DOMAIN ADAPTATION WITH CYCLEGAN FOR CHANGE DETECTION IN THE AMAZON FOREST

P. J. Soto[1,][*] G. A. O. P. Costa[2], R. Q. Feitosa[1], P. N. Happ[1], M. X. Ortega[1], J. Noa[1], C. A. Almeida[3], C. Heipke[4]

[1] Dept. of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Brazil
(psoto, raul patrick, mortega, jnoat92)@ele.puc-rio.br
[2] Dept. of Informatics and Computer Science, Rio de Janeiro State University (UERJ), Brazil
gilson.costa@ime.uerj.br
[3] Amazon Program Coordination, National Institute for Space Research (INPE), Brazil.
claudio.almeida@inpe.br
[4] Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover (LUH), Germany
heipke@ipi.uni-hannover.de

Commission III, WG III/7

KEY WORDS: Remote Sensing, Change Detection, Domain Adaptation, Cycle-Consistent Generative Adversarial Networks.

ABSTRACT:

Deep learning classification models require large amounts of labeled training data to perform properly, but the production of reference data for most Earth observation applications is a labor intensive, costly process. In that sense, transfer learning is an option to mitigate the demand for labeled data. In many remote sensing applications, however, the accuracy of a deep learning-based classification model trained with a specific dataset drops significantly when it is tested on a different dataset, even after fine-tuning. In general, this behavior can be credited to the domain shift phenomenon. In remote sensing applications, domain shift can be associated with changes in the environmental conditions during the acquisition of new data, variations of objects' appearances, geographical variability and different sensor properties, among other aspects. In recent years, deep learning-based domain adaptation techniques have been used to alleviate the domain shift problem. Recent improvements in domain adaptation technology rely on techniques based on Generative Adversarial Networks (GANs), such as the Cycle-Consistent Generative Adversarial Network (CycleGAN), which adapts images across different domains by learning nonlinear mapping functions between the domains. In this work, we exploit the CycleGAN approach for domain adaptation in a particular change detection application, namely, deforestation detection in the Amazon forest. Experimental results indicate that the proposed approach is capable of alleviating the effects associated with domain shift in the context of the target application.

1. INTRODUCTION

In the past years, global concern about climate change has risen substantially, up to a point where it is now considered the major challenge to be faced by humanity in the coming decades. Having reached an unprecedented scale, its effects directly threaten food production and natural resources, as well as a still unaccounted number of life forms.

Anthropogenic driven environmental degradation is currently believed to be one of the major causes of global warming. In this respect, the extinction of natural forests can be directly linked to climate change. Deforestation is one of the largest sources of greenhouse gas emissions, which in turn contributes to the elevation of Earth's surface temperature. Deforestation is responsible for the reduction of carbon storage and other serious environmental issues such as biodiversity losses (De Sy et al., 2015).

The tropical rainforests, particularly, store up to 140 billion metric tons of carbon, and are known to help stabilize worldwide climate. The Amazon forest alone contains 10% of all biomass on the planet and is home to 10% of the known life species (De Sy et al., 2015) (The Worldwatch Institute, 2015).

Unfortunately, the Amazon biome has faced several threats as a result of unsustainable economic development, linked to extensive cropping and cattle farming, forest fires, illegal mining, and expansion of informal settlements (Goodman et al., 2019), (Malingreau et al., 2012), (Nogueron et al., 2006).

In this context, monitoring environmental changes directly related to global warming, such the ones caused by deforestation in the Amazon, has become a priority for authorities and institutions around the world. Nevertheless, the detection of changes on Earth's surface, specially in a global or regional scale is a complex and costly process, which demands solutions that support efficient analysis of large volumes of remote sensing (RS) data.

In the past decade, artificial intelligence techniques, especially those related with deep learning (DL), have become the dominant trend in image analysis, mostly due to their capacity to learn discriminative features directly from data, when labeled samples are abundant (Bengio et al., 2009) (LeCun et al., 2015) (Krizhevsky et al., 2012).

At the same time, the availability of Earth observation (EO) data produced by RS systems has increased considerably. However, most of the RS applications still fall short in the demands imposed by DL-based techniques, basically because of the high costs required by field survey and labor-intensive visual interpretation to produce a large enough quantity of labeled data.

---

* Corresponding author

The development of wide-reaching DL-based solutions for EO problems, such as automatic change detection, therefore, remains a challenging subject.

In this sense, transfer learning (Weiss et al., 2016) (Pan, Yang, 2010) (Nogueira et al., 2017) emerged as an attractive alternative, allowing the reuse of networks already trained on large data-sets in problems in which a limited quantity of labeled data is available.

Such techniques, however, perform poorly when the domain shift phenomenon is present (Wu et al., 2019) (Zhang et al., 2019). In RS applications, domain shift can be associated with changes in the environmental conditions during the acquisition of new data, variations of objects' appearances, geographical variability and different sensor properties, among other aspects. In many applications domain shift makes it impossible to employ pre-trained classifiers on new data, even after fine-tuning, without a significant decrease in classification accuracy (Schenkel, Middelmann, 2019) (Wittich, Rottensteiner, 2019).

Domain adaptation techniques can be used to alleviate the domain shift problem (Ganin, Lempitsky, 2014) (Sun, Saenko, 2016) (Tzeng et al., 2014). In short, domain adaptation aims at minimizing the discrepancy between distributions of two different domains. One of the distributions characterizes the data used to train a classifier; the other is associated with data that the classifier has never seen, which may present several of the aforementioned variations (Zhang et al., 2019).

Among existing domain adaptation methods, those based on Generative Adversarial Networks (Goodfellow et al., 2014) represent the current state-of-the-art. Recent improvements of this technology (Hoffman et al., 2017) (Murez et al., 2018) rely on the CycleGAN (Zhu et al., 2017) approach to produce indistinguishable features from different domains. This idea has been adapted recently to remote sensing applications, such as urban land cover classification (Schenkel, Middelmann, 2019) (Wittich, Rottensteiner, 2019), cloud detection (Mateo-García et al., 2019), and multiple change detection with very high resolution (VHR) multisensor images from urban areas (Deng et al., 2019).

The present work evaluates a CycleGAN-based domain adaptation technique over RS datasets in a change detection application, namely, deforestation monitoring in the Amazon forest.

The problem this work aims at tackling is as follows. Considering two pairs of RS images from the same area, acquired at two different pairs of epochs, and considering solely reference samples (about the occurrence of deforestation) for the first pair of epochs, how can domain adaptation be used so that deforestation detection can be carried out with reasonable accuracy on the second pair of epochs?

Our results and conclusions focus on the gains brought by the proposed approach in relation to the classification accuracy obtained by a classifier trained with samples from the first pair of epochs, but tested on a second pair of epochs.

The rest of this paper is organized as follows. Section 2 briefly describes the basic techniques associated with the proposed method. A detailed description of the proposed method is the subject of the Section 3. The experimental protocol is reported in Section 4. Section 5 shows the results obtained in the evaluation experiments. Finally, Section 6 presents conclusions and indicates future research directions.

2. FUNDAMENTALS

2.1 Domain Adaptation

Domain adaptation (DA) in the machine learning context comprises methods which aim at improving the performance of models trained with a particular dataset, regarded as the *source domain*, and tested on a different, but related dataset, denoted as the *target domain* (Wang, Deng, 2018).

Among the various DA approaches proposed thus far those based on deep neural networks (DNN) constitute the current state-of-the-art, being adversarial domain adaptation techniques using generative models the most successfully in the past few years.

DNN-based adversarial DA techniques follow two main approaches. The first approach aims at generating synthetic (adapted) images that preserve the underlying structures present in the target domain, but that are somehow similar to the source domain images. In sequence, a classifier trained with the source images and corresponding references is evaluated on the adapted target images (Isola et al., 2017) (Bermudez et al., 2019) (Mateo-García et al., 2019).

The other approach aims at aligning the domains in a common (latent) feature space. The basic idea of this approach is to find representations for the different domains that are domain agnostic, i.e., representations that are associated to features that are indistinguishable with respect to their original domains (Wittich, Rottensteiner, 2019) (Hoffman et al., 2017) (Murez et al., 2018).

In this work we follow the first DA approach, i.e., the one based on the adaptation of the images latter subjected to classification.

Taking into account that the proposed DA technique relies on generative adversarial concepts, the next sections are dedicated to explaining the basis of the related models, as well as their most representative examples.

2.2 Generative Adversarial Networks

GANs (Goodfellow et al., 2014) constitute a class of unsupervised machine learning models composed by two neural networks: the *generator* and the *discriminator*. The generator model specializes in synthesizing realistic images by learning a function $\mathcal{G}$ that maps samples of a known random distribution $p(z)$ into samples of a distribution $p_{model}(x)$. The discriminator, in turn, is trained to learn a function $\mathcal{D}$ that distinguishes whether a sample comes from the real data distribution, $p_{data}(x)$, or from $p_{model}(x)$. Using a min-max procedure to train the related neural networks, the optimal mapping function $\mathcal{G}^*$ can be found by solving the following equation:

$$\mathcal{G}^* = \arg \min_{\mathcal{G}} \max_{\mathcal{D}} \mathcal{L}(\mathcal{G}, \mathcal{D}), \qquad (1)$$

where $\mathcal{L}(\mathcal{G}, \mathcal{D})$ is the GAN loss function defined as:

$$\mathcal{L}(\mathcal{G}, \mathcal{D}) = \mathbb{E}_{x \sim p_{data}(x)}[\log(\mathcal{D}(x))] \\ + \mathbb{E}_{z \sim P_{(z)}}[\log(1 - \mathcal{D}(\mathcal{G}(z)))], \qquad (2)$$

where $\mathbb{E}$ and $\log$ are the expectation and logarithmic operators, respectively, x is a real image, and z is a random noise vector, sampled from a known noise distribution $p(z)$, which is typically uniform or Gaussian.

2.3 Cycle-Consistent Generative Adversarial Networks

CycleGANs (Zhu et al., 2017) aim at capturing the essential characteristics of one image collection $\{x_i\}_{i=1}^N$, where $x_i \in X$, and translating them to another image collection $\{y_j\}_{j=1}^M$, where $y_j \in Y$. To that end, the underlying model learns the mapping functions $\mathcal{G} : X \to Y$ and $\mathcal{F} : Y \to X$, such that images produced by $\mathcal{G}$ and $\mathcal{F}$ are indistinguishable from the real set of images, Y and X respectively. Additionally, the model contains two discriminators. $\mathcal{D}_X$ is trained to discern between real images from X and the one produced with $\mathcal{F}$, while $\mathcal{D}_Y$ is trained to discriminate between real images from Y and translated images produced by $\mathcal{G}$.

The optimal mapping functions $\mathcal{G}^*$ and $\mathcal{F}^*$ can be found through an optimization procedure based on the following equation:

$$\mathcal{G}^*, \mathcal{F}^* = \arg \min_{\mathcal{G},\mathcal{F}} \max_{\mathcal{D}_X,\mathcal{D}_Y} \mathcal{L}(\mathcal{G},\mathcal{F},\mathcal{D}_X,\mathcal{D}_Y), \quad (3)$$

where, $\mathcal{L}(\mathcal{G},\mathcal{F},\mathcal{D}_X,\mathcal{D}_Y)$ is the CycleGAN loss function defined as:

$$\begin{aligned}\mathcal{L}(\mathcal{G},\mathcal{F},\mathcal{D}_X,\mathcal{D}_Y) = &\ \mathcal{L}_{GAN}(\mathcal{G},\mathcal{D}_Y,X,Y) \\ &+ \mathcal{L}_{GAN}(\mathcal{F},\mathcal{D}_X,Y,X) \\ &+ \lambda_c \mathcal{L}_{cyc}(\mathcal{G},\mathcal{F}) + \lambda_d \mathcal{L}_{idt}(\mathcal{G},\mathcal{F}).\end{aligned} \quad (4)$$

The first and second terms of Equation 4 represent *adversarial losses*, which are related to the model's capacity to match the data distribution of the generated images to the data distribution of the corresponding target collections. Both terms are defined in the same way as GAN losses (Equation 2):

$$\begin{aligned}\mathcal{L}_{GAN}(\mathcal{G},\mathcal{D}_Y,X,Y) = &\ \mathbb{E}_{y \sim p_{data}(y)}[\log(\mathcal{D}_Y(y))] \\ &+ \mathbb{E}_{x \sim p_{data}(x)}[\log(1 - \mathcal{D}_Y(\mathcal{G}(x)))],\end{aligned} \quad (5)$$

$$\begin{aligned}\mathcal{L}_{GAN}(\mathcal{F},\mathcal{D}_X,Y,X) = &\ \mathbb{E}_{x \sim p_{data}(x)}[\log(\mathcal{D}_X(x))] \\ &+ \mathbb{E}_{y \sim p_{data}(y)}[\log(1 - \mathcal{D}_X(\mathcal{F}(y)))],\end{aligned} \quad (6)$$

The third term, $\mathcal{L}_{cyc}(\mathcal{G},\mathcal{F})$, represents the *cycle consistency loss*, which serves as a constraint to the many mappings $\mathcal{G}$ and $\mathcal{F}$ that could be induced over $\hat{Y}$ and $\hat{X}$ respectively, whereby λ_c is a regularization coefficient. $\mathcal{L}_{cyc}(\mathcal{G},\mathcal{F})$ is given by:

$$\begin{aligned}\mathcal{L}_{cyc}(\mathcal{G},\mathcal{F}) = &\ \mathbb{E}_{x \sim p_{data}(x)}[||x - \mathcal{F}(\mathcal{G}(x))||_1] \\ &+ \mathbb{E}_{y \sim p_{data}(y)}[||y - \mathcal{G}(\mathcal{F}(y))||_1],\end{aligned} \quad (7)$$

The fourth term, represents the *identity loss*, which encourages the mappings to preserve the characteristics of the input images when they are already similar to those of the respective target distributions. Tuned by λ_d, $\mathcal{L}_{idt}(\mathcal{G},\mathcal{F})$ is a regularization term that forces $\mathcal{G}$ and $\mathcal{F}$ to be close to an identity mapping, such that $\mathcal{G}(Y) \approx Y$ and $\mathcal{F}(X) \approx X$. The identity loss $\mathcal{L}_{idt}(\mathcal{G},\mathcal{F})$ is given by the following equation:

$$\begin{aligned}\mathcal{L}_{idt}(\mathcal{G},\mathcal{F}) = &\ \mathbb{E}_{y \sim p_{data}(y)}[||y - \mathcal{G}(y)||_1] \\ &+ \mathbb{E}_{x \sim p_{data}(x)}[||x - \mathcal{F}(x)||_1],\end{aligned} \quad (8)$$

Whereas CycleGANs present a more complex structure, the training procedure is, in general, similar to that of basic GANs. In each training cycle, the generators are trained to improve their capacity to fool the respective discriminators; while the discriminators are trained to better recognize between real and generated samples.

Due to the CycleGAN's ability of translating image characteristics from one domain to another, recently proposed domain adaptation approaches have benefited from its underlying ideas (Hoffman et al., 2017) (Murez et al., 2018) (Deng et al., 2019). In this work, we employ CycleGANs for domain adaptation in the context of a change detection application, namely, deforestation detection in the Amazon forest using RS optical images.

3. PROPOSED METHOD

The proposed method aims at alleviating domain shift phenomenon in the context of deforestation change detection, by using a DA approach based on CycleGANs. The main idea is to learn a transformation that preserves the structural characteristics of a sequence of two images acquired at a particular pair of epochs, and adapt those images so that they match the acquisition conditions and general landscape aspects at a second pair of epochs.

To this purpose we train a CycleGAN model that learns non-linear mapping functions that takes as input a combination of two images of consecutive years and generates corresponding images, adapted to the conditions of a second set of image combination.

Formally, we consider two sets of co-registered pairs of optical images, i.e., $\{x^{t_0}, x^{t_1}\} \in X$ and $\{y^{t_2}, y^{t_3}\} \in Y$, taken at epochs t_0, t_1, t_2, and t_3 respectively, where $t_0 < t_1 \leq t_2 < t_3$. Additionally, we define X as the source domain, and Y as the target domain. The CycleGAN mapping functions are estimated following the steps described next.

First, a set of patches is extracted from each domain. The patches are extracted from the entire extents of the image combinations. The process is carried out by extracting patches using a sliding window procedure, with a fixed stride size.

Second, employing the set of patches extracted from the image combinations in each domain, the CycleGAN mapping functions $\mathcal{G}(X)$ and $\mathcal{F}(Y)$ are trained until convergence.

Third, once the models have been trained they are used to generate the adapted versions of the original image combinations in each domain. The final result is a mosaic of the generated patches. Similar to (Arkadiusz et al., 2017), we adopt the sliding window approach with overlap for patch generation. This allows building a mosaic free of artifacts by removing weak predictions close to patch boundaries.

It is important to note that each domain represents a stacked pair of images taken at a particular pair of epochs, and it is expected that the models learn to preserve the change transitions from the spectral bands of the image pairs from each domain. Indeed, as the target application has to do with change detection, such transitions constitute the most important structures to be preserved during the adaptation task.

Additionally, we observe that DA is only possible when the domains are not too dissimilar from each other, and that the more

dissimilar the domains, the harder the DA task is. In fact, in our case, preliminary results showed that when the dissimilarity between the source and target domains increases, some structures that are not found in the original images appear in the adapted images, or vice-versa. This is a well known issue named model hallucination (Cohen et al., 2018).

We believe that such behavior may be stimulated by the identity loss term in Equation 4. Although the term encourages the preservation of characteristics from the target domain, we found out that, at least in the context of this work, such term is related to the generation of artifacts.

Based on those findings, we decided to explore alternatives to control the generation of artifacts. Basically, we decided to investigate the effects of introducing a new constrain in the identity loss. The new two terms aim at regularizing the functions $\mathcal{G}$ and $\mathcal{F}$ to be closer to identity mapping by enforcing $\mathcal{G}(X) \approx (X)$ and $\mathcal{F}(Y) \approx (Y)$ in addition to $\mathcal{G}(Y) \approx (Y)$ and $\mathcal{F}(X) \approx (X)$. Therefore, Equation 8 is redefined as follows:

$$
\begin{aligned}
\mathcal{L}_{idt}(\mathcal{G}, \mathcal{F}) = & \lambda_t (\mathbb{E}_{y \sim p_{data}(y)}[||y - \mathcal{G}(y)||_1] \\
& + \mathbb{E}_{x \sim p_{data}(x)}[||x - \mathcal{F}(x)||_1]) \\
& + \lambda_s (\mathbb{E}_{x \sim p_{data}(x)}[||x - \mathcal{G}(x)||_1] \\
& + \mathbb{E}_{y \sim p_{data}(y)}[||y - \mathcal{F}(y)||_1]),
\end{aligned}
\tag{9}
$$

where, λ_t regulates the relative importance of the target domain identity loss, while λ_s regulates the importance of the identity loss considering the source domain.

4. EXPERIMENTAL ANALYSIS

The experiments aimed at verifying the effectiveness of domain adaptation in the context of deforestation change detection. We also evaluated the capacity of the proposed approach relative to domain translation through a visual analysis of the translation outcome.

4.1 Datasets

The study area is located in the Brazilian Legal Amazon, specifically in the Brazilian Rondônia state, between the following coordinates: 09°36 '51"S - 10°18'35"S latitude, and 062°56'41"W - 064°20'51"W longitude. Figure 1 (top) shows corresponding subsets of the input images, which were acquired between the years 2016 and 2019.

The images were produced by the Landsat 8-OLI sensor system, with 30m resolution, 7 spectral bands (Coastal/Aerosol, Blue, Green, Red, NIR, SWIR-1, and SWIR-2), dimensions 2550×5120 pixels, and Level-1 data processing, downloaded from the Earth Explorer web service from the United States Geological Survey (USGS). In all experiments, all individual image bands were normalized to zero mean and variance equal to one.

The reference deforestation ground truths were produced by the PRODES Deforestation Mapping project, from the Brazilian National Institute for Space Research (INPE). The data is freely available at (http://terrabrasilis.dpi.inpe.br/map/deforestation). The bottom row of Figure 1 shows the deforestation references for the target years (dark green) and also the accumulated deforestation (light blue) from 2008 until the target year minus one year. In this context, target years refer to the years where we want to detect deforested areas.

4.2 Domain Adaptation Training Setup

The training procedure for the CycleGAN method uses 256×256 pixels patches extracted from the entire extents of the images in both domains. The patches were extracted using a sliding window procedure with stride equal to 50. We randomly shuffled the sample sets in each domain to produce unpaired training samples from both domains. Moreover, each training pair of samples was resized to 286×286, randomly cropped to 256×256, and randomly flipped, following the same procedure applied in (Zhu et al., 2017).

The method was trained with a batch size of 1 sample using the Adam optimizer, with learning rate γ and momentum β_1 set to 0.002 and 0.5 respectively. The coefficients λ_{cyc} and λ_{idt} of the cycle consistency and identity loss were set to 10, while λ_s and λ_t from Equation 9 were set to 0.5. The adversarial loss function used the mean squared error instead the binary cross-entropy, used in the traditional GANs. The method was trained for 200 epochs, applying linear learning rate decay from the 100th epoch, following (Zhu et al., 2017).

4.3 Classifier Training Setup

For the deforestation detection accuracy assessment, we used the Early Fusion (EF) classifier proposed in (Ortega Adarme et al., 2020). Following that work, the *Normalized Difference Vegetation Index* (NDVI), computed before the image normalization process, was stacked along with the spectral image bands of the images, resulting in images with 8 bands for each epoch.

Additionally, also following (Ortega Adarme et al., 2020), the image space was divided into 100 tiles. Approximately 20% of the tiles were used to extract training samples/patches, 5% to extract validation patches, and the remaining 75% to extract the patches used for the evaluation of the classifier. Figure 2 shows the training, validation and test configuration tiles. Tiles 2, 6, 13, 24, 28, 35, 37, 46, 47, 53, 58, 60, 64, 71, 75, 82, 86, 88, and 93 were used for training, tiles 8, 11, 26, 49, 78 for validation and the rest for testing.

The input to EF was a tensor of size 15×15×16. The patches were extracted following the overlapping sliding windows procedure with a stride of 3, as in (Ortega Adarme et al., 2020). During the training and evaluation steps, patches with central pixels having the following characteristics were avoided: belonging to polygons that have been deforested in previous years; lying inside a buffer around the deforestation reference polygons; lying inside deforestation polygons smaller than 6.25 ha, which corresponds to 69 pixels.

Regarding the first and third conditions, we simply adopted the same procedure employed in the PRODES project. The second condition aims at avoiding the impact of inaccuracies in the ground truth, and produced by the rasterization process, which was carried out with the QGIS software. Based on visual inspection of the correspondence between the ground truth and the deforested areas in the images, the width of the buffer was set to 6 pixels: 4 outside the polygons, and 2 inside them.

Data augmentation has been applied only to patches which the central pixel is labeled as deforestation (positive samples). A 90° rotation, as well as vertical and horizontal flips, were the data augmentation transformations. Additionally, only part of the no-deforestation patches (negative samples) in the training and validation tiles were (randomly) selected (the same number

| (a) 2016 | (b) 2017 | (c) 2018 | (d) 2019 |

Deforestation maps
■ Target year
■ Past years

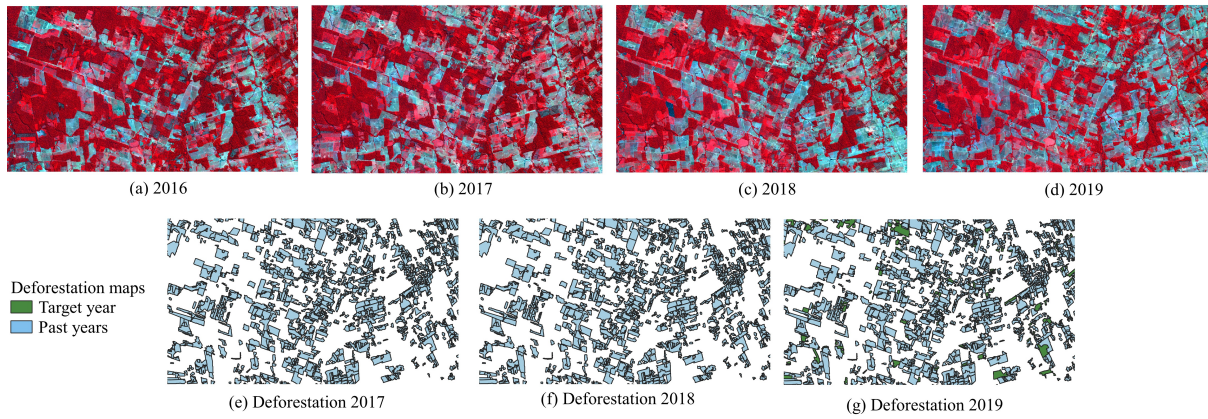| (e) Deforestation 2017 | (f) Deforestation 2018 | (g) Deforestation 2019 |

Figure 1. Images of study area in color composition (NIR, G, B). (a) to (d) Images from 2016 and 2019 respectively. (e) to (g) Deforestation maps provided by PRODES project.
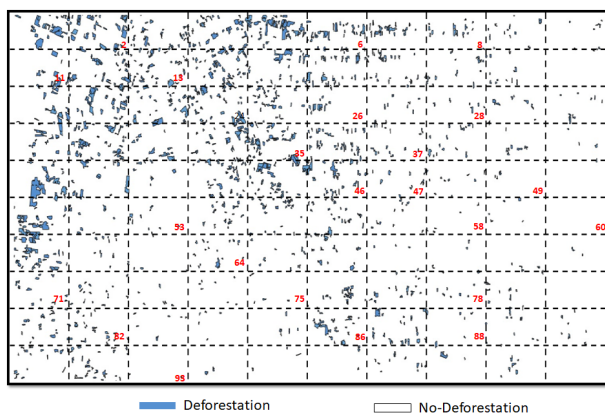

■ Deforestation   □ No-Deforestation

Figure 2. Train, Validation and Test configuration.

as positive samples), in order to balance the number of training samples per class. In the end, as in the Patch-Wise Classification Change Detection (Ortega Adarme et al., 2020), the classification outcome of an input patch is assigned to its central pixel in the evaluation step.

During training, the binary cross entropy was minimized using the Adam optimizer, with learning rate γ and momentum β_1 equal to 0.0001 and 0.9 respectively. The batch size was 32, and the early stopping procedure was used to avoid over-fitting. The patience parameter, which controls the number of epochs without improvements in the validation loss, was set to 10. The classifier was executed 50 times, each time with a different (random) initialization of the trainable parameters, and with a different set of randomly selected negative samples/patches.

4.4 Network Architecture

The network architecture of the CycleGAN's generators, generators' Resnet blocks, discriminator, and the EF classifier are described in in tables 1, 2, 3, and 4, respectively. In the tables, the symbols identifies the operations for each layer: convolution (C), deconvolution (D), instance normalization (I_n), ReLU (R_e), Leaky ReLU (L_R), and MaxPooling (M_P). The number of filters, filters' dimensions and the convolution stride are indicated in parenthesis. In the case of MaxPooling, the kernel dimension and stride are indicated adjacently, while for Reflection Padding, the parenthesis contains the number of rows and columns that will be reflected. Following (Zhu et al., 2017), the CycleGAN architecture uses the instance normalization layer

Layer	Output shape
Input	(256, 256, 14)
Reflection Padding (3,3)	(262, 262, 14)
$CI_n R_e(64, 7, 1)$	(256, 256 ,64)
$CI_n R_e(128, 3, 2)$	(128, 128, 128)
$CI_n R_e(256, 3, 2)$	(64, 64, 256)
1- ResNet block	(64, 64, 256)
$\vdots$	(64, 64, 256)
9- ResNet block	(64, 64, 256)
$DI_n R_e(128, 3, 2)$	(128, 128, 128)
$DI_n R_e(64, 3, 2)$	(256, 256 ,64)
Reflection Padding (3,3)	(262, 262 ,64)
C(14, 7, 1)	(256, 256 ,14)

Table 1. CycleGAN's generator architecture.

Layer
Reflection Padding(1,1)
$CI_n R_e(64, 3, 1)$
Reflection Padding(1,1)
$CI_n(64, 3, 1)$

Table 2. Generator's ResNet block.

Layer	Output shape
Input	(256, 256, 14)
$CL_R(64, 4, 2)$	(128, 128, 64)
$CI_n L_R(128, 4, 2)$	(64, 64, 128)
$CI_n L_R(256, 4, 2)$	(32, 32, 256)
$CI_n L_R(512, 4, 1)$	(31, 31, 512)
C(1, 4, 1)	(30, 30, 1)

Table 3. CycleGAN discriminator architecture.

Layer	Output shape
Input	(15, 15, 16)
$CR_e(128, 3, 1)$	(15, 15, 128)
$M_P(2, 2)$	(7, 7, 256)
$CR_e(256, 3, 1)$	(7, 7, 256)
$M_P(2, 2)$	(3, 3, 256)
$CR_e(512, 3, 1)$	(3, 3, 512)
Flatten(.)	(4608, 1)
Dropout(.)	(4608, 1)
Softmax(.)	(2, 1)

Table 4. EF classifier architecture.

(Ulyanov et al., 2017) instead the well known batch normalization. The difference between them is that the instance normaliz-
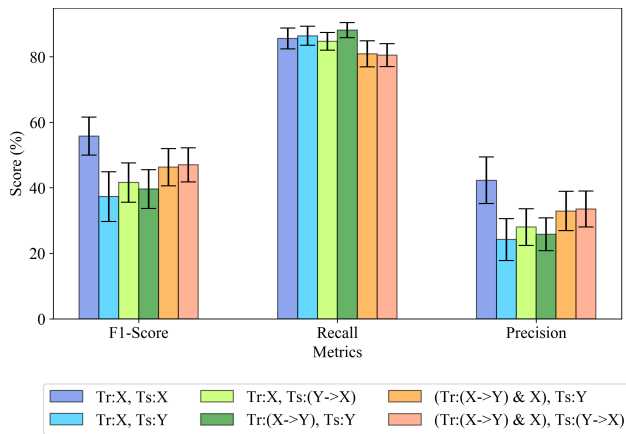
Figure 3. Temporal domain adaptation average ± standard deviation classification results in % with the identity term defined as in Equation 8.

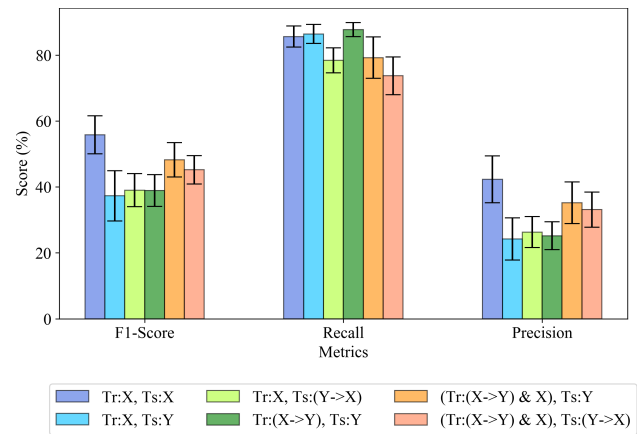

Figure 4. Temporal domain adaptation average ± standard deviation classification results in % with the identity term redefined as in Equation 9.

ation applies the standard normalization to each single image in the batch while batch normalization applies to the whole batch.

The discriminator, presented in Table 3, follows the architecture proposed in (Zhu et al., 2017), 70×70 PatchGANs (Isola et al., 2017). The α parameter in the Leaky ReLU activation function as well as the dropout rate in EF classifier were set to 0.2.

5. RESULTS

In the experiments reported in this paper, the source domain X was composed by the combination (stacking) of images from 2016 and 2017. The target domain Y was composed by the combination of images from 2017 and 2018.

Figures 3 and 4 show the results obtained by the different classification schemes evaluated in this work. In the figures, X corresponds to the source domain, Y to the target domain, $X \rightarrow Y$ represents the source domain adapted to the target, and $Y \rightarrow X$ the target domain adapted to the source domain. Each group of bars, from left to right, represents the average F1-score, Recall, and Precision for the deforestation class, respectively. The plots also show, in black bars, the standard deviation recorded in our experiments.

The classification schemes in which the classifier is trained on $X \rightarrow Y \& X$ represents a sort of data augmentation, aiming at improving the generalization capacity of the classifier by increasing the number of training samples. Moreover, the classification scheme in which the classifier is trained on X and evaluated on Y ($Tr : X, Ts : Y$) can be regarded as the baseline.

In average, the results were low in terms of the F1-score, and the standard deviation can be considered high for all classification schemes. A plausible explanation for such results is the low quantity of no-deforestation samples to train the classifier, which, in the end, determines the total number of training samples. Concerning the standard deviation, the high values may be caused by the stochastic nature of some aspects of the training process, namely, random initialization of model parameters, and random selection of no-deforestation samples due to data balancing.

As expected, the results on Y with the classifier trained on X decreased significantly in relation to the performance of the

classifier trained and tested on X in terms of the F1-score and Precision. Recall, on the other hand, remained almost equal for all schemes. Quantitatively, the baseline showed a decrease of approximately 18% in F1-score, which is evidence of the aforementioned domain shift phenomenon.

Regarding the results presented in Figure 3, which were conducted using the formulation of the identity terms as in Equation 8, the classification results obtained with schemes that use the images generated through the DA process consistently outperformed the baseline, but were lower than those obtained with the first scheme. The improvements achieved with schemes 3 to 6, were: 4.3%; 2.3%; 9%; and 9.7%, respectively. The highest improvements were brought by those schemes that employ $X \rightarrow Y$ as data augmentation.

Figure 4 shows the DA's results obtained using the proposed constrain in the Equation 9. Although they follow the same trend as in the results presented in Figure 3, a slight performance decrease in terms of F1-score can be observed in the schemes in which the classifier is tested with the adapted images $Y \rightarrow X$, specifically in $Tr : X, Ts : Y \rightarrow X$ (from 41% to 39%) and in $Tr : (X \rightarrow Y \& X), Ts : Y \rightarrow X$ (from 48% to 45%). Such behavior can be explained by the pressure that the new term puts in preserving the structures in the source domains. Is has been observed, however, that the new term helps to alleviate model hallucinations.

Figure 5 shows examples of the artifacts generated by the network with the original loss function, as well as the results obtained with the new loss term. The figure shows three images with color composition NIR, G, B. The image on the left-hand side (Figure 5(a)) represent a small subset of the real image taken in 2016. On the center (Figure 5(b)), the same region on the image that resulted from the adaptation process is shown. On the right-hand side (Figure 5(c)) the image produced using the new term is shown. Considering Figure 5(b), the transformation does not preserves the structural characteristics of the source image in the target domain. Several deforested areas have been created by the algorithm in regions where they do not exist originally. Additionally, despite the already mentioned performance decrease brought by the new loss term, it can be observed that it prevents model hallucinations (Figure 5(c)).

Similar to (Ortega Adarme et al., 2020), which evaluated the use of the classifier in an alarm system to reduce the human effort
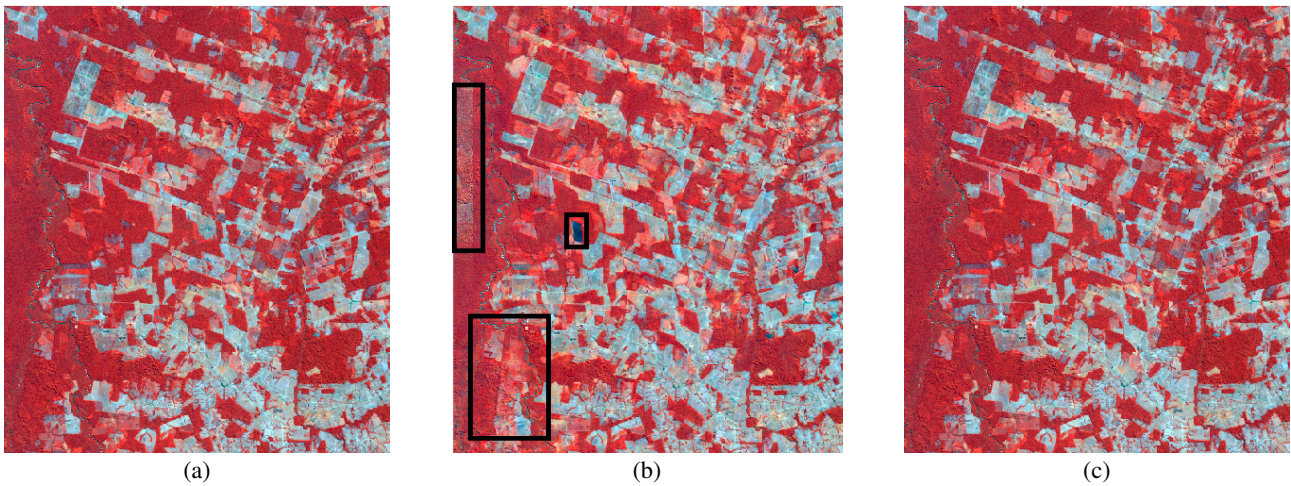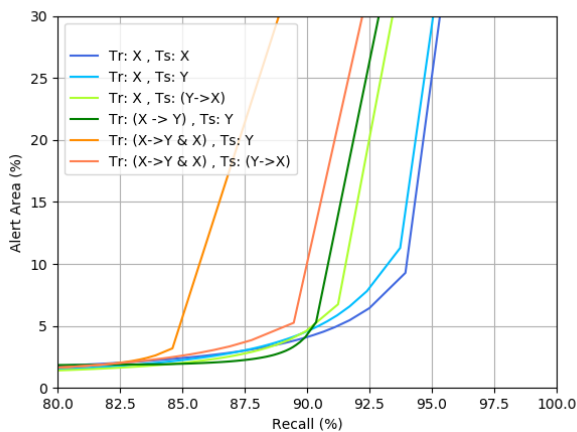
(a)   (b)   (c)

Figure 5. Resulting images from the adaptation process.(a) Original image of 2016 from source domain. (b) Adapted version of the source domain's image to the target one where artifacts where generated by the model (Compare the regions inside the black boxes).(c) Adapted version of the source domain image to the target one applying the constrain proposed in the Equation 9.

in the deforestation monitoring, we also studied the influence of DA in this context. As in this system the photo-interpreter just needs to analyze areas which the classifier indicates as likely

(a)Adaptation process performed using identity term as in Equation 8.

(b)Adaptation process performed using identity term as in Equation 9.
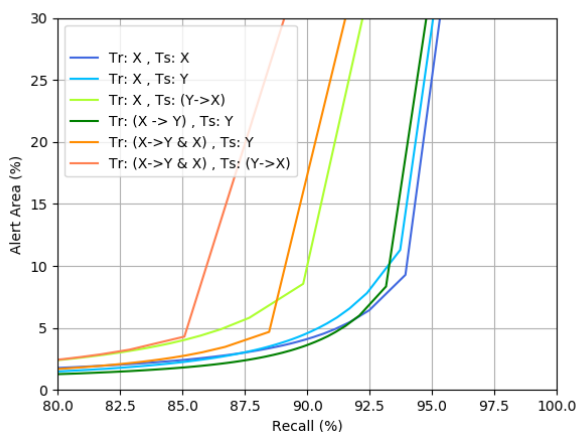
Figure 6. Recall vs Alert Area.

to have been deforested, it is important to observe the relation between the Recall and the Alert Area.

Figures 6(a) and (b) present the curves Recall versus Alert Area of each classification scheme for the domain adaptation alternatives. Each point in the curve corresponds to a threshold imposed on the deforestation probability produced by the classifier. The desirable profile is the one in which high Recall rates are obtained by observing a small area.

Compared to (Ortega Adarme et al., 2020) an identical behavior has been observed: as Recall increased the area to be observed also increased. Furthermore, practically all schemes reached 90% of Recall by observing approximately 5% of the image's extent. Surprisingly, schemes including DA, specifically the one denoted as $Tr : (X \rightarrow Y), Ts : Y$, achieved the same performance of the classifier trained and tested on X with an Alert Area of less than 5%.

6. CONCLUSIONS

In this work, we propose a domain adaptation approach based on CycleGANs in the context of deforestation change detection in the Amazon forest.

Each domain comprises a pair Landsat OLI-8 images from consecutive years: a pair of images acquired in 2016 and 2017 represents the source domain; and a pair of images from 2017 and 2018, covering the same geographic extents, represents the target domain.

The effectiveness of domain adaptation is analyzed considering the performance of a DNN-based change detection classification model, and using reference samples associated solely with the source domain.

The results showed that the domain adaptation performed with a model that is similar to the one originally proposed CycleGAN model was successful in mapping the domains (from source to target and vice-versa) in the sense that the deforestation detection accuracy obtained using the translated domains is higher than the baseline classification scheme; in which the classifier is trained using samples from the source domain and tested using samples from the target domain, without adaptation.

It was observed, however, that the outcome of the adaptation process using the original CycleGAN model produces artifacts, regarded as model hallucinations. Therefore, in an attempt to mitigate the production of such artifacts, we investigated the effect of introducing a new term in the loss function used in the training procedure of the adaptation model. This new term can be considered an extension of the so-call identity loss.

While the inclusion of the new term resulted in a slight decrease in classification accuracy, it was successful in alleviating model hallucinations in the resulting adaptation.

We note, however, that this is a preliminary investigation. In this work, for instance, we did not carry out an in-depth investigation of the relative impact of each individual term of the loss function in the final deforestation detection accuracy and in the generation of artifacts, which we plan to do in the near future.

As future research we also plan to investigate new terms in the loss function and new components in the CycleGAN model that enforce the preservation of the change transitions observed in the original domains, considering their intensities and change directions.

We also plan to further investigate the adaptation quality using different, more complex, deep learning-based change detection classification models.

ACKNOWLEDGEMENTS

REFERENCES

Arkadiusz, N., Michal, R., Adam, J., Michal, T., Konrad, C., Maksymilian, S., Kamil, K., Piotr, M., 2017. Deep learning for satellite imagery via image segmentation. https://blog.deepsense.ai/deep-learning-for-satellite-imagery-via-image-segmentation/.

Bengio, Y. et al., 2009. Learning deep architectures for AI. *Foundations and trends® in Machine Learning*, 2(1), 1–127.

Bermudez, J. D., Happ, P. N., Feitosa, R. Q., Oliveira, D. A., 2019. Synthesis of Multispectral Optical Images From SAR/Optical Multitemporal Data Using Conditional Generative Adversarial Networks. *IEEE Geoscience and Remote Sensing Letters*.

Cohen, J. P., Luck, M., Honari, S., 2018. Distribution matching losses can hallucinate features in medical image translation. *International conference on medical image computing and computer-assisted intervention*, Springer, 529–536.

De Sy, V., Herold, M., Achard, F., Beuchle, R., Clevers, J., Lindquist, E., Verchot, L., 2015. Land use patterns and related carbon losses following deforestation in south america. *Environmental Research Letters*, 10(12).

Deng, X., Yang, H. L., Makkar, N., Lunga, D., 2019. Large scale unsupervised domain adaptation of segmentation networks with adversarial learning. *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 4955–4958.

Ganin, Y., Lempitsky, V., 2014. Unsupervised domain adaptation by backpropagation. *arXiv preprint arXiv:1409.7495*.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. *Advances in Neural Information Processing Systems*, 2672–2680.

Goodman, R., Aramburu, M., Gopalakrishna, T., Putz, F., Gutiérrez, N., Alvarez, J., Aguilar-Amuchastegui, N., Ellis, P., 2019. Carbon emissions and potential emissions reductions from low-intensity selective logging in southwestern Amazonia. *Forest Ecology and Management*, 439, 18-27.

Hoffman, J., Tzeng, E., Park, T., Zhu, J.-Y., Isola, P., Saenko, K., Efros, A. A., Darrell, T., 2017. Cycada: Cycle-consistent adversarial domain adaptation. *arXiv preprint arXiv:1711.03213*.

Isola, P., Zhu, J., Zhou, T., Efros, A. A., 2017. Image-to-image translation with conditional adversarial networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1125–1134.

Krizhevsky, A., Sutskever, I., Hinton, G. E., 2012. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 1097–1105.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature*, 521(7553), 436.

Malingreau, J., Eva, H., de Miranda, E., 2012. Brazilian Amazon: a significant five year drop in deforestation rates but figures are on the rise again. *Ambio*, 41(3), 309-314.

Mateo-García, G., Laparra, V., Gómez-Chova, L., 2019. Domain adaptation of landsat-8 and proba-v data using generative adversarial networks for cloud detection. *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 712–715.

Murez, Z., Kolouri, S., Kriegman, D., Ramamoorthi, R., Kim, K., 2018. Image to image translation for domain adaptation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4500–4509.

Nogueira, K., Penatti, O. A. B., dos Santos, J. A., 2017. Towards better exploiting convolutional neural networks for remote sensing scene classification. *Pattern Recognition*, 61, 539–556.

Nogueron, R., Barreto, P., Souza Jr, C., Anderson, A., Salomão, R., 2006. *Human pressure on the Brazilian Amazon forests*. World Resources Institute, Washington, DC.

Ortega Adarme, M., Queiroz Feitosa, R., Nigri Happ, P., Aparecido De Almeida, C., Rodrigues Gomes, A., 2020. Evaluation of Deep Learning Techniques for Deforestation Detection in the Brazilian Amazon and Cerrado Biomes From Remote Sensing Imagery. *Remote Sensing*, 12(6), 910.

Pan, S. J., Yang, Q., 2010. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.

Schenkel, F., Middelmann, W., 2019. Domain adaptation for semantic segmentation using convolutional neural networks. *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 728–731.

Sun, B., Saenko, K., 2016. Deep coral: Correlation alignment for deep domain adaptation. *European conference on computer vision*, Springer, 443–450.

The Worldwatch Institute, 2015. *Vital Signs Volume 22 - The Trends That Are Shaping Our Future*. Island Press.

Tzeng, E., Hoffman, J., Zhang, N., Saenko, K., Darrell, T., 2014. Deep domain confusion: Maximizing for domain invariance. *arXiv preprint arXiv:1412.3474*.

Ulyanov, D., Vedaldi, A., Lempitsky, V., 2017. Improved texture networks: Maximizing quality and diversity in feed-forward stylization and texture synthesis. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 6924–6932.

Wang, M., Deng, W., 2018. Deep visual domain adaptation: A survey. *Neurocomputing*, 312, 135–153.

Weiss, K., Khoshgoftaar, T. M., Wang, D., 2016. A survey of transfer learning. *Journal of Big Data*, 3(1), 9.

Wittich, D., Rottensteiner, F., 2019. Adversarial Domain Adaptation for the Classification of Aerial Images and Height Data Using Convolutional Neural Networks. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 4.

Wu, Z., Wang, X., Gonzalez, J. E., Goldstein, T., Davis, L. S., 2019. Ace: Adapting to changing environments for semantic segmentation. *Proceedings of the IEEE International Conference on Computer Vision*, 2121–2130.

Zhang, Y., Tang, H., Jia, K., Tan, M., 2019. Domain-symmetric networks for adversarial domain adaptation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5031–5040.

Zhu, J.-Y., Park, T., Isola, P., Efros, A. A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. *Proceedings of the IEEE international conference on computer vision*, 2223–2232.