Semantic Segmentation with Multi-Source Domain Adaptation for Radiological Images

Hugo Oliveira†, Arnaldo de A. Araújo, Jefersson A. dos Santos
Department of Computer Science
Universidade Federal de Minas Gerais
Av. Antônio Carlos, 6627, Belo Horizonte, Brazil

***Abstract*—Differences in digitization equipment and techniques in radiology may hamper the use of data-driven deep learning approaches. In order to mitigate this limitation, in this work we merge generative image translation networks with supervised semantic segmentation architectures, yielding two semi-supervised methods for domain adaptation in medical images. We compare our methods with traditional baselines in the literature using 3 image domains, 16 datasets and 8 segmentation tasks organized into three sets of experiments. Analysis of the results showed that the proposed methods for Domain Adaptation often reached Jaccard scores of 0.9 or higher in unsupervised or semi-supervised settings. We observe that unsupervised domain adaptation performance is close to the performance of fully supervised adaptation in most cases, bridging an important gap in the efficacy of neural networks between labeled and unlabeled datasets.**

I. Introduction

Radiology has been a useful tool for assessing health conditions since the last decades of the 19^{th} century, when X-Rays were first used for medical purposes. In recent decades, Machine Learning was incorporated into the body of knowledge of Computer Aided Detection/Diagnosis (CAD) systems for biomedical image analysis. Several surveys in Biomedical Images [1]–[3] show the rapid dissemination of Deep Learning on the automated analysis of biomedical imaging during recent years.

The main limitation of Deep Learning models is the amount of data and labels available for training, as generalizing useful patterns over unstructured data can be exceptionally hard. For instance, radiological images from national agencies such as the brazilian *Sistema Único de Saúde* (SUS) are known to be gathered from multiple digitization sensors and techniques. Those differences in the acquisition process will inevitably lead to distinct visual patterns when compared to images from within the same dataset and other countries. Additionally, large annotated radiology datasets are hard to obtain due to legal/ethical problems concerning privacy and because labeling biomedical images is a highly specialized skill. These problems severely hamper the use of large datasets in order to gather useful knowledge for local medical imaging applications in smaller hospitals or regions with limited resources, wherein local data cannot be labeled.

Domain Adaptation (DA) [4] methods are often used to improve the generalization of Deep Neural Networks (DNNs) over biomedical images in an unsupervised or semi-supervised manner. The most popular method for deep DA is transfer learning via fine-tuning pretrained neural networks from larger datasets. However, fine-tuning only learns from labeled data, ignoring the larger amounts of unlabeled data available in most real-world scenarios. During the last years, several approaches have been proposed for Unsupervised Domain Adaptation (UDA) [5], [6], Semi-Supervised Domain Adaptation (SSDA) [7], [8] and Fully Supervised Domain Adaptation (FSDA) [9]. In all of these scenarios there is a reasonably large source dataset $\mathcal{S}$ that can provide labels for the training of a certain supervised task in the target dataset $\mathcal{T}$. Until recently these advances in DA were applied only to sparse labeling (classification) tasks, leaving a considerable gap in the literature for UDA and SSDA methods in dense labeling (segmentation) tasks.

Image translation [10], [11] and adversarial training [12] have been recently repurposed to perform these knowledge transfer tasks in segmentation scenarios. Following this trend, the first contribution of the doctoral work [13] summarized in this text comprised an image translation approach for pairwise UDA and SSDA for dense labeling in X-Ray images, which was a novelty at the time of publishing. This initial strategy was published in *SIBGRAPI 2018* [14] and allowed for the democratization of lung, heart and clavicle segmentation across Chest X-Ray datasets. Aiming to leverage multiple data/label sources, we modified this initial method to perform Conditional DA instead of D2D, while also adding supervision to an isomorphic representation across the datasets instead of either the source or target images directly. This was in contrast to all other Generative Adversarial Networks (GANs) for Image-to-Image Translation (I2I) applied to DA in the literature [14]. The proposal of this approach was published at *IEEE Access* [15] and is the most thoroughly evaluated part of this works, serving as a basis for multiple future works. At last, we also published a paper in *Pattern Recognition Letters* (*PRL*) employing Conditional DA to transfer knowledge between synthetic and real data [16] in order to leverage the higher dimensionality of 3D Computed Tomography (CT) to perform generalizable rib segmentation in 2D Chest X-Rays. These approaches and some of their most important results are discussed in Sections II and III of this manuscript, while current works still in development that branched from the doctoral work [13] are further described in Section IV.

†PhD. thesis presented on July 2020.

II. Proposed Methods

A. Pairwise Domain Adaptation

Let $\{\mathcal{S}, \mathcal{T}\}$ a pair of source and target datasets, with $\mathcal{S} = \{X^{\mathcal{S}}, Y^{\mathcal{S}}\}$ having both images ($X^{\mathcal{S}}$) and labels ($Y^{\mathcal{S}}$) and $\mathcal{T} = \{X^{\mathcal{T}}\}$ only containing unlabeled samples $X^{\mathcal{T}}$, the traditional D2D pipeline essentially takes into account the translation capabilities of an I2I network to learn to predict $\hat{Y}^{\mathcal{T}}$ without any label $Y^{\mathcal{T}}$ used during training. With simple modifications to the traditional Unsupervised I2I Translation pipeline, one can adapt a network as UNIT [17] or MUNIT [18] to perform cross-dataset transfer learning. This was the aim of our *SIBGRAPI 2018* paper [14], which was contemporary to multiple other similar works in the literature focused on DA for segmentation of RGB images [19]–[22]. From now on, such methods for pairwise DA using image translation will be referred to as Domain-to-Domain (D2D) approaches. D2D strategies simply combine the supervised learning from a semantic segmentation network with an I2I method to perform UDA or SSDA, attaching the supervised segmentation architecture at either end of the image translation process. A typical D2D pipeline can be seen in Figure 1. This simpler pairwise pipeline is able to perform cross-dataset transfer learning between pairs of domains, even though the generalization capability of the model is still limited to the variability in $\mathcal{S}$. The D2D strategy proposed by this work is discussed in length in Chapter 3.1 of the dissertation [13].

Pairwise Image Translation for DA presented limitations that would prevent the method to maximize its label efficiency. For instance, in the task of lung segmentation in thoracic radiographs there are four labeled large scale datasets that could have labels used for training. However, D2D only allows for one of the datasets to be used as source and another unlabeled dataset to serve as target to the translation. This was the main motivation to the development of conditional dataset encoding (Section II-B) in the more recent iteration of this work: CoDAGANs.

B. Conditional Domain Adaptation

Conditional Domain Adaptation Generative Adversarial Networks (CoDAGANs) apply a relatively similar framework to D2D in order to perform UDA, SSDA and FSDA, mixing the unsupervised learning of Cycle-Consistent GANs with the supervised pixel-wise learning of deep semantic segmentation. However, two crucial distinctions between D2D and CoDAGANs must be highlighted. Firstly, in contrast to most other works in the literature, only one Encoder ($\mathcal{G}_{\mathbb{E}}$), one Decoder ($\mathcal{G}_{\mathbb{D}}$) and one Discriminator ($\mathcal{D}$) are used in the image translation process, as the cross-translation of multiple source and/or target domains is guided by conditional encoding. In other words, while D2D approaches have two translation networks ($\mathcal{G}_{\mathcal{S}\to\mathcal{T}}$ and $\mathcal{G}_{\mathcal{T}\to\mathcal{S}}$) trained separately, CoDAGANs repurpose one single translation generator $\mathcal{G}$ capable of translating from any input domain $\mathcal{D}^{(i)}$ to any other output domain $\mathcal{D}^{(j)}$. This distinction solely is responsible for two effects: considerably lessening the GPU memory requirements of the approach due to the repurposing of the same generator for all translations and allowing for multi-source and multi-target DA.

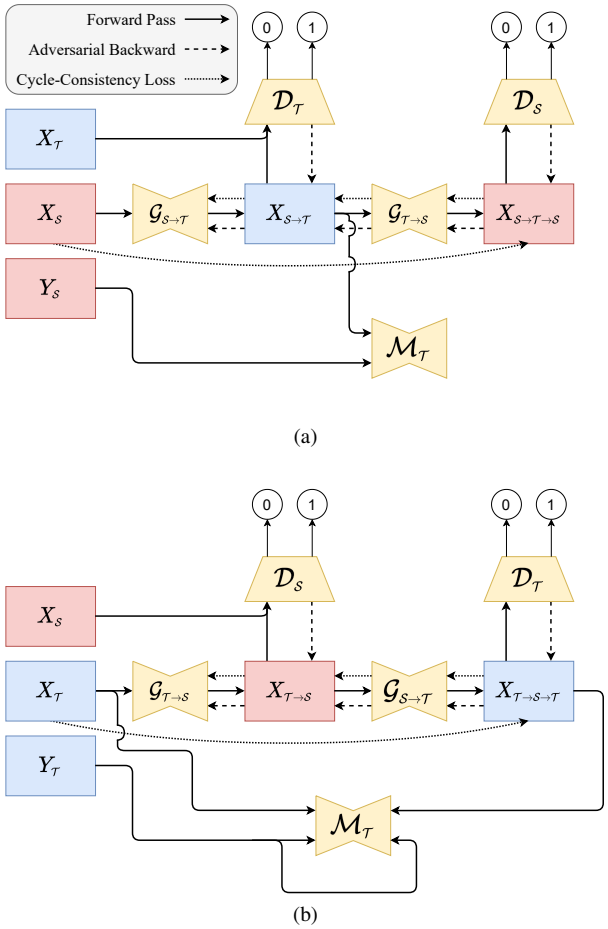

Fig. 1. Traditional pipeline of a Cycle-Consistent GAN for DA. $\mathcal{S} \to \mathcal{T} \to \mathcal{S}$ (a) and $\mathcal{T} \to \mathcal{S} \to \mathcal{T}$ (b) training are conducted simultaneously.

The second main difference between CoDAGAN and D2D regards the supervised learning model $\mathcal{M}$. In D2D, $\mathcal{M}_{\mathcal{T}}$ is capable of performing segmentation only on the target domain $\mathcal{T}$ due to its placement in the pipeline (see Figure 1). In contrast, the model $\mathcal{M}$ in CoDAGANs operates over a single space $\mathcal{I}$ trained to be isomorphic across all training domains, as depicted in Figure 2. This is achieved by splitting the generator $\mathcal{G}$ into its encoder ($\mathcal{G}_{\mathbb{E}}$) and decoder ($\mathcal{G}_{\mathbb{D}}$) halves, resulting in a bottleneck representation common to all datasets that corresponds to $\mathcal{I}$. These differences allow for drawing supervised and unsupervised knowledge from multiple datasets, depending on their label availability, as shown in Figure 3.

CoDAGAN adds a new supervised component $\mathcal{L}_{sup}$ to the loss of Unsupervised Image-to-Image Translation methods, which is already a composite loss that pairs a cycle consistency objective ($\mathcal{L}_{cyc}$) – typically L1 regression – with an adversarial component ($\mathcal{L}_{adv}$). The supervised learning component ($\mathcal{L}_{sup}$) for CoDAGANs is the default cost function for supervised classification/segmentation tasks, the Cross Entropy loss. The objective $\mathcal{L}_{CoDA}$ for CoDAGANs is, therefore, defined by:

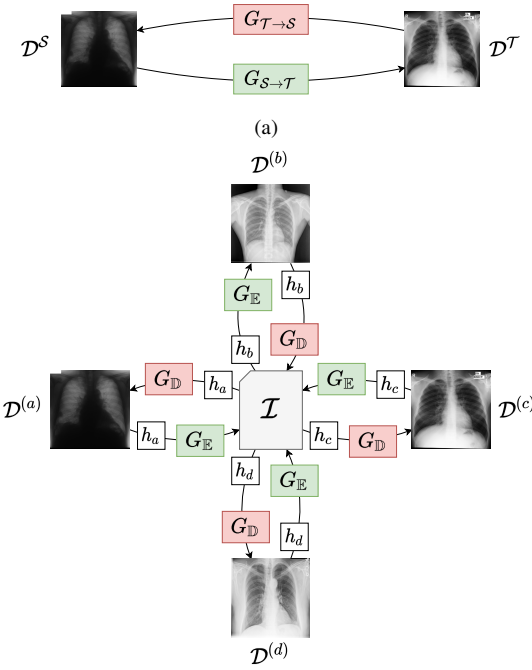

Fig. 2. Comparison between the translation processes of D2D architectures (a) and CoDAGANs (b) focusing on the generators' perspectives. While D2D uses an two generators ($\mathcal{G}_{\mathcal{S}\to\mathcal{T}}$ and $\mathcal{G}_{\mathcal{T}\to\mathcal{S}}$) for a pair of domains $\{\mathcal{S},\mathcal{T}\}$, CoDAGANs train a single generator $\mathcal{G}$ conditioned via a One-Hot-Encoding h_i code for each domain $\mathcal{D}^{(i)}$. Readers should notice that D2D does not explicitly produce an isomorphic representation $\mathcal{I}$ across domains, while this encoding naturally arises in CoDAGANs.

$$
\begin{aligned}
&\mathcal{L}_{CoDA} = \\
&\lambda_{cyc}\ [\mathcal{L}_{cyc}(X^a, X^{a\to b\to a}) + \mathcal{L}_{cyc}(X^b, X^{b\to a\to b})] + \\
&\lambda_{adv}\ [\mathcal{L}_{adv}(X^b, X^{a\to b}) + \mathcal{L}_{adv}(X^a, X^{b\to a})] + \\
&\lambda_{sup}\ [\mathcal{L}_{sup}(Y^a, \mathcal{M}(\mathcal{I}^a)) + \mathcal{L}_{sup}(Y^b, \mathcal{M}(\mathcal{I}^b)) + \\
&\qquad \mathcal{L}_{sup}(Y^a, \mathcal{M}(\mathcal{I}^{a\to b})) + \mathcal{L}_{sup}(Y^b, \mathcal{M}(\mathcal{I}^{b\to a}))], \quad (1)
\end{aligned}
$$

where a and b are two randomly drawn domains, $\mathcal{M}$ is the supervised segmentation model and λ_{cyc}, λ_{adv} and λ_{sup} are multipliers empirically set to 10, 1 and 1, respectively.

Further discussions about the details and advantages of conditional DA, network architectures, losses, training procedures and implementation details for CoDAGANs can be fully appreciated in Chapter 3.2 of the dissertation [13] or in our *IEEE Access* paper [15].

C. Leveraging Synthetic Data with CoDAGANs

The last development of conducted during the PhD. was the leverage of Conditional DA in a large domain shift scenario with synthetic data. For that, we proposed a pipeline capable of transferring knowledge from 3D CT scans to 2D CXRs using CoDAGANs [16] in order to achieve rib segmentation without requiring any per-pixel labels for ribs. We leverage the higher dimensionality of CTs to extract noisy bone segmentation masks ($Y^{\mathcal{S}}_{Bones}$) and use CoDAGANs to perform UDA from these pseudo-annotations to real CXR data. CoDAGANs are then used to compensate for the domain shift between the source dataset $\mathcal{S}$ composed of Digitally Reconstructed Radiographs (DRRs) and multiple target CXR datasets ($\mathcal{T}$) without annotations for the ribs. The full pipeline for rib segmentation from noisy labels can be seen in Figure 4.

In order to extract useful unsupervised knowledge for 2D images such as CXRs from volumetric CT-scans, the proposed pipeline for rib segmentation begins with two operations for flattening 3D volumes into 2D planes: Average Intensity Projection (AIP) and Maximum Intensity Projection (MIP) in the Posterior-Anterior (PA) axis of CT images; resulting in the images $X^{\mathcal{S}}$ and $Y^{\mathcal{S}}_{Bones}$ respectively. AIP is done by averaging all pixels in a certain location across all the PA axis, while MIP applies the max operation to this same pixel column. These flattening procedures were previously observed by the literature [23], [24] to generate useful 2D representations that could be compiled into knowledge for CXRs, especially for delineating anatomical structures such as bones and organs. The generation of DRR samples by AIP in the PA axis is delineated in green in the pipeline.

Bone masks generated by the max operation on the CT-scans ($Y^{\mathcal{S}}_{Bones}$) yield an acceptable yet noisy segmentation of the bones in the resulting DRR. Simple morphological filtering was observed to fix the noise introduced by the max operation in the labels. The computation of bone labels from the max operations in CT volumes can be seen delineated in red in Figure 4. As bones and other natural/artificial structures prominently appear in DRRs when flattening is done using the max operation, undesirable objects as scapula and humerus bones, other anatomical features, and even pacemakers are often present as False Positives in the label maps acquired for the DRRs. These artifacts are often located outside of the lung field area in the 2D projection, implying that an efficient lung segmentation could remove most of them from the label set. Thus, in order to filter all these undesirable artifacts from the labels, we first used a CoDAGAN (CoDA$_{Lungs}$) to perform UDA for lung field segmentation from labeled CXR datasets to the DRRs, as highlighted in blue in Figure 4. These networks yielded semantic prediction maps for the lung pixels and allowed for filtering the noisy labels acquired from the MIP. The resulting masks $Y^{\mathcal{S}}_{Ribs}$ are, therefore, computed according to:

$$
Y^{\mathcal{S}}_{Ribs} = Y^{\mathcal{S}}_{Bones} \ \ \& \ \ \hat{Y}^{\mathcal{S}}_{Lungs}\ , \qquad (2)
$$

where $\&$ represents the pixelwise AND operation.

More information on this project, including the DRR dataset, noisy masks for the ribs and source code can be found in the main article [16] and the project's webpage[1]. Also, details about our pipeline for knowledge transfer from DRRs can be fully appreciated in Chapter 3.3 of the dissertation [13].

III. Highlighted Results and Discussion

A total of 7 tasks were analysed: lung, heart and clavicle segmentation in CXRs; breast region and pectoral muscle seg-

[1]https://sites.google.com/view/virginiafernandes/datasets/lidc-idri-drr

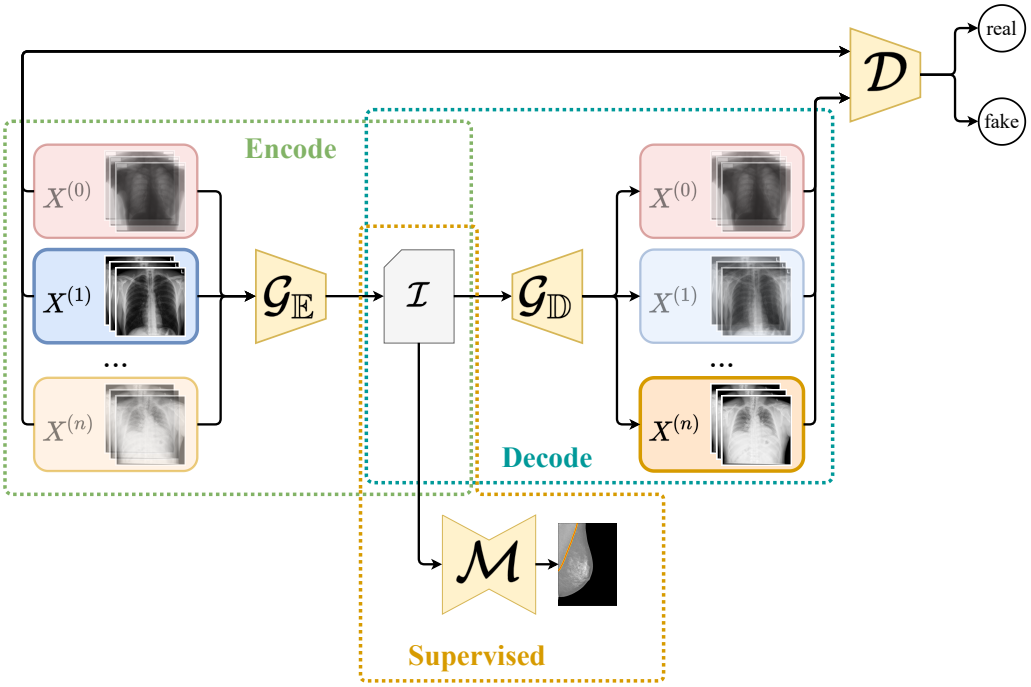

Fig. 3. Overview of multi-source and multi-target DA can be achieved via Conditional DA, only pairwise training is required. For each iteration in the DNN, one source ($\mathcal{D}^{\mathcal{S}}$, in this example, $\mathcal{D}^{(1)}$) and one target ($\mathcal{D}^{\mathcal{T}}$, in this example, $\mathcal{D}^{(n)}$) domain are sampled.

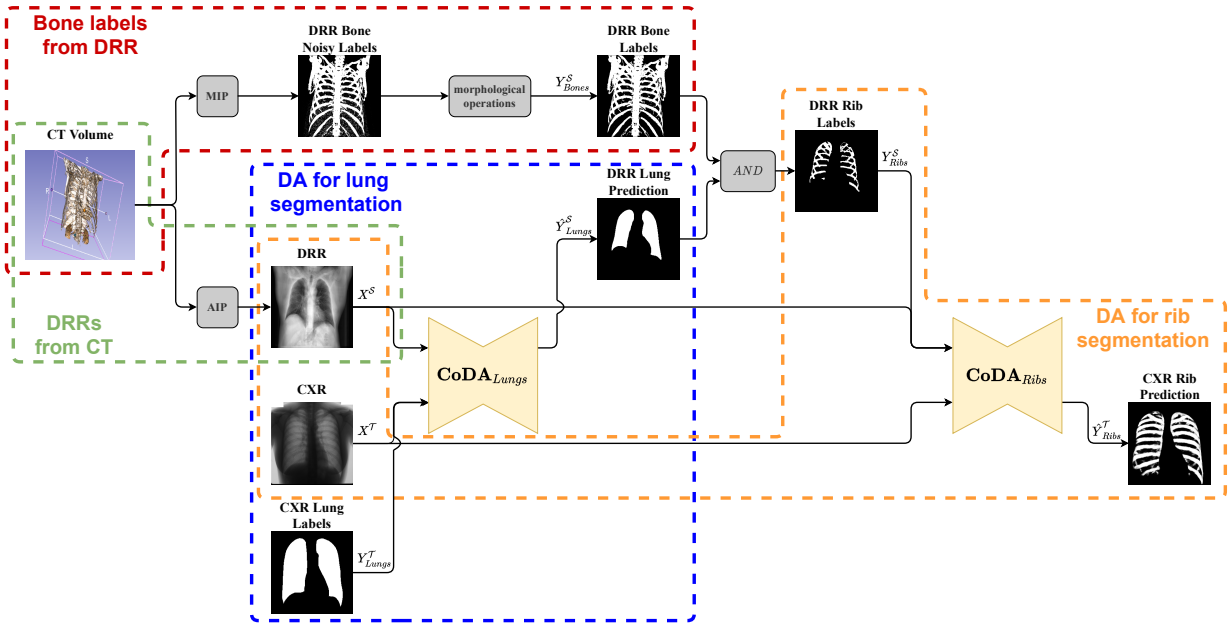

Fig. 4. Proposed rib segmentation pipeline. There are four submodules highlighted in the image: 1) the procedure for acquiring bone labels from CT-scans in red; 2) average flattening in the PA axis to produce highlighted in green; 3) Conditional DA for DRR lung segmentation from CXR labels delineated in blue; and 4) CoDAGANs for segmenting CXR ribs in orange.

mentation in MXRs; and teeth and mandible segmentation in DXRs. Additionally, we use a public CT dataset as source for computing the synthetic data for rib segmentation according to the methodology presented in Section II-C. These datasets vary widely according to label availability, number of samples and standardization in data acquisition. Here, we focus on two subsets of experiments: 1) CXR lung segmentation with CoDAGANs and the D2D baseline (Figure 5); and 2) rib segmentation for transferring knowledge from synthetic data (Figure 7). We simulate UDA and SSDA on the labeled target datasets by ignoring most or all reference annotations during training, only fixing one source dataset among in each

domain/task pair. All throughout the results, an experiment referenced to as $E_{\rho\%}$ indicates that only $\rho\%$ of the labels from the target dataset are used.

A. D2D Proof of Concept

Table I shows the results obtained by the proposed D2D method [14] compared with Fine-tuning and From Scratch training with the limited labels in the case of SSDA. It is clear that D2D significantly surpasses the effectiveness of Fine-tuning when using between 0% and 20% of the labels from the target training set. When using 50% and 100% of the target labels, Fine-tuning marginally surpassed our method, even though the difference was not statistically significant. While Fine-tuning and From Scratch yield poor performance particularly when label scarcity is more prevalent, the D2D approach is able to learn from both labeled and unlabeled samples, achieving a Jaccard score of 88% even in the fully unsupervised case. From Scratch results only reach Jaccard scores above 90% when 100% of the labels from $\mathcal{T}$ are used.

TABLE I
D2D [14] RESULTS WITH $\mathcal{S}$ = JSRT [25] AND $\mathcal{T}$ = MONTGOMERY [26] FOR MULTIPLE PERCENTAGES OF LABELED SAMPLES ON THE TARGET DATASET. STANDARD U-NETS FOR SEMANTIC SEGMENTATION TRAINED ON THE TARGET DOMAIN WITH AND WITHOUT PRETRAINING ON THE SOURCE ARE USED AS BASELINES IN THIS EARLY EXPERIMENT. BOLD VALUES INDICATE THE BEST RESULTS FOR EACH LINE.

$\mathcal{T}$ Label %	D2D	Fine-Tuning	From Scratch
$E_{0\%}$	**88.20 ± 9.80**	4.30 ± 4.13	–
$E_{5\%}$	**90.79 ± 7.05**	83.46 ± 8.60	55.10 ± 14.42
$E_{10\%}$	**89.18 ± 9.18**	83.66 ± 9.69	87.80 ± 6.78
$E_{20\%}$	**91.26 ± 7.20**	88.71 ± 8.73	89.50 ± 7.65
$E_{50\%}$	92.15 ± 5.90	**93.78 ± 5.42**	89.82 ± 4.34
$E_{100\%}$	93.18 ± 5.47	**94.81 ± 5.15**	94.16 ± 4.57

The preliminary results presented in Table I, although promising, proved to be highly unstable, as can be observed via the very large confidence intervals of D2D approaches in Figure 5. This limitation of pairwise DA fomented the development of CoDAGANs and their validation experiments presented in Section III-B.

B. Domain Adaptation Across Multiple Domains

As shown in Figure 5, UDA ($E_{0\%}$) using CoDAGANs reaches quantitative above 80% (Figure 5a) or even 90% (Figure 5b), depending on the target dataset. These values are close to the performance of a supervised lung field segmentation model trained directly on the target dataset close to FSDA, as can be appreciated in Chapter 5.2 of the dissertation [13]. In other words, CoDAGANs are able to correctly compensate for cross-dataset domain shifts without requiring any target dataset labels, effectively democratizing segmentation annotations for novel unlabeled datasets.

Additionally to the good performance on cross-dataset DA, CoDAGANs also allow for Domain Generalization due to their inherently multi-source and multi-target nature. This, again, contrasts with previous approaches that rely on pairwise D2D adaptation. A qualitative assessment of Domain Generalization from CoDAGANs and baselines can be appreciated in Figure 6. This discussion is expanded on the original proposal of CoDAGANs, published on *IEEE Access* [15].

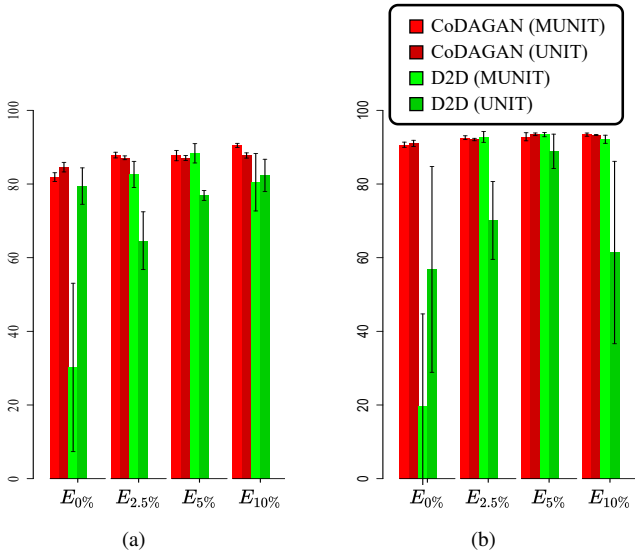

Fig. 5. Jaccard scores for two target datasets in CXR lung segmentation obtained from CoDAGANs and D2D approaches. (a) $\mathcal{T}$ = Montgomery. (b) $\mathcal{T}$ = OpenIST. Best viewed in color and with zoom.

Once Conditional DA was successfully validated, as shown in Figures 5 and 6, it was possible to try more out-of-the-box experimental setups that took advantage of CoDAGANs. In this context, we evaluated the method in a setup with synthetic data and labels for rib segmentation (Section III-C).

C. Synthetic DRR Experiments

Figure 7 shows the performance of conditional DA on Domain Generalization of rib segmentation from automatically computed labels from DRRs. It is evidenced mainly in Figure 7a that CoDAGANs achieve considerably higher AUCs than Pretrained U-Nets in large domain shift scenarios for rib segmentation. Results shown in Figure 7c stress the generalization capabilities of CoDAGANs to segment ribs in multiple target CXR domains, as the source DRRs are quite smoother than the real CXRs. This is due to the AIP procedure that flattens the 3D space into a 2D projection, resulting in a large domain shift between DRR and CXR samples. A full discussion regarding DA from synthetic data is found in Chapter 5.3 of the dissertation [13] or on the *PRL* manuscript [16].

IV. CONCLUSION AND CURRENT DEVELOPMENTS

This manuscript summarized the strategies for democratizing segmentation labels in radiology via UDA, mainly focusing on conditional DA, which can be applied to Domain Generalization proposed in Hugo Oliveira's PhD. dissertation [13]. Thorough analyses of CoDAGANs and pairwise DA baselines are shown in the main text of the dissertation [13], evidencing the advantages and possible limitations of conditional DA.

Published manuscripts [14]–[16] and the dissertation text address aspects of the proposed approaches that are not

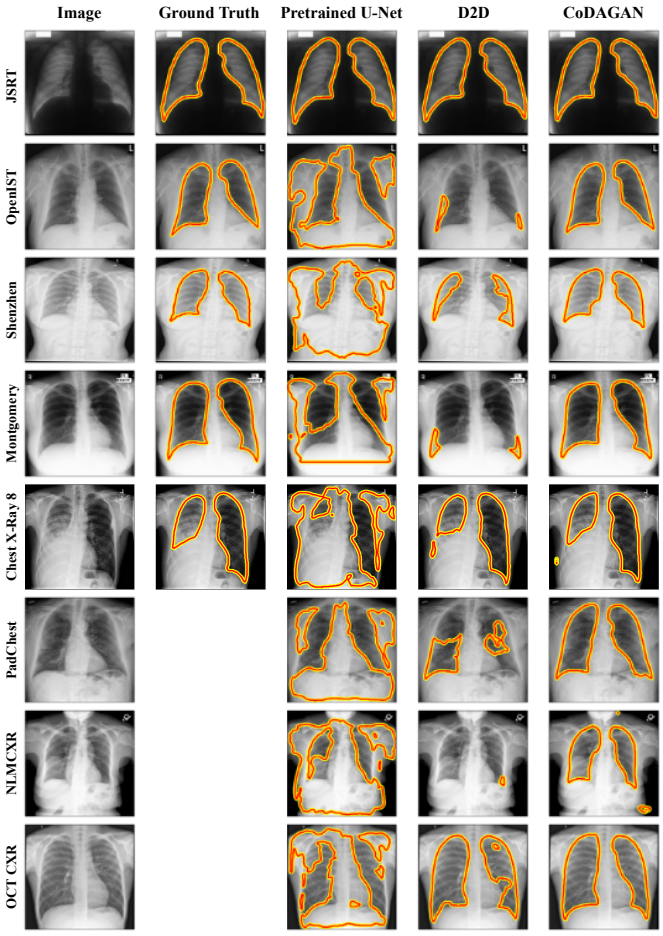

Fig. 6. Segmentation predictions of CoDAGANs and baseline methods in one source (JSRT) and seven target CXR datasets without using any labels from the target domains (UDA).

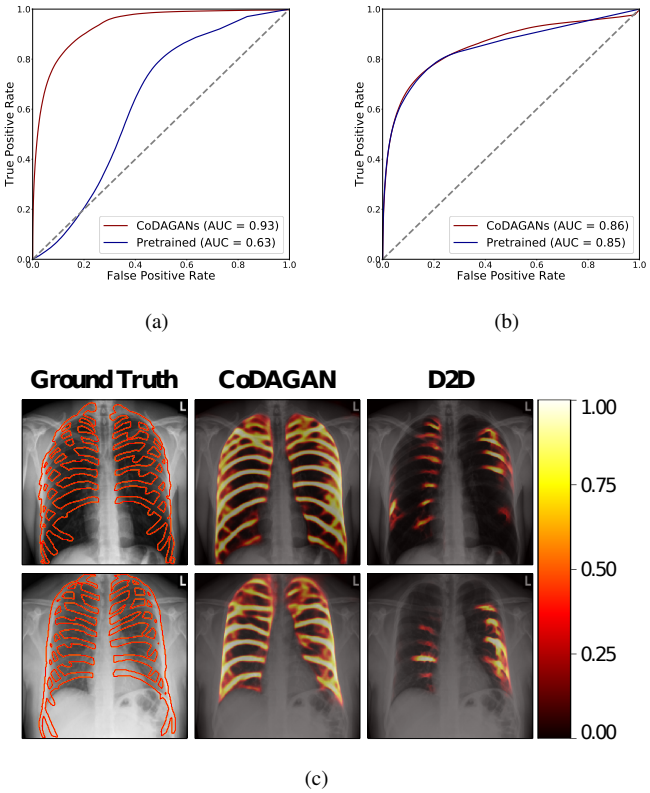

Fig. 7. ROC curves for the JSRT (a) and OpenIST (b) target datasets on rib segmentation from noisy labels using CoDAGANs and Pretrained U-Nets. We also present qualitative results for the target OpenIST dataset (c).

fully described in this manuscript. Additionally, we highlight multiple other related papers [27]–[31] that were published either in conferences or journals in the fields of Deep Learning, Domain Adaptation and/or Medical Imaging during this research project.

We highlight three derived works branching from the original CoDAGANs [15] already under development:

1) the CAD-COVID-19 project[2], approved by FAPEMIG, with multiple professors, post-doctoral researchers, graduate and undergraduate students aiming to adapt CoDAGANs to aid in the diagnosis of COVID-19 and other pulmonary illnesses;
2) undergoing results soon to be submitted to a journal that apply 2D CoDAGANs to zero-shot domain adaptation. This new research direction is an explicit effort to democratize medical annotations even when they are done in private datasets. This project will result in an MSc. dissertation of a graduate student from Universidade Federal de Minas Gerais co-advisored by Hugo Oliveira; and
3) the adaptation of CoDAGANs to generalize segmentation predictions in volumetric data, with promising results in lung and liver segmentation in CT volumes, reaching preliminary Jaccard scores of 0.90 or higher. This branch of development of 3D CoDAGANs is being conducted by an undergraduate student at Universidade Federal de Minas Gerais, resulting in her bachelors dissertation, also co-advisored by Hugo Oliveira.

We highlight that all these current research projects branching from the original CoDAGANs aim to democratize the access of labels in medical imaging. We argue that both the doctoral work described in this text and the aforementioned research branches considerably advanced the literature of Domain Generalization in medical image segmentation.

ACKNOWLEDGMENTS

The authors would like to thank CAPES for the scholarship granted all throughout the PhD., as well as CNPq (424700/2018-2) and FAPEMIG (APQ-00449-17) for partially funding this research. Additionally, we acknowledge the GPU grants from NVIDIA that provided a part of the computational resources used in our experiments.

[2]http://www.cadcovid19.dcc.ufmg.br/

REFERENCES

[1] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez, "A Survey on Deep Learning in Medical Image Analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[2] T. Zhou, S. Ruan, and S. Canu, "A Review: Deep Learning for Medical Image Segmentation Using Multi-Modality Fusion," *Array*, vol. 3, p. 100004, 2019.

[3] G. Haskins, U. Kruger, and P. Yan, "Deep Learning in Medical Image Registration: A Survey," *Machine Vision and Applications*, vol. 31, no. 1, p. 8, 2020.

[4] J. Zhang, W. Li, and P. Ogunbona, "Transfer Learning For Cross-Dataset Recognition: A Survey," 2017.

[5] Z. Cao, L. Ma, M. Long, and J. Wang, "Partial Adversarial Domain Adaptation," in *ECCV*, 2018, pp. 135–150.

[6] W. Zhang, W. Ouyang, W. Li, and D. Xu, "Collaborative and Adversarial Network for Unsupervised Domain Adaptation," in *CVPR*, 2018, pp. 3801–3809.

[7] M. Yamada, L. Sigal, and Y. Chang, "Domain Adaptation for Structured Regression," *International Journal of Computer Vision*, vol. 109, no. 1-2, pp. 126–145, 2014.

[8] Y. Wu and Q. Ji, "Constrained Deep Transfer Feature Learning and its Applications," in *CVPR*, 2016, pp. 5101–5109.

[9] P. Koniusz, Y. Tas, and F. Porikli, "Domain Adaptation by Mixture of Alignments of Second-or Higher-Order Scatter Tensors," in *CVPR*, vol. 2, 2017.

[10] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," in *CVPR*. IEEE, 2017, pp. 5967–5976.

[11] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," in *ICCV*, 2017.

[12] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets," in *NIPS*, 2014, pp. 2672–2680.

[13] H. Oliveira, "Semantic segmentation with multi-source domain adaptation for radiological images," Ph.D. dissertation, Universidade Federal de Minas Gerais, 2020.

[14] H. N. Oliveira and J. A. dos Santos, "Deep Transfer Learning for Segmentation of Anatomical Structures in Chest Radiographs," in *SIBGRAPI*. IEEE, 2018.

[15] H. N. Oliveira, E. Ferreira, and J. A. Dos Santos, "Truly Generalizable Radiograph Segmentation With Conditional Domain Adaptation," *IEEE Access*, vol. 8, pp. 84 037–84 062, 2020.

[16] H. Oliveira, V. Mota, A. M. Machado, and J. A. dos Santos, "From 3d to 2d: Transferring knowledge for rib segmentation in chest x-rays," *PRL*, vol. 140, pp. 10–17, 2020.

[17] M.-Y. Liu, T. Breuel, and J. Kautz, "Unsupervised Image-to-Image Translation Networks," in *NIPS*, 2017, pp. 700–708.

[18] X. Huang, M.-Y. Liu, S. Belongie, and J. Kautz, "Multimodal Unsupervised Image-to-Image Translation," in *ECCV*, 2018, pp. 172–189.

[19] J. Hoffman, E. Tzeng, T. Park, J.-Y. Zhu, P. Isola, K. Saenko, A. Efros, and T. Darrell, "CyCADA: Cycle-Consistent Adversarial Domain Adaptation," in *ICML*, 2018, pp. 1994–2003.

[20] Z. Murez, S. Kolouri, D. Kriegman, R. Ramamoorthi, and K. Kim, "Image to Image Translation for Domain Adaptation," in *CVPR*, 2018, pp. 4500–4509.

[21] Z. Wu, X. Han, Y.-L. Lin, M. Gokhan Uzunbas, T. Goldstein, S. Nam Lim, and L. S. Davis, "DCAN: Dual Channel-wise Alignment Networks for Unsupervised Scene Adaptation," in *ECCV*, 2018, pp. 518–534.

[22] Y. Zou, Z. Yu, B. Vijaya Kumar, and J. Wang, "Unsupervised Domain Adaptation for Semantic Segmentation via Class-Balanced Self-Training," in *ECCV*, 2018, pp. 289–305.

[23] S. Candemir, S. Jaeger, S. Antani, U. Bagci, L. R. Folio, Z. Xu, and G. Thoma, "Atlas-based Rib-Bone Detection in Chest X-Rays," *Computerized Medical Imaging and Graphics*, vol. 51, pp. 32–39, 2016.

[24] Y. Zhang, S. Miao, T. Mansi, and R. Liao, "Task Driven Generative Modeling for Unsupervised Domain Adaptation: Application to X-Ray Image Segmentation," in *MICCAI*. Springer, 2018, pp. 599–607.

[25] J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K.-i. Komatsu, M. Matsui, H. Fujita, Y. Kodera, and K. Doi, "Development of a Digital Image Database for Chest Radiographs with and without a Lung Nodule: Receiver Operating Characteristic Analysis of Radiologists' Detection of Pulmonary Nodules," *American Journal of Roentgenology*, vol. 174, no. 1, pp. 71–74, 2000.

[26] S. Jaeger, S. Candemir, S. Antani, Y.-X. J. Wáng, P.-X. Lu, and G. Thoma, "Two Public Chest X-Ray Datasets for Computer-Aided Screening of Pulmonary Diseases," *Quantitative Imaging in Medicine and Surgery*, vol. 4, no. 6, p. 475, 2014.

[27] H. N. Oliveira, J. A. Dos Santos, M. C. de Melo, T. G. do Rêgo, and L. V. Batista, "Information theory-based detection of noisy bit planes in medical images," in *SIBGRAPI*. IEEE, 2016, pp. 32–39.

[28] E. Ferreira, H. Oliveira, M. S. Alvim, and J. A. dos Santos, "A Comparative Study on Unsupervised Domain Adaptation for Coffee Crop Mapping," in *CIARP*. Springer, 2018, pp. 72–80.

[29] H. N. Oliveira, C. S. Avelar, A. M. C. Machado, A. A. Araujo, and J. A. dos Santos, "Exploring Deep-Based Approaches for Semantic Segmentation of Mammographic Images," in *CIARP*. Springer, 2018.

[30] C. C. da Silva, K. Nogueira, H. N. Oliveira, and J. A. dos Santos, "Towards open-set semantic segmentation of aerial images," in *LAGIRS*. IEEE, 2020, pp. 16–21.

[31] V. F. Mota, H. N. de Oliveira, S. Scalzo, D. Dittz, R. J. Santos, J. A. dos Santos, and A. d. A. Araújo, "From Video Pornography to Cancer Cells: A Tensor Framework for Spatiotemporal Description," *MTAP*, pp. 1–31, 2020.