

Machine Learning based sampling of X-Ray images for a computer-aided detection of Tuberculosis*

*

Fernando Ferreira
*Signal Processing Laboratory
COPPE/POLI*

Federal University of Rio de Janeiro
Rio de Janeiro, Brazil
fferreira@lps.ufrj.br

Philipp Gaspar
*Signal Processing Laboratory
COPPE/POLI*

Federal University of Rio de Janeiro
Rio de Janeiro, Brazil
philipp.gaspar@lps.ufrj.br

Lukas Müller de Oliveira
*Signal Processing Laboratory
COPPE/POLI*

Federal University of Rio de Janeiro
Rio de Janeiro, Brazil
lukasmullerdeoliveira@poli.ufrj.br

Rodrigo Torres
*Signal Processing Laboratory
COPPE/POLI*

Federal University of Rio de Janeiro
Rio de Janeiro, Brazil
torres@lps.ufrj.br

Micael Veríssimo de Araújo
*Signal Processing Laboratory
COPPE/POLI*

Federal University of Rio de Janeiro
Rio de Janeiro, Brazil
micael.verissimo@lps.ufrj.br

Carlos Eduardo Covas
*Signal Processing Laboratory
COPPE/POLI*

Federal University of Rio de Janeiro
Rio de Janeiro, Brazil
kaducovas@gmail.com

Mayara Bastos
*Research Institute of the
McGill University Health Centre*
Quebec, Canada
atrajman@gmail.com

Anete Trajman
*Faculty of Medicine
Federal University of Rio de Janeiro*
Rio de Janeiro, Brazil
atrajman@gmail.com

José Manoel de Seixas
*Signal Processing Laboratory
COPPE/POLI*
Federal University of Rio de Janeiro
Rio de Janeiro, Brazil
seixas@lps.ufrj.br

Abstract—Computer Aided Detection software relies on an annotated data set of X-rays to be developed. The annotation task requires extensive know-how and it is very time-consuming. This work presents a sampling method to select the most relevant images which will be annotated for the development of Tuberculosis screening platform based on machine learning algorithms. The sampling task optimizes the annotation process by reducing the number of images to be analyzed without compromising the diversity and the significance power of the images in the dataset. In this context, the image relevance is based on similarity and dissimilarity measurements. The experiment consisted in a deep learning feature engineering step, followed by topological analysis based on Self-Organizing Map and K-Means.

Index Terms—Deep Learning, CNN, SOM, Clustering, CAD

I. INTRODUCTION

In recent years, Computer Aided Detection (CAD) software packages have been used to support radiologists during clinical practice in detecting potential abnormalities on diagnostic exams [1].

In some applications, the CAD software is also used to produce a numerical score indicating the likelihood that a chest X-ray (CXR) image represents a specific disease process [2]. Recently, the World Health Organization (WHO)

has recommended for the first time that CAD applications could be used as an alternative to human interpretation of digital CXR for screening and triage of subjects with Tuberculosis (TB) [3].

CAD applications based on Artificial Intelligence (AI) algorithms rely on an annotated database of images that will be used to train those algorithms [4]. However, in most cases, the X-ray images are not properly annotated, while the process to generate new annotations requires extensive domain knowledge to catalog and index the images manually.

Therefore, the process of sampling images to enter the new annotation phase must follow a selection criterion that prioritizes high diversity images to maximize the data information. In this context, this work presents a sampling method to select images based on similarity and dissimilarity measurements to decrease redundancy and maximizing representativeness in the CXR images selected for the annotation process.

In the first part of the analysis, we trained a Convolutional Neural Network (CNN) [5] to perform a binary classification of CXR images into TB or not TB. After the training phase, we discard the last layer of the network in order to get its embeddings, found during the training process, of the second to last layer.

Those embeddings are the fed into a Self Organizing Map (SOM) [6] that will be responsible for projecting those samples

The authors would like to thank CAPES and CNPq (Brazil) for their support during this work.

into a 2D grip map. The final part of the analysis consists of a clustering algorithm used to sampling CXR images based on similarity and dissimilarity criteria.

II. BACKGROUND MATERIAL

This section overviews the machine learning techniques for developing the aimed data sampling method for picking chest x-ray images. The developed method, shown in Figure 1, consists in the following 3-steps pipeline:

- extract a set of discriminative features in order to exploit different aspects of x-ray images.
- project data in a non-linear way for both reducing dimensionality and representing the data set in fewer elements, preserving the topology of the observations.
- Group observed samples (or projections) into clusters according to similarity measures.

A. Feature Extraction

Feature extraction is one of the most important steps of medical image processing which requires extensive domain knowledge [7]. For many unsupervised learning techniques, replacing raw images with features extracted by convolutional neural networks (CNN) leads to better results as this kind of architecture can efficiently produce a set of discriminative features without any expert guidance [7], [8].

1) *Convolutional Neural Networks (CNN)*: CNN is a class of deep neural network architecture widely used in image analysis and it consists of an input layer, hidden layers and an output layer. This specialized type of neural network performs convolution operations on images due the way weights are shared throughout the network. After passing through a convolutional layer, the image becomes abstracted to a feature map, also called an activation map. At each layer, the input image is convoluted with a set of K kernel $\mathcal{W} = \{\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_k\}$ and added biases $\mathcal{B} = b_1, \dots, b_2$, each generating a new feature map \mathcal{X}_i . This features are subjected to an element wise non-linear transform $\sigma(\cdot)$ [9]. The model does not need to learn separate detectors for the same object occurring at different locations of a image. For this reason, CNN provides the flexibility of extracting intrinsic and discriminative features from X-ray images [7].

B. Data projection methods

Projection methods are mainly meant for reducing the data dimensionality by representing the observation in a subspace (lower number of directions) that describes the data structure in a concise way and features are preserved faithfully [10]. Principal Component Analysis (PCA) [11] is widely used due to its simplicity and the possibility to reduce the data dimensionality in a controlled way: each extracted component (dimension) is ranked by how much data variance it represents.

In summary, PCA searches in an unsupervised way for uncorrelated directions. The eigenvectors of the data covariance matrix represent the principal directions for which the variance of data has the maximum values. The corresponding eigenvalues

define how much energy the component retains. In Equation 1, if the principal directions are represented by u_j , we can define a linear orthogonal transformation of data x as:

$$a_j = u_j^T x = x^T u_j, j = 0, \dots, n - 1 \quad (1)$$

where a_j is the data projection onto a principal component. In order to reduce the original data dimension, one may use only the major n projections, discarding the projections of smaller variance. It is important to notice that PCA only considers the second-order statistics. Here, the PCA is used among with the Self-Organizing Map (SOM), a computational data analysis method which produces nonlinear mappings of data to lower dimensions [12].

1) *The Self-Organizing Map*: The SOM algorithm consists of an unsupervised trained neural network with the Kohonen layer [6]. The algorithm uses a similarity measure (usually based on the Euclidean distance) between data samples to embed a low dimensional space (typically two dimensions) into the original data. The 2D grid embedded into the data space is called the code-book. The result is a topological organization of the input data where their most relevant aspects are preserved, revealing hidden structures throughout the non-linear mapping.

The SOM computes the Euclidean distance of the input vector to each neuron, and find the winning neuron, denoted neuron, using the nearest-neighbor rule (competition phase). The winning node is called the excitation center and it determines a neighborhood of excited nodes (cooperation phase). All the input vectors that are closest to the winning neuron adjust the weights in order to strengthen its response (adaptation phase) [13].

The input data space is fully connected to each neuron from the Kohonen layer, and the weights are computed iteratively using:

$$w_j(n+1) = w_j(n) + \eta(n)h_{ij}(n)(x(n) - w_j(n)) \quad (2)$$

where $\eta(n)$ stands for the learning rate and $h_{ij}(n)$ is the neighborhood kernel. In this application, we used a Gaussian kernel:

$$h_{ij}(n) = \exp\left(\frac{-d_{ij}^2}{2\sigma^2(n)}\right) \quad (3)$$

where d_{ij}^2 denotes the similarity measure, and $\sigma(n)$ is the monotonically decreasing width of the kernel. In the end, after the training procedure, the network outputs are calculated for each neuron using the Equation 4.

$$u_i = x^T w_i \quad (4)$$

where the vector w_i represents the weight that connects the input data to the neuron i . Hence, the vectors $w_{i,i=1,m}$ form the SOM code-book.

The obtained map offers powerful tools for data exploration due its bi-dimensional (on most of cases) nature, which allows to identify cluster borders, projection directions and possible dependencies between variables [10].

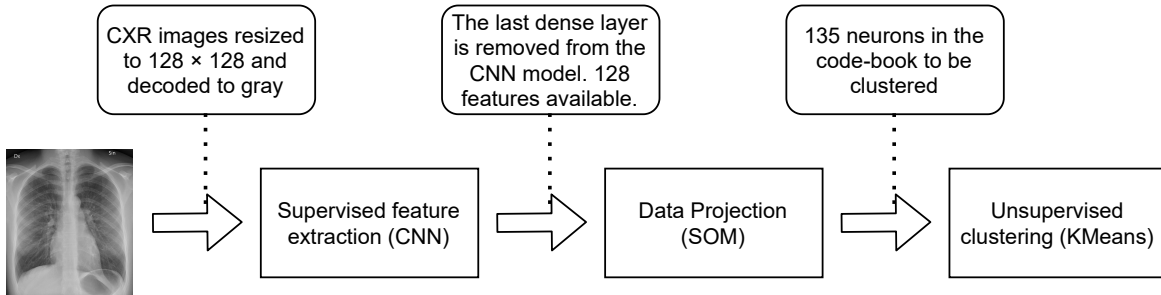


Figure 1: Proposed unsupervised sampling pipeline.

C. Data clustering

Clustering methods are widely used for grouping observations into smaller subsets (clusters) defined by a similarity measure. In other words, samples belonging to a cluster should be as much as similar between each other and quite differ from observations from other clusters.

The most common clustering technique is the k -Means algorithm, which is described in [14]. The SOM can be viewed as a clustering algorithm which produces a set of clusters organized on a regular grid [13].

D. Clustering consistency measures

The k -Means clustering procedure does not provide any measure of how consistent the estimated clusters. Hence, the silhouette index [15] provides a measurement of consistency of the estimated clusters through the following equation:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

where $a(i)$ is the average dissimilarity between the i -th sample from a given cluster and the other samples in the same cluster and $b(i)$ is the lowest average dissimilarity of the i -th sample to the remaining clusters. The above expression can be rewritten as:

$$s(i) = \begin{cases} 1 - a(i)/b(i), & \text{if } a(i) < b(i) \\ 0, & \text{if } a(i) = b(i) \\ a(i)/b(i) - 1, & \text{if } a(i) > b(i) \end{cases}$$

It can be seen that $-1 < s(i) < 1$. If the i -th sample is well-represented in its cluster, then $s(i)$ is closer to 1. Otherwise, $s(i)$ is closer to -1 . Data samples with silhouette values close to zero are in the border of two or more clusters. The average silhouette index, considering all data samples, is used to measure the performance of the clustering configuration.

III. EXPERIMENT SETUP

This section describes the proposed pipeline to deliver the most representative CXR images from an extensive database, which will be used in a later annotation task. This pipeline is designed to select images based on similarity and dissimilarity measurements.

Firstly, the CXR images are resized to 128×128 and decoded to gray-scale single channel, where each value varies

between 0.0 (black) and 1.0 (white). The resulting vectors are normalized by the standard deviation. Then, a deep learning-based feature engineering is performed using a CNN architecture.

The CNN setup uses a 6-layer network structure, which consists of four convolutional layers and two fully-connected layers, as shown by the schematic diagram in Figure 2. Dropout and ReLU are deployed to address over-fitting and convergence issues [16]. The model was trained using Adam optimization algorithm [17] with 10 epochs and batch size equals to 64.

The available dataset was divided into 10 separated partitions using a stratified cross-validation split [18] procedure assuring that we have well-balanced classes in each train and validation subset. Also, an early stopping strategy was used by monitoring the validation set using the SP (Equation 5) index [19] in each iteration.

$$SP = \sqrt{\sqrt{P \times (1 - F)} \times \left(\frac{P + (1 - F)}{2} \right)} \quad (5)$$

where P denotes the true positive rate and F the false positive samples.

After the training process, the last dense layer is removed from the CNN model and the CXR images are fed forward through the network. The representations found by this procedure will have size $n_{\text{samples}} \times 128$, since the dense layer has 128 neurons, and will be used in subsequent models: PCA and SOM.

The SOM model built for exploring the nonlinear statistical relationships among the extracted features was optimized to a 15×9 (135 neurons) grid using a rectangular topology. The number of neurons is calculated from the number of data points of the training dataset using the Equation 6, as proposed by [20]:

$$N \approx \sqrt{5M} \quad (6)$$

where N represents the number of neurons, rounded to the nearest integer, and M is the number of CXR images.

The SOM network weights were initialized by spanning the first two principal components. This initialization does not depend on random processes, making the training process much faster [20]. As training parameters for building the map, the learning rate was set to 0.5, the spread of neighborhood

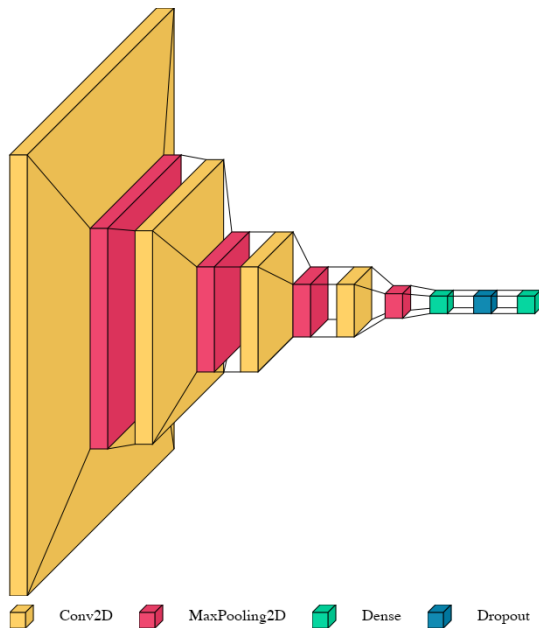


Figure 2: CNN architecture proposed using 6 layers: 1) Conv2D (32 filters), kernel (3x3), ReLU, Max Pooling (2x2) 2) Conv2D (64 filters), kernel (3x3), ReLU, Max Pooling (2x2) 3 and 4) Conv2D (128 filters), kernel (3x3), ReLU, Padding Same, Max Pooling (2x2) 5) Fully-connected (128 units), ReLU, DropOut (0.5) 6) Fully-connected (1 unit), Sigmoid

function was equal to 2 and the maximum number of 1000 epochs. Euclidean distance was used as activation function and the Gaussian function was used to weight the neighborhood of a position in the map.

Later, the SOM code-book was clustered using the K-Means algorithm varying the number of centroids from 2 to 20 centroids and the results were compared using the silhouette index. At this point, we are not interested on finding the very best clustering model (with higher silhouette values). In fact, the clustering configurations are exploited in order to understand the intrinsic structures evidenced by performed non-linear mapping. This task aims to select samples according to similarity and dissimilarity parameters and the silhouette index will help to understand how well-fitted the given model is and how well the clustered samples are conformed inside it. In other words, how similar a given sample is to its neighbors. The next section will show the obtained analytic results.

IV. DATASET

Some of the difficulties in building CAD software for lung diseases are due to the small datasets ($<10^3$) publicly available, which also presents some additional artifacts created by radiologists, like text and symbols written in different parts of the image. Nevertheless, this work used training data, provided by the U.S. National Library of Medicine [21], containing normal and abnormal chest X-rays with manifestations of TB, also including associated radiologist readings and initial classification labels used to train the CNN III.

Model	SP	PD (NOT TB)	FA (NOT TB)	PD (TB)	FA (TB)
1	81,70%	88,04%	24,70%	75,30%	11,96%
2	81,44%	84,36%	21,73%	78,27%	15,64%
3	80,64%	87,12%	25,89%	74,11%	12,88%
4	81,44%	84,66%	22,02%	77,98%	15,34%
5	80,52%	84,97%	24,11%	75,89%	15,03%
6	80,19%	86,20%	25,89%	74,11%	13,80%
7	82,64%	87,12%	22,02%	77,98%	12,88%
8	83,51%	89,88%	22,92%	77,08%	10,12%
9	84,61%	87,12%	18,15%	81,85%	12,88%
10	81,12%	86,20%	24,11%	75,89%	13,80%

Table I: Results after passing all data (train and valid) throughout the trained CNN.

The used data set contains 662 frontal chest X-rays, of which 326 are normal cases and 336 are cases with manifestations of TB, including pediatric X-rays (anteroposterior). The chest X-rays are from outpatient clinics and were collected from the Shenzhen No.3 People’s Hospital, Shenzhen, China. The images are provided in PNG format as 12-bit gray level images and their size is approximately 3000×3000 pixels.

V. RESULTS AND DISCUSSION

In order to obtain the performance measurements of the CNN, the entire dataset (train and validation) is passed throughout the models trained in each phase of the cross-validation procedure III. Table I shows the results of SP index, Probability Detection (PD) and False Alarm (FA) for each class. The CNN trained in the partition #9 was chosen to be put into production due to its better results in the SP index compared with the other cross validation folds.

The two first principal components were computed and together they represented 80 % of the PCA explained variance and the ratio between them is approximately 3.6. As said before, the two components are used for initializing the SOM neuron weights.

Then, the features are projected over the SOM map and the result can be seen as the graphical representation displayed in Figure 3. Each neuron is represented by a circle (neurons with no hits are hidden from the graphic) which shows the ratio between CXR images of patients with and without TB that hit the neuron. Fully filled circles represent neurons hit by vectors extracted from CXR of TB patients. On the other hand, hollow circles demonstrate neurons reached only by vectors from CXR images with no TB detected.

The graph clearly shows a cross-sectional band (from lower left to upper right) which divides two regions, one with mostly CXR images of TB-infected patients and other with non-TB patients. Yet, both regions presented neurons hit by both kind of patients. These “mixed neurons” might represent images with high degree of dissimilarity, as they were excited by images from distinct classes; while neurons hit only one type of patient might represent a degree of similarity, specially when located in region with similar neighbors.

At this point, K-Means is applied in order to surface the intrinsic relations. Figure 4 shows the silhouette index values for several clustering initialization. The number of centroids were varied from 2 until 20 and each setup was initialized

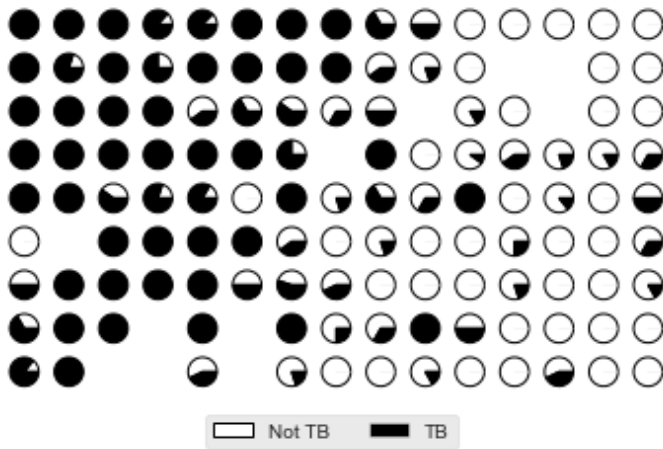


Figure 3: Graphical representation of the SOM topological map. Each neuron is represented by a circle and colored according to the ratio of “not TB” and “TB” hits.

100 times. In the box plot, the line within the box marks the median. The boundaries of the box indicate the 25th and 75th percentiles. Whiskers above and below the box indicate the 10th and 90th percentiles. Points above and below the whiskers indicate outliers. The best fitted models have 2 centroids (as it could be foreseen from the cross-sectional band analysis described earlier) with silhouette index slighted over 0.50, with minimal variance. Models with 3 and 4 centroids present a silhouette over 0.42 for all initializations, with picks of 0.48 and 0.47, respectively. These three configurations were considered for the exploitation analysis since from 5 clusters the silhouette value drops to below 0.40 on average, with little variation as we increase the number of centroids.

At this stage, the idea is to exploit the intrinsic relations uncovered by all nonlinear methods applied in sequence. For this purpose, the following characteristics will be observed: the samples best conformed within a cluster (indicating similarity); samples with higher silhouette values, but placed in a cluster with neighbors labeled differently (indicative of dissimilarity); and samples located at the borders of the clusters, in transition areas with little activated neurons (indicative of dissimilarity). This analysis is performed on the best models with 2, 3 and 4 centroids. In this way, samples were chosen until completing 50 samples. Figure 5 displays four examples sampled after the proposed procedure.

Figure 6 shows the SOM Maps for different clustering settings and each one allowed the detailed analysis according to different aspects, describe as follows:

- A) Figure 6a displays 2 clusters and TB patients are pointed by hollow circles varying their sizes according the silhouette value. Bigger the circle, better the CXR is fitted inside the cluster. The predominance of circles in the cluster on the left of the map is visible, mostly with values between 0.4 and 0.7 (with the exception of objects located in the border area between the clusters). This region is mostly composed of neurons that were excited by images of

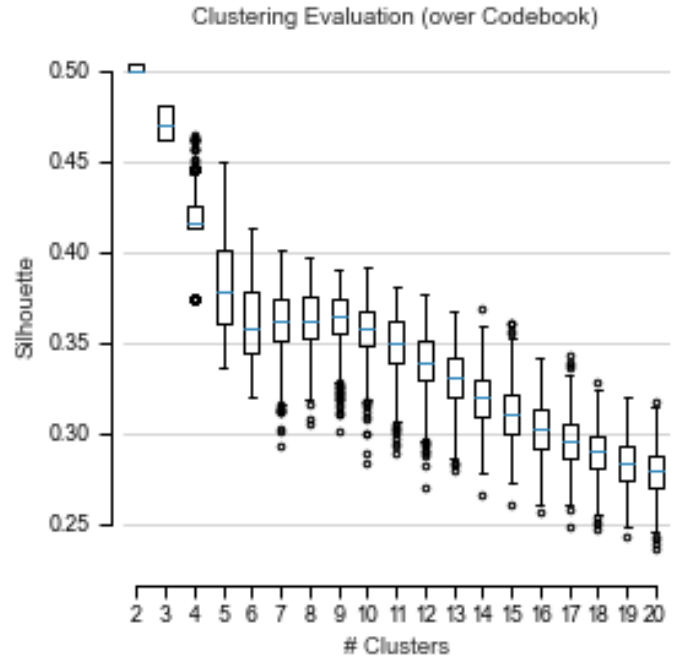


Figure 4: K-Means clustering evaluation over the SOM codebook. Each clustering setup was initialized 100 times. In the box plots, the silhouette index variance is shown for each clustering model.

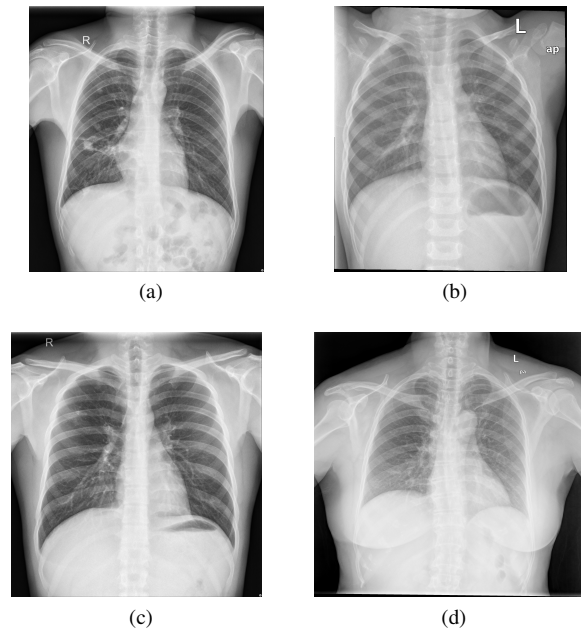
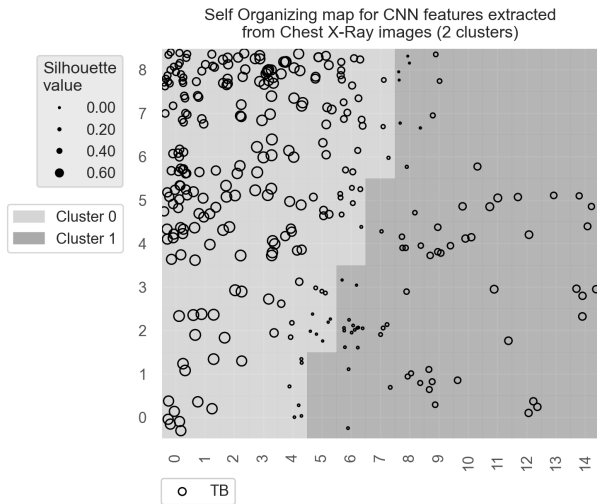
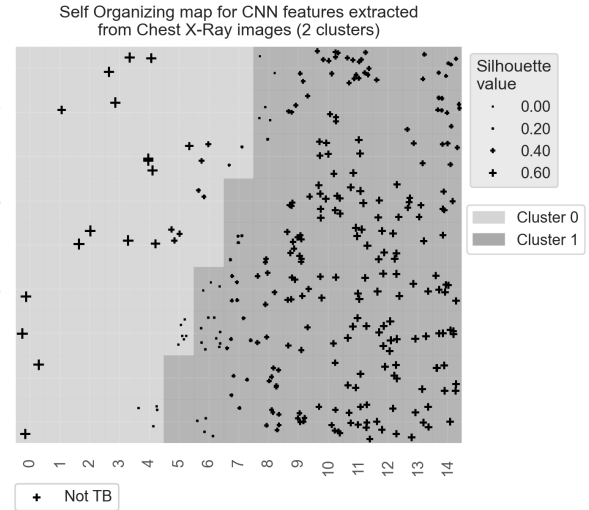


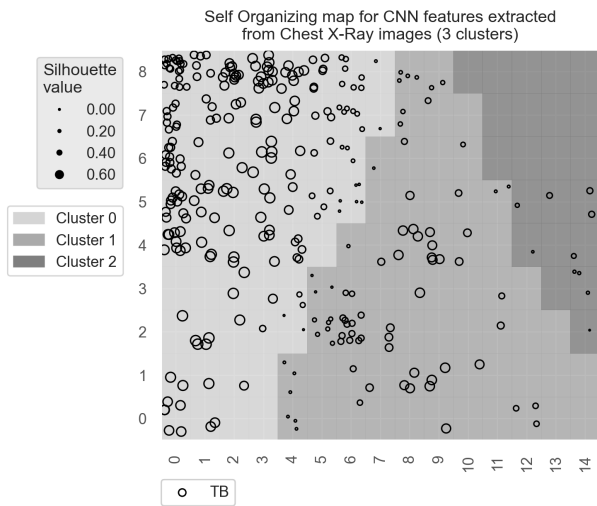
Figure 5: Examples of Chest X-Ray images selected by similarity and dissimilarity patterns.



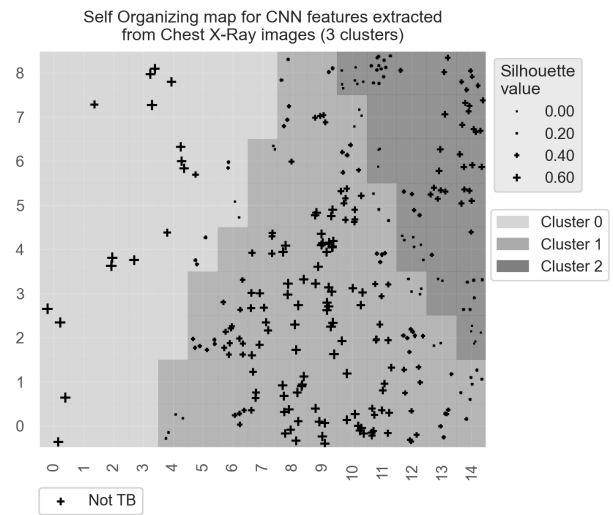
(a) Two clusters - TB patients



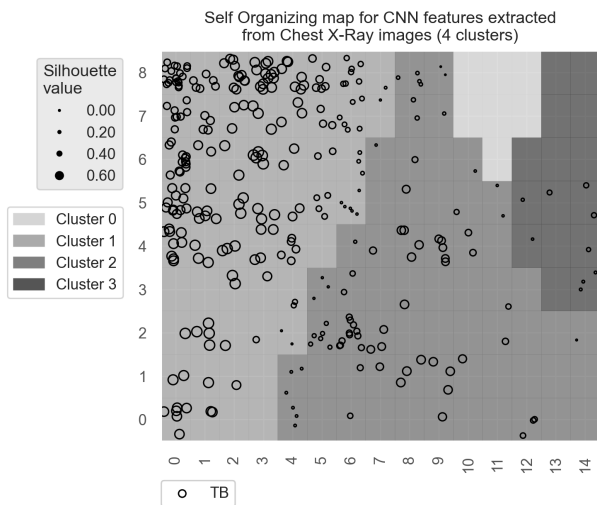
(b) Two clusters - Non-TB patients



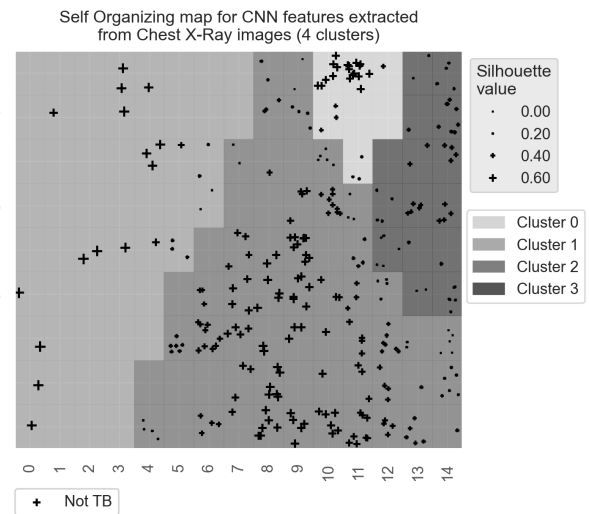
(c) Three clusters - TB patients



(d) Three clusters - Non-TB patients



(e) Four clusters - TB patients



(f) Four clusters - Non-TB patients

Figure 6: Self-Organizing map for the features extracted using the CNN network. The images shows how TB and non-TB patients are projected onto the map and how different clustering models divide the SOM code-book.

patients diagnosed with tuberculosis. Therefore, objects (20 images) with higher silhouette values with this class can be chosen as good representatives for the criterion of similarity with TB. Figure 5a displays an example of this group of images.

- B) Figure 6b also displays 2 clusters, but it shows only non-TB patients (cross marker sized proportional to the sample's silhouette values). As expected, the crosses are mostly located in the right side of the map. There are some objects located on the Cluster 0 with high silhouette values. These are interesting cases, as it would be expected that only images of TB patients were in the center of this cluster. The cluster result shows there are features that turn non-TB patients similar to infected ones. Hence, these images (like the one displayed by Figure 5b) are good examples of dissimilarity for non-TB images and, for this reason, will be picked for the final sampling. (20 images)
- C) Figures 6c and 6d exhibit the SOM map split onto three clusters. It is possible to notice neurons presenting well fitted samples for both TB and non-TB patients, specially in the center part of the *Cluster*1. Those samples are interesting for representing images that don't share similarity with the target classes. This way, they were picked as dissimilarity examples. (40 images)
- D) Figure 6e shows the map divided into four clusters. Cluster 3 shows some samples of TB patients in a region mainly activated by non-TB ones. Therefore, those are sampled as exemplary of images with high dissimilarity to other TB images. One example may be seen in Figure 5c. (14 images)
- E) Finally, Figure 6f presents the non-TB images split into the same four clusters. Cluster 0 is located in a region where no CXR image labeled as a TB patient was projected. All samples are from non-TB patients, high silhouette values around 0.6. These samples might present relevant features to distinguish and differentiate them from positive-TB X-Rays. Hence, they are picked for final sampling by their similarity to non-TB samples. An example may be seen in Figure 5d(36 images).

VI. CONCLUSION

This paper described how images from a popular CXR dataset were sampled using a machine learning pipeline. In the end, 120 images were chosen after an extensive exploitation of the non-linear relations among the entire dataset due the deployment feature engineering, data projection and clustering techniques working in conjunction. Section V describes step-by-step how the final images were sampled.

REFERENCES

[1] E. Kotter and M. Langer, "Computer aided detection and diagnosis in radiology," *European Radiology*, vol. 21, 2011.

[2] J. Burt, N. Torosdagli, and N. Khosravan, "Deep learning beyond cats and dogs: recent advances in diagnosing breast cancer with deep neural networks," *The British Journal of Radiology*, vol. 91, 2018.

[3] W. H. Organization, "Who consolidated guidelines on tuberculosis," 2021. [Online]. Available: <https://apps.who.int/iris/bitstream/handle/10665/340255/9789240022676-eng.pdf>

[4] H.-P. Chan, R. K. Samala, and L. M. Hadjiiski, "Cad and ai for breast cancer - recent development and challenges," *The British Journal of Radiology*, vol. 93, 2019.

[5] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[6] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464–1480, 1990.

[7] M. Srinivas, D. Roy, and C. K. Mohan, "Discriminative feature extraction from x-ray images using deep convolutional neural networks," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 917–921.

[8] J. Guérin and B. Boots, "Improving image clustering with multiple pretrained cnn feature extractors," *arXiv preprint arXiv:1807.07760*, 2018.

[9] 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.

[10] F. Corona, M. Mulas, R. Baratti, and J. A. Romagnoli, "On the topological modeling and analysis of industrial process data using the SOM," *Computers & Chemical Engineering*, vol. 34, no. 12, pp. 2022–2032, 2010.

[11] I. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philos Trans A Math Phys Eng Sci.*, vol. 374, 2016.

[12] S. Kaski, *Self-Organizing Maps*. Boston, MA: Springer US, 2010, pp. 886–888. [Online]. Available: https://doi.org/10.1007/978-0-387-30164-8_746

[13] K.-L. Du, "Clustering: A neural network approach," *Neural networks*, vol. 23, no. 1, pp. 89–107, 2010.

[14] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 28, no. 1, pp. 100–108, 1979.

[15] P. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987.

[16] C. Liu, Y. Cao, M. Alcantara, B. Liu, M. Brunette, J. Peinado, and W. Curioso, "TX-CNN: Detecting tuberculosis in chest X-ray images using convolutional neural network," *Proc. - Int. Conf. Image Process. ICIP*, vol. 2017-Sept, pp. 2314–2318, 2018.

[17] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017.

[18] M. Ojala and G. C. Garriga, "Permutation tests for studying classifier performance," *Journal of Machine Learning Research*, pp. 1833–1863, 2010.

[19] T. Ciodaro, D. Deva, J. de Seixas, and D. Damazio, "Online particle detection with neural networks based on topological calorimetry," in *Journal of Physics: Conference Series*, vol. 368, no. 1. IOP Publishing, 2012, p. 012030.

[20] G. Vettigli, "Minisom: minimalistic and numpy-based implementation of the self organizing map," 2018. [Online]. Available: <https://github.com/JustGlowing/minisom/>

[21] S. Jaeger, S. Candemir, S. Antani, Y.-X. J. Wang, 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.