2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)
April 4-7, 2018, Washington, D.C., USA

EMPHYSEMA CLASSIFICATION USING A MULTI-VIEW CONVOLUTIONAL NETWORK

David Bermejo-Peláez[⋆] *Raúl San José Estépar*[†] *M. J. Ledesma-Carbayo*[⋆]

[⋆]Biomedical Image Technologies, Universidad Politécnica de Madrid & CIBER-BBN, Madrid, Spain
[†]Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

ABSTRACT

In this article we propose and validate a fully automatic tool for emphysema classification in Computed Tomography (CT) images. We hypothesize that a relatively simple Convolutional Neural Network (CNN) architecture can learn even better discriminative features from the input data compared with more complex and deeper architectures. The proposed architecture is comprised of only 4 convolutional and 3 pooling layers, where the input corresponds to a 2.5D multiview representation of the pulmonary segment tissue to classify, corresponding to axial, sagittal and coronal views. The proposed architecture is compared to similar 2D CNN and 3D CNN, and to more complex architectures which involve a larger number of parameters (up to six times larger). This method has been evaluated in 1553 tissue samples, and achieves an overall sensitivity of 81.78 % and a specificity of 97.34%, and results show that the proposed method outperforms deeper state-of-the-art architectures particularly designed for lung pattern classification. The method shows satisfactory results in full-lung classification.

Index Terms— Computed Tomography, Emphysema, Tissue Classification, Convolutional Neural Networks

1. MOTIVATION

Chronic Obstructive Pulmonary Disease (COPD) is the fourth leading cause of death worldwide, and it is estimated to be the third one in a few years, which has led to an increasing interest of population studies that allow a better understanding of its patophysiology. COPD is defined as persistent obstruction of the airways. It can be divided into two main phenotypes: chronic bronchitis and pulmonary emphysema that causes an overall decrease in the lung elasticity affecting the lung tissue. The progression of emphysema has traditionally been evaluated by spirometric pulmonary function tests (PFTs), although nowadays it is considered as a non-specific parameter and non-sensitive to very early stages of the disease.

This study was supported by projects TEC2013-48251-C2(MINECO/FEDER). DBP was supported by a FPU grant by the Spain's Ministry of Education. RSJ was supported by NHLBI grant 1R01HL116931. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

Densitometric analysis in chest CT is widely accepted as an alternative measurement of pulmonary emphysema more specific and sensitive than PTFs for in vivo studies [1]. This method is based on the choice of a Hounsfield threshold inside the lung mask to discriminate between emphysematous and non-emphysematous tissue. Although a densitometric approach may be able to quantify the disease extension, it may not be able to classify it into different subtypes. That is why other methods based on texture information [2–4] or local histograms [5] have been proposed to carry out an emphysema subtype classification.

Recently, two approaches have been published to tackle the parenchyma patterns classification problem employing CNNs [6, 7]. In [6], a specific CNN to classify 2D patches into interstitial lung patterns was proposed, whereas [7] used the same CNN proposed in [6], but pre-trained with a variety of texture data.

In this work we aim to compare these previously proposed CNN architectures in the emphysema subtyping problem with respect to 3D, 2.5D (multi-view) and 2D simpler architectures.

2. METHODS

In this work, we proposed a CNN for objective and automated classification of emphysema patterns on chest CT images, considering a total of six radiographic tissue patterns: normal parenchyma (NP) and five emphysematous subtypes (paraseptal (PS), panlobular (PL) and mild, moderate and severe centrilobular (CL1, CL2, CL3) emphysema).

Deep architectures usually include hundreds of thousands of parameters, imposing considerable computation and a large number of training samples. In this work, we propose a relative simple architecture with a reduced number of parameters that tackle the emphysema classification problem considering a 2.5-dimensional representation of the input data. To justify the usage of the proposed architecture, we compared it to different 2D- and 3D- CNNs, as well as the CNN proposed in [6]. We also compared the proposed method with a state-of-the-art work for emphysema subtypes discrimination that uses local intensity probability distribution functions subsequently classified with a KNN classifier [5].

978-1-5386-3636-7/18/$31.00 ©2018 IEEE

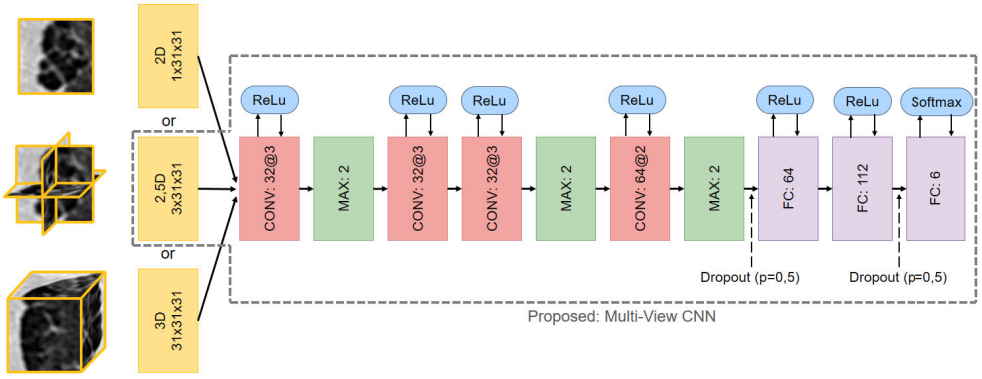

Fig. 1: Proposed multi-view CNN (2.5D CNN) for emphysema classification. 2D and 3D architectures are the same as the 2.5D network, except for the input dimensionality, and the convolutional operations in the 3D case. Filter size and pooling size in convolutional (CONV) and max-pooling (MAX) layers specify their size on each dimension (2D or 3D depending on the architecture).

2.1. Multi-view Convolutional Network (2.5D CNN)

2.1.1. Architecture

The proposed multi-view CNN architecture is shown in Figure 1, and it is composed of 4 convolutional, 3 max-pooling layers and 3 fully-connected layers. The input of the network consists of a multi-view (2.5D) representation of pulmonary segment tissue under study. Therefore, the input is comprised of three 2D patches of size 31x31 pixels, corresponding to the axial, sagittal and coronal views.

Emphysematous patterns in CT images are mainly characterized by local texture patterns. To capture these local patterns and to reduce the complexity of the network, and hence the number of parameters, we use relatively small kernels in the convolutional layers and reduce the number of pooling operations through the network compared to standard convolutional networks.

The network training is based on an optimization problem that minimizes the loss function. In this work we use the Stochastic Gradient Descent (SDG) method updated with Nesternov momentum [8] to minimize the categorical cross entropy.

2.1.2. Overfitting prevention

Convolutional Neural Networks easily overfit the training data. In this work we apply 4 different techniques to prevent and reduce the overfitting. First, we have applied a regularization of the loss function. L2 regularization penalizes the square magnitude of the parameters (w) in the loss function by adding the term $1/2\lambda w^2$, where λ is the regularization strength. This method penalizes sharp changes in the parameters preferring soft ones. The second technique employed is data augmentation. We apply data augmentation using 7 different spatial transformations by rotating 90, -90 and 180 degrees, flipping along horizontal and vertical axes and combining both transformations over the training dataset. Another prevention technique is early stopping, where the training stage is stopped before overfitting process begins. Finally, we have applied dropout to prevent the overfitting, by randomly dropping units with a given probability (p=0,5) from the network during training.

2.2. Comparative methods

2.2.1. 2D and 3D CNN

For comparative purposes, we also designed the 2D and 3D versions of the proposed architecture. Both networks have the same configuration as the one proposed with the only difference that the convolutional operations in the 3D CNN are made in three dimensions.

For the 2D CNN, the input consist of an image patch of size 31x31 pixels in the axial plane, whereas for the 3D CNN, the input corresponds to a 3D segment of size 31x31x31 pixels (see Figure 1).

2.2.2. State-of-the-art CNN

To justify the usage of the proposed CNN we also re-implemented the CNN proposed in [6], where the authors focused on detecting patterns of interstitial lung diseases from 2D patches. The input of the CNN is a 32x32 image patch which is convolved by a series of 5 convolutional layers with 2x2 kernels, and followed by one average pooling with size equal to the size of the final features maps. Finally, the classification stage is comprised of 3 fully-connected layers.

The authors proposed the use of LeakyReLU as a non-linear activation function for both convolutional and fully-

Table 1: Description of the dataset and distribution along different emphysema patterns.

Lung pattern	NP	PL	PS	CL1	CL2	CL3
Num. Samples	447	250	152	223	299	182

connected layers, and Adam optimizer for minimizing the categorical cross-entropy during training.

3. EXPERIMENTS AND RESULTS

3.1. Dataset and Evaluation setup

267 CT scans were selected from the COPDGene study. CT scans were acquired with equipment from 3 vendors and a total of 9 different models. Through these scans, an experienced pulmonologist manually placed 1553 points corresponding to six different radiographic patterns (see Table 1). This database has been previously used for automated emphysema classification in [9]. Of all of these points, 390 were randomly selected for testing (around 25%), while the remaining points were randomly split using a 10-fold cross validation scheme leading to the training and validation sets. The final evaluation of the proposed method as well as a comparison with other CNNs and state-of-the-art methods were carried out on the test set. Training sets was used for training the algorithms, while the validation sets were used to hyperparameter tuning of the proposed architecture.

To evaluate the methods, we consider the sensitivity (SN), specificity (SP), geometric mean (GM) and balanced accuracy (BA), defined as:

$$SN = \frac{TP}{TP+FN} \qquad SP = \frac{TN}{TP+FP}$$
$$GM = \sqrt{TP*TN} \qquad BA = \frac{SN+SP}{2} \tag{1}$$

, where TP, FP, TN and FN stand for true positives, false positives, true negatives and false negatives rates respectively.

3.2. Validation

The proposed multi-view CNN, as well as other CNNs are trained on Regions of Interest (ROIs) extracted around the manually labeled points. The size of the ROIs are selected in accordance with each network architecture as described in Sections 2.1.1 and 2.2.1.

The only pre-processing performed on the dataset was subtracting the mean and dividing by the standard deviation computed on the training and validation sets to normalize the image ROIs.

The framework used in this work to train the CNNs is based on Theano and Lasagne libraries using a PC with GPU NVIDIA TITAN X Pascal 12GB, CPU Intel Core i7 3.6 GHz and 32GB of RAM.

Table 2: Comparison of the proposed CNN (bold) with other CNNs architectures. Results are computed on the independent test set. Evaluation metrics (SN, SP, GM, BA), the number of parameters involved and the time needed to train each CNN are reported.

Method	2D	**2.5D**	3D	Lung-CNN [6]
SN(%)	78.52	**81.78**	82.52	77.20
SP(%)	96.79	**97.34**	97.10	96.45
GM(%)	87.18	**89.22**	89.51	86.33
BA(%)	87.66	**89.56**	89.81	86.86
#Params	68918	**69494**	157558	467426
Time(s)	19.68	**20.48**	240.4	137.3

3.2.1. Comparision to other CNNs

Table 2 provides a comparison of the proposed multi-view CNN with other CNNs, including the 2D and 3D version of the latter, as well as the network proposed in [6]. All the networks were trained with the same samples, and tested in the same test set. As derived from the Table 2, the proposed 2.5D CNN achieves analogous results to those obtained using the 3D CNN, while keeping the same small number of parameters and the same training time as in the 2D CNN case.

The proposed method also outperforms the CNN proposed in [6], and considerably reduces the computational cost by a factor around 6 to 7 in terms of number of parameters involved in the architecture and training time.

3.2.2. Comparison to Local Histogram

As an additional comparative study, we also compared our proposed method against a reference method in emphysema classification. This method has been previously detailed in [5]. Briefly, the local intensity distribution of a two-dimensional ROI of size 31x31 pixels is computed using a kernel density estimator (KDE), and a k nearest neighbors (kNN) classifier assigns each ROI local histogram to its corresponding emphysema pattern based on a majority consensus from the k nearest training samples.

Our proposed method outperforms the results achieved by [5], which obtained a sensitivity (SN) of 44.75%, specificity (SP) of 88.39%, geometric mean (GM) of 60.65% and balanced accuracy (BA) of 66.57%. Although the training time of [5] is lower than the time needed to train the proposed CNN, the classification time is much higher due to the need to compute the KDE features.

3.2.3. Analysis of the method's performance

Figure 2 shows the confusion matrix of the proposed 2.5D CNN. As can be seen, misclassifications generally occurred within the centrilobular emphysema classes, due to their common nature that differ only by the disease stage. Confusion error also occurred between severe centrilobular (CL3) and pan-

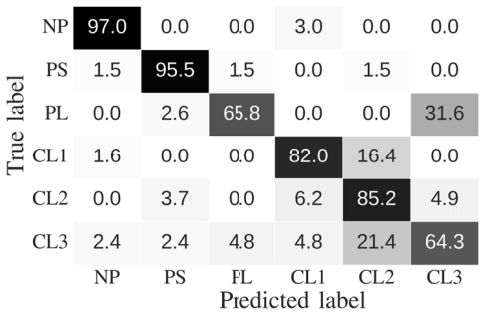

Fig. 2: Confusion matrix of the proposed method.

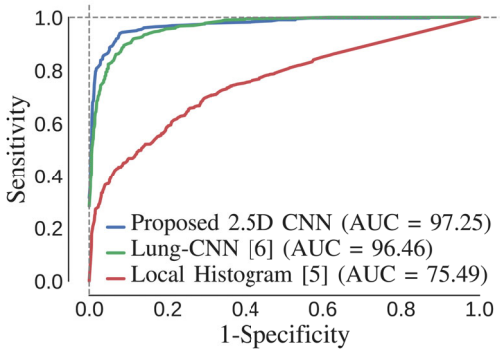

Fig. 3: ROC analysis for the proposed CNN and state-of-the-art approaches. The ROC curves are obtained by macro-averaging over all classes. The area under the curve (AUC) are also reported.

lobular (PL) emphysema, since both classes involve a large parenchyma destruction.

Figure 3 shows the macro-average ROC curves and the area under the curve (AUC) for the proposed CNN and state-of-the-art methods, where it can be seen that the proposed method achieved the highest AUC.

For a further validation, we performed full-lung classification on a severe emphysema CT scan (Figure 4). The classification was carried out at a sampling grid of 5x5x5 pixels in each axial image slice. The rest of the voxels were classified using nearestneighbor interpolation. An expert evaluator confirmed the good agreement of the resultant classification.

4. CONCLUSION

In this work, we present a relatively simple Convolutional Neural Network (CNN) architecture for automated emphysema classification in CT images, considering six different tissue classes, including normal parenchyma and five emphysematous patterns. The proposed architecture, which is composed of 4 convolutional and 3 pooling layers, is considerably simpler than other deeper state-of-the-art architectures proposed in the literature that are specific for lung pattern

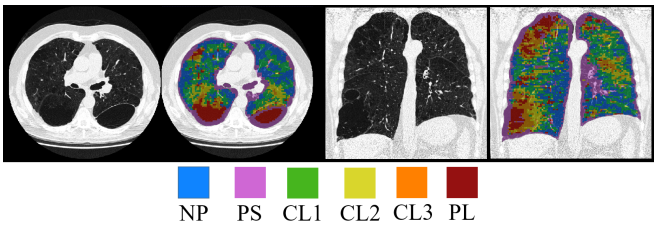

Fig. 4: Full-lung classification results for a severe emphysema case.

classification [6], reducing the computational cost by a factor around six to seven, and outperforming the results.

Additionally, we showed that the 2.5D proposed approach is optimally cost effective, since it is able to reflect 3D information and achieve similar results to those obtained with 3D CNNs, and have similar computation cost compared to 2D approaches.

To train and test the proposed method, we used 1553 manually labelled tissue samples corresponding to six different tissue classes. Visual inspection of full-lung classification and results in the test set (sensitivity of 81.78%, specificity of 97.34%) prove the potential of the method for emphysema detection and classification.

5. REFERENCES

[1] E. Cavigli, et al., "Whole-lung densitometry versus visual assessment of emphysema," *Eur Radiol*, vol. 19, no. 7, pp. 1686–1692, 2009.

[2] R. Uppaluri, et al., "Quantification of pulmonary emphysema from lung computed tomography images," *Am J Resp and Crit Care*, vol. 156, no. 1, pp. 248–254, 1997.

[3] A. Depeursinge, et al., "Lung tissue classification using wavelet frames," *IEEE EMBC*, 2007.

[4] Y. S. Park, et al., "Texture-based quantification of pulmonary emphysema on high-resolution computed tomography: Comparison with density-based quantification and correlation with pulmonary function test," *Invest Radiol*, vol. 43, no. 6, pp. 395–402, 2008.

[5] Peter J. Castaldi, et al., "Distinct quantitative computed tomography emphysema patterns are associated with physiology and function in smokers," *Am J Resp and Crit Care*, vol. 188, no. 9, pp. 1083–1090, 2013.

[6] M. Anthimopoulos, et al., "Lung pattern classification for interstitial lung diseases using a deep convolutional neural network," *IEEE Trans Med Imaging*, vol. 35, no. 5, pp. 1207–1216, 2016.

[7] S. Christodoulidis, et al., "Multisource transfer learning with convolutional neural networks for lung pattern analysis," *IEEE J Biomed Health*, vol. 21, no. 1, pp. 76–84, 2017.

[8] Y. Bengio, et al., "Advances in optimizing recurrent networks," *IEEE ICASSP*, pp. 8624–8628, 2013.

[9] C. S. Mendoza, et al., "Emphysema quantification in a multi-scanner HRCT cohort using local intensity distributions," *IEEE ISBI*, pp. 474–477, 2012.