

早稲田大学大学院情報生産システム研究科

博士論文審査結果報告書

論 文 題 目

**Study on Efficient Pedestrian Trajectory Prediction Using
Social-Aware Deep Learning Framework**

申 請 者

Dongchen LI

情報生産システム工学専攻
ニューロコンピューティング研究

2026 年 2 月

本論文は、歩行者軌道予測問題に焦点を当てたものである。歩行者軌道予測は歩行者が未来にどのような経路を辿るかを過去の軌跡から予測するタスクであり、自動運転車やモバイルロボットなどの人間社会をサポートする技術として注目されている。一方、歩行者の意思決定プロセスは柔軟かつ主観的であり、全く同じ状況でも異なる個人が異なる判断を下すため、このような人間の機動性と柔軟性が問題の難易度を大幅に高めている。そのため、歩行者軌道予測において、歩行者間相互の影響をモデル化し、物理的制約と社会的規範の両方に適合するように歩行者軌道を予測する必要がある。さらに、合理的な軌道を単一でなく複数予測する必要性もある。近年、深層ニューラルネットワークアーキテクチャの適用および深層学習技術の急速な進展に伴い、歩行者軌道予測も著しい進歩を遂げてきた。しかしながら、歩行者軌道予測は依然として大きな課題に直面している。特に、ほとんどの深層学習モデルはブラックボックスとして機能しているため、モデルが複雑であり大きな冗長性を持っている。深層学習に大規模データが必要となるが、実際の応用では、データ不足が深刻な問題となる。さらに、歩行者軌道予測はリアルタイム性が必要であるために、予測モデルの効率性を向上させることも課題として残されている。

そこで本論文では、これらの課題に対処するため、2ステップからなる効率的な歩行者軌道予測フレームワークを提案している。具体的には、観測情報は主に空間情報を含む周囲情報と主に時間情報を含むエージェント（歩行者）情報に分けられ、2ステップで歩行者軌道の予測を行っている。ステップ1では、周囲情報を地図等の事前情報と組み合わせて、深層ニューラルネットワークモデルを用いて歩行者向けの先行歩行可能区域等の情報を抽出する。特に、訓練済み大規模言語モデル(Large Language Models: LLMs)を利用すれば、歩行可能区域だけでなく区域内どのような歩行者軌道が有り得るかという情報も得られる。ステップ2では、歩行可能区域等の情報を条件として利用し、対象エージェントの過去の軌跡や向き等の情報で歩行者の軌道を予測している。また、提案法では、ステップ1において訓練済み LLMs や転移学習を利用することでデータ不足問題を解決しており、ステップ2において軽量モデルを利用することで効率性が高い歩行者軌道予測を実現している。

以下に、本論文の構成と各章の概略について述べ、評価を与える。

第1章では、歩行者軌道予測の基礎とその関連研究を紹介し、本論文の動機・目的を明らかにしている。

第2章では、効率的なソーシャルフォースモデルについて述べている。従来の社会的認識型深層学習に基づいた手法では、観測情報は深層ニューラルネットワークモデルを直接ビデオ信号に適用して抽出され、対象エージェントと周囲とのすべての情報を埋め込んでおり、大きな冗長性がある。このような冗長性を持つ特徴量に基づいた歩行者軌道予測は良い性能が得られるが、モデルが複雑になり効率性が低下している。そこで本論文では、まず観測情報を周囲情報と対象エージェント情報に分ける。次に、人工ポテンシャル場(Artificial Potential Field: APF)に基づいた2ステップからなる効率的な歩行者軌道予測フレームワークを提案している。APF は引力場と斥力場から構

成される。すなわち、周囲情報を地図等の事前情報と組み合わせた社会的特徴により斥力場を構成し、歩行者の過去の軌跡や向き等の対象エージェント情報により引力場を計画的視点から構築する。また、歩行者軌道は歩行者が APF の深い谷を通るように予測している。このように構成することで、従来法に見られる社会的特徴の冗長性を回避し有用な情報のみを保持し、効率性を大幅に向上した歩行者軌道予測を実現している。性能評価の数値例では、提案法を ETH、HOTEL、STUDENT、ZARA01、ZARA02 という 5 つのベンチマークデータセットに適用して、複数の最先端の従来法と比較した。平均変位誤差 (Average Displacement Error: ADE) による評価では、提案法が平均順位 3 位で、3 つのデータセットにおいて従来法より優れていた。例えば、ETH データセットに対して、従来の最良法である TPNMS (R. Liang et al., 2021) の ADE 評価は 0.710、モデルサイズと推論時間はそれぞれ 34.1K と 0.034s であるのに対して、提案法では ADE 評価は 0.690、モデルサイズと推論時間はそれぞれ 5.1K と 0.004s であり、7 分の 1 のモデルサイズと 9 分の 1 の推論時間でほぼ同等の ADE 評価性能を達成した。これらのことから、提案法の有効性が明らかとなっている。

第 3 章では、生成モデルによる効率的な歩行者軌道予測法について述べている。歩行者軌道予測において、歩行者のゴールという情報がないので、全く同じ状況でも異なる個人が異なる軌道を選択することが考えられる。そのため、物理的制約と社会的規範の条件を満たすような合理的な複数の軌道から歩行者軌道を予測する必要がある。第 2 章で述べた確定的モデルによる歩行者軌道予測法では、単一的な歩行者軌道の予測しかできない。本論文では、まず、確率的生成モデルで可能な歩行者軌道を複数生成し予測を行う。具体的には、対象エージェントの過去の軌跡や向き等の情報に基づいて、変分オートエンコーダー (Variational Autoencoder: VAE) という生成モデルにより歩行者の軌道を複数生成する。次に、事前注釈付き歩行可能区域マップを介して、周囲情報を地図等の事前情報と組み合わせた社会的特徴量を生成モデルバックボーンの後処理として追加し、最も可能性ある歩行者軌道を識別して、効率的な歩行者軌道の予測を実現している。性能評価の数値例では、提案法を第 2 章で述べた 5 つのベンチマークデータセットに適用して、複数の最先端の従来法と比較した。ADE による評価では、提案法が平均順位 3 位で、2 つのデータセットにおいて従来法より優れていた。例えば、ETH データセットに対して、従来の最良法である Social-VAE (P. Xu et al., 2022) の ADE 評価は 0.410、モデルサイズと推論時間はそれぞれ 8M と 0.450s であるのに対して、提案法では ADE 評価は 0.530、モデルサイズと推論時間はそれぞれ 93K と 0.004s であり、ADE 評価性能が低下したが、モデルサイズが 80 分の 1、推論時間が 100 分の 1 となった。これらのことから、提案法の効率性が高いことが明らかとなっている。

第 4 章では、訓練済み LLMs を利用した歩行者軌道予測法について述べている。第 3 章で述べた手法では、周囲情報としての歩行可能区域マップの事前注釈には、多くの事前情報が必要であり膨大な時間がかかる。本論文では、訓練

済み LLMs を利用して歩行可能な区域情報を抽出する。具体的には、まず周囲に関する空間情報に対して訓練済みセグメンテーションモデル (Segmentation Anything Model: SAM) を適用し、プロンプト設計により歩行可能な区域マスクや歩く可能性が高い歩行者軌道等の情報を抽出する。次に、これらの情報を条件として利用し、対象エージェントの過去の軌跡や向き等の情報により第 3 章で述べた VAE 生成モデルに基づいて複数の歩行者軌道を生成し効率的な歩行者軌道予測を実現する。さらに、対象エージェントの過去の軌跡や向き等の情報に対して二次計画最適化法を適用することで、訓練不要な歩行者軌道予測を行うことが可能となる。訓練済み LLMs を利用した提案法は従来の深層学習アプローチと同等の予測精度を維持しつつ、効率性、可解釈性、適応性という明確な利点がある。性能評価の数値例では、提案法を第 2 章で述べた 5 つのベンチマークデータセットに適用して、複数の最先端の従来法と比較した。ADE と最終変位誤差 (Final Displacement Error: FDE) による評価では、提案法が平均順位 1 位で、3 つのデータセットにおいて従来法より優れていた。例えば、従来の最良法である TrajEVO (Zhao et al., 2025) の ADE/FDE 評価の平均は 0.360/0.710 であるのに対して、提案法では ADE/FDE 評価の平均は 0.350/0.560 である。さらに、提案法は推論効率が高く、軌道予測モデルを構成する SAM モジュールと計画モジュールの合計推論時間は 0.200s 未満である。これらのことから、提案法の有効性と実用性があることが明らかとなっている。

第 5 章では、結論として、本論文の成果について総括し、今後の研究課題を論じている。

以上を要約すると、本論文では、まず観測情報を周囲情報とエージェント情報に分けて考える。次に、周囲情報とエージェント情報により 2 ステップからなる歩行者軌道予測を行い、APF ソーシャルフォースモデルや VAE 生成モデルおよび SAM 訓練済みセグメンテーションモデルに基づいた効率性が高い予測法の提案を行っている。また、提案法を複数のベンチマークデータセットに適用しその有効性を明らかにしている。これらの成果は深層学習、ニューロコンピューティング分野の発展に寄与するところ大である。よって、本論文は、博士 (工学) の学位論文として価値あるものと認める。

2026 年 1 月 19 日

審査員

主査	早稲田大学	教授	博士(情報工学)(九州工業大学)	古月敬之
副査	早稲田大学	教授	工学博士(早稲田大学)	吉江 修
	早稲田大学	教授	博士(国際情報通信学)(早稲田大学)	伍 軍