Classification Tasks on 3D Lung CT Scans Datasets Using Deep Transfer Learning

Candidate: Emilio Cecchini

Supervisors: Maria Cecilia Verri [1], Lucio Anderlini [2]

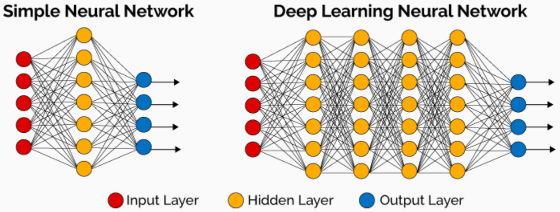
Friday 12th February, 2021

[1]Università degli Studi di Firenze

[2]INFN Firenze

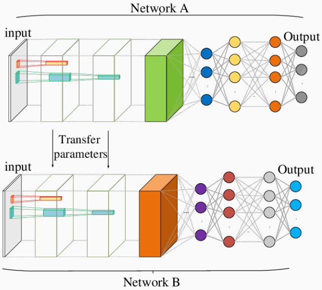
Deep Learning

- **Deep learning** is the artificial intelligence field where complex statistical models are trained on large datasets to enhance their performance with respect to some task.
- It is characterized by the use of **deep neural networks** (**DNN**), which are artificial neural networks that have two or more hidden **layers** between input and output.
- In the last decade, the computer vision community has gradually shifted its main focus to deep learning for most applications, including **medical data processing**.



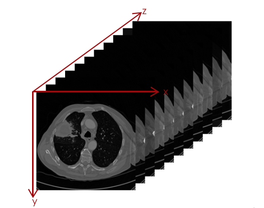
Transfer Learning

- DNN training often relies on Transfer learning, which refers to the situation where what has been learned in one setting is exploited to improve the generalization in another setting.
- If such a pre-training relies on a much larger dataset, the DNN may learn a **representation** which can then be applied to different tasks, with the benefit of quickly generalize from only very few examples from a second dataset.



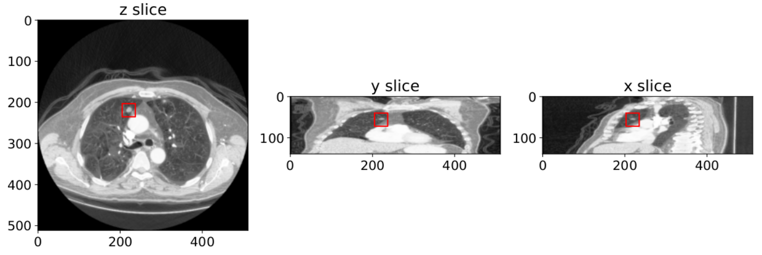
CT Scan

- A **computerized tomography** (CT) scan consists of a series of X-ray images stacked together to create a volumetric image of the internal structure of the body.
- 3D CT scan data can be processed in two ways: one slice at the time (2D CNNs) or a series of 2D images together (3D CNNs).
- Processing one image at the time there is the risk of loosing a lot of **spatial information**, which can be critical to identify important features.
- Computer vision techniques are widely employed in research on automatic diagnosis processing CT scans, but annotating CTs scans requires clinical experience and it is therefore very expensive.



Lung Nodule Classification

- Lung cancer is the leading cause of cancer-related death worldwide. Early diagnosis of pulmonary nodules in CT chest scans provides an opportunity for designing effective treatment and making financial and care plans.
- The problem of diagnostic classification between **benign and malignant lung nodules** in CT can be addressed with a DNN modeling a direct mapping from 3D images to class labels.



Datasets

To represent the classical scenario in which pre-training on a large dataset is functional to the training on a small dataset we consider:

- **LIDC-IDRI dataset** [2]: 754 lung nodules with unique annotations (50.3% benign and 49.7% malignant).
- **LUNGx Challenge dataset** [1]: 73 lung nodules (37 begnin and 36 malignant). This dataset, besides being very small, it is much more difficult to classify than the LIDC-IDRI dataset, this is mainly due to the very similar nodules size.

(a) Benign **(b)** Malignant

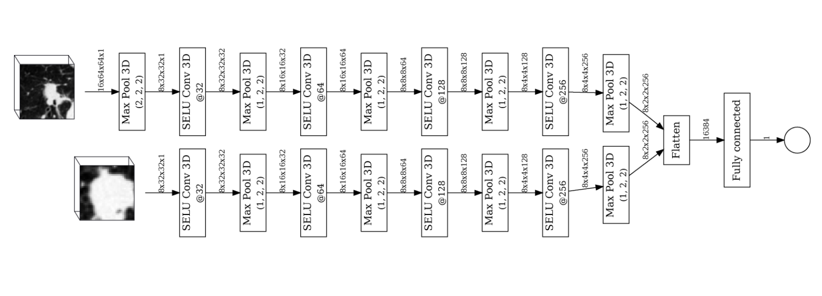
Examples from the LIDC dataset

(a) Benign **(b)** Malignant

Examples from the LUNGx Challenge dataset

Architecture

- Here is the proposed network architecture for the lung nodule classification.
- It is **3D convolutional neural network**, which has two almost identical paths for the same input nodule: one with a 3D patch containing the nodule and its surrounding and one with a zoomed version of the same image.



Computing Platforms

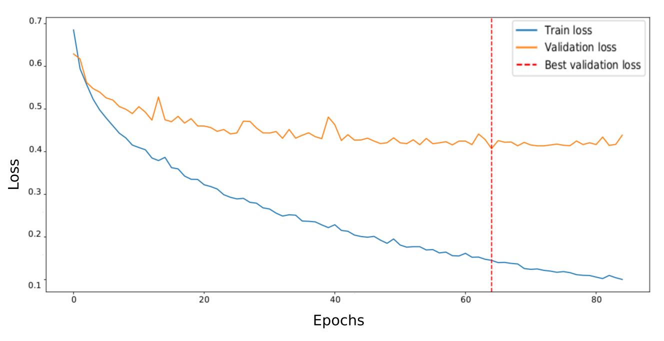
Large computing resources are necessary to train 3D CNNs:

- **acer-1**: CPU-only platform, not suitable for deep learning and 3D high resolution images.
- **Marconi100**: CINECA supercomputer equipped with 4 GPUs. Very powerful platform but not ideal for model design and testing.
- **Lb7**: computing platform equipped with a high memory GPU. It was acquired by INFN-Florence in order to have a platform dedicated to deep learning model development.

	Batch size = 2			
	Filters = 8		**Filters** = 16	
	Exec. time	**Speed-up**	**Exec. time**	**Speed-up**
acer-1	1194.03s	-	2027.96s	-
M100 (1 GPU)	5.34s	$\times$224	6.4s	$\times$317
M100 (4 GPUs)	7.95s	$\times$150	8.82s	$\times$230
Lb7	6.32s	$\times$189	7.19s	$\times$282

LIDC Pre-training

- The first step is to train the neural network on the full LIDC dataset.
- 20% of the elements as validation set.
- Batch size was set to 16.
- The model reached the optimal loss value on the validation set after 63 epochs, with accuracy 0.78 and AUC 0.9 on the same validation set.



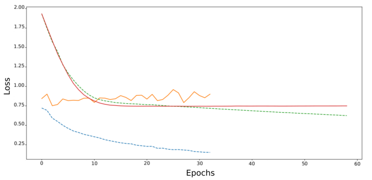
LUNGx Training

Then 3 different models are evaluated on the LUNGx Challenge dataset (reducing the batch size to 8):

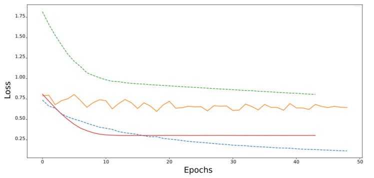
- **LIDC only**: a model trained only on the LIDC dataset.
- **w/o pre-training**: a model trained from scratch on the LUNGx dataset.
- **w/ pre-training**: a model pre-trained on LIDC and fine-tuned on LUNGx.

LUNGx Training Visualization

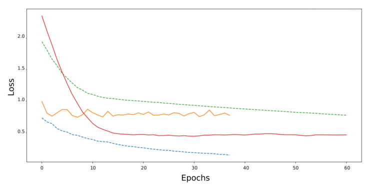
3-fold cross-validation: one subset for testing, 10% of the two subsets for validation and the remaining data samples for training.



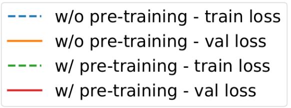
(a) Fold 1



(b) Fold 2



(c) Fold 3



AUC table

AUC scores of the 3-fold cross-validation on the LUNGx Challenge dataset. The obtained AUC is comparable with the best score among the partecipants of the challenge.

	AUC			
LIDC only	0.65			
	fold 1	**fold 2**	**fold 3**	**average**
w/o pre-training	0.55	0.69	0.46	0.57
w/ pre-training	0.78	0.74	0.49	**0.67**

Confusion matrix

		True diagnosis		
		Benign	Malignant	Total
Predictions	Benign	8	1	9
	Malignant	29	35	64
	Total	37	36	73

LIDC only

		True diagnosis		
		Benign	Malignant	Total
Predictions	Benign	28	9	37
	Malignant	26	10	36
	Total	54	19	73

w/o pre-training

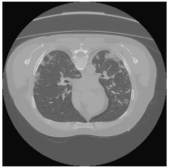
		True diagnosis		
		Benign	Malignant	Total
Predictions	Benign	22	15	37
	Malignant	12	24	36
	Total	34	39	73

w/ pre-training

Covid Classification

- Covid's patients, in few months, show evidences of the infection in CT scans as damages to the lung tissues.
- I tested the model development in my thesis on this challenging task using the MosMedData dataset [5], which contains 1110 3D lung CT scans distributed in 5 categories according to the severity of the lung infection.
- With current technology, 3D NNs are not competitive with deeper 2D networks (as proposed in [3]) for such large data.

(a) Healthy lung



(b) Moderate infections

(c) Severe infections

Conclusions

- This study clearly indicates that, even if computationally intensive, the adoption of 3D convolutional neural networks pre-trained on larger, possibly different datasets makes it possible to employ deep learning techniques for clinical studies with datasets as small as few hundreds of cases, or even less.
- In the future, the technological advancements in computing will open to the application to the CT scans analysis of pre-trained 3D convolutional networks with hundreds of layers, which is currently possible with 2D images only.

Thanks for the attention

S. G. Armato III, K. Drukker, F. Li, L. Hadjiiski, G. D. Tourassi, R. M. Engelmann, M. L. Giger, G. Redmond, K. Farahani, J. S. Kirby, et al.
Lungx challenge for computerized lung nodule classification.
Journal of Medical Imaging, 3(4), 2016.

S. G. Armato III, G. McLennan, L. Bidaut, M. F. McNitt-Gray, C. R. Meyer, A. P. Reeves, B. Zhao, D. R. Aberle, C. I. Henschke, E. A. Hoffman, et al.
The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans.
Medical physics, 38(2):915–931, 2011.

N. Gupta, A. Kaul, D. Sharma, et al.
Deep learning assisted covid-19 detection using full ct-scans.
2020.

J. Lemley, S. Bazrafkan, and P. Corcoran.
Transfer learning of temporal information for driver action classification.
In *MAICS*, pages 123–128, 2017.

S. Morozov, A. Andreychenko, N. Pavlov, A. Vladzymyrskyy, N. Ledikhova, V. Gombolevskiy, I. A. Blokhin, P. Gelezhe, A. Gonchar, and V. Y. Chernina.
Mosmeddata: Chest ct scans with covid-19 related findings dataset.
arXiv preprint arXiv:2005.06465, 2020.

S. P. Singh, L. Wang, S. Gupta, H. Goli, P. Padmanabhan, and B. Gulyás.
3d deep learning on medical images: A review.
arXiv preprint arXiv:2004.00218, 2020.