IEEE Geoscience and Remote Sensing Letters
2022, Volume 19, Pages 2504505 (5p.)
https://doi.org/10.1109/LGRS.2022.3163575
https://archimer.ifremer.fr/doc/00762/87439/

Archimer
https://archimer.ifremer.fr

Domain-Adversarial Neural Networks for Deforestation Detection in Tropical Forests

Soto Vega Pedro Juan [1], Costa Gilson A. [2], Feitosa Raul Q. [3], Ortega Mabel X. [3], Bermudez Jose D. [3], Turnes Javier N. [4]

[1] French Research Institute for Exploitation of the Sea, Brest, France.
[2] State University of Rio de Janeiro, Rio de Janeiro, Brazil.
[3] Department of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, Brazil.
[4] Department of System Design Engineering, University of Waterloo, Waterloo, Canada

Abstract :

Many deep-learning-based, domain adaptation methods for remote sensing applications rely on adversarial training strategies to align features extracted from images of different domains in a shared latent space. However, the performance of such representation matching techniques is negatively impacted when class occurrences in the target domain, for which no labelled data is available during training, are highly imbalanced. In this work, we propose a deep-learning-based representation matching approach for domain adaptation in the context of change detection tasks. We further evaluate the approach in a deforestation mapping application, characterized by a high-class imbalance between the deforestation and no-deforestation classes. The domains represent different sites in the Amazon and Brazilian Cerrado biomes. To mitigate the class imbalance problem, we devised an unsupervised pseudo-labeling scheme based on Change Vector Analysis that prevents the feature alignment to be biased towards the over-represented class. The experimental results indicate that the proposed approach can improve the accuracy of cross-domain deforestation detection.

Keywords : Deforestation Detection, Change Detection, Domain Adaptation, Deep Learning, Remote Sensing

Please note that this is an author-produced PDF of an article accepted for publication following peer review. The definitive publisher-authenticated version is available on the publisher Web site.

I. Introduction

CHANGES in environmental conditions, geographical variability, and different sensor properties, make it virtually impossible to employ previously trained deep-learning (DL) based classifiers on new data without a significant decrease in classification accuracy. This is known as the domain shift (DS) problem, which has to do with the dissimilarity between the distributions derived from the data acquired at the different sites or epochs. Considering the high demand for labeled data, DS may seriously impair the operationalization of DL-based classification approaches in real-world applications. Learning domain agnostic features through domain adaptation (DA) techniques represents a way to mitigate the DS problem, and also lessen the demand for labeled data in RS applications.

To date, a few different DA approaches have been employed for CD on urban areas, e.g., [1], [2], [3], and [4]. To the best of our knowledge, only two DA approaches have focused on the deforestation detection task, namely: [5] and [6]. Such DA approaches can be categorized as *representation matching* and *appearance adaptation* techniques [7]. Appearance adaptation is mainly supported by so-called Image-to-Image (I2I) translation methods, which aim at transforming images from a target domain so that the generated images have a style that is similar to the images from a source domain. Representation matching aims at aligning features extracted from the domains in a common latent feature space, and involves either creating a different classifier for each domain [5], or a single classifier that is able to properly classify data from both domains [8].

In this work we introduce a representation matching DA approach tailored for deforestation detection, which is based on the Domain Adversarial Neural Network (DANN) method [8], and we evaluate the approach on three sites located in two Brazilian forest biomes. The sites cover different forest types and are characterized by different deforestation practices and patterns. The sites are also characterized by a high class imbalance, as the deforestation areas cover only a small fraction of total areas.

Such critical class imbalance is challenging for DA techniques, especially for semi-supervised representation matching methods such as DANN. In semi-supervised DA [9], labelled training samples are only available for the source domain. Therefore simple balancing strategies, such as under-sampling of the over-represented class, can be used for source training samples. However, as no class labels are available for the target domain samples, such balancing procedures cannot be employed directly. We propose an unsupervised procedure that handles this problem by balancing the target domain samples for DA training, relying on pseudo-labels generated by Change Vector Analysis (CVA).

Considering that the pseudo-labels are noisy, in the sense that they might be incorrect, we do not use them for training the classifier (*Label Predictor*) that composes the DANN architecture, as their quality can decisively affect classification accuracy [10]. In this work, the pseudo-labels are solely used to provide for a weak supervision in the balancing procedure of the target domain samples for the feature alignment DA process carried out by DANN. The problem tackled here is that in a highly imbalanced problem the feature discriminator (*Domain Classifier* in terms of the DANN terminology) tends to be biased towards the over-represented class. Our solution uses the noisy labels to balance the target domain samples during training, so to prevent such bias and produce features that are representative of both the over- and the under-represented classes. The main contribution of this work is a proposal for

Pedro J. Soto is with the French Research Institute for Exploitation of the Sea, Brest, France.(e-mail: pjsotove@ifremer.fr).

Gilson A. Costa is with the State University of Rio de Janeiro, Rio de Janeiro, Brazil.

Mabel X. Ortega, Raul Q. Feitosa, and José D. Bermudez are with the Department of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, Brazil.

Javier N. Turnes is with the Department of System Design Engineering, University of Waterloo, Waterloo, Canada.

1545-598X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

balancing the target domain samples used in representation matching based DA for deforestation detection. It is worth noticing that the proposed procedure is not limited to the DANN method, as potentially any representation matching DA techniques can adopt it.

In the experiments carried out to verify the effectiveness of the proposed procedure, we conducted an ablation study, in which the proposed DANN+CVA solution is compared with the original DANN training strategy. Additionally, we compare the results obtained with the proposed approach with those achieved by state-of-the-art DA approaches employed in deforestation detection.

II. Fundamentals

Basically, the CVA [11] technique computes the magnitude and direction of change between two corregistered multispectral images acquired at different epochs. Formally, let $\mathbf{x}_{t_0}(i,j)$ and $\mathbf{x}_{t_1}(i,j)$ represent a pixel's spectral vector at a given pixel location (i,j) in a pair of coregistered images acquired at t_0 and t_1, respectively. The magnitude M and direction ϕ of changes are computed as follows:

$$M = ||\mathbf{x}_{t_1}(i,j) - \mathbf{x}_{t_0}(i,j)||_2 \tag{1}$$

$$\cos\phi = \frac{\mathbf{x}_{t_1}(i,j) \cdot \mathbf{x}_{t_0}(i,j)}{||\mathbf{x}_{t_1}(i,j)||_2 ||\mathbf{x}_{t_0}(i,j)||_2} \tag{2}$$

where $||\cdot||_2$ denotes the L_2 norm.

CVA components have been used in several ways, for different change detection purposes. In this work, we binarize each component separately using thresholds calculated with the OTSU [12] algorithm. Then, we considered the change transition between $\mathbf{x}_{t_0}(i,j)$ and $\mathbf{x}_{t_1}(i,j)$ as positive if both the corresponding magnitude and phase are greater than the respective thresholds. A detailed explanation of that procedure is given in Section III.

Proposed in [8], rather than a particular method, DANN introduces an adversarial strategy for learning domain invariant latent representations, which aim at minimizing the divergence between two latent probability distributions parameterized by a deep neural network. As shown in Fig. 1, DANN comprises three modules: a *Feature Extractor*, $G_f(\cdot,\theta_f)$, which maps both the source and target input feature space into a common latent space Z; a *Label Predictor*, $G_l(\cdot,\theta_l)$, which estimates class labels for the input samples; and a *Domain Classifier*, $G_d(\cdot,\theta_d)$, used only during training, which discriminates between source and target samples based on the features produced by the Feature Extractor.

In the DANN strategy, the features derived by the Feature Extractor for the target domain samples are not forwarded through the Label Predictor during training, as their corresponding class labels are unknown. The features derived from both source and target domain samples are, however, forwarded through the Domain Classifier, as their domain labels are known. The optimal parameters θ_f, θ_l, and θ_d can be found by solving the equations 3 and 4:

$$(\theta_f^*, \theta_l^*) = \text{argmin}_{\theta_f,\theta_l} E(\theta_f, \theta_l, \theta_d^*) \tag{3}$$

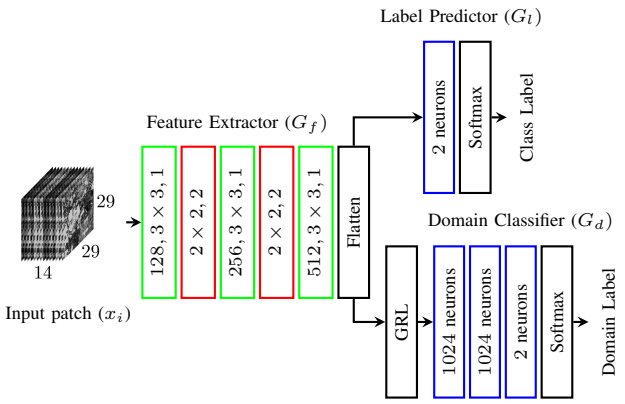

Fig. 1. Overview and architecture details of the proposed DANN-based approach. The model follows a patch-wise classification scheme, aimed at predicting the class of the central pixel in the input patch, similar to an image classification task. The operations performed at each layer are described as: Convolution followed by ReLU activation function in green; MaxPooling in red; and Fully Connected in blue. The values inside the boxes indicate: for the Convolutional layers, the number of filters, filter dimensions and stride; for the MaxPooling layers, the kernel dimensions and stride; and for Fully Connected layers, the number of neurons. GRL stands for *Gradient Reversal Layer*.

$$(\theta_d^*) = \text{argmax}_{\theta_d} E(\theta_f^*, \theta_l^*, \theta_d) \tag{4}$$

where $E(\theta_f, \theta_l, \theta_d)$ is the total loss function, defined as:

$$E(\theta_f, \theta_l, \theta_d) = \mathcal{L}_l(\theta_f, \theta_l) - \lambda \mathcal{L}_d(\theta_f, \theta_d) \tag{5}$$

The first term of Equation 5 represents the Label Predictor loss for the source domain training samples, while the second represents the Domain Classifier loss for the source and target training samples. The parameter λ controls the influence of the Domain Classifier loss on the update of the Feature Extractor parameters, which, according to [8] should start at zero, and gradually increase through the training epochs. Thus, the learning procedure updates the model parameters by applying the rules described in lines 17 and 18 of the algorithm presented in the next section. We observe that the positive update term (multiplied by the λ coefficient), in line 18 of the algorithm represents an adversarial term in the training of the Feature Extractor network. The term is intended to penalize the Feature Extractor when the Domain Classifier is able to discern the domain a sample belongs to. This operation is implemented by a so-called *gradient reversal layer* (GRL) (see Fig. 1), which acts as an identity mapping in the forward step, but reverses the gradient (multiplying it by -1) coming from the Domain Classifier during backpropagation.

III. Proposed Method

Let $x_r^s = [x_{r_{t_0}}^s : x_{r_{t_1}}^s]$ and $x_q^t = [x_{q_{t_2}}^t : x_{q_{t_3}}^t]$ represent two coregistered, concatenated pairs of multispectral remote sensing image patches of dimensions $w \times h \times b$ pixels, where w is the width, h is the height, b the number of spectral bands, and $[:]$ is the concatenation operator. S and T denote the sets of domains samples x_r^s and x_q^t belong to, i.e., *source* or *target*; y_r^s represents the class label of x_r^s taking a value from the set $\{0,1\}$, where 1 means *deforestation*, and 0 means *no-deforestation*; and $\hat{y}_q^t$ represents the pseudo-label of x_q^t, either 0 or 1, predicted by a function $C(\cdot)$. We denote as $R = |S|$

Algorithm: DANN+CVA Training

Inputs :

- $S = \{(x_r^s, y_r^s)\}_{r=1}^{R}$ // Labelled source samples
- $T = \{x_q^t\}_{q=1}^{Q}$ // Unlabelled target samples
- N // Number of training samples from each domain
- $\{\theta_f, \theta_l, \theta_d\}$ // Initial weights of G_f, G_l and G_d
- $\{\mu_0, \alpha, \beta, \gamma\}$ // Hyper-parameters

Output: $\{\theta_f, \theta_l\}$ // Final weights of G_f, G_l

// Step 1: Compute pseudo-labels
$T' \leftarrow C(T)$ // Pseudo-labeled target sample set
// Step 2: Select N samples from the source domain
$S^{tr} = \{(x_n^{str}, y_n^{str})\}_{n=1}^{N} \leftarrow U(S, N)$, where $N < R$
// Step 3: Select N samples from the target domain with the pseudo-labels
$T^{tr} = \{x_n^{ttr}, \hat{y}_n^{ttr}\}_{n=1}^{N} \leftarrow U(T', N)$, where $N < Q$
// Step 4: Training the DANN model
while $j <$ # *Epochs* **do**
  $p \leftarrow \frac{j}{\#Epochs}$; $\mu \leftarrow \frac{\mu_0}{(1+\alpha p)^{\beta}}$; $\lambda \leftarrow \frac{2}{(1+e^{-\gamma p})} - 1$
  // Step 4.1: Building the Batch
  $Batch \leftarrow SelectBatch(S^{tr}, T^{tr})$
  $S_{batch}^{tr} \leftarrow Batch \cap S^{tr}$; $T_{batch}^{tr} \leftarrow Batch \cap T^{tr}$
  // Step 4.2: Forward processing
  $\mathcal{L}_l \leftarrow \frac{1}{|S_{batch}^{tr}|} \sum_{(x_i,y_i)\in S_{batch}^{tr}} y_i \log(G_l(G_f(x_i))) + (1 - y_i) \log(1 - G_l(G_f(x_i)))$
  $\mathcal{L}_d \leftarrow \frac{1}{2|S_{batch}^{tr}|} \sum_{(x_i,y_i)\in S_{batch}^{tr}} \log(G_d(G_f(x_i))) + \frac{1}{2|T_{batch}^{tr}|} \sum_{(x_i)\in T_{batch}^{tr}} \log(1 - G_d(G_f(x_i)))$
  // Step 4.3: Backpropagation
  $\theta_l \leftarrow \theta_l - \mu \frac{\partial \mathcal{L}_l}{\partial \theta_l}$; $\theta_d \leftarrow \theta_d - \mu \frac{\partial \mathcal{L}_d}{\partial \theta_d}$
  $\theta_f \leftarrow \theta_f - \mu \frac{\partial \mathcal{L}_l}{\partial \theta_f} + \lambda \mu \frac{\partial \mathcal{L}_d}{\partial \theta_f}$
  $j \leftarrow j + 1$
end while
// Pseudo-labelling function
Function $C(\{x_q\})$**:**
  // Extract patches' central pixels
  $\{\mathbf{x}_q^{t0}\} \leftarrow \{x_{q_{t_0}}\}$; $\{\mathbf{x}_q^{t1}\} \leftarrow \{x_{q_{t_1}}\}$
  // Compute magnitude and phase of central pixels
  $\{M_q\} \leftarrow \{||\mathbf{x}_q^{t1} - \mathbf{x}_q^{t0}||_2\}$
  $\{\phi_q\} \leftarrow \left\{\arccos \frac{\mathbf{x}_q^{t1} \cdot \mathbf{x}_q^{t0}}{||\mathbf{x}_q^{t1}||_2 ||\mathbf{x}_q^{t0}||_2}\right\}$
  // OTSU threshold for magnitude and phase
  $T_{h_M} \leftarrow OTSU(\{M_q\})$, $T_{h_\phi} \leftarrow OTSU(\{\phi_q\})$
  // Generate pseudo-labels
  $\{\hat{y}_q\} \leftarrow \left\{ \begin{array}{ll} 1 & \text{if } (M_q > T_{h_M}) \wedge (\phi_q > T_{h_\phi}) \\ 0 & \text{otherwise} \end{array} \right\}$
  return $\{(x_q, \hat{y}_q)\}$
// Balancing function
Function $U(Samples, N)$**:**
  // Select positive samples and perform data augmentation to get N samples
  $\{x_n^{pos}, y_n^{pos}\}_{n=1}^{N} \leftarrow SelectPos\&Augment(Samples, N)$
  // Select N negative samples randomly
  $\{x_n^{neg}, y_n^{neg}\}_{n=1}^{N} \leftarrow SelectNeg(Samples, N)$
  return $\{(x_n^{neg}, y_n^{neg})\} \cup \{(x_n^{pos}, y_n^{pos})\}$

and $Q = |T|$ the number of images in the S and T sets, respectively. The initial learning rate and the hyper-parameters used to compute λ are denoted as μ_0, α, β and γ.

As represented in the DANN+CVA Training Algorithm, the DA process begins by selecting balanced sets of training samples from both domains. The actual balancing procedure follows the creation of pseudo-labels $\hat{y}_q^t$ for the target domain samples x_q^t (Step 1 of the algorithm). The sample balancing procedure (steps 2 and 3) is implemented in a function $U(\cdot)$, which performs traditional data augmentation operations, i.e., rotations and reflections, to increase the number of samples of the underrepresented positive class (*deforestation*), and randomly selects an equal number of samples of the over-represented negative class (*no-deforestation*). Both resulting sets comprise N samples, with $N < R$ and $N < Q$.

Since class labels of target domain are not available during training, the function in $U(\cdot)$ uses $\{\hat{y}_q^t\}_{q=1}^{Q} \leftarrow C(\{x_q^t\}_{q=1}^{Q})$ as pseudo-labels for the target samples x_q^t. Such labels are produced by the function $C(\cdot)$, which implements the pseudo labeling procedure of the algorithm (Step 2). The procedure is based on change vector analysis (CVA) [11], and on an unsupervised thresholding technique, namely the OTSU [12] method. First, CVA delivers the magnitude M_q and phase ϕ_q of each patch central pixel. Then, the OTSU procedure computes the thresholds T_{h_M} and T_{h_ϕ} for each set of CVA components $\{M_q\}$ and $\{\phi_q\}$. Finally, a binarization procedure, in which the outcome of magnitude and phase are combined through the AND (&) logic operation, delivers the pseudo-label set.

Then, the selected samples from both domains, represented in the sets S^{tr} and T^{tr}, are used to train the DANN model (Step 4) until convergence by simultaneously updating the set of parameters $\{\theta_f, \theta_l, \theta_d\}$, being the objective function the one given by Equation 5, and the loss components $\mathcal{L}_l$ and $\mathcal{L}_d$ given by lines 14 and 15 of the Algorithm.

IV. Experimental Setup

The study areas were the same ones used in [5] and [6]. Two of the sites are located in the Amazon biome, specifically in the Brazilian states of Rondônia (RO) and Pará (PA). The third is located in the Brazilian Cerrado (Savannah) biome, in Maranhão (MA) state. Table I shows information regarding image dates, vegetation typology and class distribution.

TABLE I
Detailed information of each domain: image acquisition dates, classes distribution, and vegetation typology.

Domains	RO	PA	MA
Vegetation	Open Ombrophyll	Dense Ombrophyll	Seasonal Deciduous and Semi-Decidous
Date 1	July 18, 2016	August 2, 2016	August 18, 2017
Date 2	July 21, 2017	July 20, 2017	August 21, 2018
Deforested pixels	225 635 (3%)	82 970 (3%)	71 265 (3%)
No-deforested pixels	3 816 981 (29%)	1 867 929 (65%)	1 389 844 (57%)
Previously deforested pixels	9 013 384 (69%)	903 901 (32%)	986 891 (40%)

The images were acquired by the Landsat 8-OLI system, with 30m spatial resolution and 7 spectral bands. Level-1 data processing was applied for all images, and in all of the experiments, each image band was normalized to zero mean and unit variance. The PRODES [13] project conducted by the Brazilian National Institute for Space Research (INPE) produced the reference deforestation maps. The data is freely available at Terrabrasilis website[1].

Analogous to [5], [6], we split the images pairs from the different sites into non-overlapping tiles, from which we selected the training, validation and test samples/patches. We

[1]http://terrabrasilis.dpi.inpe.br/map/deforestation

TABLE II
MEAN AVERAGE PRECISION (MAP), F1-SCORE (F1), PRECISION (P) AND RECALL (R) OF DIFFERENT DA METHODS.

Methods	S	PA				PA				RO				RO				MA				MA			
	T	RO				MA				PA				MA				PA				RO			
	Metrics	mAP	F1	P	R	mAP	F1	P	R	mAP	F1	P	R	mAP	F1	P	R	mAP	F1	P	R	mAP	F1	P	R
No Domain Adaptation		29.8	20.9	37.3	14.8	66.8	32.6	20.3	83.6	56.9	44.9	77.2	32.5	40.2	29.6	17.9	86.6	**86.8**	**75.7**	68.4	85.3	63.8	61.9	58.4	66.5
ADDAm[5]		*56.0*	**57.2**	52.9	62.2	63.7	**49.2**	35.1	82.4	**67.7**	**63.1**	72.4	55.9	**84.2**	**77.4**	79.9	75	82.5	74.6	82.7	67.9	64.5	**63.8**	61.3	66.5
CycleGAN DN[6]		47.2	48	39.2	62	**79.2**	*48.0*	33.4	88.2	60.2	51	55.8	48.5	*81.2*	*67.0*	68.7	66.1	86.1	72	82.2	65	64.7	45.4	55	40
DANN[8]		19.8	19.2	13.4	64.7	70.7	21	11.9	89.8	43.3	22.7	15.3	71.81	49.1	26.3	15.8	87.5	81.7	48.4	40.9	83.9	*65.3*	51.6	44.8	75.2
DANN+CVA		**65.8**	*50.7*	46.4	58.4	*78.3*	23.1	12.5	91.5	*62.4*	*54.5*	44.9	65.1	58.9	42.9	34.3	85.3	*86.6*	*74.7*	67.3	82	**76.1**	*63.7*	50.9	79.3

split the larger images from RO into 100 tiles and took approximately 20%, 5% and 75% for training, validation and testing, respectively. We split the comparatively smaller images from PA and MA into 15 tiles, whereby 26% of which for training, 13% for validation, and 60% for testing.

We adapted the classification branch of DANN method (Feature Extractor and Label Predictor) for deforestation detection using the Early Fusion procedure proposed in [14] (EF-CNN), which follows a patch-wise classification scheme, aimed at predicting the class of the central pixel of each patch. The samples extracted from the image pairs in all sites correspond to patches with 29×29 pixels and 14 channels (two times the number of the Landsat8-OLI images). Training and validation patches were extracted following a sliding window procedure with a stride of 3 pixels, avoiding patches with central pixels covering areas deforested in previous years. Data augmentation was applied to the minority class samples, for both the source and target domains. In particular, we applied 90° rotation, vertical and horizontal flips.

Similar to [8], we used the Momentum optimizer and learning rate decay during training. We set the initial learning rate μ_0 and momentum β_1 equal to 0.01 and 0.9, respectively. Additionally, we followed the same updating procedure for the λ parameter. The batch size was set to 32, and the early stopping procedure was used to avoid over-fitting. The model was executed 10 times, each time with a different (random) initialization of the trainable parameters. Fig. 1 shows a detailed description of the DANN model implemented in our experiments. The source code is available at `http://www.lvc.ele.puc-rio.br/wp/?cat=41`.

V. RESULTS

Tables II shows the results obtained with the domain adaptation approaches under different cross-domain scenarios, in terms of mean average precision (mAP), F1-score (F1), Precision (P) and Recall (R). In the table, each column represents a scenario: S corresponds to the source domain and T to the target domain. The *No Domain Adaptation row*, regarded as the *base-line classifier*, corresponds to a classification scheme in which the Feature Extractor and the Label Predictor are trained on the source domain and evaluated directly on the target domain. The remaining rows show the results of the different domain adaptation methods, specifically: the original DANN method [8]; the ADDA-based approach [5]; the CycleGAN DN approach [6]; and the extension of DANN method proposed in this work (DANN+CVA). We observe that the mAP values correspond to the area under the curve obtained when computing pairs of Precision and Recall values for different classification thresholds, in the range of 0 to 1, over the average of probability maps delivered by each classifier. Complementing the results shown in Table II, the mAP values obtained with the base-line model trained and tested in the same domains were: 82.7% for the PA domain; 84% for RO; and 94.1% for MA.

For the cross-domain classification schemes, the best performances were obtained when training on the MA domain, whereby the overall best cross-domain result was obtained in the [S: MA|T: PA] scenario. On the other hand, training on the PA domain led to the lowest mAP values. Additionally, the classifier trained on RO reached performances that are between those obtained when training on PA and MA.

In general, the results show that the proposed DANN+CVA approach brought significant performance improvements relative to the respective cross-domain baselines in five of the six scenarios evaluated. However, the proposed approach led to a negative transfer in [S: MA, T: PA] (-0.2% and -1% for mAP and F1 respectively), which is also not surprising because in that scenario, the base-line classifier obtained the best results among all cross-domain evaluations.

Regarding the classification results obtained with the variants of the DANN model, the one trained with balanced source and target samples (DANN+CVA) was consistently better than its counterpart, i.e., the original DANN method, trained with an imbalance set of target domain samples. The improvements brought by the proposed DANN+CVA approach in relation to the plain DANN method were 46%, 7.6%, 19.1%, 9.1%, 4.9% and 10.8%, respectively for each domain combination in terms of mAP, while in terms of F1-score the gains were 31.5%, 2.1%, 31.8%, 16.6%, 26.3%, 12.1% respectively.

Fig. 2 shows qualitative results. The classification maps in figures 2(g)(h)(i) show that the performance of the base-line classifier is hindered by the domain shift problem. In the selected cross-domain scenarios, the plain DANN method obtains results (figures 2(p)(q)(r)) that are visually inferior to the ones obtained with both the base-line classifier, and with the DANN+CVA approach (figures 2(s)(t)(u)). Such results are consistent with the mAP and F1-score values in Table II.

Finally, we compared the proposed method with recent, state-of-the-art domain adaptation approaches applied to deforestation detection, namely, ADDAm [5] and CycleGAN-DN [6]. In terms of mAP, the proposed DANN+CVA approach outperformed ADDAm [5] and CycleGAN-DN [6] in three of the six studied scenarios. The CycleGAN-DN [6] approach was slightly superior to the proposed approach in the [S: PA|T: MA] domain combination, and ADDAm [5] was better in [S: RO|T: PA]. Additionally, both ADDAm [5] and CycleGAN-DN [6] outperformed the proposed approach in the [S: RO|T: MA] domain combination. At this point, it is

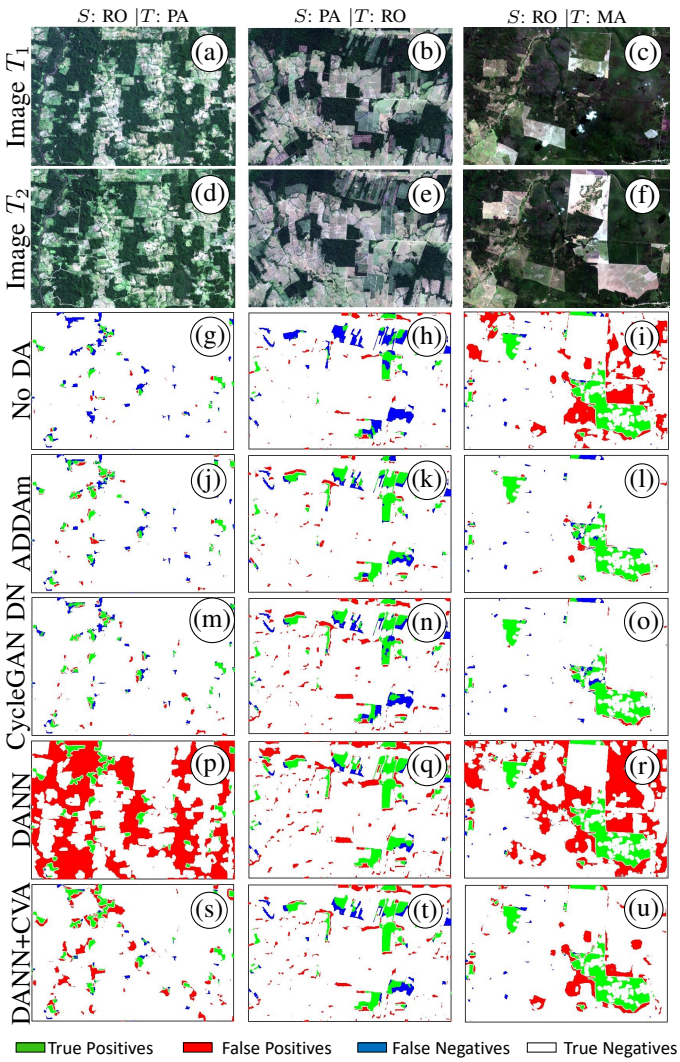

Fig. 2. Qualitative results of the cross domain of the deforestation detection task. Results of each domain are represented in the columns while each row contains visual representation of the images and the results when no DA was employed, as well when DA with ADDAm, CycleGAN DN, DANN and DANN+CVA were employed.

not clear what causes the lower performance of the proposed method in this particular domain combination, such behaviour deserves further investigation. On the other hand, ADDAm [5] outperformed DANN+CVA and CycleGAN DN [6] in four of six scenarios in terms of F1-Score.

VI. Conclusions

In this letter, we proposed a domain adaptation approach based on the DANN method for change detection, applied to deforestation detection in Brazilian forest biomes. The approach relies on an unsupervised pseudo-labeling procedure based on Change Vector Analysis, which aims at mitigating the negative effects brought by class imbalance in the training of the DA model.

We compared the performance of the proposed aproach with that of the original DANN method, trained with unbalanced target samples, on six different domain combinations. In all cases, the proposed balancing scheme alleviated the consequences of class imbalance in the domain adaptation process. Additionally, the proposed approach outperformed, in most cases, state-of-the-art domain adaptation methods proposed for the same application, over the same dataset.

We believe that the reported results encourage the use of the proposed balancing scheme in combination with other domain adaptation approaches employed in change detection applications that consider vegetation classes.

In the future, we plan to evaluate the effect of using pseudo-labels in training not only the Domain Classifier component of the DANN model, but also the Label Predictor. We also plan to implement the components of the DANN strategy with fully convolutional architectures.

Acknowledgment

The authors would like to thank the founding provided by CAPES, CNPq, FAPERJ and NVIDIA corporation.

References

[1] H. Chen, C. Wu, B. Du, and L. Zhang, "Dsdanet: Deep siamese domain adaptation convolutional neural network for cross-domain change detection," *arXiv preprint arXiv:2006.09225*, 2020.

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

[3] S. Saha, F. Bovolo, and L. Bruzzone, "Unsupervised deep change vector analysis for multiple-change detection in vhr images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 6, pp. 3677–3693, 2019.

[4] S. Saha, Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "Unsupervised deep transfer learning-based change detection for hr multispectral images," *IEEE Geoscience and Remote Sensing Letters*, 2020.

[5] J. Noa, P. Soto, G. Costa, D. Wittich, R. Feitosa, and F. Rottensteiner, "Adversarial discriminative domain adaptation for deforestation detection," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 3, pp. 151–158, 2021.

[6] P. J. S. Vega, G. A. O. P. da Costa, R. Q. Feitosa, M. X. O. Adarme, C. A. de Almeida, C. Heipke, and F. Rottensteiner, "An unsupervised domain adaptation approach for change detection and its application to deforestation mapping in tropical biomes," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 181, pp. 113–128, 2021.

[7] D. Wittich and F. Rottensteiner, "Appearance based deep domain adaptation for the classification of aerial images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 180, pp. 82–102, 2021.

[8] Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation," *arXiv preprint arXiv:1409.7495*, 2014.

[9] D. Tuia, C. Persello, and L. Bruzzone, "Domain adaptation for the classification of remote sensing data: An overview of recent advances," *IEEE geoscience and remote sensing magazine*, vol. 4, no. 2, pp. 41–57, 2016.

[10] Y. Li, Y. Zhang, and Z. Zhu, "Error-tolerant deep learning for remote sensing image scene classification," *IEEE transactions on cybernetics*, vol. 51, no. 4, pp. 1756–1768, 2020.

[11] W. A. Malila, "Change vector analysis: an approach for detecting forest changes with landsat," in *LARS symposia*, 1980, p. 385.

[12] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[13] C. A. Almeida, L. E. P. Maurano, D. D. M. Valeriano, G. Camara, L. Vinhas, A. R. Gomes, A. M. V. Monteiro, A. A. A. Souza, C. D. Renno, D. E. Silva, M. Adami, M. I. S. Escada, M. Mota, and S. A. Kampel, "Methodology for Forest Monitoring used in PRODES and DETER projects." INPE, São José dos Campos, Tech. Rep., 2021. [Online]. Available: http://urlib.net/rep/8JMKD3MGP3W34R/443H3RE

[14] M. Ortega Adarme, R. Queiroz Feitosa, P. Nigri Happ, C. Aparecido De Almeida, and A. Rodrigues Gomes, "Evaluation of deep learning techniques for deforestation detection in the brazilian amazon and cerrado biomes from remote sensing imagery," *Remote Sensing*, vol. 12, no. 6, p. 910, 2020.