

論文 / 著書情報  
Article / Book Information

題目(和文)	深層学習による画像認識における推論根拠と未知不均衡ドメイン学習に関する研究
Title(English)	
著者(和文)	桑島洋
Author(English)	Hiroshi Kuwajima
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第12361号, 授与年月日:2023年3月26日, 学位の種別:課程博士, 審査員:田中 正行,奥富 正敏,蜂屋 弘之,中臺 一博,川上 玲
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12361号, Conferred date:2023/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

## 論文審査の要旨及び審査員

(2000字程度)

報告番号	乙 第 号	学位申請者	桑島洋	
	氏 名	職 名	氏 名	職 名
論文審査員	主査 田中 正行	准教授	川上 玲	准教授
	奥富 正敏	教授		
	蜂屋 弘之	教授		
	中臺 一博	教授		

本論文は「深層学習による画像認識における推論根拠と未知不均衡ドメイン学習に関する研究」と題して、全5章から構成されている。

第1章「緒言」では、本研究の目的と背景について述べている。深層学習に代表されるデータ駆動に基づく複雑なモデル学習は、高い性能を示しており、自動運転や医療分野などセーフティクリティカルなシステムにおいても利用されつつあるものの、データ駆動の限界と複雑なモデルの限界が存在することを指摘している。データ駆動に基づく深層学習では、学習データの分布とテストデータの分布が等しいことを仮定しているが、一般には、テストデータは未知不均衡ドメインと呼ぶ、複数のデータドメインから構成されており、各ドメインにおける性能は大きな偏りがあることが述べられている。深層学習のモデルは複雑な特徴量抽出部を含み、内部処理はブラックボックスであり、従来のソフトウェアのように原因を解析し修正することは困難であることを指摘している。このような背景のもと、深層学習による画像認識における推論根拠の解析、未知不均衡ドメインを考慮した機械学習方法ならびに能動学習を、新しく提案することを研究目的とすると述べられている。

第2章「深層学習における推論根拠」では、深層学習モデルの中間特徴量に着目した推論根拠の解析方法と、クラウドソーシングを活用した推論根拠解析結果を評価する方法を提案している。頻出する特徴量をクラス毎に解析し、入力画像に対して活性化している特徴量を組み合わせることにより、推論根拠を解析している。さらに解析結果に基づく、特徴量を人間に理解可能な言葉としてアノテーションを行っている。提案手法により得られた解析結果を、クラウドソーシングを用いた評価により、人間にも理解可能な解析結果になっていることが示されている。

第3章「未知不均衡ドメイン機械学習」では、異なるドメインが含まれる学習データを用いて、各ドメインの認識性能を向上させる機械学習方法を提案している。一般に、学習データには複数の異なるドメインが含まれており、各サンプルがどのドメインに属しているかは未知である。このような未知不均衡ドメインが含まれる学習データに基づき学習モデルを評価する際には、ドメイン毎に性能を評価することが重要であると述べられている。特に多数のデータが属している多数派ドメインと、少数のデータしか属していない少数派ドメインが存在する場合、従来の学習方法で学習されたモデルは、少数派ドメインのテスト性能が多数派ドメインのテスト性能に比べて、著しく低いことを指摘している。このような未知不均衡ドメイン機械学習の問題設定に対して、センターロスと呼ばれる損失関数と、サンプルのセントロイド距離に基づく重点サンプリングアルゴリズムを提案している。提案手法を用いることにより、多数派ドメインのテスト性能は従来手法と同程度のまま、少数派ドメインの性能を向上できることが示されている。

第4章「未知不均衡ドメイン能動学習」では、未知不均衡ドメインデータに対する能動学習が提案されている。多くの場合、未知不均衡ドメインデータをランダムサンプリングすると、サンプリングされた学習データにはサンプル数に大きな偏りが生じてしまい、その結果、学習されたモデルのドメイン毎のテスト性能にも大きな差が生じることが指摘されている。まず、未知不均衡ドメインデータを模擬するために必要なデータの条件を検討し、現実的な実験設定が示されている。また、これまでの能動学習の研究においては、データのサンプリング方法に注目されており、機械学習手法は簡便な手法が用いられていることが指摘されている。信頼度のマージンに基づく従来の能動学習に基づくサンプリングと第3章の手法を組み合わせることにより、各ドメインのテスト性能を少ないサンプルの学習データであっても、効率的に向上可能であることを示している。

最後に、第5章「結言」では、研究成果のまとめと将来展望について述べている。

以上のように、本論文は、深層学習による画像認識における推論根拠解析手法と、未知不均衡ドメインデータに対する機械学習手法および能動学習手法を提案し、実験によりその有用性を示している。その成果は工学上・工業上貢献するところが大きい。よって、本論文は博士(工学)の学位論文として、十分な価値があると認められる。