

# Class-wise Feature Distribution Matching for Domain Generalization

Seungmin Lee  
Seoul National University  
profile2697@gmail.com

## Abstract

*Domain generalization (DG) aims to learn a model that generalizes well to an unseen domain (target domain), which has a different distribution than known domains (source domain). Many of the previous works try to learn domain-invariant features. These methods adopt a loss that tries to match the whole distributions of the source domains. However, these works are sub-optimal because they rarely utilize task-specific information, such as class labels. For concerning the information, we propose a simple but effective regularizing method. The proposed method attempts to match the feature distributions of the same classes. By doing this, a model is expected to learn more task-specific and robust features than the previous works.*

## 1. Introduction

Deep learning has been remarkably successful in many areas []. However, many studies find out deep learning methods are hard to generalize when they encounter an unseen domain, which has a different data distribution than domains used for training []. This problem called *domain shift* []. For alleviating the domain shift problem, many studies have been carried out on different assumptions. Domain Adaptation (DA) assumes there are two domains []. The first is a fully-labeled source domain, and the other is a sparsely labeled or totally unlabeled target domain. Otherwise, domain generalization (DG) assumes there are some fully-labeled source domains, but the target domain is totally unavailable. DG is a challenge but important research area because generalization to other domains is crucial to make a safe artificial intelligence. Moreover, it is helpful to understand how deep neural networks see the world.

Existing DG studies can be classified into several categories depending on their strategies. Some methods proposed novel model architectures that are robust to domain shift []. Others suggested learning algorithms aim to induce a model to fit in a more robust minimum []. The others adopted losses to learn domain-invariant features by matching feature distributions of the source domains []. Although

these domain-invariant feature learning methods work well, the methods are sub-optimal because they do not utilize task-specific information such as class labels explicitly.

Our approach also aims to learn domain-invariant features, but the proposed method explicitly adopts the task-specific information. More specifically, we add a simple consistency loss between features that share the same label but have different domains. By doing this, the model is expected to learn features that are adequately semantic and constant across domains.

To demonstrate the proposed method works, we are planning to experiments on the standard DG benchmarks such as PACS [] or VLCS []. Specifically, we first show the consistency loss is helpful to improve the performance of the baseline, which simply aggregates all the source domains and uses it as a training set. Then, we will show that the proposed method is comparable to or better than many previous works.

## 2. Related Works

### 2.1. Multi-Source Domain Adaptation

### 2.2. Domain Adaptation and Generalization

## References

- [1] FirstName Alpher. Frobnication. *Journal of Foo*, 12(1):234–778, 2002.
- [2] FirstName Alpher and FirstName Fotheringham-Smythe. Frobnication revisited. *Journal of Foo*, 13(1):234–778, 2003.
- [3] FirstName Alpher, FirstName Fotheringham-Smythe, and FirstName Gamow. Can a machine frobnicate? *Journal of Foo*, 14(1):234–778, 2004.
- [4] Authors. The frobnicable foo filter, 2014. Face and Gesture submission ID 324. Supplied as additional material `fg324.pdf`.
- [5] Authors. Frobnication tutorial, 2014. Supplied as additional material `tr.pdf`.