Unsupervised Patient Re-identification in Chest X-Ray Images

Fábio Rodrigues[1,2]
fabioorodrigues03@gmail.com
Sofia Pereira[1,2]
sofia.c.pereira@inesctec.pt
João Pedrosa[1,2]
joao.m.pedrosa@inesctec.pt

[1] Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciência
Porto, PT
[2] Faculdade de Engenharia da Universidade do Porto
Porto, PT

Abstract

Patient re-identification in medical imaging is essential for maintaining data integrity, reducing redundancy, and enhancing the performance of artificial intelligence (AI)-based diagnostic tools. However, in clinical practice, chest X-ray images often lack proper labeling due to privacy concerns or errors during data extraction from Picture Archiving and Communication Systems (PACS). This issue can lead to the presence of images from the same patient across different datasets, which compromises the validity of AI model evaluations. This project proposes an unsupervised learning approach to address the challenge of patient re-identification in chest X-ray images when identification information is incomplete or missing. It was developed a neural network model utilizing a contrastive learning framework to learn discriminative image representations without relying on labeled data. This aproach was compared with a supervised contrastive learning model that used patient identification (ID) labels to explicitly guide the learning process. The results demonstrated that the supervised approach outperforms the unsupervised method, achieving higher precision, recall, and area under the curve (AUC) metrics. Nonetheless, the unsupervised model demonstrated significant potential, achieving a 66.7% top-1 precision rate in identifying the correct patient match from a large test set. This performance highlights its usefulness in real-world scenarios with limited or low-quality labeled data.

1 Introduction

Early disease diagnosis is critical for initiating treatment before irreversible damage occurs. Clinical imaging methods like chest X-rays (CXR), ultrasounds, and computed tomography are essential for extracting diagnostic features. Advances in AI have led to tools that automatically detect diseases, improving diagnostic accuracy, reducing healthcare professionals' workload, fast processing and minimizing diagnostic variability.

In order to train and validate AI models, patient identification is essential to avoid patient information leakage, i.e. that images from the same patients are present in both training and validation sets, artificially boosting validation performance. However, in spite of the existence of medical image standards, this information is often erroneously registered in clinical practice or is not available once a dataset is extracted from PACS due to privacy issues. As such, in order to enable the use of these images, automatic patient re-identification tools, that analyse a dataset and identify pairs of images that likely belong to the same patient are needed.

Previous work by Packhäuser *et al*. [4] demonstrated the effectiveness of supervised learning for patient re-identification using a balanced artificial subset of data. However, their approach may not be suitable for real-world scenarios, where datasets are often unbalanced and incomplete. In this work, an unsupervised neural network for patient re-identification in chest X-ray images is explored. The results obtained are compared to those obtained using a supervised aproach.

2 Metodolody

2.1 Dataset Description

The public ChestX-ray14 [3] dataset was utilized, comprising a collection of frontal-view chest X-ray images, totaling 112,120 images from unique patients collected between 1992 and 2015. These patients were diagnosed with various conditions and there are many patients associated with multiple X-ray images. The dataset was split into a training set (approximately 90%) used during the model training phase, and a test set (approximately 10%) used for evaluating the model's performance. Special care was taken to ensure that all images from the same patient were included in the same set to prevent model bias.

2.2 Patient Reidentification Framework

The ResNet-18 architecture [2], a specific type of residual convolutional neural network, was employed for patient reidentification. The fully connected layer of ResNet-18 was modified to include an MLP projection head, consisting of a linear layer, a ReLU activation, and another linear layer leading to the network output.

To train the ResNet-18, a contrastive learning aproach as proposed by Chen *et al*. [1] was used and both unsupervised and supervised training was tested as described below. For unsupervised contrastive learning, positive and negative pairs were generated indirectly from the image set by applying transformations to the images to generate two views of each image, without using existing labels such as patient IDs. For supervised contrastive learning, patient ID labels were used to generate positive pairs from different images of the same patient. Both aproaches are represented in Figure 1. The network was trained for 100 epochs with a batch size of 256, an initial learning rate of 0.0003, and 128-dimensional feature vectors.

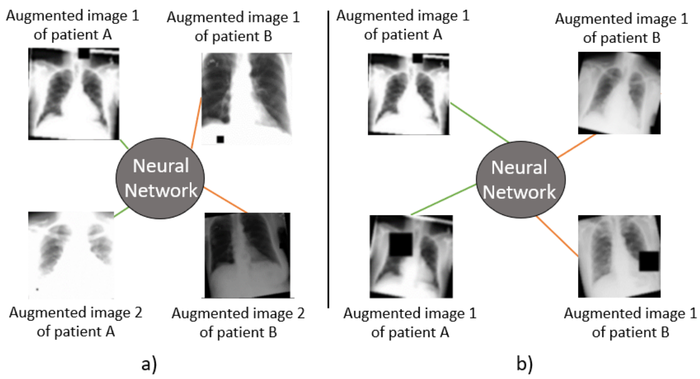

Figure 1: a) Supervised Contrastive Learning b) Unsupervised Contrastive Learning

Chest X-ray images can exhibit slight variations in angles, lighting conditions, and zoom levels. To account for these differences and improve the robustness of the neural network, several image augmentations were applied. Initially, a small number of low-intensity augmentations were used to avoid significant distortion. As the study progressed, different augmentation strategies were tested to determine their impact on model performance. Table 1 shows the different augmentation strategies tested. Figure 2 illustrates examples of these transformations. During training, multiple augmentations were applied simultaneously with randomly selected parameters to simulate diverse conditions and improve model robustness.

Table 1: Applied augmentations

Training 0	Training 1	Training 2
No augmentations (only for supervised training)	Random Cropping	Random Cropping Random Brightness Changes Rotation - 0 to 15º
Training 3	**Training 4**	**Training 5**
Random Cropping Random Brightness Changes Random Rotation - 0 to 15º Random Occlusions - 0 to 1/3	Random Cropping Random Brightness Changes Random Rotation - 0 to 15º Random Occlusions - 0 to 1/3 Gaussian Blur	Random Cropping Random Brightness Changes Random Rotation - 0 to 15º Gaussian Blur

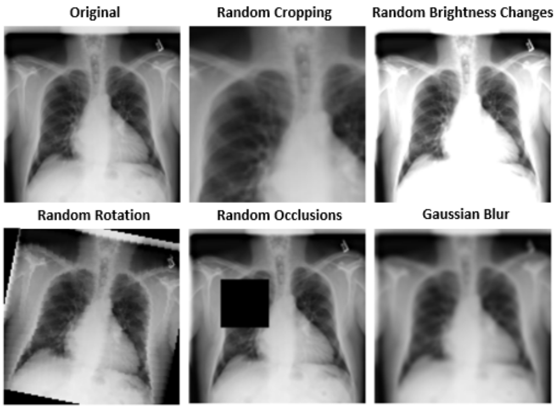

Figure 2: Examples of the augmentations applied

2.3 Evaluation

To evaluate whether the neural networks could effectively determine if two images correspond to the same patient, performance metrics were calculated on the test set. Each image was first passed through the network to generate a feature representation. Afterwards, the cosine similarity was used to compare feature representations, resulting in a similarity score indicating whether any two images belong to the same patient,

ROC (Receiver Operating Characteristic) and PR (Precision-Recall) curves were plotted, allowing for the calculation of the area under the curve (AUC) values. Similarity score was used as the separability metric, and each image was classified as positive or negative based on whether it belonged to the same patient as the image it was compared against.

Additionally, top precision and top recall values were computed. Top-X precision is the percentage of positive pairs ranked within the top X elements. Top-K precision was also calculated, where K is the total number of positive pairs per patient, ensuring consistent evaluation across patients with varying numbers of images. Top-X recall is the percentage of positive pairs in the top X elements, relative to total positive pairs available.

3 Results

The results presented in Table 2 show the values of the various performance evaluation metrics for the unsupervised learning model. Applying more transformations generally led to better results which can be attributed to the model being forced to learn more specific features from the images to identify positive pairs, as the views become more distinct. However, it was observed that removing the Random Occlusions transformation resulted in improved outcomes. Therefore, the best set of transformations were those of Training 5: Random Cropping, Random Brightness Changes, Rotation (0 to 15º), and Gaussian Blur.

The high AUC ROC values can be explained by the imbalanced nature of the problem, where the number of positive examples is significantly lower than the number of negative examples. Therefore, a high AUC ROC value can be misleading, making it essential to also analyze the AUC PR. Given that the training was unsupervised, the results obtained were quite satisfactory. When comparing a given image of a certain patient with all the other images in the test set, the most similar image corresponded to another image of the same patient 66.7% of the time. This is an excellent result, especially considering that the test set contains 10,032 images.

Table 2: Evaluation of unsupervised learning (Pr-*Precision*, Rc-*Recall*). Best results for each metric are highlighted in bold.

Training	Rc Top1000	Pr Top1	Pr TopK	AUC ROC	AUC PR
1	0.760	0.641	0.297	0.908	0.291
2	0.736	0.641	0.289	0.902	0.281
3	0.778	0.649	0.296	0.916	0.293
4	0.781	0.660	0.308	0.920	0.305
5	**0.785**	**0.667**	**0.312**	**0.921**	**0.311**

In Table 3, the evaluation metrics for the model using a supervised training approach are presented. Since training 5 achieved the best results in unsupervised training, only this set of augmentations and training 0 were tested. The application of transformations improved model performance, likely due to increased variability among image pairs, which forced the model to identify more specific characteristics. The results achieved were quite high; notably, for the training with transformations, 93.5% of the time, the most similar image to the one being compared belongs to the same patient. Other metrics also showed high values, with special emphasis on the AUC of the PR curve, crucial in imbalanced problems. Although the results obtained by Packhäuser *et al.* [4] were superior, the two studies are not directly comparable, as they used a balanced artificial subset, which was not the case in this project.

Table 3: Evaluation of supervised learning. Best results for each metric are highlighted in bold.

Training	Rc Top1000	Pr Top1	Pr TopK	AUC ROC	AUC PR
0	0.984	0.928	0.735	0.99315	0.77438
5	**0.987**	**0.935**	**0.739**	**0.99444**	**0.78113**

The results obtained with the supervised approach were better than those obtained with the unsupervised approach, achieving significantly higher values in precision, recall, and AUC, indicating a superior model performance in solving the presented problem. However, it is important to note that in scenarios with limited data quantity and quality, where the data may be improperly labeled, the unsupervised approach provides a viable solution, enabling the re-identification of patients from X-ray images, which would be impossible using a supervised approach.

4 Conclusion

This project explored and compared supervised and unsupervised learning approaches for patient re-identification using X-ray images, a significant challenge due to frequent mislabeling in hospital settings. The supervised approach, which relies on large amounts of high-quality labeled data, demonstrated superior performance in this study, achieving significantly higher precision, recall, and AUC metrics. However, this approach is limited by its dependence on labeled data, as manual labeling is time-consuming and prone to errors. Conversely, the unsupervised approach, which does not require labeled data, is particularly useful when accurate labeling is difficult or impractical. It allows learning discriminative image representations to identify similar images from the same patient, albeit with slightly lower performance compared to the supervised approach.

Future work could refine image transformations and extend training epochs to improve unsupervised results. Testing on different datasets may reveal cases where unsupervised learning surpasses supervised. Despite data limitations, this project shows that effective patient re-identification is possible with an unsupervised approach, offering flexibility for real-world applications with limited or low-quality data.

Acknowledgments

This work is financed by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia, within project LA/P/0063/2020. João Pedrosa and Sofia Cardoso Pereira are supported by Fundação para a Ciência e Tecnologia (FCT) through application reference 2022.06138.CEECIND and grant 2020.10169.BD respectively.

References

[1] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations, 2020. doi: 10.48550/arXiv.2002.05709.

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. volume 9908, pages 630–645, 10 2016. ISBN 978-3-319-46492-3. doi: 10.1007/978-3-319-46493-0_38.

[3] Yongcheng Liu, Bin Fan, Lingfeng Wang, Jun Bai, Shiming Xiang, and Chunhong Pan. Semantic labeling in very high resolution images via a self-cascaded convolutional neural network. *ISPRS Journal of Photogrammetry and Remote Sensing*, page 13, 2018. doi:10.1016/j.isprsjprs.2017.12.007.

[4] K. Packhäuser, S. Gündel, N. Münster, et al. Deep learning-based patient re-identification is able to exploit the biometric nature of medical chest x-ray data. *Scientific Reports*, 12:14851, 2022. doi:10.1038/s41598-022-19045-3.