

氏 名 豊田 祥史

学位(専攻分野) 博士(統計科学)

学位記番号 総研大甲第 2406 号

学位授与の日付 2023 年 3 月 24 日

学位授与の要件 複合科学研究科 統計科学専攻  
学位規則第6条第1項該当

学位論文題目 A Study Toward Practical Application of Domain Invariance  
Learning: Domain Invariance Estimation with Coarse Labels  
and its Hyperparameter Selection

論文審査委員 主 査 日野 英逸  
統計科学専攻 教授  
吉田 亮  
統計科学専攻 教授  
福水 健次  
統計科学専攻 教授  
山田 誠  
京都大学 大学院情報学研究科 准教授

(様式3)

## 博士論文の要旨

氏名 豊田 祥史

論文題目: A Study Toward Practical Application of Domain Invariance Learning: Domain Invariance Estimation with Coarse Labels and its Hyperparameter Selection

Training data used in machine learning may contain features that are spuriously correlated to the labels of data. Machine learning models often learn such spurious correlations embedded in training data and hence may fail to predict desired labels of test data generated by a different distribution from one to provide training data. In classification of animal images, models tend to misclassify cows on sandy beaches since most training pictures are taken in green pastures and models inherit context information in training. Another example is learning from medical data. Systems trained with data collected in one hospital do not generalize well to other hospitals; models unintentionally extract factors specific to a particular hospital in training.

To address the problem of spurious correlations, "out-of-distribution (o.o.d.) generalization" has been recognized as an important issue for the future of machine learning; here, the term "o.o.d. generalization" means generalization to data generated by a distribution outside training ones.

Domain Invariance Learning (DIL) is a rapidly developed approach for o.o.d. generalization. We use the term domain to specify the bias in the distribution. Using training data from multiple domains, IL estimates a predictor invariant to the change of domains, aiming at keeping good performance in unseen domains as well as in the training domains.

While the DIL approach has attracted much attention, they have two shortcomings in practice. Firstly, requiring training data from multiple domains often involves expensive data annotation. In real-world data, labels may be missing or incomplete; in some cases, data may only specify classes to which the image does not belong. Such data with insufficient annotation are not directly applicable to the standard DIL methods; they must be re-annotated accurately, often at great financial or human expense. The high cost drives a strong need to establish a new DIL framework without or with lower annotation costs. The second drawback is hyperparameter selection. Most DIL methods involve hyperparameters to balance the classification accuracy and the degree of invariance. It is known that most DILs give high predictive performance only when a hyperparameter is selected by using unseen test data; without them, simple hyperparameter selection methods fail to find a preferable hyperparameter.

In the Ph.D. thesis, we propose a novel domain invariance learning framework to mitigate the annotation cost problem of DIL. In the new DIL framework, we consider the situation where the training data of target classification is given in only one domain, while the task with coarser labels than those of the target classification, which needs lower annotation cost, has data from multiple domains. Consider the case where a target classification has 300 labels {bird1,...,bird100, snake1,..., snake100, turtle1,..., turtle100} corresponding to 300 species. Then, the binary labels {bird, reptile} are an example of coarser labels. The decrease in the number of classes reduces annotation time per image. Moreover, annotating the binary labels does not require expert knowledge, while annotating the original 300 labels would require expert knowledge about birds, snakes, and turtles. Hence, the new IL framework significantly reduces the annotation cost in comparison with previous DIL methods; we need exhausting annotation of 300 classes only for one domain and just binary labels for other domains.

Moreover, we propose two methods of cross-validation (CV) for hyperparameter selection in our new DIL framework. Since we assume training data of a single domain for the target task, it is impossible to estimate the deviation of the risks over the domains. Our CV methods mitigate the difficulty by using additional coarser labeled data from multiple domains.

The theoretical properties of our framework, including two CV methods, are investigated. We mathematically prove that our framework can estimate a correct domain invariant predictor with a hyperparameter fixed correctly, and such a preferable hyperparameter is selected under some settings. In the result, the difference between the two CV methods is also revealed; the second CV method can select an optimal hyperparameter with milder conditions than the ones of the first method.

We numerically demonstrate that the proposed framework extracts a domain invariant predictor more effectively than other existing methods. MNIST and ImageNet experiments show the effectiveness of our framework, even when there is a large difference in the class numbers between the target and coarser classifications. We compare two CV methods with synthetic data, to numerically highlight the theoretical observations; we can demonstrate that the second method is more applicable, which accords with the theoretical result.

## 博士論文審査結果

Name in Full  
氏名 豊田 祥史

Title  
論文題目 A Study Toward Practical Application of Domain Invariance Learning:  
Domain Invariance Estimation with Coarse Labels and its Hyperparameter Selection

2023年1月23日午後3時から約2時間にわたり豊田祥史氏の博士論文審査委員会を開催した。出願者による1時間の公开发表による概要説明と質疑応答、さらに約1時間の審査委員のみによる審査を行った結果、審査委員会は本論文が学位の授与に値すると判断した。

### [論文の概要]

深層学習などデータからの学習において課題となっている、訓練データとは異なる分布を持つテストデータに対する汎化性能を向上させるための分布外汎化の枠組みに関して論じたもので、英文で書かれており全7章101頁からなる。

1章は本論文の序章であり、分布外汎化の概要と、その対処法であるドメイン不変学習に関して論じ、本論文の貢献がまとめられている。

2章は準備であり、分布外汎化の数理的な枠組みを論じ、分布が異なる複数のドメインから収集されたデータセットを用いて分布外汎化を達成しようとするドメイン不変学習に関して詳述し、現在のドメイン不変学習の問題点として、アノテーションコストの問題と、ハイパーパラメータ選択における交差検証の困難に関して論じている。

3章は本論文の主題であり、新しいドメイン不変学習の枠組みとして、複数ドメインの訓練データを用意する際に、ターゲットとしている識別問題で想定されるクラスよりも粗いクラスラベルを付与してドメイン不変学習を行う方法を提案している。二値などの粗いラベルを手で付与したり高性能の学習済み識別機で自動識別したりすることによりアノテーションコストが下げられることが論じるとともに、一定の条件の下、提案した目的関数が最適な分布外汎化を達成することを理論的に示している。また、本来のラベルよりも粗いラベルを用いた状況で、ドメイン不変学習法におけるハイパーパラメータ選択のための交差検証法を2種類提案し、それらの方法が分布外汎化を最小にするための十分条件を理論的に与えている。

4章では本論文の主結果を構成する定理を含め、理論的な結果の証明が詳述されている。

5章は関連手法を論じており、学習した深層モデルのファイン・チューニング、深層学習により得られた特徴写像のドメイン適合、分布ロバストな最適化、few-shot学習の既存手法などが論じられている。

6章は数値実験結果を論じており、人工的なデータ、カラー付きMNIST、ImageNetの3種類のデータセットに対する実験結果により、提案手法が分布外汎化の既存手法に対して分布外汎化性能に関して優位性を持っていることを示している。

7章は論文のまとめである.

[論文の評価]

本論文は, ドメイン不変学習による分布外汎化に関し, 従来の枠組みが持つアノテーションコストとハイパーパラメータ選択の問題に対して, 新しい枠組みを提案することにより解決をはかっており, 理論的な結果に加え, 実験的に従来法に対する優位性を示している点から, 統計科学の博士論文として十分な意義を持つと考える.

なお, 論文誌等への採択状況として, 3章から5章までの内容をまとめた論文が査読付き国際会議 Neural Information Processing Systems 2022 (第一著者) に採択されている.