Subspace-based Domain Adaptation Using Similarity Constraints for Pneumonia Diagnosis Within a Small Chest X-ray Image Dataset

Conference Paper · April 2021

DOI: 10.1109/ISBI48211.2021.9434173

8 authors, including:

Karen Sanchez
Industrial University of Santander
10 PUBLICATIONS **16** CITATIONS
SEE PROFILE

Carlos Hinojosa
Industrial University of Santander
16 PUBLICATIONS **114** CITATIONS
SEE PROFILE

Henry Arguello
Industrial University of Santander
311 PUBLICATIONS **2,038** CITATIONS
SEE PROFILE

Olivier Meyrignac
Hôpital Bicêtre (Hôpitaux Universitaires Paris-Sud)
55 PUBLICATIONS **396** CITATIONS
SEE PROFILE

SUBSPACE-BASED DOMAIN ADAPTATION USING SIMILARITY CONSTRAINTS FOR PNEUMONIA DIAGNOSIS WITHIN A SMALL CHEST X-RAY IMAGE DATASET

Karen Sanchez[1], *Carlos Hinojosa*[1], *Henry Arguello*[1], *Simon Freiss*[2], *Nicolas Sans*[3], Denis Kouamé[4], Olivier Meyrignac[5], and Adrian Basarab[4].

[1] HDSP research group, Universidad Industrial de Santander, Bucaramanga, Colombia.
[2] Department of Radiology - Toulouse Purpan University Hospital, France
[3] Department of Neuroradiology - Toulouse Purpan University Hospital, France
[4]IRIT, University of Toulouse, CNRS 5505, Université Paul Sabatier Toulouse 3, France.
[5]Department of Radiology, Bicêtre University Hospital, Assistance publique des hôpitaux de Paris, Biomaps, France.

ABSTRACT

Recent advances in deep learning have led to an accurate diagnosis of pneumonia from chest X-ray images. However, these models usually require large labeled training datasets, not always available in practice. Furthermore, combining images from different medical centers does not preserve the accuracy of the results mainly because of differences in image acquisition settings. In this work, we propose an approach aiming to overcome this challenge, consisting of a subspace-based domain adaptation technique to increase pneumonia detection accuracy using a small training dataset. This dataset is augmented with automatically selected images from a large dataset acquired in a different medical center. This is performed by computing a subspace basis of the target domain dataset on which is projected the source dataset to find the most representative images. Augmenting the training set using the proposed method allows achieving an improvement from 90.03% to 96.18% in overall accuracy using the Xception neural network.

***Index Terms*—** Domain adaptation, Deep learning, Chest X-ray, Pneumonia.

1. INTRODUCTION

Pneumonia affects 7% of the world population and is the leading cause of death worldwide in children under the age of five years, with about 1.4 million deaths per year. Among the available imaging techniques, chest X-rays are the most commonly used for screening and diagnosis of pneumonia, due to their relatively low cost and easy access. However, accurately diagnosing lung diseases on chest X-rays remains a challenging and time consuming task even for experienced radiologists [1].

Recent advances in the application of deep learning and computer vision in healthcare allowed to attain radiologist-level performance for pneumonia diagnosis from chest X-ray images [1, 2, 3]. However, the overall accuracy of deep learning models generally, and in particular for pneumonia application, strongly depends on the size of the training dataset. For instance, in [1, 2, 3], the authors used a publicly available pneumonia dataset of $5,849$ labeled images acquired in Guangzhou Women and Children's Medical Center, China [4], that allowed them to achieve an overall accuracy of 98%, 98.43% and 99.41%, respectively. Despite these very impressive results, deep learning classifiers are well-known to over-fit to a particular data domain source in medical imaging. Precisely, medical images from different clinical centers often vary in appearance due to different acquisition protocols and/or device technologies. Consequently, a deep learning model trained on a large dataset originated from one medical center does not conserve the same accuracy when tested on a dataset from another medical center, mainly because of image distribution discrepancies. For instance, a model trained on the previously mentioned dataset with an overall accuracy of 98%, only achieves 88% accuracy when classifying 176 non-public chest X-ray images acquired at a clinical center in Toulouse, France. This represents a major challenge limiting the clinical applicability of such technologies.

Domain adaptation (DA) has emerged as a learning transfer alternative to address the lack of massive amounts of labeled images and the difficulty of deep learning methods to obtain high performance when applied to small datasets (target domain) different from the one used during the training phase [5]. Precisely, DA is the area of machine learning that enables knowledge to be transferred from one source domain to a different but related target domain to increase the learning models' capability on the latter. Its applications mainly focus on natural images and, to a lesser extent, on medical applications.

This paper proposes a new subspace-based DA method to improve the performance of a pneumonia diagnosis neural

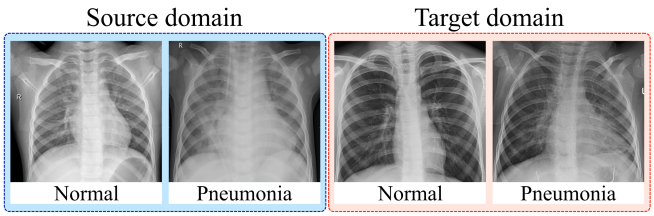

Fig. 1. Images of normal and pneumonia chest X-ray images from source and target domain datasets used in this work.

network over a small chest X-ray image dataset. Specifically, the main idea is to use automatically chosen images from the large dataset in [4], in order to improve the training procedure over a small image dataset from a different medical center (Toulouse Hospital University, France). Figure 1 shows images of normal and pneumonia cases from both source [4] and target (Toulouse Hospital University) domains. In particular, the proposed method uses principal component analysis (PCA) to obtain a subspace basis for each target domain class, *i.e.*, pneumonia and healthy. Images from the source dataset are then projected onto the target domain PCA subspaces, in order to select the ones that match the best the target dataset based on the projection errors. Finally, a Xception convolutional neural network [6] was trained using the resulting augmented dataset containing the images from the target dataset and the ones automatically selected from the source dataset. The experimental results show an important accuracy gain due to incorporating images from the source dataset, and better performances than standard transfer learning.

The remainder of the paper is organized as follows. The proposed method consisting of subspace generation, similarity functions, and data selection, is detailed in Section 2. Simulation results and conclusions are presented in Section 3, and 4, respectively.

2. PROPOSED METHOD

This section introduces the proposed subspace-based DA method to improve pneumonia diagnosis neural networks' performance within a small training dataset. Let us denote by S a set of medium to large labeled data (e.g., chest X-ray images), and by T a small labeled or unlabeled dataset lying in a given dimensional space, drawn i.i.d. according to a fixed but unknown source $\mathcal{D}_S$ and target $\mathcal{D}_T$ distributions. Furthermore, we denote by $S_P \subseteq S$, and $S_N \subseteq S$ the subsets of images labelled as pneumonia and respectively healthy, in the source domain. Similarly, we denote by $T_P \subseteq T$ and $T_N \subseteq T$ the subsets of images labeled pneumonia and healthy respectively in the target domain. We assume that m images from T, n images from S, divided in m_1 images from T_P and m_2 images from T_N, and n images from S and n_1 images from S_P, are available. Given that dataset T is assumed much larger than S, it results that $m \gg n$. The aim of this method is to propose a strategy of incorporating automatically images from T to S, in order to improve the training process of a chest X-ray image classification neural network. Section 2.1 explains how the target subspace of size d is generated, used to project the images from the source S. These projections are used within four similarity functions designed to select the most suitable images from S to augment the training dataset T further used to train a classification network.

2.1. Subspace generation

First, every source and target image is reshaped in the form of a normalized vector, *i.e.*, of zero mean and unit standard deviation. Second, using PCA, for each subset of the target domain, d eigenvectors corresponding to the d largest eigenvalues are selected. These eigenvectors are used as bases vectors of T_P and T_N subspaces, respectively denoted by $X_{T_P} \in \mathbb{R}^{D\times d}$ and $X_{T_N} \in \mathbb{R}^{D\times d}$. Note that X_{T_P} and X_{T_N} are orthonormal, *i.e.*, $X'_{T_P}X_{T_P} = I_d$ and $X'_{T_N}X_{T_N} = I_d$ where I_d is the identity matrix of size d^2. X_{T_P} and X_{T_N} are used, as explained in the next subsection, to define four shift transformations between source and target domains and to define four similarity metrics.

2.2. Similarity functions

Two approaches are considered here to define distance metrics. The first consists in projecting the source image over the target subspaces within the same class of images, *i.e.*, images $\mathbf{y}_{S_N} \in \mathbb{R}^D$, and $\mathbf{y}_{S_P} \in \mathbb{R}^D$ are projected onto the subspaces X_{T_N} and X_{T_P}, respectively. The resulting projection errors are considered as similarity functions. The second approach consists in projecting the source images onto the target subspaces within the other class of images, *i.e.*, images $\mathbf{y}_{S_N}$ and $\mathbf{y}_{S_P}$ are projected onto the subspaces X_{T_P}, and X_{T_N}, respectively. The resulting projection errors are used to form the two other similarity functions. The four projection matrices are defined as follows:

$$\mathbf{A}_{S_P|T_P} = X_{S_P}X'_{T_P}, \quad \mathbf{A}_{S_N|T_N} = X_{S_N}X'_{T_N} \tag{1}$$

$$\mathbf{B}_{S_P|T_N} = X_{S_P}X'_{T_N}, \quad \mathbf{B}_{S_N|T_P} = X_{S_N}X'_{T_P}, \tag{2}$$

where A' denotes the transpose of A.

From the four previous projections, four similarity functions are defined as the projection errors obtained by projecting images from source domain $(\mathbf{y}_{S_N}, \mathbf{y}_{S_P})$ to the target domain subspaces X_{T_N}, X_{T_P}. Specifically, the similarity functions are defined as follows:

$$sim_{X_{T_P}}(\mathbf{y}_{S_P}) = \|\mathbf{A}_{S_P|T_P}\mathbf{y}_{S_p} - \mathbf{y}_{S_P}\|_F, \tag{3}$$

$$sim_{X_{T_N}}(\mathbf{y}_{S_N}) = \|\mathbf{A}_{S_N|T_N}\mathbf{y}_{S_N} - \mathbf{y}_{S_N}\|_F, \tag{4}$$

$$cross\_sim_{X_{T_N}}(\mathbf{y}_{S_P}) = \|\mathbf{B}_{S_P|T_N}\mathbf{y}_{S_P} - \mathbf{y}_{S_P}\|_F, \tag{5}$$

$$cross\_sim_{X_{T_P}}(\mathbf{y}_{S_N}) = \|\mathbf{B}_{S_N|T_P}\mathbf{y}_{S_N} - \mathbf{y}_{S_N}\|_F, \tag{6}$$

where $||*||_F$ stands for the Frobenious norm.

2.3. Data selection

Using (3), (4), (5) and (6), four error vectors $\mathbf{E}_1 \in \mathbb{R}^{n_1}$, $\mathbf{E}_2 \in \mathbb{R}^{n_2}$, $\mathbf{Q}_1 \in \mathbb{R}^{n_1}$, and $\mathbf{Q}_2 \in \mathbb{R}^{n_2}$ are formed by projecting all images $\mathbf{y}_{S_P}$, and $\mathbf{y}_{S_N}$ from source domain S onto the target domain subspace. Vectors $\mathbf{E}_1$ and $\mathbf{E}_2$ are sorted in ascending order, and vectors $\mathbf{Q}_1$ and $\mathbf{Q}_2$ in descending order. Finally, considering the first k values from each error vector, the corresponding images from the source domain S are selected to augment the training set T, further used to train a neural network to classify chest X-ray images as healthy and pneumonia.

3. RESULTS

3.1. Domain adaptation datasets

As explained previously, the proposed method requires two datasets: one large source dataset from which a limited number of images will be extracted to augment the small target dataset. In this work, we selected the publicly available $5,849$ labeled chest X-ray images acquired in the Guangzhou Women and Children's Medical Center in China [4] as the source dataset S. Specifically, S consist of $1,583$ normal (S_N) and $4,266$ pneumonia (S_P) images. The target small dataset T consists of 573 chest X-ray images acquired at the Toulouse University Hospital, in France, labeled as healthy or pneumonia by two expert radiologists. The T dataset is divided into 275 normal and 298 pneumonia images. In the following, 70% of T was used for training and the rest for testing. Specifically, the training set of the target domain consisted of 200 normal (T_N) and 200 pneumonia (T_P) chest X-ray images.

3.2. Deep neural network architecture

The proposed DA method is not specific to a particular deep learning classifier. To illustrate its interest, we used in this work the convolutional neural network called Xception [6]. All chest X-ray images were resized to 299×299 pixels. Moreover, to accelerate the learning process, the pre-trained network weights with the ImageNet dataset available in Tensorflow was used as initialization. While the weights corresponding to the low-level feature extraction layers were not retrained, the weights from block convolutional 10 to the top layer of the network were fine-tuned, to adapt them to the specific task of pneumonia detection. The network was trained using the Adam optimizer with a learning rate of 0.0001, a batch size of 16, a dropout of 0.2 before the decision layer, and 100 epochs. All simulations were implemented in Python with the Tensorflow 2.3 and ran on an Nvidia Tesla T4 GPU provided by Google Colab.

3.3. Classification results

Figure 2 displays the results of the projection errors, $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{Q}_1$ and $\mathbf{Q}_2$, obtained after projecting images from source domain ($\mathbf{y}_{S_N}$, $\mathbf{y}_{S_P}$) to the target domain subspaces X_{T_N} and X_{T_P}. Because of the imbalance of the source dataset between the two classes, a considerably greater number of images are observed in (b) and (c) corresponding to S_P. The results are displayed such as the lowest error in (a) and (b) corresponds to the chest X-rays of S most similar to their corresponding classes in T, and such as the the lowest values in (c) and (d) correspond to the images of S_P and S_N most different from their opposite classes in the target domain T.

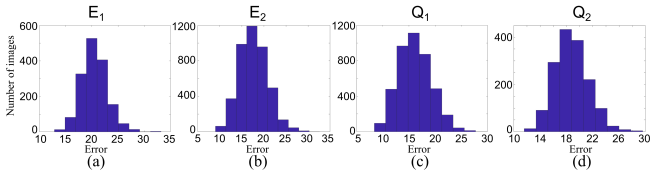

Fig. 2. Results of the error vectors $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{Q}_1$ and $\mathbf{Q}_2$.

Based on these error vectors, 0 (no augmentation), 100 and 200 images from S domain were selected to augment the target training dataset. In each case, 60% of the images were taken from the first positions in Q_1 and Q_2 vectors, and 40% from the first positions in E_1 and E_2 vectors. The classification results obtained with the proposed DA PCA-based method are compared to three different approaches. The first one consists of using standard transfer learning and all images from S, *i.e.*, to load the ImageNet pre-trained weights into the Xception network, to refine the setting of the weights from convolutional block 10 onwards with S, and once again to adjust the weights of block 10 to a fully connected layer not trained with T. The second uses no data augmentation, in which the Xception network weights previously trained by the ImageNet dataset were refined, from convolutional block 10 onwards, with only images from T. The third evaluates the classification in the case where randomly chosen images from S domain are used to augment the target training dataset. As explained previously, in all experiments, the same 400 images from T were used (alone if no data augmentation or for transfer learning, or in addition to images selected or with the proposed method from S) to train the Xception neural network to classify images in two classes, *i.e.*, healthy or pneumonia. Accuracy classification (ACC), precision (PRE), recall (REC), specificity (SPEC), F1-score (F1) and area under the ROC curve (AUC) were used to evaluate the performance of the classification. The classification results are regrouped in Table 1, with the best values shown in bold. Within the first approach, although the complete dataset S was used to refine the weights, the overall accuracy is around 88%. Within the second one, using only 400 images for fine-tuning, the overall accuracy increases to 90.03%. Furthermore, one may

Table 1. Classification results with transfer learning, no data augmentation, and data augmentation with randomly selected images, and selected with the proposed method.

Methods	Images from S	ACC	PREC	REC	SPEC	F1	AUC
Transfer Learning	5216	88.36	88.26	88.45	88.65	88.47	0.87
No data augment.	0	90.03	91.12	90.80	90.60	90.03	0.88
Random S selection	100	89.52	88.16	88.93	88.06	88.19	0.90
	200	89.98	89.65	89.45	89.35	89.68	0.89
Proposed method	100	93.25	92.56	92.68	92.27	92.54	0.92
	200	**96.18**	**95.78**	**95.86**	**95.23**	**95.96**	**0.95**

remark the different classification results depend on how the target training data set was augmented with images from S dataset. Specifically, in the case where images from S are randomly selected, the results do not improve compared to transfer learning or to no data augmentation approach. In contrast, in the case where additional images from S were selected with the proposed method, the classification accuracy increased up to 93.25% and 96.18%, with only 100 and 200 images added, respectively. The last result is remarkable given that it is obtained by training the network with only 600 images (400 from T and 200 from S), compared to roughly 90% in the case of no data augmentation or with randomly selected images. Finally, Figure 3 shows the class activation maps of the last convolutional layer of the Xception network using the proposed data selection method. These maps illustrate which pixels in the image contribute the most to the model's classification of a pneumonia class.

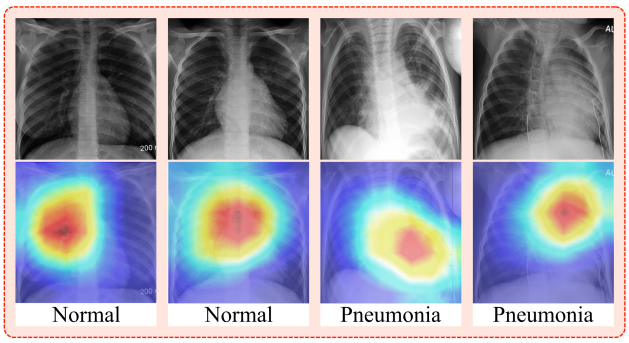

Fig. 3. Class activation maps output of four randomly chosen chest X-ray images from the test subset.

4. CONCLUSION

This work studied the problem of automatic pneumonia diagnosis with a convolutional neural network trained on small datasets. In particular, we proposed a subspace-based domain adaptation method to carefully select images from a large and publicly available dataset to augment the small target data set acquired in a different clinical center. Training the Xception CNN using the augmented data set with the proposed method allows to achieve an overall precision of 96.18%. This represents an increase in precision of 7.82% compared to transfer learning, and 6.15% compared to training with images only from the target domain. Future work will be devoted to study the interest of generative adversarial networks within this application, in order to transform the images from the source dataset to fit the appearance of images from the target dataset.

5. COMPLIANCE WITH ETHICAL STANDARDS

This research study was conducted retrospectively using human subject data made available in open access by [4]. Ethical approval was not required as confirmed by the license CC BY 4.0 attached with the open access data. The acquisition and use of Toulouse data were performed in line with the Declaration of Helsinki principles. Ethics approval was granted by the Hospital of Toulouse.

6. ACKNOWLEDGMENTS

The authors declare no conflicts of interest, not relevant financial or non-financial interests to disclose.

7. REFERENCES

[1] Tawsifur Rahman, Muhammad EH Chowdhury, Amith Khandakar, Khandaker R Islam, Khandaker F Islam, Zaid B Mahbub, Muhammad A Kadir, and Saad Kashem, "Transfer learning with deep convolutional neural network (cnn) for pneumonia detection using chest x-ray," *Applied Sciences*, vol. 10, no. 9, pp. 3233, 2020.

[2] Mohammad Farukh Hashmi, Satyarth Katiyar, Avinash G Keskar, Neeraj Dhanraj Bokde, and Zong Woo Geem, "Efficient pneumonia detection in chest xray images using deep transfer learning," *Diagnostics*, vol. 10, no. 6, pp. 417, 2020.

[3] M Toğaçar, B Ergen, Z Cömert, and F Özyurt, "A deep feature learning model for pneumonia detection applying a combination of mrmr feature selection and machine learning models," *IRBM*, vol. 41, no. 4, pp. 212–222, 2020.

[4] Daniel Kermany, Kang Zhang, and Michael Goldbaum, "Labeled optical coherence tomography (oct) and chest x-ray images for classification," *Mendeley data*, 2018.

[5] Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell, "Adversarial discriminative domain adaptation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7167–7176.

[6] François Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.