

Article

Adversarial Hard Attention Adaptation

Hui Tao ¹, Jun He ^{1,*} , Quanjie Cao ¹ and Lei Zhang ²

¹ School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing 210044, China; taohui_95@163.com (H.T.); caodada1118@163.com (Q.C.)

² School of Electrical and Automation Engineering, Nanjing Normal University, Nanjing 210023, China; leizhang@njnu.edu.cn

* Correspondence: jhe@nuist.edu.cn; Tel.: +86-25-58235967

Received: 18 March 2020; Accepted: 16 April 2020; Published: 18 April 2020



Abstract: Domain adaptation is critical to transfer the invaluable source domain knowledge to the target domain. In this paper, for a particular visual attention model, saying hard attention, we consider to adapt the learned hard attention to the unlabeled target domain. To tackle this kind of hard attention adaptation, a novel adversarial reward strategy is proposed to train the policy of the target domain agent. In this adversarial training framework, the target domain agent competes with the discriminator which takes the attention features generated from the both domain agents as input and tries its best to distinguish them, and thus the target domain policy is learned to align the local attention feature to its source domain counterpart. We evaluated our model on the benchmarks of the cross-domain tasks, such as the centered digits datasets and the enlarged non-centered digits datasets. The experimental results show that our model outperforms the ADDA and other existing methods.

Keywords: domain adaptation; attention learning; adversarial training; transfer learning

1. Introduction

In recent years, deep convolutional neural networks have achieved state-of-the-art performance in many visual tasks, such as image classification, object detection and semantic segmentation [1,2]. However, in many practical cases, there exists problems of distribution mismatch [3] and domain shift [4] between different visual tasks, which results in poor generalization performance to the new task. To some extent, transfer learning can solve this challenge by fine-tuning the model in the target domain. However, such a transfer learning method is hindered when only few or even no labels are available in the target domain.

Unsupervised domain adaptation has a wide range of applications, such as speech recognition, natural language processing and image vision. For example, Ref. [5] proposed a Bayesian adaptation learning framework to obtain the maximum posterior (MAP) estimation of hidden activation function parameters in deep models based automatic speech recognition (ASR) system, which solved the unsupervised adaptation work in the field of speech recognition. Moreover, in image visual task, several methods in unsupervised domain adaptation have been proposed to reduce the harmful effects of domain shift and distribution mismatch. In particular, since the emergence of the seminal work of generative adversarial networks (GAN) [6], adversarial adaptation methods have been attracting great attention [7–11], which explore the performance advantages of pitting two networks against each other to reduce the distribution difference between source and target domain. Specifically, the adversarial discriminative domain adaptation (ADDA) approach [12] trains a feature encoder S and a classifier C using the label information in the source domain at the first stage; then a critic network D and another target feature encoder T are trained by competing to each other by adversarial training technique; then finally the classifier C can correctly classify target images by mapping the target domain features to the common space of the source domain.

Previous adversarial adaptation works mainly aim to match the global features extracted from the entire images across domains. However, few works regarding attention adaptation have been well studied, even though attention methods have been dominating state of the arts in vision tasks [13–21]. It is worth noting that for this interesting attention adaptation challenge, Ref. [22] proposed two types of transferable attention: local and global attention by considering the transferability of different regions of images. In this paper, we focus on the transferability of hard attention inspired by [23,24] which also has wide applications such as long document classification [25] and weakly labeled sensor data recognition [26], etc. Notably, a hard attention model usually consists of a recurrent structure to select the most discriminative local features from image patches, and thus hard attention is non-differentiable due to sampling and cropping operations, which poses a challenge to adapt such hard attention from the source domain to the target domain.

To tackle the challenge of hard attention adaptation, in this paper, we propose a novel adversarial hard attention adaptation framework shown in Figure 1. Leveraging the popular ADDA [12] framework, our work mainly consists of three components, saying that a source agent \mathcal{S} , a target agent \mathcal{T} , and a discriminator D . Different from the classical ADDA framework, to overcome the non-differentiable nature of hard attention, the most important contribution of this work is that we design an adversarial reward strategy to train the target agent \mathcal{T} via reinforcement learning technique. Furthermore, the target agent \mathcal{T} and the discriminator D naturally constructs the adversarial training counterparts. On one hand, the target domain policy $\pi_{\mathcal{T}}$ is trained to extract the target attention feature $\mathcal{T}(x_t)$ closing enough to the source model; and on the other hand, the discriminator D is trained to distinguish the target attention feature $\mathcal{T}(x_t)$ from the source attention feature $\mathcal{S}(x_s)$.

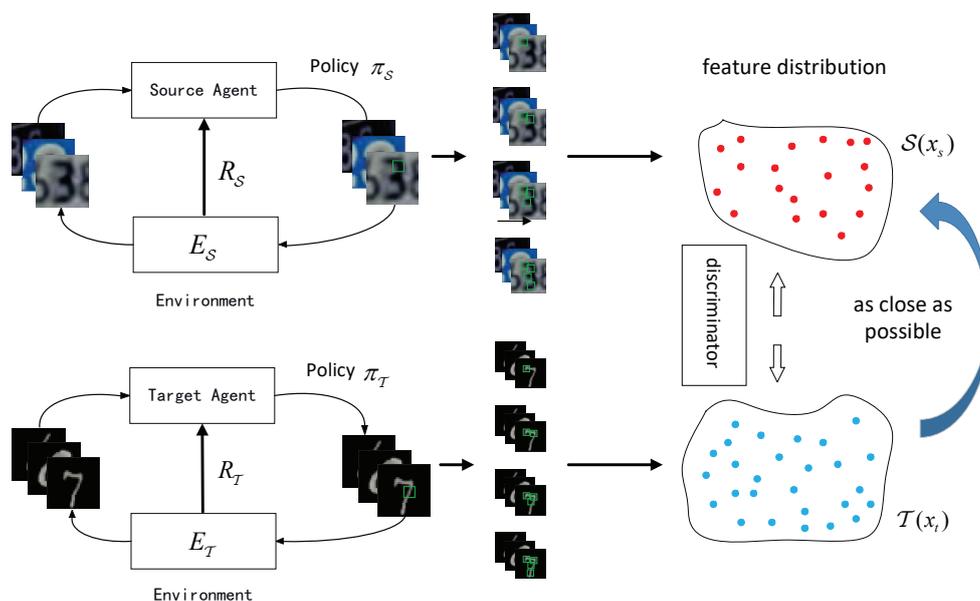


Figure 1. Illustration of the adversarial hard attention adaptation framework. Source images and target images are represented as the source environment E_S and target environment E_T respectively in this reinforcement learning framework. Policy π_S is learned based on the classical hard attention method in the source domain. Policy π_T is trained via our proposed adversarial reward strategy to extract the target attention feature $\mathcal{T}(x_t)$ closing enough to the source model; on the other hand, the discriminator D is trained to distinguish the target attention feature from the source attention feature.

The rest of the paper is organized as follows: Section 2 reviews the related works of domain adaptation and attention methods. In Section 3, we propose the adversarial hard attention adaptation framework and describe how to train it in detail. In Section 4, we experiment the proposed framework on

the standard cross-domain datasets and further evaluate its capability on a cross-domain non-centered digits datasets. Section 5 concludes the paper and discusses our future work.

2. Related Work

2.1. Domain Adaptation

Recent unsupervised domain adaptation can be divided into two main categories: instance-based adaptation and feature representation adaptation. For instance-based adaptation, these methods directly compute the re-sampling weights by matching distributions between source and target domains in a non-parametric way [27,28]. For the latter, many researches attempt to reduce the distribution discrepancy between source domain and target domain by projecting two domains into a common feature space [29–32]. For example, deep domain confusion method (DoC) [31] uses the Maximum Mean Discrepancy (MMD) [33] loss to learn representation from labeled source domain to unlabeled target domain.

Recently, enlightened by generative adversarial networks (GAN) [6], approaches utilizing adversarial objectives were proposed to align the distribution across domains [7–12]. Specially, Ref. [12] introduced a novel adversarial domain adaptation framework in which the discriminator D and the target domain feature generator T are used to compete with each other, so that T can generate the feature distribution close to the source counterpart. Inspired by CycleGAN [34,35] proposed a discriminatively-trained cycle-consistent adversarial domain adaptation model that focuses on the representation adaptation at both pixel-level and feature-level. Unlike the usual GAN-style domain adaptation approaches, Ref. [11] utilizes a very different adversarial training technique for unsupervised domain adaptation, in which the classifier with Dropout regularization is used to detect the classification boundary, and the target generator is encouraged to generate discriminative features far away from the boundary.

2.2. Attention Mechanism

In recent years, attention mechanism plays an important role in many fields such as machine translation [13], speech recognition [36], and image caption [15]. Like human visual system, attention does not need to focus on the whole image, but only on the key areas of the image. For a large image, focusing on the salient areas of the image makes it possible to process fewer pixels and save more computing resources. Attention models can be divided into two categories: soft (deterministic) and hard (stochastic).

For the former, soft attention computes the weight vector as the attention probability, which is differentiable and can be embedded in the vanilla model for end to end training. Soft attention is used in a variety of visual tasks, such as image caption [16–18] and visual question answering task (VQA) [19–21].

For the latter, hard attention is non-differential due to the sampling and cropping operations, and thus it is usually optimized by reinforcement learning technique [37]. Recurrent attention model (RAM) [23] and its extension deep recurrent attention model (DRAM) [24] are the representatives for hard attention, all of which consist of a controller module controlling where to glimpse the next image patch. Comparing to the usual convolutional neural network models for visual tasks, these approaches significantly reduce the computation complexity for large images. Furthermore, Ref. [14] proposed a hard attention based image caption method, which was the state of the arts model for COCO image caption task in 2015. Besides successes in the visual tasks, hard attention models have also been introduced into natural language processing [25] and weakly human activity recognition [26].

Few works regarding attention adaptation have been well studied though, it is worth noting that [22] presents a transferable attention for domain adaptation (TADA) which focuses on two complementary transferable local and global attention. On one hand, the local attention generated by the region-level domain discriminators is to highlight the transferable regions, and on the other

hand the global attention generated by a single image-level domain discriminator is to highlight the transferable images.

3. Model

For unsupervised domain adaptation, it is supposed that a source image x_s and its corresponding label y_s can be accessed from the rich labeled source domain $\{X_s, Y_s\}$; and on the other hand, from the target domain $\{X_t\}$, a target image x_t can only be accessed without any label information. In our adversarial hard attention adaptation work, there are two reinforcement learning agents \mathcal{S} and \mathcal{T} which correspond to the source domain agent and the target domain agent respectively; and $\mathcal{S}(x_s)$ represents the extracted attention feature by \mathcal{S} and $\mathcal{T}(x_t)$ represents the extracted attention feature by \mathcal{T} . The policies of the two agents are denoted as $\pi_{\mathcal{S}}$ and $\pi_{\mathcal{T}}$.

3.1. Problem Formulation

For classical hard attention models, for example the recurrent attention model (RAM) [23], the core idea is to train a reinforcement learning agent interacting with a visual environment to learn the policy $\pi_{\mathcal{S}}$ in the source domain. Such learned policy $\pi_{\mathcal{S}}$ can be safely treated as the target domain policy $\pi_{\mathcal{T}}$, because there is not significant discrepancy between the source domain and the target domain. However, in this work, we break this assumption, that is to say there is a large discrepancy between the source domain and the target domain. For example, in one of following domain adaptation tasks, we consider to adapt the hard attention policy $\pi_{\mathcal{S}}$ trained on the SVHN datasets to its target domain attention policy $\pi_{\mathcal{T}}$ which should be suitable for the very different MNIST dataset.

Unlike ADDA [12] in which the feature encoder S only utilizes the conventional CNN architecture, the hard attention model usually consists of a recurrent structure to select the most discriminative local features from image patches according to its learned policy $\pi_{\mathcal{S}}$. It is well-known that hard attention is non-differentiable due to such sampling and cropping operations, which poses a challenge to directly leverage the ADDA framework to hard attention adaptation. For hard attention adaptation, not only should the local feature extractor be adapted to the source domain, but most importantly the target domain attention policy $\pi_{\mathcal{T}}$ should be aligned to the source domain attention policy $\pi_{\mathcal{S}}$. Then the critical problem for hard attention adaptation is how to train the target domain attention policy $\pi_{\mathcal{T}}$ without any label information. In this work, we propose to train the target domain policy $\pi_{\mathcal{T}}$ via a novel adversarial reward strategy in which $\pi_{\mathcal{T}}$ is encouraged by positive reward if the collected hard attention feature is close enough to the source domain attention feature. Since our work leverages the seminal ADDA framework but introduces such a novel adversarial reward strategy, this work is referred as adversarial hard attention adaptation.

3.2. Adversarial Hard Attention Adaptation

In this adversarial hard attention adaptation work, we mainly focus on the RAM model [23] which consists of four subnetworks as shown in Figure 2: the Glimpse network, the Localization network, the Context network, and the Controller network. To apply RAM to domain adaptation, following ADDA [12], on one hand a discriminator D is used to distinguish by which agent the attention feature is extracted from, the source or the target. On the other hand, the target agent \mathcal{T} tries to generate the “real” source domain attention feature to fool the discriminator D despite the images are from the target domain. Then naturally, the discriminator D and the target agent \mathcal{T} construct the adversarial training counterparts. For the discriminator, it receives the attention feature from both domains and then predict where it is from. Then in this sense, the discriminator D plays a role of a binary classifier without providing any labeled target samples. Then, naturally, the discriminator D can be trained via the standard discriminator loss function L_{advD} as ADDA [12], shown in Equation (1).

$$L_{advD} = -E_{x_s \sim X_s} [\log D(\mathcal{S}(x_s))] - E_{x_t \sim X_t} [\log(1 - D(\mathcal{T}(x_t)))] \quad (1)$$

For the target agent \mathcal{T} , since our adaptation work follows ADDA [12], \mathcal{T} is cloned from the source agent \mathcal{S} , and thus the network structure of \mathcal{T} is the same as \mathcal{S} shown in Figure 2. Because the target images do not have any labels, \mathcal{T} can not be trained in the same way as \mathcal{S} . Moreover, as the non-differentiable characteristic of the Localization network of \mathcal{T} due to such sampling and cropping operations, the target agent \mathcal{T} should also be trained via reinforcement learning technique. Here a novel adversarial reward strategy is proposed to train the target agent \mathcal{T} . Specifically, the discriminator D is leveraged as the critic to classify whether $\mathcal{T}(x_t)$ is from the target domain or the source domain. If $\mathcal{T}(x_t)$ is wrongly classified as an attention feature from \mathcal{S} by the discriminator D , a positive reward +1 is given to the target agent \mathcal{T} because \mathcal{T} is encouraged to beat the D in this adversarial training framework. Otherwise, zero reward should be given to \mathcal{T} .

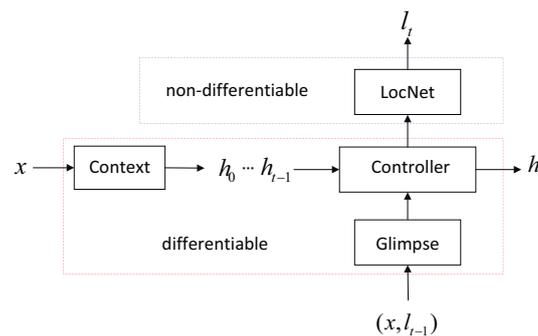


Figure 2. Flow chart of RAM, the hard attention model used in our adaptation work.

At each step, the reward is given after the target agent \mathcal{T} takes a glimpse following its current policy $\pi_{\mathcal{T}}$ and the goal of the agent is to maximize the sum of the reward: $R_{\mathcal{T}} = \sum_{i=1}^T r_i$. Note that to simplify the reward calculation, following the reward strategy of RAM [23], the positive reward +1 can only be given after the T_{th} glimpse is taken, saying that $r_i = 0$ ($i = 1, \dots, T - 1$) and $r_T = 1$ if $\mathcal{T}(x_t)$ is recognized as a source domain attention feature by D . As Figure 2 shows that the target agent \mathcal{T} is composed of not only the non-differentiable part, saying the Localization network, but also the differentiable parts, saying the Context network, the Glimpse network, and the Controller network. Then to adversarial train the target agent \mathcal{T} against the discriminator D , the loss function of \mathcal{T} should also consists of two parts as shown in Equation (2):

$$L_{\mathcal{T}} = -L_{\log ll \mathcal{T}} + L_{adv \mathcal{T}} \tag{2}$$

Here $L_{\log ll \mathcal{T}}$, referring to Equation (3), represents the log likelihood optimized for the non-differentiable part of \mathcal{T} , and $L_{adv \mathcal{T}}$, referring to Equation (4), represents the classical adversarial loss optimized for the differentiable parts of \mathcal{T} .

$$L_{\log ll \mathcal{T}} = \sum_{t=1}^T \log \pi_{\mathcal{T}}(A_i; \theta) * (R_{\mathcal{T}} - baseline) \tag{3}$$

$$L_{adv \mathcal{T}} = -E_{x_t \sim X_t} [\log D(\mathcal{T}(x_t))] \tag{4}$$

Figure 3 shows the complete framework of our proposed adversarial hard attention adaptation. We discuss the training details of this attention adaptation model in the next sub-section. Through this kind of adversarial training, the hard attention can be successfully transferred to the target domain.

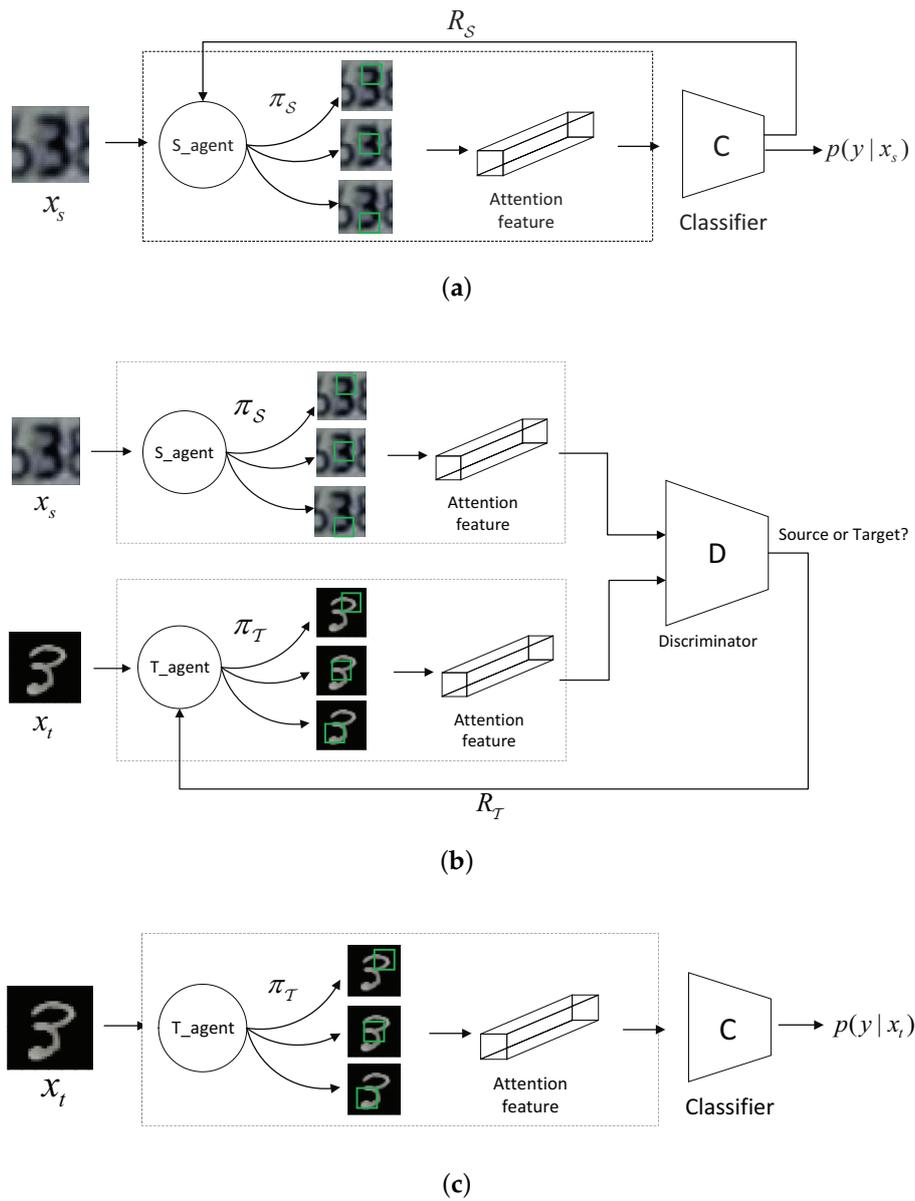


Figure 3. The complete framework of adversarial hard attention adaptation. (a) Pre-training the source agent \mathcal{S} and the Classifier \mathcal{C} in the source domain; (b) Adversarial learning of the target agent \mathcal{T} and the Discriminator \mathcal{D} in the target domain; (c) Testing in the target domain.

3.3. Training Procedure

The complete training procedure of unsupervised adversarial hard attention adaptation consists of the following three stages.

Stage 1: Referring to Figure 3a, we aim to pre-train the source agent \mathcal{S} and the classifier \mathcal{C} in the source domain for K -class classification. Training \mathcal{S} and \mathcal{C} is typically the same as the classical RAM, which is trained by minimizing a hybrid loss function L_S , shown in Equation (5). L_{logllS} , referring to Equation (6), represents the log likelihood loss optimized for the non-differentiable part of \mathcal{S} . L_{xent} , referring to Equation (7), represents the classical cross-entropy loss optimized for the differentiable parts of \mathcal{S} and the classifier.

$$L_S = -L_{logllS} + L_{xent} \tag{5}$$

$$L_{logllS} = \sum_{i=1}^T \log \pi_S(A_i; \theta) * (R_S - baseline) \tag{6}$$

$$L_{xent} = -E_{(x_s, y_s) \sim (X_s, Y_s)} \sum_{k=1}^K \mathbb{I}_{[k=y_s]} \log C(S(x_s))_k \quad (7)$$

Stage 2: This stage performs the adversarial unsupervised domain adaptation, referring to Figure 3b. By alternating between optimizing the discriminator D and the target agent \mathcal{T} , D tries to distinguish where the attention feature comes from, saying that “real” is from the source domain, and \mathcal{T} tries to make its attention feature as “real” as possible. Specifically, D is optimized by minimizing its loss function L_{advD} shown in Equation (1), and \mathcal{T} is optimized by minimizing its loss function $L_{\mathcal{T}}$ shown in Equation (2) respectively.

Stage 3: In the previous two stages, the source agent \mathcal{S} , the source classifier C , and the target agent \mathcal{T} have been well trained accordingly. To evaluate the adapted hard attention extracted by the target agent \mathcal{T} , following ADDA [12], the target attention feature $\mathcal{T}(x_t)$ is simply classified by the source classifier C which is trained in the source domain in stage 1. Figure 3c demonstrates this evaluation stage.

4. Experiments

We evaluated our proposed adversarial hard attention adaptation model on the popular unsupervised domain adaptation datasets. We first focused on the typical domain adaptation tasks between three centered digits datasets, SVHN [38], MNIST [39] and USPS. To further evaluate the capability of hard attention adaptation, we then explored a challenging adaptation task from the enlarged non-centered SVHN datasets to the enlarged non-centered MNIST datasets, which can be regarded as weak datasets. For both types of experiments, ADDA [12] was used as the baseline adaptation model.

4.1. Experiments Setup

We describe the specific network configuration and hyper-parameters of our model as following.

Hard Attention: Since the hard attention model considered in our work is RAM [23], the network structure of the source domain and the target domain follows the protocol used in [23]. Specifically, the Glimpse network has a two-layer fully connected network which has 128,256 hidden units along with a RELU activation function. The Localization network consists of a fully connected hidden layer that emits the location tuple l_t . The recurrent structure of the Controller is LSTM [40], and the size of the LSTM cell is set to 256. Different from the original RAM structure, a Context network is used to provide the global overview of the image as the initial state for the Controller network. The Context network consists of 3 convolutional layers with pooling layers, the convolution kernel size is 5 for each layer, and the filter numbers are 32, 64, 64 respectively.

Discriminator: The discriminator D in this adversarial training framework is composed of a three-layer fully connected network which has 128 hidden units for the first two layers along with a Leaky-RELU activation function. The last layer predicts the probability of TRUE or FALSE.

Classifier: The classifier C shared by the source agent \mathcal{S} and the target agent \mathcal{T} is composed of a fully connected layer followed by a Softmax layer.

The whole model is trained by the ADAM [41] with batch size of 128. In stage 1, saying that the typical supervised learning stage on the source domain, the learning rate is set 1×10^{-3} as initial and then decays by 0.97 for every epoch. During the adversarial training of stage 2, the learning rate is set 2×10^{-4} and 5×10^{-5} for optimization of D and \mathcal{T} respectively. All our experiments are implemented by Tensorflow 1.4 and are trained on a workstation with NVIDIA Titan X GPU and 32Gb system RAM.

4.2. Adaptation between Centered Digits Datasets: SVHN, MNIST and USPS

We validated the performance of our model among the following adaptation tasks: SVHN to MNIST, MNIST to USPS and USPS to MNIST. All of these settings follow ADDA [12]. Specifically, for adaptation between SVHN and MNIST, full datasets are used and the images of MNIST are

resized to the same scale of SVHN, that is 32×32 . For adaptation between MNIST and USPS, 2000 images are sampled from MNIST and 1800 images are sampled from USPS. The images of MNIST are resized to 16×16 as the scale of USPS images. For both the source agent \mathcal{S} and the target agent \mathcal{T} , 6 glimpses are taken for better performance, and the glimpse size is 8×8 for the adaptation between SVHN and MNIST, 4×4 for the adaptation between USPS and MNIST respectively.

The results of these basic adaptation experiments are shown in Table 1. From the table, it can be seen that the proposed adversarial hard attention adaptation approach achieves better performance than the baseline ADDA [12] and also outperforms other existing methods. In particular, for the task of adaptation from SVHN to MNIST, our method surpasses others by a large margin. It is reasonable that our adaptation method outperforms other competitors, since our approach considers both the global feature provided by the Context network and the local attention feature glimpsed by the Glimpse network. Naturally the adaptation of our approach not only projects the global feature of the two domain into a common space as ADDA [12] does, but also aligns the local attention features among the two domains.

Table 1. Experimental results on unsupervised adaptation among 3 centered digits datasets.

Method	SVHN to MNIST	MNIST to USPS	USPS to MNIST
DANN [7]	73.9	77.1 ± 1.8	73.0 ± 2.0
DoC [31]	68.1 ± 0.3	79.1 ± 0.5	66.5 ± 3.3
CoGAN [42]	did not converge	91.2 ± 0.8	89.1 ± 0.8
ADDA [12]	76.0 ± 1.8	89.4 ± 0.2	90.1 ± 0.8
Ours	84.06 ± 1.89	95.10 ± 1.23	90.36 ± 0.86

Figure 4 visualizes the target policy $\pi_{\mathcal{T}}$ adapted from the source policy $\pi_{\mathcal{S}}$. For each sub-figure, the left two rows correspond to the actions (glimpses) taken by the source domain agent \mathcal{S} , and the right two rows correspond to the actions (glimpses) taken by the target domain agent \mathcal{T} . It is obvious that the target policy $\pi_{\mathcal{T}}$ emits the discriminative hard attention as the source policy $\pi_{\mathcal{S}}$ does, which verifies the effectiveness of our model. Furthermore, Figure 5 shows the T-SNE [43] visualization of the original image distribution of both domains (SVHN and MNIST), the feature distribution generated by the source agent \mathcal{S} , and the adapted feature distribution generated by the proposed adversarial hard attention adaptation model. It clearly shows that our model does align the global and local attention features to the same space after hard attention adaptation.

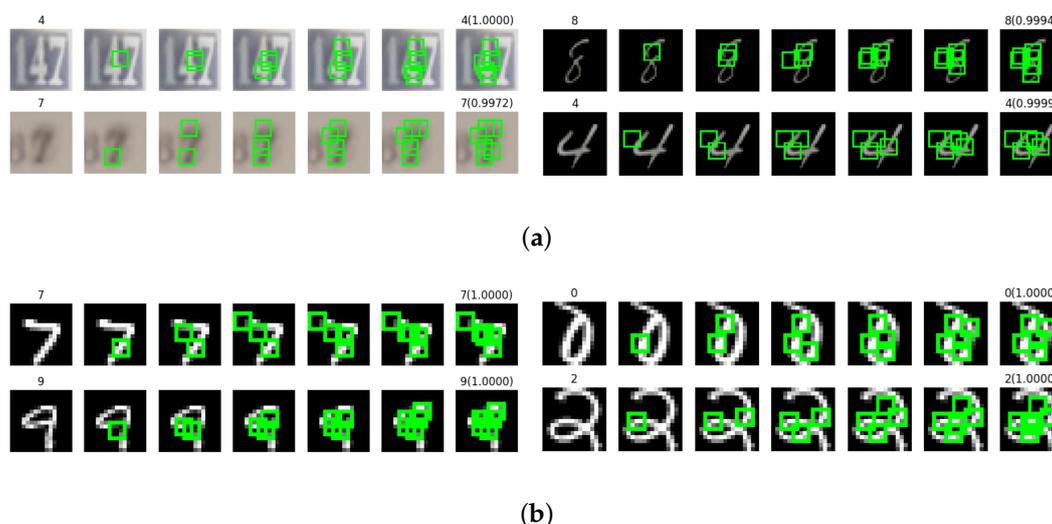


Figure 4. Cont.

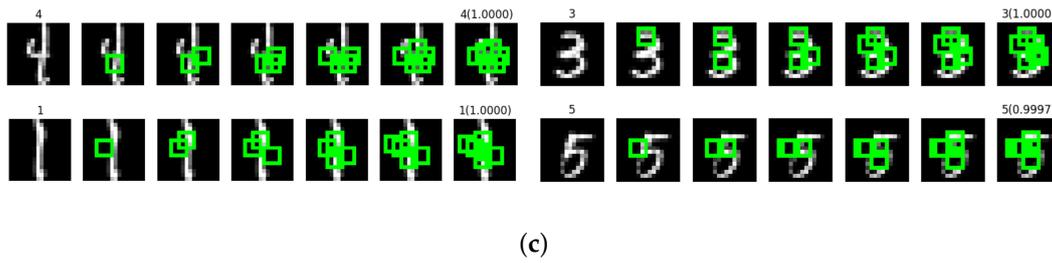


Figure 4. Visualization of the learned target policy $\pi_{\mathcal{T}}$ by adversarial hard attention adaptation. The green boxes represents the locations selected by reinforcement learning agents at each time. For each sub-figure, the left two rows correspond to the actions (glimpses) taken by the source domain agent \mathcal{S} , and the right two rows correspond to the actions (glimpses) taken by the target domain agent \mathcal{T} . For both sides, the first column is the image to be recognized with its true label, and the last column shows the predicted label and its corresponding probability. (a) hard attention adaptation from SVHN to MNIST; (b) hard attention adaptation from MNIST to USPS; (c) hard attention adaptation from USPS to MNIST.

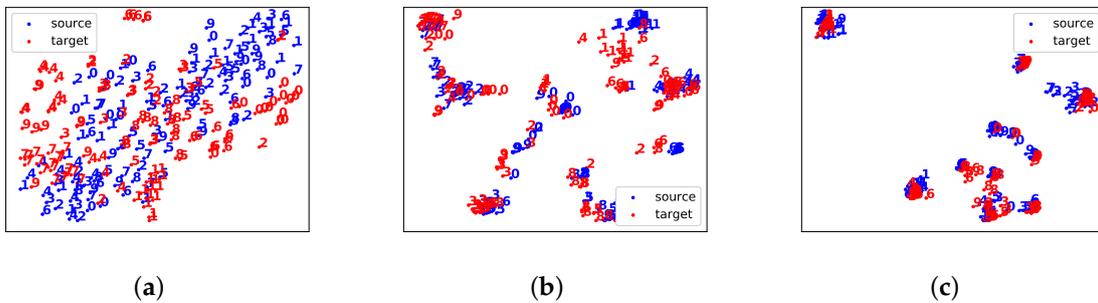


Figure 5. Visualization of images and feature distributions of source and target domains using T-SNE [43], the blue and red digits represent source SVHN dataset and target MNIST dataset respectively. (a) original images distributions of both domains; (b) feature distributions of both domains generated by the source agent \mathcal{S} ; (c) feature distribution of both domains generated the proposed adversarial hard attention adaptation model.

4.3. Adaptation between the Enlarged Non-Centered Datasets: SVHN to MNIST

To further examine the capability of hard attention adaptation, enlightened by the experiments done in RAM [23], we considered a challenging task of adaptation between enlarged non-centered datasets. Specifically, we created the enlarged SVHN dataset by placing a SVHN digit on a random location of a large blank canvas. The enlarged MNIST dataset was constructed in the same way. In this adaptation experiment, we only considered the adaptation from the enlarged SVHN to the enlarged MNIST. The size of the canvas is 60×60 and the size of the original digits is 32×32 in both domains. Figure 6a shows the sample image of the source domain, the enlarged non-centered SVHN dataset; and Figure 6b shows the sample image of the target domain, the enlarged non-centered MNIST dataset.



Figure 6. Examples of non-centered SVHN and MNIST datasets. (a) Random sample of non-centered labeled SVHN datasets; (b) Random sample of non-centered unlabeled MNIST datasets.

In this experiment, since the images are relatively large in both domains, two scales of glimpse are adopted by the Glimpse network, 12×12 and 24×24 respectively. For the Context network, it takes the coarse global overview of the down-sampled 32×32 images. To compare with ADDA for this enlarged weak datasets, we validated two versions of ADDA. One is the down-sampled version ADDA, saying that the images of both domains are all resized to 32×32 which corresponds to the coarse scale of our Context network; and the other is the typical version ADDA, saying that ADDA manipulating the original size of images for both domains.

Table 2 lists the accuracy validated in the target domain for each model. The consumed training and testing time for processing each mini-batch and total training time for each model are also reported in this table. Note that the temporal cost comparative of the different models only includes the time cost of the training process of adversarial unsupervised domain adaptation and the test process, that is to say it does not include the time of training a source model directly on the source domain dataset. The results in the table demonstrate that our model obtains better performance than ADDA. For ADDA with downsampling, because the image is down-sampled to 32×32 , it is obvious that the coarse scale of image does deteriorate the adaptation performance of ADDA even if it only takes the least time to train the model. However, for our model, the Context network requires the coarse scale image to provide the global overview for the agents of both domains, which can help the agents to locate the discriminative hard attention. Since our model needs to deal with coarse global features and local attention features at the same time, it takes more time to train to converge to the best accuracy. Specifically, for typical ADDA, since the input images contain large blank background, aligning features cross-domain extracted from whole images does pay much attention to the useless background. On the contrary, for our model, the target agent \mathcal{T} learns its policy through adversarial training, which can highlight the important areas on the target images despite the large blank background. Figure 7 visualizes the source policy π_S and the target policy π_T on the same MNIST digits: 7, 8, 6, and 4. Figure 7a shows that the source policy π_S can not locate the MNIST digits well, and thus the predicted probability for the true label is low. However, Figure 7b shows that the adapted target policy π_T could nicely locate the MNIST digits and predicts the correct digit label with high probability. Therefore our model is more suitable for cross-domain adaptation on such enlarged non-centered datasets.

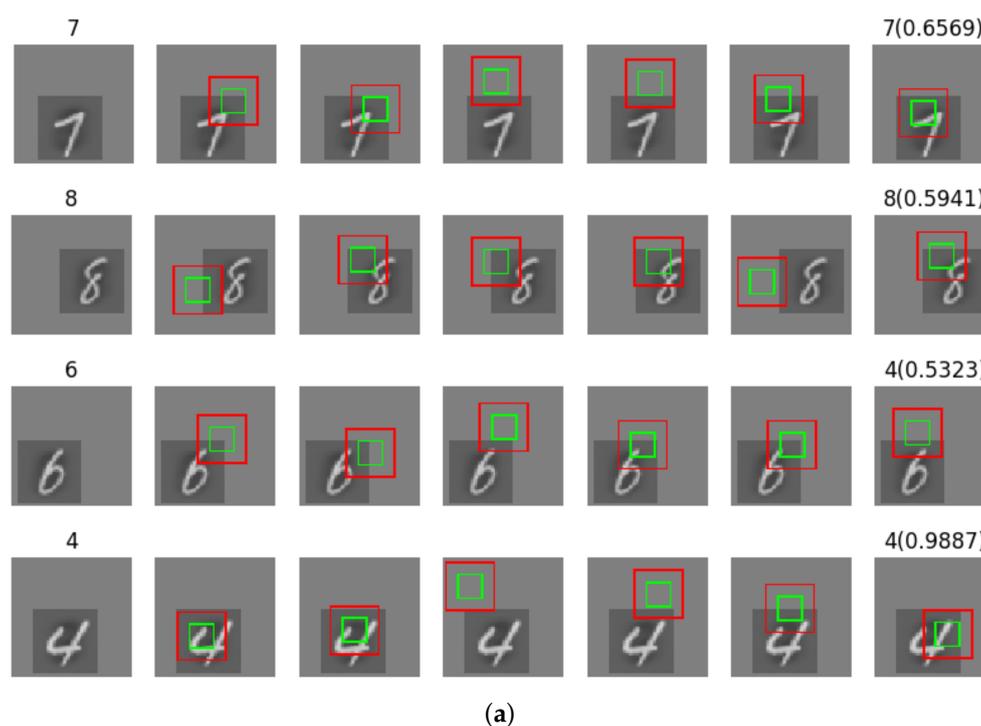


Figure 7. Cont.

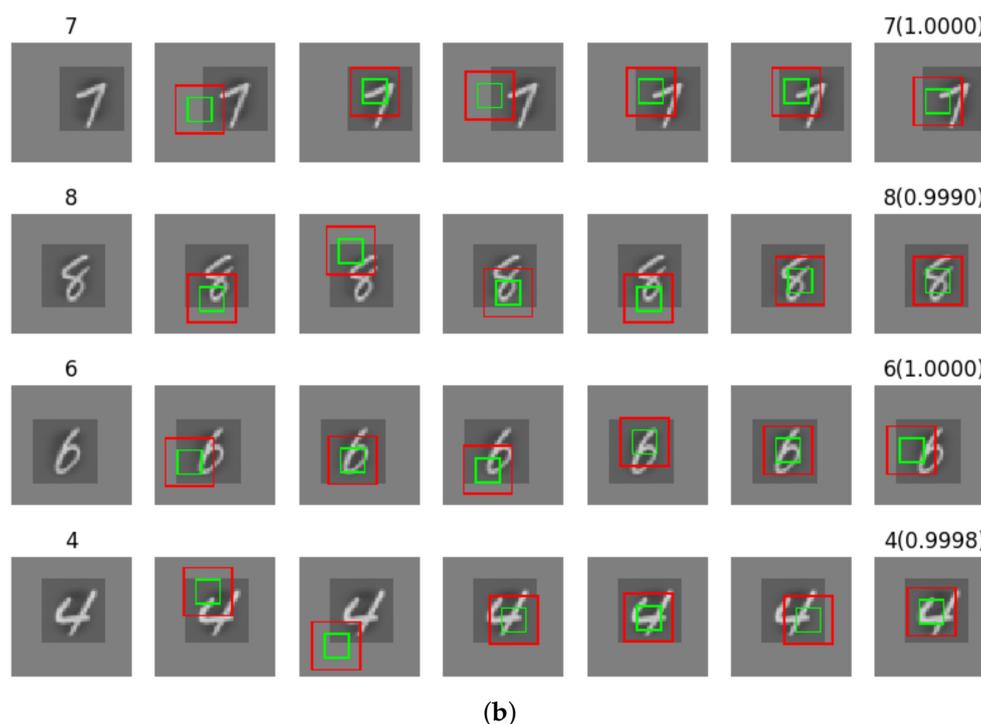


Figure 7. Visualization of locations selected on the non-centered unlabeled target domain datasets by source and target policy respectively. The green box is the first scale of glimpse and the red box represents the second scale. (a) Locations selected on the target datasets by source policy trained on the source domain; (b) Locations selected on the target domain by target domain policy learned after adversarial training.

Table 2. Experimental results on unsupervised adaptation from the enlarged non-centered SVHN datasets to the enlarged non-centered MNIST datasets.

Model	Training Time Per Mini-Batch (S)	Testing Time Per Mini-Batch (S)	Training time in Total (Min)	Accuracy
ADDA with downsampling [12]	0.09	0.06	36.5	68.12 ± 0.59
ADDA without downsampling [12]	0.11	0.08	57.9	74.23 ± 1.60
Ours	0.23	0.18	131.5	78.02 ± 1.43

5. Conclusions

Adapting attention from label-rich source domain to unlabeled target domain would inevitably improve the overall performance of transfer learning. In this paper, we proposed a novel adversarial hard attention adaptation framework to transfer the learned source policy π_S to the target domain policy π_T . By introducing a novel adversarial reward strategy, we construct two adversarial training counterparts, a reinforcement learning agent and a discriminator: the reinforcement learning agent works in the target domain to generate the “real” source domain attention feature to fool the discriminator, and the discriminator tries its best to distinguish where the attention feature is extracted from, the source or the target.

Though the popular recurrent attention model (RAM) has been mainly discussed for our adversarial hard attention adaptation framework, we believe that other kinds of hard attention, saying that attention generated via non-differentiable operations could also benefit from our framework, in particular the proposed adversarial reward strategy. In future work, we plan to explore other interesting hard attention methods for domain adaptation and other challenging visual tasks, such as transferring attention knowledge for fine-grained classification.

Author Contributions: Conceptualization, H.T., J.H. and L.Z.; Methodology, Q.C.; Software, H.T.; Supervision, J.H.; Validation, Q.C.; Visualization, H.T.; Writing—original draft, H.T. and J.H.; Writing—review & editing, L.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China grant number 61601230.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? In Proceedings of the Advances in Neural Information Processing Systems 27, Montreal, QC, Canada, 8–13 December 2014; pp. 3320–3328.
2. Oquab, M.; Bottou, L.; Laptev, I.; Sivic, J. Learning and transferring mid-level image representations using convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 24–27 June 2014; pp. 1717–1724.
3. Gretton, A.; Smola, A.; Huang, J.; Schmittfull, M.; Borgwardt, K.; Schölkopf, B. Covariate shift and local learning by distribution matching. In *Dataset Shift in Machine Learning*; MIT Press: Cambridge, MA, USA, 2009; pp. 131–160.
4. Quionero-Candela, J.; Sugiyama, M.; Schwaighofer, A.; Lawrence, N.D. *Dataset Shift in Machine Learning*; The MIT Press: Cambridge, MA, USA, 2009.
5. Huang, Z.; Siniscalchi, S.M.; Lee, C. Bayesian Unsupervised Batch and Online Speaker Adaptation of Activation Function Parameters in Deep Models for Automatic Speech Recognition. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2017**, *25*, 64–75. [[CrossRef](#)]
6. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the Advances in Neural Information Processing Systems 27, Montreal, QC, Canada, 8–13 December 2014; pp. 2672–2680.
7. Ganin, Y.; Lempitsky, V. Unsupervised Domain Adaptation by Backpropagation. In Proceedings of the 32nd International Conference on International Conference on Machine Learning, ICML'15, Lille, France, 6–11 July 2015; Volume 37, pp. 1180–1189.
8. Ganin, Y.; Ustinova, E.; Ajakan, H.; Germain, P.; Larochelle, H.; Laviolette, F.; Marchand, M.; Lempitsky, V. Domain-adversarial training of neural networks. *J. Mach. Learn. Res.* **2016**, *17*, 2096–2030.
9. Tzeng, E.; Hoffman, J.; Darrell, T.; Saenko, K. Simultaneous deep transfer across domains and tasks. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 4068–4076.
10. Saito, K.; Watanabe, K.; Ushiku, Y.; Harada, T. Maximum classifier discrepancy for unsupervised domain adaptation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 3723–3732.
11. Saito, K.; Ushiku, Y.; Harada, T.; Saenko, K. Adversarial Dropout Regularization. In Proceedings of the International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018.
12. Tzeng, E.; Hoffman, J.; Saenko, K.; Darrell, T. Adversarial discriminative domain adaptation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7167–7176.
13. Bahdanau, D.; Cho, K.; Bengio, Y. Neural Machine Translation by Jointly Learning to Align and Translate. *arXiv* **2014**, arXiv:1409.0473.
14. Xu, K.; Ba, J.; Kiros, R.; Cho, K.; Courville, A.; Salakhudinov, R.; Zemel, R.; Bengio, Y. Show, attend and tell: Neural image caption generation with visual attention. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 2048–2057.
15. Vinyals, O.; Toshev, A.; Bengio, S.; Erhan, D. Show and tell: A neural image caption generator. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3156–3164.

16. You, Q.; Jin, H.; Wang, Z.; Fang, C.; Luo, J. Image captioning with semantic attention. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 4651–4659.
17. Chen, L.; Zhang, H.; Xiao, J.; Nie, L.; Shao, J.; Liu, W.; Chua, T.S. Sca-cnn: Spatial and channel-wise attention in convolutional networks for image captioning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 5659–5667.
18. Anderson, P.; He, X.; Buehler, C.; Teney, D.; Johnson, M.; Gould, S.; Zhang, L. Bottom-up and top-down attention for image captioning and visual question answering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 6077–6086.
19. Xu, H.; Saenko, K. Ask, attend and answer: Exploring question-guided spatial attention for visual question answering. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Cham, Switzerland, 2016; pp. 451–466.
20. Lu, J.; Yang, J.; Batra, D.; Parikh, D. Hierarchical question-image co-attention for visual question answering. In Proceedings of the Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, Barcelona, Spain, 5–10 December 2016; pp. 289–297.
21. Zhu, Y.; Groth, O.; Bernstein, M.; Fei-Fei, L. Visual7w: Grounded question answering in images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 4995–5004.
22. Wang, X.; Li, L.; Ye, W.; Long, M.; Wang, J. Transferable attention for domain adaptation. In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), Honolulu, HI, USA, 27 January–1 February 2019.
23. Mnih, V.; Heess, N.; Graves, A. Recurrent models of visual attention. In Proceedings of the Advances in Neural Information Processing Systems 27, Montreal, QC, Canada, 8–13 December 2014; pp. 2204–2212.
24. Ba, J.; Mnih, V.; Kavukcuoglu, K. Multiple Object Recognition with Visual Attention. In Proceedings of the International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015.
25. He, J.; Wang, L.; Liu, L.; Feng, J.; Wu, H. Long Document Classification From Local Word Glimpses via Recurrent Attention Learning. *IEEE Access* **2019**, *7*, 40707–40718. [[CrossRef](#)]
26. He, J.; Zhang, Q.; Wang, L.; Pei, L. Weakly Supervised Human Activity Recognition From Wearable Sensors by Recurrent Attention Learning. *IEEE Sens. J.* **2018**, *19*, 2287–2297. [[CrossRef](#)]
27. Huang, J.; Gretton, A.; Borgwardt, K.; Schölkopf, B.; Smola, A.J. Correcting sample selection bias by unlabeled data. In Proceedings of the Advances in Neural Information Processing Systems 20, Vancouver, BC, Canada, 3–6 December 2007; pp. 601–608.
28. Yan, H.; Ding, Y.; Li, P.; Wang, Q.; Xu, Y.; Zuo, W. Mind the class weight bias: Weighted maximum mean discrepancy for unsupervised domain adaptation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2272–2281.
29. Sun, B.; Saenko, K. Deep coral: Correlation alignment for deep domain adaptation. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Cham, Switzerland, 2016; pp. 443–450.
30. Sun, B.; Feng, J.; Saenko, K. Return of frustratingly easy domain adaptation. In Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016.
31. Tzeng, E.; Hoffman, J.; Zhang, N.; Saenko, K.; Darrell, T. Deep Domain Confusion: Maximizing for Domain Invariance. *arXiv* **2014**, arXiv:1412.3474.
32. Long, M.; Cao, Y.; Wang, J.; Jordan, M.I. Learning Transferable Features with Deep Adaptation Networks. In Proceedings of the 32nd International Conference on International Conference on Machine Learning, ICML'15, Lille, France, 6–11 July 6 2015; Volume 37, pp. 97–105.
33. Borgwardt, K.M.; Gretton, A.; Rasch, M.J.; Kriegel, H.P.; Schölkopf, B.; Smola, A.J. Integrating structured biological data by kernel maximum mean discrepancy. *Bioinformatics* **2006**, *22*, e49–e57. [[CrossRef](#)] [[PubMed](#)]
34. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.
35. Hoffman, J.; Tzeng, E.; Park, T.; Zhu, J.Y.; Isola, P.; Saenko, K.; Efros, A.; Darrell, T. CyCADA: Cycle-Consistent Adversarial Domain Adaptation. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018.

36. Graves, A.; Mohamed, A.R.; Hinton, G. Speech recognition with deep recurrent neural networks. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 6645–6649.
37. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*; MIT Press: Cambridge, MA, USA, 2018.
38. Netzer, Y.; Tao, W.; Coates, A.; Bissacco, A.; Bo, W.; Ng, A.Y. Reading Digits in Natural Images with Unsupervised Feature Learning. In Proceedings of the NIPS Workshop on Deep Learning and Unsupervised Feature Learning, Sierra Nevada, Spain, 16–17 December 2011.
39. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [[CrossRef](#)]
40. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
41. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. In Proceedings of the International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015.
42. Liu, M.Y.; Tuzel, O. Coupled generative adversarial networks. In Proceedings of the Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, Barcelona, Spain, 5–10 December 2016; pp. 469–477.
43. Maaten, L.V.D.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).