

審査の結果の要旨

氏名 張 确軒

本論文では、深層学習を用いデータにおける潜在的知識を顕在化するフレームワークを提案した。まず、従来に比較してモデルの精度と汎化性能を向上する深層学習手法を開発した。次に、学習済みのモデルに対し新たに開発したモデル解釈手法を適用し、解釈の結果を人間のユーザに可視化で提示する。ユーザがその可視化結果より今まで気付いていなかった事象を見出し、新しい知識として定着できるようにするためである。主に、モデル解釈手法と深層学習による特徴抽出と異常検知手法の開発に着目した。

第一章では、研究全体の構造を概括した。深層学習 (DL) は、さまざまな領域で問題を解決するために多様なデータへの適応性を有する高性能ツールとして広く使用されているが、高度なデータ利活用をめざすデータジャケットワークショップにおいて、参加者から DL の解釈可能性が強く求められる適応領域が明らかになった。レーザー加工プロセスにおけるフーリエ変換画像などのデータは人力で特徴を抽出したりターゲットにラベルを付けたりすることが困難であり、適用した DL の解釈可能性のなさは致命的である。ここでは本研究の究極の目標として、精緻で一般性を有するモデルに基づき深い知識を説明するための可視化を位置付け、モデル解釈を経て知識獲得に至るためのアプローチと、異常検出のためのマルチタスク自己監視アプローチを呈示した。

第二章では、モデル解釈について本研究の核心となる新手法に踏み込んだ。モデル解釈 (MI) は、人 (ユーザ) のために機械学習モデルの予測プロセスを説明することを意味する。ディープモデルのブラックボックス性により、何を学習したかを理解したり、モデルの信頼性と改善を検討したりするためには MI が不可欠となる。おおむね MI は、非モデル依存 (MA) とモデル依存 (MS) の2つのカテゴリに分類できる。非モデル依存のアプローチは、モデルの内部構造なしで入力と出力の間の関連性を順方向に検証するのに対し、モデル依存の場合は、モデルの内部構造を使用して出力から入力に関連性を逆方向に伝播させる。

本章では、独自に開発したディープモデルの解釈可能性を高めるため、MS手法である非線形関連性伝搬 (NRP) と MA手法である Key Input Subset Sampling (KISS) [8] の2つの MI メソッドを示した。NRP を適用し、質問応答モデルである Attentive Pooling Network (APN) を解釈し従来の線形関連性伝搬 (LRP) と比較評価を行った。適切に非線形関数を選んで適用すると、要な入力まで補足する説明が可能となった。一方の KISS は、エネルギーベースモデル (EBM) 理論に基づく手法で、深層モデルと入力、出力を一つの閉じた系とみなし、各入力要素が各出力の分類に対する影響量を重要度スコア (EBIS) として計算する。ピクセルの一部を削除した画像分類モデルの評価から、KISS の性能が比較対照モデルを上回ることを確認した。また、6枚の画像に対する KISS の可視化機能を使用して、MI が知識発見を

支援することによる効果を調査するアンケート調査を実施し、可視化により人が目立ちにくい重要部分に気付くのに役立つことがわかった。MS手法とのハイブリッド化も実現した。

第三章では、ケーススタディとしてレーザー加工の表面画像を対象とした。画像のフーリエ変換パターンに対し、探索的分析のため主成分分析（PCA）を適用した。次にスペクトルパターンのデータに対して適用した特徴抽出のパフォーマンスを評価するために、シングルタスクまたはマルチタスクの多種のモデルにAlexNetとResNetを採用し、特徴抽出にAlexNetを使用する方が、2つのタスクの両方で比較手法よりも優れていることがわかった。

第四章は、悲観的対照学習の異常検出への適用研究である。系列データにおいて異常ラベル付けが困難であるという問題を解決するため、自己教師の悲観的対照学習（PCL）を提案した。PCLでは、シーケンスに異常なデータポイントがあると想定し、コンテキストのウィンドウ内でデータポイントを相互に対比して異常を認識する。また、異常を考慮した関係を学習するためのコンテキスト指向相対エントロピー（CARE）損失と組み合わせマルチタスク学習アルゴリズムを開発した。実験では、PCLを1次元の合成データで一般的に使用される2つの異常検出方法と比較し、PCLを適用した場合に判読しづらい手書き数字を最も高精度で検出することができた。レーザー加工の過程を監視するため、加工状態についての系列画像データを対象とする応用においても、深層マルチタスク学習手法による特徴抽出を用いてモデル解釈手法と異常検知手法を適用し有意な成果を得た。

第五章で述べるように、機械学習の研究は、進めるほどに新たな課題への地平線が開ける。今後に向けて、各手法のパラメーター選択、ハイパーパラメーター最適化などの問題を解決するだけでなく、知識発見のための人間と深層学習のコラボレーションを促進し、一層広く深く独自技術の適用範囲を広めることが期待される。

これまでに審査委員からは提案手法における応用を中心として、特に人の理解プロセスとの相互作用についてエビデンスを示すべきとの指摘事項があった。これらのうち、第二章を中心として指摘に応えた成果を盛り込んだ充実した博士論文となった。以上を総括すると、新規性、有用性に加え、研究の完成度と外部発表実績についても評価できる。学術上は工学的手法をもって新規性と実用性を備えた機械学習および可視化の手法を提案・実装しその効果を検証したものであり、実務上も有益な成果であると認められる。

よって本論文は博士（工学）の学位請求論文として合格とする。

以上