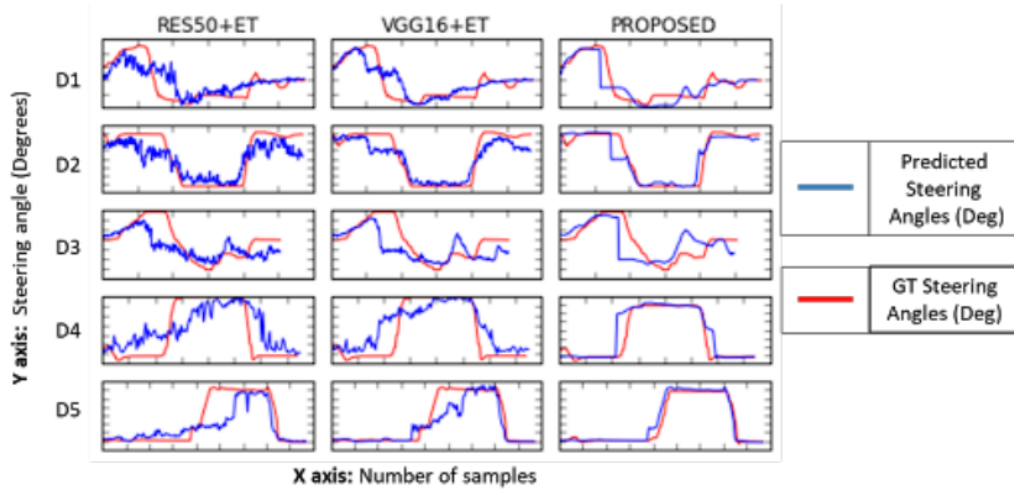


図5: RES50+ET、VGG16+ET、Proposedを用いた各データセットの予測値 \hat{s}_t の軌跡(青線)とグラントトゥールの軌跡 s_t (赤線)。

表1: 異なるデータセットにおける、予測されたステアリング角とグラントトゥールのステアリング角の平均ユークリッド距離(度)。

Data.	RES50	RES50+ET	VGG16	VGG16+ET	Prop.
D1	93.74	23.44	157.35	17.08	9.22
D2	64.02	56.00	30.13	41.0	7.89
D3	93.74	105.66	212.78	131.11	11.93
D4	64.47	43.75	157	65.35	38.61
D5	166.17	56.29	29.37	41.46	17.70

表II: 異なるステアリング角の予測値とグラントトゥールのとの間の平均ユークリッド距離(度)。

datasets

Data.	Prop.	CNN-LSTM-スイッチ	Dual CNN-LSTM
D1	9.22	29.49	22.28
D2	7.89	28.28	31.69
D3	11.93	31.98	155.86
D4	38.61	54.54	133.25
D5	17.70	143.30	92.13

角度とグラントトゥールのステアリング角度。

図5と表1で得られた結果は、提案ネットワークが異なるデータセットにおいて、ベースラインアルゴリズムよりも優れていることを示している。ベースラインと比較して、性能の向上は以下のことに起因する：

- ステアリング角の初期推定値がLSTMによって精緻化される2段アーキテクチャ。
- エンコーダ・デコーダのアーキテクチャを利用して、U-Netの出力マップが完全連結ネットワークへの入力として与えられる、操舵角の初期推定値を得る。
- 予測のために2台のカメラを切り替える。

C. パラメータ解析

提案モデルとそれぞれのベースラインモデルの比較の後、パラメトリック依存性を研究するために、第一段階での提案モデルアーキテクチャを変化させた。

パラメトリック解析では、提案したフレームワークの2つのバリエーションを使用した(図4)。最初のバリエーションでは、カメラ切り替えメカニズムを排除し、U-Netをフロントカメラ画像とバックカメラ画像用の2つのCNNブランチに置き換える。CNNの枝によって抽出された特徴は連結され、第一段階で洗練ネットワークに与えられる。第2ステージでは、LSTMによって初期ステアリング角が精緻化される。このモデルをデュアルCNN-LSTMモデルと呼ぶ。

2つ目のバリエーションでは、提案フレームワークのU-NetをCNNに置き換え、切り替えメカニズムを保持したまま、画像特徴を抽出する。このモデルをCNN-LSTMスイッチングモデルと呼ぶ。2つの変形モデルのCNNアーキテクチャを図に示す。

4.

パラメトリック解析で得られた結果を図6と表2に示すが、提案モデルの方がパラメトリックバリエーションよりも優れていることがわかる。パラメトリック解析により、U-Net、スイッチング機構、LSTMフレームワークを用いることの利点が示された。U-Netの利点はCNN-LSTM-Switchingモデルとの比較で明らかであり、提案モデルのU-Netと変形モデルのCNNを除いて、これらのモデルはどちらも類似している。

提案モデルでは、スイッチング機構を用いて前方の ”または ”後方 ”を選択するため、U-Netとスイッチング機構の優位性は、Dual CNN-LSTMモデルとの比較において明らかである。一方、変形モデルでは、両方のカメラ画像が使用される。

V. CONCLUSION

本論文では、前方または後方に取り付けた単眼カメラを用いて、ステアリング角度とギアの状態を予測するために、ビジョンとDRを組み合わせた新しい2段エンコーダ・デコーダアーキテクチャを提案する。提案モデルは、限られた学習データセットにおいて、ベースラインのエンドツーエンドモデルと比較して、より良い性能を示す。

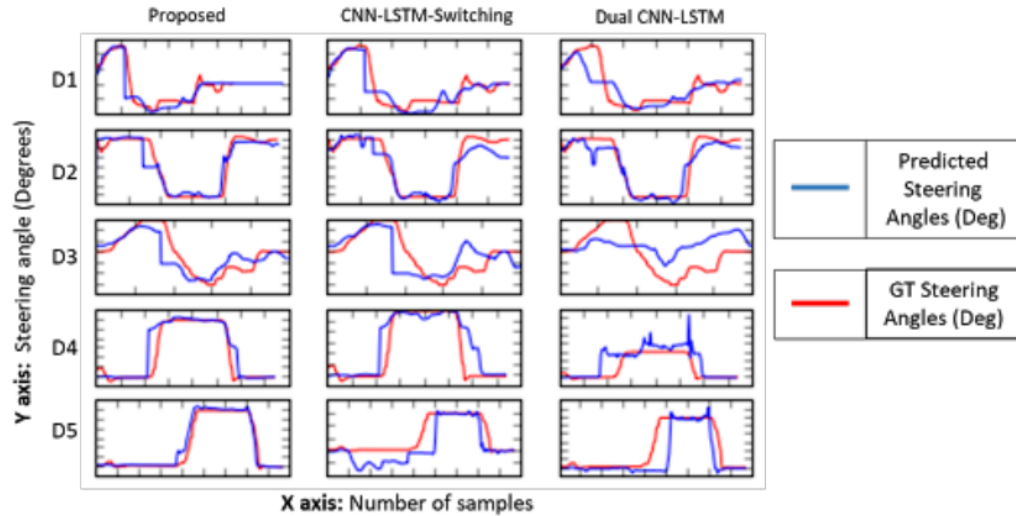


図6: 提案モデル、分離モデル、複合モデルを用いた各データセットの予測値 $s \sim_f$ の軌跡とグラントゥールースの軌跡 s 。

提案モデルの有効性を検討するため、提案ネットワークの第1段階でもパラメータ変動はほとんど行わなかった。提案モデルは、ステアリング軌道とギアの状態を予測するのに有効であり、システムを完全に自律化し、マルチマンオブを行うことができることがわかった。提案するディープラーニングに基づくエンドツーエンドの駐車場は、完全に見える、部分的に見える/見えない駐車場の白線がある非構造駐車場や、駐車場に駐車場がない場合でも使用できる。

REFERENCES

- [1] W. Wang, Y. Song, J. Zhang, and H. Deng, "Automatic parking of vehicles: A review of literature," *International Journal of Automotive Technology*, vol. 15, no. 6, pp. 967-978, 2014.
- [2] The Hybrid That Started it All, Mar. 2014. [Online]. Available: <http://www.toyota.com/prius/>
- [3] BMW 7 Series. Park Assist, Mar. 2013. [Online]. Available: <http://www.bmw.com/>
- [4] X. Du and K. K. Tan, "Autonomous reverse parking system based on robust path generation and improved sliding mode control," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1225-1237, Jun. 2014.
- [5] J. Pohl, M. Sethsson, P. Degerman, and J. Larsson, "A semi-automated parallel parking system for passenger cars," *Proc. Inst. Mech. Eng. D, J. Autom. Eng.*, vol. 220, no. 1, pp. 53-65, Jan. 2006.
- [6] H. G. Jung, Y. H. Cho, P. J. Yoon, and J. Kim, "Scanning laser radar based target position designation for parking aid system," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 3, pp. 406-424, Sep. 2008.
- [7] P. Degerman J. Pohl M. Sethsson "Ultrasonic sensor modeling for automatic parallel parking systems in passenger cars," *SAE 2007 World Congress & Exhibition*, Detroit, MI, U.S.A., 16th19th April, 2007.
- [8] K. Fintzel R. Bendahan C. Vestri S. Bougnoux T. Kakinami "3D parking assistant system," *Proc. IEEE Intell. Veh. Symp.*, pp. 881-886 2004.
- [9] N. Kaempchen U. Franke R. Ott "Stereo vision based pose estimation of parking lots using 3d vehicle models," *Proc. IEEE Intell. Veh. Symp.*, pp. 459-464 2002.
- [10] C. Wang, Hengrun Zhang, Ming Yang, Xudong Wang, Lei Ye and Chunzhao Guo. "Automatic parking based on a bird's eye view vision system," *Advances in Mobility Theories, Methodologies, and Applications*, vol. 2014 pp. 847406-1-847406-13 2014.
- [11] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems (NIPS)*, 2012, pp. 1097-1105.
- [12] M. Thoma, "A survey of semantic segmentation," *CoRR*, vol. abs/1602.06541, 2016.
- [13] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, and J. G. Rodriguez, "A review on deep learning techniques applied to semantic segmentation," *CoRR*, vol. abs/1704.06857, 2017.
- [14] C. Chen, A. Seff, A. Kornhauser, and J. Xiao. Deepdriving: Learning affordance for direct perception in autonomous driving. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2722-2730, 2015.
- [15] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, et al. End to end learning for self-driving cars. *arXiv preprint arXiv:1604.07316*, 2016.
- [16] H. Xu, Y. Gao, F. Yu, and T. Darrell. End-to-end learning of driving models from large-scale video datasets. *arXiv preprint arXiv:1612.01079*, 2016.
- [17] Lu Chi and Yadong Mu. "Deep Steering: Learning End-to-End Driving Model from Spatial and Temporal Visual Cues". In: *arXiv preprint arXiv:1708.03798* (2017).
- [18] Y. LeCun, U. Muller, J. Ben, E. Cosatto, and B. Flepp. Offroad obstacle avoidance through end-to-end learning. In *NIPS*, pages 739-746, 2005.
- [19] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *MIC- CAI*, pages 234-241. Springer, 2015.
- [20] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770-778, 2016.
- [21] Vijay John, Seiichi Mita, Hossein Tehrani Niknejad, Kazuhisa Ishimaru, "Automated driving by monocular camera using deep mixture of experts," *IV 2017*, 10.1109/IVS.2017.7995709.