様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　５ 年　 ６ 月　 ７ 日現在

機関番号：１７１０２
研究種目：基盤研究(B)（一般）
研究期間：2019～2021
課題番号：１９Ｈ０４０８６

研究課題名（和文）深層学習システムの自動テスト技術の確立

研究課題名（英文）Automated Testing of Deep Learning Systems

研究代表者
趙　建軍（Zhao, Jianjun）

九州大学・システム情報科学研究院・教授

研究者番号：２０２９９５８０

交付決定額（研究期間全体）：（直接経費）　　13,440,000円

研究成果の概要（和文）：本研究では、深層学習システムにおける系統的な自動テスト技術を確立することを目的としている。具体的な研究成果は以下の通りである：(1) 深層学習システムの総合的なテストカバレッジ基準を設計・開発した。(2) 深層学習システムにおける不具合の自動テスト生成フレームワークを構築した。(3) 深層学習システムにおける不具合の自動修正と性能向上技術を開発した。(4) 提案した手法の有効性を実用的な深層学習システムへの適用によって検証した。

本研究の進展により、深層学習システムにおける系統的な自動テスト技術とそのテスト支援環境が整備され、信頼性の高い深層学習システムの構築が期待される。

研究成果の学術的意義や社会的意義
【学術意義】本研究では、深層学習システムのテストカバレッジ基準設計、不具合の自動テスト生成フレームワーク構築、不具合の自動修正と性能向上技術の開発を行った。提案手法は実用的なシステムへの適用によって検証され、深層学習システムの評価と検証手段の整備に貢献した。

【社会意義】本研究の進展により、深層学習システムにおける系統的な自動テスト技術と支援環境が整備され、信頼性の高いシステム構築が期待される。これにより、深層学習技術は医療、交通、金融など多様な領域において高品質かつ安全なシステムとして社会にポジティブな影響を与えることが期待される。

研究成果の概要（英文）：In this research, our objective is to establish systematic automated testing techniques for deep learning systems. The specific research accomplishments are as follows: (1) We designed and developed comprehensive test coverage criteria for deep learning systems.(2) We constructed a framework for the automatic test generation of defects in deep learning systems.(3) We developed techniques for automated defect correction and performance improvement in deep learning systems.(4) We validated the effectiveness of the proposed methods by applying them to practical deep learning systems.

The progress of this research is expected to lead to the establishment of systematic automated testing techniques and supporting environments for deep learning systems, thereby enabling the construction of reliable deep learning systems.

研究分野：ソフトウエア工学

キーワード：ソフトウエアテスト　深層学習システム　安全性と信頼性

１．研究開始当初の背景

最近、深層学習（deep learning、DL）は画像処理、音声認識、囲碁などの応用領域で驚異的な成功を収めており、自動運転車やロボットなど、社会インフラに関わる重要な分野での成果がますます期待されている[1,2]。一方で、DL システムに障害が発生した場合、社会や自然に大きな災害をもたらす可能性があるため、その信頼性への要求はますます高まっている。信頼性の高い DL システムを系統的な方法で効率的に開発することは、情報化社会からの重要な要請であり、情報化社会の安定と発展にとって急務な課題となっている。

DL システムでは、優れた機能を持つ一方で、偏った訓練データや過適合、過小評価などの理由により、特定のケースで誤った挙動を示すことが多々ある。また、現在の DL システム開発に応用されている工学的な方法は未熟であり、開発を支援するツールも不十分である。そのため、信頼性の高い DL システムを開発することは非常に困難である。特に、セーフティクリティカルなシステムやミッションクリティカルなシステムに適用される場合、DL システムの信頼性と安全性に大きな疑念が生じている。最近では、DL システムの問題によって深刻な事故や多大な損失が発生している。例えば、Google/Tesla の自動運転車事故[3,4]や、1 ピクセル攻撃による DL の堅牢性の問題[5]などがある。これらの事故の原因は、稀な条件下で DL システムに不具合などの問題が発生し、予期せぬ動作が引き起こされたためである。このようなケースはテストセットの一部ではなかったため、テスト中にこのような挙動が現れなかった。そのため、従来のソフトウェアと同様に、信頼性と安全性が重要な DL システムにおいて、潜在的な不具合や望ましくない動作を検出するために、コーナーケースごとに系統的なテストが重要である。しかし、数千のニューロンと数百万のパラメータを持つ実世界の DL システムでは、完全自動化による潜在的な不具合の究明は非常に困難である。また、検出された不具合の修正も一層困難である。

２．研究の目的

本研究は、上記の課題を解決するため、DL システムにおける自動テスト技術の確立を目指している。具体的には、以下の項目に取り組んでいます：(1) DL システムの総合的なテストカバレッジ基準の設計と開発、(2) DL システムの自動テスト生成フレームワークの構築、(3) DL システムにおける不具合の自動修正と性能向上技術の開発、(4) 提案手法の有効性を実用 DL システムへの適用によって検証する。

本研究の進展により、DL システムにおける系統的な自動テスト技術とその支援環境が整備され、信頼性の高い DL システムの開発が期待される。さらに、自動運転システムに関連するテストソリューションを活用することで、高品質かつ信頼性の高い自動運転知能ソフトウェアシステムの開発が可能と考えられる。

３．研究の方法

本研究では、2019 年度から 2021 年度までの 3 年間計画で、以下の研究項目(1)、(2)、(3)、(4)に分けて実施した。

研究項目(1)　DL システムの総合的なテストカバレッジ基準の設計と開発。
テストカバレッジ基準はソフトウェア品質を評価するための基本的かつ最も重要な方法であり、従来のソフトウェアのテスト効果を評価するための共通の基準となっている。しかし、DL ソフトウェア［6,9］は従来のソフトウェアと比べ、基本的な違いがあるため、本項目では、DL システムのテストカバレッジ基準を調査し、新しい DL システムのテストカバレッジ基準を提案し、その支援ツール群を開発する。これらの基準は主にディープニューラルネットワーク（deep neural network, DNN）における制御フロー及びデータフローを考慮し、テストの有効性を分析することを可能にする。また、DL システムにおける実行時の挙動の理解を容易にすることや DL システム開発者にフィードバックを与えてソフトウェアおよびテストスイート（test suite）の改善のヒントを提供することも可能にする。

研究項目(2)　DL システムの自動テスト生成フレームワークの開発。
(a)　テスト生成のためのメタモルフィック関係に関する研究
従来の DL システムをテストする際、メタモルフィック関係（metamorphic relation, MR)を用いてテストを生成している[6,7,8]。しかし、各 MR が不具合検出に及ぼす影響はまた不明であり、MR の間にどのような関係を持つのかについても研究されていない。このため、本項目では、厳密な統計解析による MR の関係を系統的に調査し、不具合検出能力との関係を明らかにするとともに、テスト生成のための最小限の MR ベースのミューテーション演算子(mutation operator)を形成する新しい MR を提案する(図 1）。ここでは、効率を上げるため、任意の関係の代わりに、

物理的な世界 (光条件、コントラスト、回転など)で発生する可能性のある MR 関係のみを考慮する。

(b) 自動テスト生成フレームワークの開発
本項目では、図 1 に示されているように、完全自動化する DL システムのグレイボックス誘導（grey-box guided）テスト生成フレームワークを開発する。最初のシード（seed）は、DL 開発者から準備された最小シードデータである。最初のシードテストは、手動で選択するか自動的に作成するかが選択できるが、対象タスクのアプリケーションのドメインに配置する必要がある。テスト生成の繰り返しが開始されると、テストフレームワークは対象 DL システムとシードの入力を受け付け、入力されたシードをバッチプールに前処理してから、大規模な並列処理を行う。次に、テストデータをバッチプールからランダムに選択し、バッチ内のテストデータをサンプリングして（項目 B1 から）ミューテーション演算子を実行し、新しいテストを生成する。生成されたテストを提案された MR 関係で検証し、MR 関係に違反すれば、さらなる調査が必要である。パスしたテストに対して、得られたカバレッジ（A で提案されたもの）について分析し、ミューテーションのさらなる選択確率を優先させてバッチプールに入れる。テストの反復は、計算リソースが使い果たされるまで続く。テストフレームワークは、完全自動化よりスケーラビリティがあり、系統的なカバレッジガイダンスで新しいテストを生成することができる。

研究項目(3) DL システムにおける不具合の自動修正と性能向上技術の開発。
(a) データ駆動型 DL システムにおける不具合の修正
項目 B で開発された自動テスト生成フレームワークを用いて、自動的に生成されたテストにより対象 DL ソフトウェアの不具合を引き起こすことを可能にした。次のステップでは、検出された DL モデルの不具合を自動的に修正することである。本項目では、項目 A で提案されたテストカバレッジ基準を組み合わせ、DL モデルの内部の役割や振る舞いを自動的に分析する。特に、異なるニューロンとレイヤーの役割がどのようになっている、正しいデータと検出された不具合を引き起こしたデータに対してどのように決定しているかを分析する方法を提案する。これに基づいて、ニューロンの重要度を分類し、一般的に活性化されたニューロンおよび行動発散ニューロンに、強い、中、および軽いのように定量的に示す。このような情報に基づいて、主要な共通ニューロンを凍結させ、正しいデータと不具合のあるデータのギャップを減らすことができるようなニューロンの動作をトリガーするようにオリジナルのトレーニングセットからデータを選択するようにガイドする。このようにデータを活用して、実行時 DL の動作を継続的にトレーニングし、微調整することで、DL システムの不具合を自動的に修正する。

(b) DL システムのカスタマイズと最適化
DL システムの開発、テスト（項目 B）および修正フェーズ（項目 C1）に成功すると、実用的なプラットフォームとデバイスに展開する準備が整った（図 2 参照）。しかし、異なるデバイスの計算能力およびリソースの制限があるため、DL ソフトウェアは、しばしばターゲットプラットフォームのカスタマイズおよび最適化をする必要がある。本項目では、主に適切なトレーニングを受けた DL モデルを GPU 対応サーバーからモバイルデバイスプラットフォーム(Android や IOS など）に導入することを検討する。精度の低下なしに、計算効率、メモリ消費、エネルギー効率を大幅に向上させる量子化と DL ソフトウェア圧縮技術を提案する。項目 B で生成されたテストの一部は回帰テストであり、これらのテストを再利用し、特定のプラットフォームのカスタマイズや最適化の手順で予測精度の問題が発生するかどうかを分析する。もし元の DL ソフトウェアでのテストが成功し、カスタマイズされたバージョンの展開（deployment）されたモバイルデバイスでは失敗する場合、一定の品質が満たされるまで修正手順を繰り返す必要がある。さらに、トレーニング、予測、展開の各段階から DL ソフトウェアの品質を系統的に評価する。

研究項目(4) 実用システムへの適用より本研究で提案した手法の有効性の検証。
本項目では、項目 A、B、C で提案された DL システム自動テスト関連のソリューションを実用的な DL ベースの自動運転ステアリング（steering）制御システムへ適用し、その有効性を検証する。まず、Udacity の自動運転試験環境をセットアップし、自動化された DL テストおよび修正ソリューションを適用する。 この段階では、Udacity の自動運転挑戦で競争力のあるパフォーマンスを得ることができる現在の最先端の DL モデル（例えば、Chauffeur、Rambo）を選択する。自動運転環境を設定した後、項目 A、B、C で提案された解決策を自動運転のシナリオに適用し、大規模なテストを生成し、DL モデルの潜在的な問題を検出する。 エラートリガテストが行われた後、問題があるモデルに対して自動修正を行い、品質と堅牢性を向上させる。さらに、修正前後の DL モデルの頑健性（敵対的攻撃による）を比較する大規模な分析を行う。 本研究で提案されたソリューションでは、DL ソフトウェアの潜在的な問題を系統的に検出し、これらの問題を解決まで完全自動化することができることを実験で検証する。

４．研究成果

(1) DL システムのテストカバレッジ基準を調査し、新しい DL システムのテストカバレッジ基準を提案し、その支援ツール群を開発した。

(2) 厳密な統計解析による MR の関係を系統的に調査し、不具合検出能力との関係を明らかにするとともに、テスト生成のための最小限の MR ベースのミューテーション演算子(mutation operator)を形成する新しい MR を提案した。また、完全自動化する DL システムのグレイボックス誘導（grey-box guided）テスト生成フレームワークを開発した。

(3) 提案されたテストカバレッジ基準を組み合わせ、DL モデルの内部の役割や振る舞いを自動的に分析した。特に、異なるニューロンとレイヤーの役割がどのようになっている、正しいデータと検出された不具合を引き起こしたデータに対してどのように決定しているかを分析する方法を提案した。このようにデータを活用して、実行時 DL の動作を継続的にトレーニングし、微調整することで、DL システムの不具合を自動的に修正する手法を提案した。

(4) 提案された解決策を自動運転のシナリオに適用し、大規模なテストを生成し、DL モデルの潜在的な問題を検出し、エラートリガテストが行われた後、問題があるモデルに対して自動修正を行い、品質と堅牢性を向上させることができた。さらに、修正前後の DL モデルの頑健性（敵対的攻撃による）を比較する大規模な分析を行い、本研究で提案されたソリューションでは、DL ソフトウェアの潜在的な問題を系統的に検出し、これらの問題を解決まで完全自動化することができることを実験で検証した。

＜引用文献＞

[1] Ian Goodfellow and Nicolas Papernot. The Challenge of Verification and Testing of Machine Learning. http://www.cleverhans.io/security/privacy/ml/2017/06/14/verification.html. (2017).

[2] Dario Amodei, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, and Dan Mane. Concrete Problems in AI safety. arXiv preprint arXiv:1606.06565, 2016.

[3] Google-accident 2016. A Google Self-driving Car Caused a Crash for the First Time. http://www.theverge.com/2016/2/29/11134344/google-selfdriving-car-crash-report. (2016).

[4] BBC News, Tesla in fatal California crash was on Autopilot, 2018, https://www.bbc.com/news/world-us-canada-43604440

[5] BBC News, AI image recognition fooled by single pixel change, 2018, AI image recognition fooled by single pixel change.

５．主な発表論文等

〔雑誌論文〕　計0件

〔学会発表〕　計20件（うち招待講演　1件／うち国際学会　20件）

１．発表者名
Zhuo Li, Derui Zhu, Yujing Hu, Xiaofei Xie, Lei Ma, Yan Zheng, Yan Song, Yingfeng Chen, and Jianjun Zhao

２．発表標題
Neural Episodic Control with State Abstraction

３．学会等名
The 11th International Conference on Learning Representations (ICLR 2023) (Spotlight)（国際学会）

４．発表年
2023年

１．発表者名
Zhenya Zhang, Deyun Lyu, Paolo Arcaini, Lei Ma, Ichiro Hasuo, Jianjun Zhao

２．発表標題
FalsifAI: Falsification of AI-Enabled Hybrid Control Systems Guided by Time-Aware Coverage Criteria

３．学会等名
IEEE Transactions on Software Engineering, Vol.49, No.4, pp.1842-1859（国際学会）

４．発表年
2023年

１．発表者名
Bing Yu, Hua Qi, Qing Guo, Felix Juefei-Xu, Xiaofei Xie, Lei Ma, Jianjun Zhao

２．発表標題
DeepRepair: Style-Guided Repairing for Deep Neural Networks in the Real-World Operational Environment

３．学会等名
IEEE Transactions on Reliability, p.1401-1416, December 2022（国際学会）

４．発表年
2022年

１．発表者名
Ziyi Cheng, Xuhong Ren, Felix Juefei-Xu, Wanli Xue, Qing Guo, Lei Ma, Jianjun Zhao

２．発表標題
Deepmix: Online auto data augmentation for robust visual object tracking

３．学会等名
2021 IEEE International Conference on Multimedia and Expo (ICME 2021), pp.1-6（国際学会）

４．発表年
2021年

１．発表者名
Qing Guo, Ziyi Cheng, Felix Juefei-Xu, Lei Ma, Xiaofei Xie, Yang Liu, and Jianjun Zhao

２．発表標題
Learning to Adversarially Blur Visual Object Tracking

３．学会等名
International Conference on Computer Vision, (ICCV 2021), p.10839-10848（国際学会）

４．発表年
2021年

１．発表者名
Qing Guo, Jingyang Sun, Felix Juefei-Xu, Lei Ma, Xiaofei Xie, Wei Feng, Yang Liu, and Jianjun Zhao

２．発表標題
EfficientDeRain: Learning Pixel-wise Dilation Filtering for High-Efficiency Single-Image Deraining

３．学会等名
The 35th AAAI Conference on Artificial Intelligence (AAAI 2021)（国際学会）

４．発表年
2021年

１．発表者名
Xiaoning Du, Yi Li, Xiaofei Xie, Lei Ma, Yang Liu, Jianjun Zhao

２．発表標題
Marble: Model-Based Robustness Analysis of Stateful Deep Learning Systems

３．学会等名
The 35th IEEE/ACM International Conference on Automated Software Engineering (ASE 2020)（国際学会）

４．発表年
2020年

１．発表者名
David Berend, Xiaofei Xie, Lei Ma, Lingjun Zhou, Yang Liu, Chi Xu, and Jianjun Zhao

２．発表標題
Cats Are Not Fish: Deep Learning Testing Calls for Out-Of-Distribution Awareness

３．学会等名
The 35th IEEE/ACM International Conference on Automated Software Engineering (ASE 2020)（国際学会）

４．発表年
2020年

１．発表者名
Xuhong Ren, Bing Yu, Hua Qi, Felix Juefei-Xu, Zhuo Li, Wanli Xue, Lei Ma, and Jianjun Zhao

２．発表標題
Few-Shot Guided Mix for DNN Repairing

３．学会等名
Proc. 36th IEEE International Conference on Software Maintenance and Evolution (ICSME 2020)（国際学会）

４．発表年
2020年

１．発表者名
Hua Qi, Qing Guo, Felix Juefei-Xu, Xiaofei Xie, Lei Ma, Wei Feng, Yang Liu, and Jianjun Zhao

２．発表標題
DeepRhythm: Exposing DeepFakes with Attentional Visual Heartbeat Rhythms

３．学会等名
The 28th ACM International Conference on Multimedia (ACM MM 2020)（国際学会）

４．発表年
2020年

１．発表者名
Xiyue Zhang, Xiaofei Xie, Lei Ma, Xiaoning Du, Qiang Hu, Yang Liu, Jianjun Zhao, and Meng Sun

２．発表標題
Towards Characterizing Adversarial Defects of Deep Learning Software from the Lens of Uncertainty

３．学会等名
The 42nd International Conference on Software Engineering (ICSE 2020)（国際学会）

４．発表年
2020年

１．発表者名
Xiongfei Wu, Liangyu Qin, Bing Yu, Xiaofei Xie, Lei Ma, Yinxing Xue, Yang Liu, and Jianjun Zhao

２．発表標題
How are Deep Learning Models Similar? An Empirical Study on Clone Analysis of Deep Learning Software

３．学会等名
The 28th International Conference on Program Comprehension (ICPC 2020)（国際学会）

４．発表年
2020年

１．発表者名
Yang Liu, Lei Ma, and Jianjun Zhao

２．発表標題
Secure Deep Learning Engineering: A Road towards Quality Assurance of Intelligent Systems

３．学会等名
Proc. 21st International Conference on Formal Engineering Methods (ICFEM 2019)（招待講演）（国際学会）

４．発表年
2019年

１．発表者名
Qiang Hu, Lei Ma, Xiaofei Xie, Bing Yu, Yang Liu, and Jianjun Zhao

２．発表標題
DeepMutation++: a Mutation Testing Framework for Deep Learning Systems

３．学会等名
Proc. 34th IEEE/ACM Conference on Automated Software Engineering (ASE 2019)（国際学会）

４．発表年
2019年

１．発表者名
Xiaoning Du, Xiaofei Xie, Yi Li, Lei Ma, Yang Liu, and Jianjun Zhao

２．発表標題
A Quantitative Analysis Framework for Recurrent Neural Network

３．学会等名
Proc. 34th IEEE/ACM Conference on Automated Software Engineering (ASE 2019)（国際学会）

４．発表年
2019年

１．発表者名
Xiaofei Xie, Hongxu Chen, Yi Li, Lei Ma, Yang Liu, and Jianjun Zhao

２．発表標題
DeepHunter: A Coverage-Guided Fuzzer for Deep Neural Networks

３．学会等名
Proc. 34th IEEE/ACM Conference on Automated Software Engineering (ASE 2019)（国際学会）

４．発表年
2019年

１．発表者名
Qianyu Guo, Sen Chen, Xiaofei Xie, Lei Ma, Qiang Hu, Hongtao Liu, Yang Liu, Jianjun Zhao, Xiaohong Li

２．発表標題
An Empirical Study towards Characterizing Deep Learning Development and Deployment across Different Frameworks and Platforms

３．学会等名
Proc. 34th IEEE/ACM Conference on Automated Software Engineering (ASE 2019)（国際学会）

４．発表年
2019年

１．発表者名
Xiaoning Du, Xiaofei Xie, Yi Li, Lei Ma, Yang Liu and Jianjun Zhao

２．発表標題
DeepStellar: Model-Based Quantitative Analysis of Stateful Deep Learning Systems

３．学会等名
Proc. 27th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2019)（国際学会）

４．発表年
2019年

１．発表者名
Xiaofei Xie, Lei Ma, Felix Juefei-Xu, Minhui Xue, Hongxu Chen, Yang Liu, Jianjun Zhao, Bo Li, Jianxiong Yin, and Simon See

２．発表標題
DeepHunter:A Coverage-Guided Fuzz Testing Framework for Deep Neural Networks

３．学会等名
Proc. 28th ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA 2019)（国際学会）

４．発表年
2019年

１．発表者名
Chao Xie, Hua Qi, Lei Ma, and Jianjun Zhao

２．発表標題
DeepVisual: A Visual Programming Tool for Deep Learning Systems

３．学会等名
Proc. 27th IEEE/ACM International Conference on Program Comprehension (ICPC 2019)（国際学会）

４．発表年
2019年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

―

6．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
研究分担者	鵜林　尚靖 (Ubayashi Naoyasu) (80372762)	九州大学・システム情報科学研究院・教授 (17102)	
研究分担者	亀井　靖高 (Kamei Yasutaka) (10610222)	九州大学・システム情報科学研究院・准教授 (17102)	
研究分担者	馮　尭楷 (Feng Yaokai) (60363389)	九州大学・システム情報科学研究院・助教 (17102)	
研究分担者	馬　雷 (Ma Lei) (70842061)	九州大学・システム情報科学研究院・准教授 (17102)	

7．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計2件

国際研究集会 The 11th Asia-Pacific Symposium on Internetware (Intertware 2019)	開催年 2019年～2019年
国際研究集会 The 8th Asian-Pacific Workshop on Advanced Software Engineering (AWASE 2019)	開催年 2019年～2019年

8．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関