

## 審査の結果の要旨

氏名 小林 由弥

この学位請求論文「深層生成モデルによる環境の構成的な認識と生成」は、深層学習における構成性に焦点をあて、画像を対象として教師なしで物体認識と各物体の表現学習を行う手法について述べたものである。深層生成モデルにおいて、さまざまな画像の生成手法が提案されているが、現状では簡単なトイデータや図形の組み合わせなどに適用範囲が限られており、実画像のように複雑な対象はうまく扱うことができない。本研究では、物体を認識するために必要な帰納バイアスをモデルに導入することでこの問題の解決を試みる。帰納バイアスとは、学習アルゴリズムの補助となるものであり、普遍的に真であるものではないが、多くの場合で有効なルールや前提知識である。学習の不安定性や適用範囲の狭さといった現状のシーン解釈手法の課題を解決することや、物体の認識に関する性能向上を主な目的とする。

本論文では、まず第1章で本研究の背景と目的を述べた後、第2章では、深層生成モデルに関する前提知識を述べる。第3章では、深層生成モデルによる構成的な認識を行う手法である、シーン解釈モデルとその関連技術について説明する。これらの背景を踏まえ、第4章では本研究の新規性や意義について述べている。

そして、第5章では、まず、背景情報を利用したシーン解釈手法の提案を行う。既存のシーン解釈手法では、物体や背景のテクスチャや形状が複雑になった場合、うまく認識することができなかった。未知の物体に対する認識のための帰納バイアスとして、背景の情報を用いる。学習時に背景の情報を利用することによって、物体として認識すべき対象についての帰納バイアスを与え、シーン解釈手法の適用範囲を拡張する。そして、背景と物体の情報が異なる潜在変数に保持されるような手法を構築した。背景と数字に対して複雑なテクスチャを適用したデータセットなどを利用し、評価実験により、背景と前景に分けた表現を獲得できることを示した。また、背景と複数物体を含む場合について、既存のシーン解釈手法を拡張する手法として提案手法が利用可能であることを確認した。

第6章では、自己教師あり学習と transformer を用いたシーン解釈手法を提

案する。自己教師あり学習は、クラス分類やセグメンテーションといった、目的とする課題そのものではなく、適切に設計された何らかの事前課題を解くことによって様々な課題に汎用的な特徴を獲得する手法である。自己教師あり学習の利用によって、追加データを用いずに物体に関する帰納バイアスを与える手法を提案した。また、既存手法は、モデル構造への過剰な制約があると解釈でき、これを解消するための **transformer** を用いたネットワーク構造の提案を行った。評価実験の結果、自己教師あり学習による学習の安定性を確認できた。

第7章では、複数視点の画像を入力として与える問題設定について考える。これは3次元空間を仮定することに相当し、階層的な表現学習による3次元空間の認識と生成を行う手法を提案する。既存手法では、シーン中の物体は潜在変数が従う確率分布から独立に生成されるものと仮定されており、物体間の関係性やシーン全体の空間的な構成についての表現学習は行われていなかった。提案手法では、シーン全体の空間的な構成を表現するグローバルな潜在変数と、各物体を表現するローカルな潜在変数に分割するモデルを用いる。既存手法と比較し、観測にない新たなシーンの生成が可能になること、汎化性能の向上が得られることを確認した。

第8章では、第3章から第7章までの研究内容を踏まえて考察を行った。その上で、今後の課題、構成的な認識技術の応用可能性、人間の認知機構との関係等を述べた。第9章で本論文の結論をまとめている。

本論文は、深層学習の活用において重要な役割を果たす帰納バイアスに着目し、それを統一的な視点から複数の手法を提案している。環境の認識、あるいは構成性は、今後の知能研究においても重要な基盤となるものであり、本論文で提案した手法は新規性が高く、その意義は大きい。また、本研究で得られた知見は、今後の深層生成モデルに関する多くの研究に対して有用なものである。

よって本論文は博士（工学）の学位請求論文として合格と認められる。