様式３

論　文　内　容　の　要　旨

氏　名　（西田圭吾）	
論文題名	Stable Learning Algorithm of Bayesian Deep Neural Networks towards Low Power Dedicated Processor （専用プロセッサによる低電力実行に向けたベイジアン深層ニューラルネットワークの安定した学習アルゴリズム）

論文内容の要旨

Deep learning has been studied by researchers interested in understanding the basic principles of information engineering, such as those pertaining to computer vision and language processing. In addition, deep learning has been investigated for application to fields such as automated driving, medical diagnostics, and drug discovery. Recent advances in deep learning have resulted in vast performance improvements, because continuous improvements in the computational power according to Moore's Law have made it possible to handle large-scale neural network models. However, existing deep learning models tend to lead to overconfident inferences, and this tendency becomes even stronger for inputs with distributions that differ from those of the training data, as is often the case in practical applications. Reliable inferences via deep learning are essential to precisely capture predictive uncertainty. In this regard, Bayesian neural networks (BNNs) are expected to capture predictive uncertainty to a high degree of accuracy because of their ability to model parameter uncertainty. However, in large-scale NNs such as those employed for deep learning, BNNs are problematic in that the stable convergence of parameters is difficult and the learning process of these NNs is computationally costly. Further development of Bayesian deep learning would therefore need to focus on improving the learning method and developing a dedicated processor with high power efficiency. Attempts to address these shortcomings are expected to improve Bayes by Backprop (BBB), an algorithm for BNNs that optimizes the variational parameter $\boldsymbol{\theta} = \{\boldsymbol{\mu}, \boldsymbol{\rho}\}$. The advantage of BBB is its capability to deliver efficient inference with dedicated processors; however, because the variational parameters do not easily converge, the efficiency of learning with dedicated processors has not been sufficiently discussed in previous studies.

The objective of the research presented in this thesis was to develop algorithms for stable parameter convergence with BBB. The aim was to design these algorithms such that they are suitable for efficient learning with dedicated Bayesian deep learning processors. These algorithms and hardware are expected to enable reliable predictions of complex phenomena by BNNs. BBB operates by sampling neural network weights from variational parameters, which makes the loss function noisy. Although the parameters of noisy loss functions can be made to successfully converge using Adam as the optimizer, for BBB this optimizer prevents stable parameter convergence. To overcome this problem, I proposed using Adam with decoupled Bayes by Backprop (AdamB), which decouples the log-likelihood terms of the prior and posterior distributions in the BBB cost function from Adam. Using a covariate shift benchmark of the image classification tasks with a shifted distribution from the training data, I evaluated the accuracy and reliability of the models trained by AdamB. In addition, I discuss dedicated hardware architectures that enable AdamB to conduct training efficiently.

My study revealed that the difficulty of learning with BBB lies in the rapid and excessive update of parameter $\boldsymbol{\rho}$. This problem was fundamentally solved by AdamB using a Gaussian scale mixture prior (SM prior). In this research, I demonstrated that the rapid updates of parameter $\boldsymbol{\mu}$ can be decoupled from Adam, whereas the excessive increase can be suppressed by using an SM prior. Experiments also showed that parameter $\boldsymbol{\mu}$ takes a sparse distribution and, by using the SM prior, is strongly robust against noise type corruption. Currently, the robustness of AdamB has only been proven for covariate shift in image identification tasks, and its applications to other tasks would have to be assessed in future. Additionally, by examining the basic architecture of Movitan, a system-on-a-chip (SOC) generator system capable of enabling AdamB to efficiently train neural networks, I provided guidelines for a dedicated processor capable of efficient and stable Bayesian deep learning. This study makes a fundamental contribution to the existing knowledge base by proposing a novel approach to solve the difficulties associated with handling BNNs because of their unstable convergence of parameters and increased computational cost.

様式7

論文審査の結果の要旨及び担当者

氏　名　（　西田　圭吾　）			
論文審査担当者		（職）	氏　名
	主 査	招へい教授	泰地 真弘人
	副 査	教授	上田 昌宏
	副 査	教授	北澤 茂
	副 査	教授	西本 伸志
	副 査	産学連携教授	加納 敏行

論文審査の結果の要旨

近年の深層学習の発展は、コンピュータビジョンや言語処理などの分野を中心に進み、医療・創薬や生命科学研究にも応用が始まっている。既存の多くの深層学習モデルは自信過剰な推論を行う傾向があり、予測の不確実性を的確に捉えた信頼性の高い推論が求められている。Bayesian Neural Network (BNN) は、パラメータの不確実性をモデル化することにより、予測の不確実性をより高い精度で表現することができる。しかし、深層BNNはパラメータの安定的な収束が難しく、またその学習に計算コストがかかるという問題がある。そのため、ベイズ深層学習の発展には、学習方法の改善と電力効率の高い専用プロセッサの開発が必要である。
本論文ではまず、BNNの典型的な学習手法であるBayes by Backprop (BBB)の改良を行った。BBBは専用プロセッサによる効率的な推論を実現できる手法であるが、変分パラメータが容易に収束しない問題があった。BBBでは損失関数がノイジーになる。一般的には、adaptive moment estimator (Adam)をoptimizerとして用いると、ノイジーな勾配でも安定な収束が得られるが、BBBの場合にはAdamが安定な収束を妨げていること、その理由を解明した。解決策として、BBBのコスト関数に含まれる事前分布・事後分布の対数尤度の項をAdamから切り離した“Adam with decoupled Bayes by Backprop (AdamB)”を提案した。AdamBを画像識別タスクの共変量シフトベンチマークに対して適用し、その精度と信頼性の評価を行った。現在信頼性の高い学習手法として使われているアンサンブル手法と比較し、優れた性能を確認した。
ここまでの成果は、IEEE Access誌に発表されている。
さらに、AdamBによる学習を実現するシステムオンチップ (SoC) 生成システム“Movitan”の基本アーキテクチャを検討し、効率的かつ安定的なベイズ深層学習向け専用プロセッサの指針を示した。
これらの成果は、収束が不安定かつ計算コストが高いというBNNの困難を解決するための新しいアプローチを提案するもので、深層ベイズ学習の本質に迫る要素を多く含む重要な成果であり、かつ高いオリジナリティを有すると認められた。論文の内容は数理アルゴリズムからハードウェア開発までを含むもので、多分野に跨る境界領域研究としても評価される。また、ノイズの多い系に対して信頼度の高い学習を可能とする本手法は、今後の生命科学における画像解析等に広く適用可能であると期待される。これらを総合し、博士の学位を授与するに値するものと認める。