



Title	Data- and Parameter-Efficient Adaptation for Foundation Models and Applications
Author(s)	Zhang, Jiahao
Citation	大阪大学, 2026, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/105827
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

論文内容の要旨

氏 名 (Jiahao Zhang)	
論文題名	Data- and Parameter-Efficient Adaptation for Foundation Models and Applications (基盤モデルおよびその応用に向けたデータ効率的・パラメータ効率的適応)
論文内容の要旨 Foundation models have revolutionized artificial intelligence. While they possess strong generalization capabilities, tailoring them to specialized domains typically requires task adaptation beyond off-the-shelf prompting to approach optimal performance, which can be resource-intensive due to annotation costs and computational demands. To accommodate real-world constraints, this thesis explores data- and parameter-efficient adaptation, establishing a spectrum of adaptation strategies ranging from parameter tuning to input-level optimization and training-free interaction. We present three complementary projects that advance this premise across different levels of model access. First, regarding parameter-efficient fine-tuning for volumetric medical data, we propose ILD-Slider, a framework integrating anatomy-guided slice selection with a novel Slice-level 3D Adapter for efficient diagnosis. Second, shifting to input-level optimization for Visual In-Context Learning (VICL), we propose E-InMeMo, which introduces learnable perturbations to input prompts to enhance few-shot performance without backbone tuning. Third, investigating training-free interaction for proprietary generative models where gradients are inaccessible, we propose VASCAR, a visual-aware self-correction framework that steers black-box models through iterative visual feedback for superior layout generation. Collectively, these studies demonstrate effective strategies for adapting foundation models across varying degrees of parameter access and data availability, offering scalable solutions for specialized applications.	

論文審査の結果の要旨及び担当者

氏 名 (Jiahao Zhang)			
	(職)	氏 名	
論文審査担当者	主 査	教授	中島悠太
	副 査	教授	瀬尾茂人
	副 査	教授	長原一
	副 査	准教授	大倉史生

論文審査の結果の要旨

本本学位論文は、様々なダウンストリームタスクに適用可能な人工知能モデルである基盤モデルについて、追加学習に必要となるデータ数および学習対象パラメータ数を抑えつつ、特定タスクでの精度向上を実現する方法を論じる。具体的には、少数のパラメータをチューニングして、基盤モデルの学習で用いられる一般的な画像とは大きく異なるドメインである医用画像へ適応する手法（第2章）、基盤モデルのパラメータをチューニングせずに入出力例を併せて入力することで、例示された物体検出・セグメンテーションタスクの精度を向上させる手法（第3章）、さらにパラメータも入出力例も用いず、大規模言語モデルが持つ知識のみを利用してレイアウトデザインを実現する手法（第4章）を扱い、これらの研究成果に基づいて構成される。

第2章では、3次元CT画像を用いた間質性肺疾患の診断モデルを対象に、基盤モデルを3次元CT画像へ適応させる手法を論じる。医用画像はプライバシー等の理由から大量収集が難しく、人工知能モデルの学習が困難なドメインの一つである。本論文では、CTスキャンから得られる3次元データに対し、すべての2次元スライスを基盤モデルの入力として扱うことで、少数パラメータのアダプタであっても基盤モデルの適応を可能にし、中間表現においてスライス間の依存関係を扱うことによって、2次元スライが入力であっても3次元構造を考慮した判断を可能にしている。3次元データを、2次元画像を入力とする基盤モデルで効率よく扱う手法を提案しており、学術的価値が高いと考える。

第3章では、画像生成モデルに対し、例となる入力・出力画像と対象画像を一枚にまとめて入力する（プロンプト画像と呼ぶ）ことで、例により示されるタスクを実現する **In-context Learning** に着目する。提案手法は、プロンプト画像に学習可能なパラメータを追加することで、タスク精度を向上させるものである。具体的には、プロンプト画像の周囲に学習可能なパラメータを配置し、タスクごとにそれらを調整することで、基盤モデルに対してタスク固有の条件付けを行う。これらのパラメータは、タスク特有の出力が得られやすいよう中間表現をシフトさせる役割を担うと考えられる。基盤モデル本体のパラメータを調整する代わりに、入力側画像にパラメータを付与する点で、**In-context Learning** の新しい活用方法を提案するものと言える。

第4章では、大規模視覚言語モデルを用いてレイアウトデザインを出力する手法を提案する。本手法はモデルのパラメータを調整せず、基盤モデルは生成能力よりも評価能力の方が相対的に高い（生成はうまくできなくても、評価はできる）という点に着目する。提案手法は、基盤モデルが生成したレイアウトを評価し、修正点を言語で記述して基盤モデルに再入力する自己修正を繰り返すことで、出力されるレイアウトデザインを段階的に改善する。この方法によって、パラメータを一切調整せずとも出力の質を高められる。これは、新しいプロンプティング戦略を提案するものと言える。

以上より、本学位論文は、基盤モデルを少数パラメータの調整、またはパラメータを全く調整しない形で各種タスクへ適応させる方法論を提案しており、情報科学分野に大きく寄与するものである。よって、博士（情報科学）の学位論文として価値のあるものと認める。