An application of deep learning to chest diseases detection using images and clinical data
Un'applicazione del deep learning al riconoscimento delle patologie toraciche utilizzando immagini e dati clinici

Federica Crobu and Agostino Di Ciaccio

Abstract In the last few years, computer-assisted diagnosis systems have obtained a growing interest from researchers thanks to the use of deep learning techniques. We propose a deep neural network based on a multi-input architecture that allows to use all the information available to physicians during the diagnosis. The results obtained show an interesting improvement in performance in terms of predictive skill compared to the results in the literature.

Abstract *Negli ultimi anni, i sistemi di diagnosi assistita da computer hanno ottenuto un crescente interesse da parte dei ricercatori grazie all'uso di tecniche di deep learning. Nell'articolo proponiamo un'architettura di rete neurale basata sull'uso di una struttura multi-input che consente di utilizzare tutte le informazioni a disposizione dei medici durante la diagnosi. I risultati ottenuti mostrano un interessante miglioramento delle prestazioni in termini di abilità predittiva rispetto ai risultati presenti in letteratura.*

Key words: deep learning, convolutional neural networks, chest x ray

Introduction

In the classical context of image recognition, our task is to classify what is contained in an image, since all the information are inside the picture. However, in the analysis of medical imaging the challenge is quite different. Indeed, what the model should be able to do, at least ideally, is to emulate the role of the doctor. Thus, we would like to consider not only images like X-ray but also other clinical information of the patients. The goal of including multiple inputs of a different nature, images and

Federica Crobu, Sapienza Università di Roma; email: federicacrobu@gmail.com

Agostino Di Ciaccio, Sapienza Università di Roma; email: agostino.diciaccio@uniroma1.it

numerical values, was achieved by us using a multi-input neural network architecture.

Data and Models

The data chosen for our application belongs to the ChestX-ray14 dataset, released by the U.S. *National Institutes of Health* (NIH). It contains over 112,000 frontal chest X-ray images of 30,805 patients. Each of them can be healthy or sick, with one or more of the following 14 pathologies: Atelectasis, Cardiomegaly, Consolidation, Edema, Effusion, Emphysema, Fibrosis, Hernia, Infiltration, Mass, Nodule, Pleural Thickening, Pneumonia, Pneumothorax. Moreover, a category "No Finding" accounts for the images in which no one of the previous mentioned diseases has been detected.

For each image there is one label, corresponding to the detected diseases, extracted from the radiological reports using a natural language processing technique. The authors declared that the accuracy is above 90% [5].

1.1 Related works

The ChestX-ray14 dataset has already been used by other researchers. Surely, the best-known work is from the Stanford's team (including prof. Ng) [4]. They proposed an architecture called CheXNet based on a convolutional neural network architecture called DenseNet121 [3]. With their network they have outperformed other analyzes such as that of Yao et al. [6] and Wang et al. [5] who explored different CNN architectures. Other interesting and more recent works are those of Baltrushat [1] and Guendel [2].

1.2 Model's architecture

Starting with the work of Stanford's team, we have decided to improve the model by exploiting the other information available together with the X-ray images. We considered 17 new inputs: the age, the gender, the view position and 14 new variables containing the information of the previous pathological anamnesis available from the same data. Thus, our idea was to enhance the DenseNet121 model with another parallel neural network (based on two hidden layers with 128 neurons each) that processes the non-image features. The two independent branches are then concatenated and connected to the final layer of the network: an output layer based on 14 sigmoid neurons whose task is to estimate the probability of presence of each disease on the X-ray image.

In order to solve this multi-label multi-class problem, we have employed a binary cross-entropy loss function, weighted to take into account the imbalance of the diseases' frequencies. Moreover, to make the training feasible we have decided to work with 224x224 images with 3 RGB channels, in order to exploit the initial weights pretrained on ImageNet. Finally, to avoid overfitting, data augmentation has been applied by horizontal flipping the images.

The split in training and test has been made according to the official no patient overlap subdivision, released by the NIH, with approximately 80% of training and 20% of test.

Three training processes have been performed: one for the CNN, one for the non-image features and the last one with the entire network using as initial weights the best ones obtained in the first two nets. In all the cases we have used an Adam optimization (β_1=0.9, β_2=0.999, learning rate of 0.001 for the first and 0.0001 for the last two) with a reduction of the learning rate in the test loss pleateau. Weights that provide the maximum mean AUC on the test-set were considered optimal.

Results

In the following table we can see the comparison of the performances of our model with the best results obtained by other researchers in terms of the mean AUC scores.

Table 1: AUC scores comparison

	Wang et al. (2017)	Yao et al. (2018)	CheXNet (2017)	Multi-input
Official split	Yes	No	No	Yes
Atelectasis	0.716	0.772	0.809	**0.816**
Cardiomegaly	0.807	0.904	0.925	**0.925**
Effusion	0.784	0.859	0.864	**0.867**
Infiltration	0.609	0.695	0.735	0.731
Mass	0.706	0.792	0.868	**0.897**
Nodule	0.671	0.717	0.780	**0.827**
Pneumonia	0.633	0.713	0.768	**0.776**
Pneumothorax	0.806	0.841	0.889	**0.927**
Consolidation	0.708	0.788	0.790	**0.801**
Edema	0.835	0.882	0.888	**0.893**
Emphysema	0.815	0.829	0.937	**0.946**
Fibrosis	0.769	0.767	0.805	**0.881**
Pleural Thickening	0.708	0.765	0.806	**0.827**
Hernia	0.767	0.914	0.916	**0.963**
Average	0.738	0.803	0.841	**0.863**

We have chosen the split suggested by the authors, after having verified that the previously proposed approaches, which had used different divisions, obtain the same results with this split. In fact, the size of test set, over 25,000 images, guarantees a great stability of the results with respect to different splits. It is evident in the table that the average AUC has been significantly improved by our method and, for most of the classes, we clearly outperformed the previous works.

4. Conclusions

The results of this application confirmed our intuition: for these data, a multi-input neural network architecture can significantly improve predictions. Clearly, the idea of combining different heterogeneous sources of information can be applied in other fields of medicine, as in the analysis of MRI scans. Whenever the patient's clinical and/or demographic information are available, it is possible and fruitful to apply this approach.

References

1. Baltruschat, I.M., Nickisch, H., Grass, M., Knopp, T., Saalbach, A.: Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification, arXiv:1803.02315v2 (2019).
2. Guendel, S., Grbic, S., Georgescu, B., Zhou, K., Ritschl, L., Meier, A., Comaniciu, D.: Learning to recognize Abnormalities in Chest X-Rays with Location-Aware Dense Networks, preprint arXiv:1803.04565 (2018).
3. Huang, G., Liu, Z., van der Maaten, L., Weinberger, K. Q.: Densely Connected Convolutional Networks, In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017). DOI 10.1109/cvpr.2017.243.
4. Rajpurkar, P.,Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Ball, R. L., Langlotz, C., Shpanskaya, K., Lungren, M. P., Ng, A.: CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning, preprint arXiv:1711.05225v3 (2017).
5. Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R. M.: ChestX-ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3462–3471 (2017). DOI 10.1109/CVPR.2017.369.
6. Yao, L., Poblenz, E., Dagunts, D., Covington, B., Bernard, D., Lyman, K.: Learning to diagnose from scratch by exploiting dependencies among labels, preprint arXiv:1710.10501v2 (2018).