Possible Ways to Solve Insufficiency of Data: A Survey

Fei Jie
hfut_jf@aliyun.com

March 17, 2019

Abstract

It is common knowledge that the more data an ML algorithm has access to, the more effective it can be. Rather than starting with an extremely large corpus of unstructured and unlabeled data, we can instead take a small, curated corpus of structured data and augment in a way that increases the performance of models trained on it. This article explores possible ways to solve the insufficiency of dataset and compare different solutions. We try to find the best way to improve the performance of models for structure prediction.

1 Introduction

We first introduce some terms related to our target, which includes:

Data Augmentation [3] is routinely used in classification problem. Often it is non-trivial to encode known invariances in a model. It can be easier to encode those invariances in the data instead by generating additional data items through transformations from existing data items. For example, the labels of handwritten characters should be invariant to small shifts in location, small rotations or shears, changes in intensity, changes in stroke thickness, changes in size etc. Almost all cases of data augmentation are from a prior known invariance.

Transfer Learning is a research problem in machine learning that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem. Research on transfer learning has attracted more and more attention since 1995 in different names: **learning to learn**, life-long learning, knowledge transfer, inductive transfer, multitask learning, knowledge consolidation, context-sensitive learning, knowledge-based inductive bias, **metalearning**, and incremental/cumulative learning [4].

Few-shot Learning is tasks that study the ability to learn from few examples [1]. When the classes covered by training instances and the classes we aim to classify are disjoint, this paradigm are called **zero-shot learning** [5].

Generative Adversarial Networks (GAN) [2], and specifically Deep Convolutional GANs (DCGAN) use of the ability to discriminate between true and generated examples as an objective. GAN approaches can learn complex joint densities. Recent improvements in the optimization process have reduced some of the failure modes of the GAN learning process.

2 Data Augmentation

The problem with small datasets is that models trained with them do not generalize well data from the validation and test set. Hence, these models suffer from the problem of overfitting. To reduce overfitting, we can add regularization, dropout, or batch normalization in the model or training process. Data augmentation is another way we can reduce overfitting on models, where we increase the amount of training data using information only in our training data.

A very generic and accepted current practice for augmenting image data is to perform geometric and color augmentations, such as reflecting the image, cropping and translating the image, and changing the color palette of the image. All of the transformation are affine transformation of the original image that take the form:

$$y = Wx + b$$

3 Transfer Learning

Definition 1 (Transfer Learning)**.** *Given a source domain $\mathcal{D}_S$ and learning task $\mathcal{T}_S$, a target domain $\mathcal{D}_T$ and learning task $\mathcal{T}_T$, transfer learning aims to help improve the learning of target predictive function $f_T(\cdot)$ in $\mathcal{D}_T$ using the knowledge in $\mathcal{D}_S$ and $\mathcal{T}_S$, where $\mathcal{D}_S \neq \mathcal{D}_T$, or $\mathcal{T}_S \neq \mathcal{T}_T$*

Based on the definition of transfer learning, we can categorize transfer learning under three subsettings, based on different situations between the source and target domains and tasks [4]

- Inductive transfer learning. In the inductive transfer learning setting, the target task is different from the source task, no matter when the source and target domains are the same or not.
- Transductive transfer learning. In the transductive transfer learning setting, the source and target tasks are the same, while the source and target domains are different.
- Unsupervised transfer learning. In the unsupervised transfer learning setting, the target task is different from but related to the source task. There are no labeled data available in both source and target domains in training.

Specifically, for neural networks, transfer learning is a technique in which we take pre-trained weights of neural net trained on some similar or more comprehensive data and fine tune certain parameters to best solve a more specific problem.

4 Few-shot Learning

5 Generative Adversarial Networks

6 memo

DATA AUGMENTATION GENERATIVE ADVERSARIAL NETWORKS

traditional transformations not affect the class, not only for low-data cases, train a GAN in source domain, apply it in the low-data domain / target domain, DAGAN does not depend on the classes themselves, data augmentation from a single novel data point?

dataset, source domain, validation domain, target domain (test domain)

Matching Networks for One Shot Learning

one-shot learning, learning a class from a single labeled example.

2017, DATA AUGMENTATION GENERATIVE ADVERSARIAL NETWORKS, 65

2016, Matching Networks for One Shot Learning, 526

2018, META-LEARNING FOR SEMI-SUPERVISED FEW-SHOT CLASSIFICATION, 39

2017, Few-Shot Learning Through an Information Retrieval Lens, 44

2018, Few-shot learning with graph neural networks, 49

2016, OPTIMIZATION AS A MODEL FOR FEW-SHOT LEARNING, 279

2017, The Effectiveness of Data Augmentation in Image Classification using Deep Learning, 111

2017, Learning to Compare Relation Network for Few-Shot Learning, 75

2017, Zero-Shot Learning - A Comprehensive Evaluation of the Good, the Bad and the Ugly, 92

2017, Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks, 451

2016, Meta-Learning with Memory-Augmented Neural Networks, 228

2017, Prototypical Networks for Few-shot Learning, 286

2017, One-Shot Imitation Learning, 141

2016, One-shot learning with memory-augmented neural networks, 153

2015, Siamese neural networks for one-shot image recognition, 433

References

[1] Garcia, V., and Bruna, J. Few-shot learning with graph neural networks. *arXiv preprint arXiv:1711.04043* (2017).

[2] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. In *Advances in neural information processing systems* (2014), pp. 2672–2680.

[3] Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (2012), pp. 1097–1105.

[4] Pan, S. J., and Yang, Q. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering 22*, 10 (2010), 1345–1359.

[5] Wang, W., Zheng, V. W., Yu, H., and Miao, C. A survey of zero-shot learning: Settings, methods, and applications. *ACM Trans. Intell. Syst. Technol. 10*, 2 (Jan. 2019), 13:1–13:37.