TDLPAAN: Target Data-Label Pair Agnostic Adversarial Network for Domain Adaptation

Anonymous CVPR submission

Paper ID ****

Abstract

The abstract goes here.

1. Introduction

Here, the necessary points are documented in brief just to give an overview of how we got the motivation for this work. Further elaborations and systematic explanations will follow up in subsequent iterations. Section 1 will have discussions over the following points, presented in order of their arrangement:

- Annotation cost is high for constructing datasets for supervised image classification.
- Need to come up with strategies that circumvent/reduce the annotation process.
- This can be done by leveraging off-the-shelf labeled data from a different but related source domain. However, this transfer learning paradigm suffers from the shift in data distributions across different domain poses a major obstacle in adapting classification models to target tasks.
- A solution to this is using existing transfer learning methods to bridge different domains by learning domain-invariant feature representations without using target labels, and the classifier learned from source domain can be directly applied to target domain.
- These techniques can be used to develop an image classifier on target domain, but during training, they require target images corresponding to the labels from the label set of the classifier. To tackle unknown data from the domain, models assign images of unknown types to an extra unknown class which it has not encountered during training.
- We might encounter situations where obtaining data points corresponding to specific labels is a cumbersome task. A low representation of a specific class is not conducive to capturing the underlying distribution governing the target domain.
- this motivates us to present a new domain adaptation problem, where the label set of the target data is unknown and we have to build a classifier that has a label set, which is a subset with the sources'.
-

2. Related Work

Place filler

3. Domain Adaptation

Imagine a standard classification problem of mapping a datapoint $x \in \mathbb{R}^D$ to its label y from the label set C_T, where x is an instance of domain T and is drawn from an unknown distribution p_T. The task becomes conceivable and, correspondingly, a supervised classification framework can be learnt if we have access to a labelled dataset $D_T = \{(x_i^T, y_i^T)\}_{i=1}^{n_T}$, with $x_i^T \in \mathbb{R}^D$ drawn from p_T and $y_i \in C_T$. However, the problem turns out to be more challenging if the learning task is thwarted by making D_T unavailable during training, i.e., the only information present during designing the classification model is the label set C_T. In such a scenario, relevant information can be acquired by looking into:

- a labelled dataset D_S comprising of samples from a different domain S ($S \neq T$), where $D_S = \{(x_j^S, y_j^S)\}_{j=1}^{n_S}$, $x_j^S \in \mathbb{R}^D$) is drawn from a distribution p_S ($p_S \neq p_T$), $y_j \in C_S$ and $C_T \subseteq C_S$.
- an unlabelled dataset $D_T^x = \{x_k^T\}_{k=1}^{n'_T}$, with x_k ($x_j \in \mathbb{R}^D$) drawn from p_T with no prior information about the label types present in the dataset.

While D_S can be utilized to extract features from the classification point of view, D_T^x forms a repository of relevant information from the domain discriminatory standpoint.

3.1. Neural Network

4. Experiments

4.1. Setup

4.2. Results

4.3. Analysis

5. Conclusion

6. Acknowledgements

References