1 Summary

Authors take inspiration from few shot learning and tackle the joint problem of adaptation across another domain which has new tasks and very little labeled data. The paper proposes a novel multi-layered unsupervised adversarial domain adaptation network and domain dependent class similarity objective which is defined even when the domains have non-overlapping classes. SOTA classification accuracy for SVHN to MNIST on non-overlapping classes and action recognition accuracy for ImageNet to UCF-101 videos show the effectiveness of this approach.

2 Strengths

- (i) The paper tackles DA by proposing a method that directly aligns the joint feature space distributions which works even when source and target labels are non-overlapping. (ii) The use of a single multi-layered domain discriminator adversary instead of multiple discriminators for each layer is clever. This results in improved target classification performance and more stable adversarial learning. (iii) Introduces semantic transfer loss which makes use of labeled data for class specific domain adaptation.

3 Weaknesses

- (i) For training, the authors first learn a source feature extractor and classification network using labeled source data and freeze it before doing DA techniques. The standard technique is to let this network train concurrently with DA framework. The paper does not explicitly mention why they chose to do this. (ii) The choice of temperature coefficient τ is not clearly mentioned. Does it make sense to choose a larger value of τ when there is a large domain shift from source to target ? If so, how sensitive is the model to this parameter. (iii) Semantic alignment will not be possible in purely unsupervised domain adaptation setting using this method.

4 Critique of Experiments

- (i) Matching networks perform poorly when adapting from SVHN to MNIST for no class overlap because of domain shift. This method beats other SOTA by a huge margin, and the performance gain is most significant when there very little labeled target data. This shows the model is effectively doing few-shot learning. (ii) Using Multi-layered domain adversarial training increase the model performance by more than 10% compared to the usual domain adversarial loss. This shows the effectiveness of proposed multi-layered architecture.
- Problem of domain shift in action classification from images (ImageNet) to videos (UCF-101) is tackled. Per frame performance is decreased compared to fine-tuned model, however, average performance across frames increases. This is because model is making more confident predictions among key frames, even when per frame prediction is similar.
- Ablation studies for unsupervised domain adaptation setting clearly reveal the dominance (5% improvement) of multi-layered domain adversarial network/loss over standard domain adversarial methods like ADDA/Domain Confusion.

5 Follow Ups/Extensions

- Semantic transfer for unsupervised DA setting can be tried using pseudo-label prediction on unlabeled target data using a classifier and defining semantic loss on class centroids. Expecting that classifier accuracy increases as domain alignment is achieved using multi-layered adversarial loss.
- The paper proposes to pre-train source classifier using labeled source data and freeze the network. It would be interesting to see if there are any convergence issues if this is not the case. Will the model perform sub-optimally.