Radiologist-Level Pneumonia Detection with Deep Learning

Gaëtan Dissez et Guillaume Duboc

Analyser des radios des poumons

Multi-label classification

13 pathologies différentes identifiables

Plan

I. Le dataset CheXpert
II. Architecture : *Densely Connected Convolutional Network* (DenseNet)
III. Résultats et comparaison avec les radiologues
IV. Heatmap : localiser les pathologies

I. Le dataset CheXpert

Input: Radio des poumons

Output: Vecteur binaire de taille 14

Dans le dataset train, 4 labels possibles:

- 1 (pathologie présente)
- 0 (pathologie absente)
- ‘’ (pathologie absente)
- -1 (incertitude)

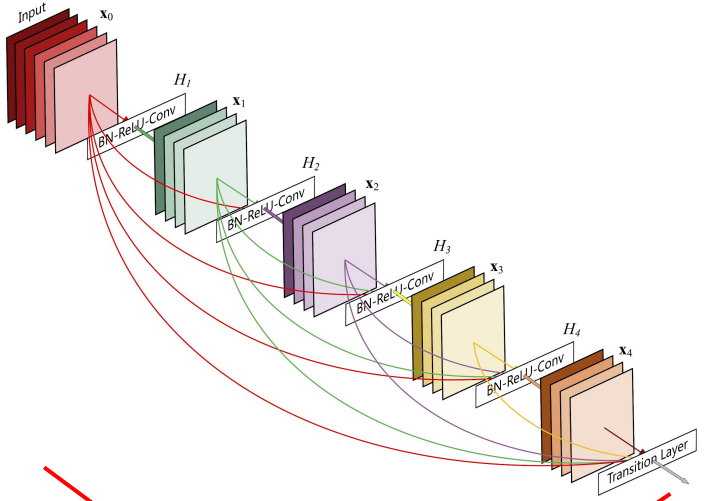
No Finding	Enlarged Cardiomediastinum	Cardiomegaly	Lung Opacity	Lung Lesion	Edema	Consolidation	Pneumonia	Atelectasis	Pneumothorax	Pleural Effusion	Pleural Other	Fracture	Support Devices
1.0									0.0				1.0
		-1.0	1.0		-1.0	-1.0		-1.0		-1.0		1.0	
			1.0			-1.0						1.0	
			1.0			-1.0						1.0	

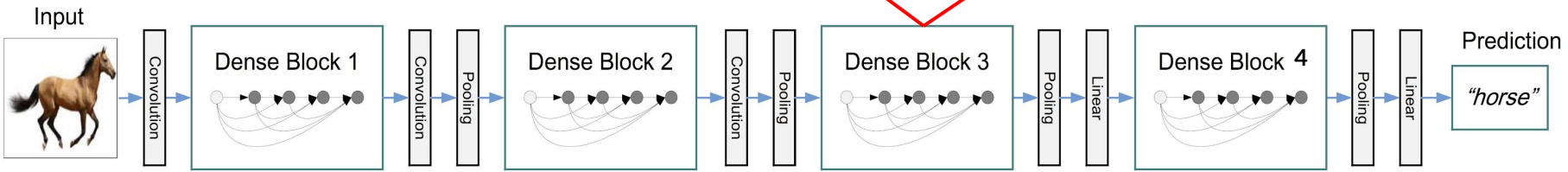
I. Le dataset CheXpert

Train set: 223 414 images
Validation set: 234 images

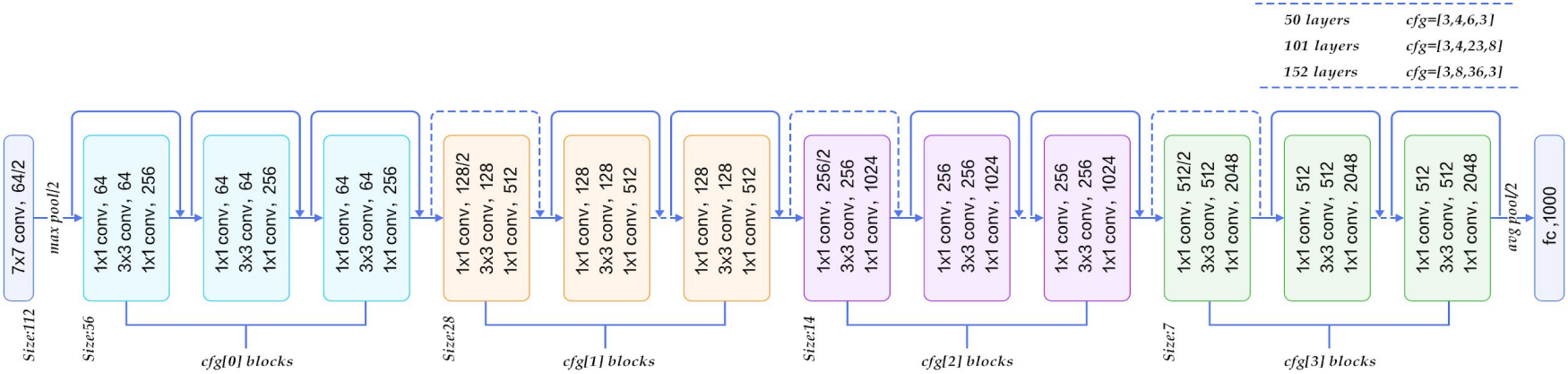
	Positif	Incertain	Negatif
No Finding	10,0%	0,0%	90,0%
Enlarged Cardiomediastinum	4,8%	5,6%	89,6%
Cardiomegaly	12,1%	3,6%	84,3%
Lung Opacity	47,3%	2,5%	50,2%
Lung Lesion	4,1%	0,7%	95,2%
Edema	23,4%	5,8%	70,8%
Consolidation	6,6%	12,4%	81,0%
Pneumonia	2,7%	8,4%	88,9%
Atelectasis	14,9%	15,1%	70,0%
Pneumothorax	8,7%	1,4%	89,9%
Pleural Effusion	38,6%	5,2%	56,2%
Pleural Other	1,6%	1,2%	97,2%
Fracture	4,0%	0,3%	95,7%
Support Devices	51,9%	0,5%	47,6%

DenseNet


Input
$\mathbf{x}_0$
H_1
BN-ReLU-Conv
$\mathbf{x}_1$
H_2
BN-ReLU-Conv
$\mathbf{x}_2$
H_3
BN-ReLU-Conv
$\mathbf{x}_3$
H_4
BN-ReLU-Conv
$\mathbf{x}_4$
Transition Layer


Input
Convolution
Dense Block 1
Convolution
Pooling
Dense Block 2
Convolution
Pooling
Dense Block 3
Pooling
Linear
Dense Block 4
Pooling
Linear
Prediction
"horse"

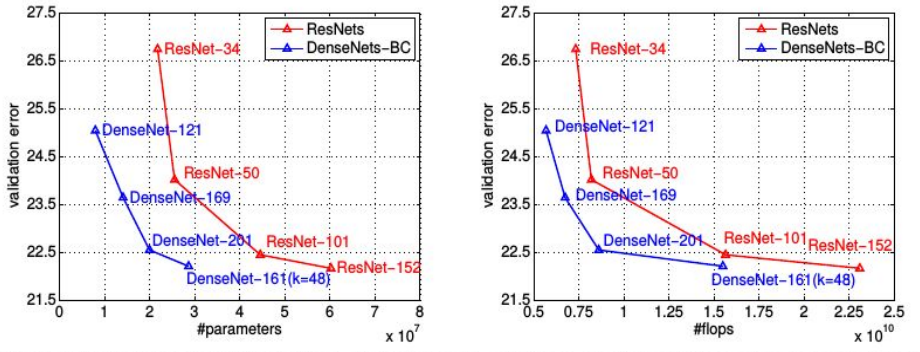
ResNet


50 layers cfg=[3,4,6,3]
101 layers cfg=[3,4,23,8]
152 layers cfg=[3,8,36,3]
7x7 conv, 64/2
max pool/2
1x1 conv, 64
3x3 conv, 64
1x1 conv, 256
1x1 conv, 128/2
3x3 conv, 128
1x1 conv, 512
1x1 conv, 128
3x3 conv, 128
1x1 conv, 512
1x1 conv, 256/2
3x3 conv, 256
1x1 conv, 1024
1x1 conv, 256
3x3 conv, 256
1x1 conv, 1024
1x1 conv, 512/2
3x3 conv, 512
1x1 conv, 2048
1x1 conv, 512
3x3 conv, 512
1x1 conv, 2048
avg pool/2
fc ,1000
Size:112
Size:56
Size:28
Size:14
Size:7
cfg[0] blocks
cfg[1] blocks
cfg[2] blocks
cfg[3] blocks

Comparaison ResNet vs. Densenet

Sur notre Dataset :

- validation_loss = 0.358 pour DenseNet 121
- validation_loss = 0.39 pour ResNet 50



Comparison of the DenseNets and ResNets top-1 error rates (single-crop testing) on the ImageNet validation dataset as a function of learned parameters (left) and FLOPs during test-time (right).
source : G. Huang et al., Densely Connected Convolutional Networks, 2018

III. Résultats et comparaison avec les radiologues

Évolution de la *loss* sur un premier modèle

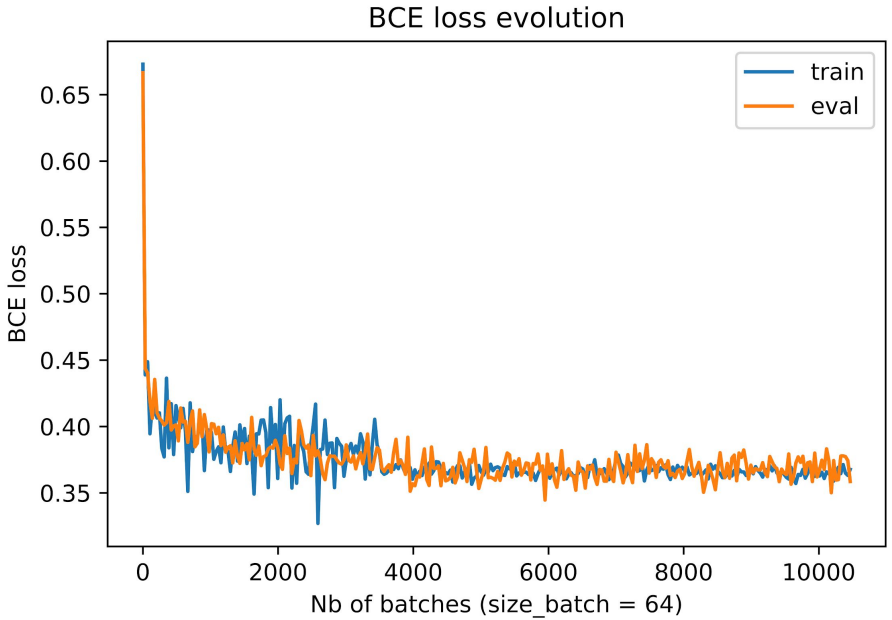
model: DenseNet 121

policy: U-Ones

epochs: 3

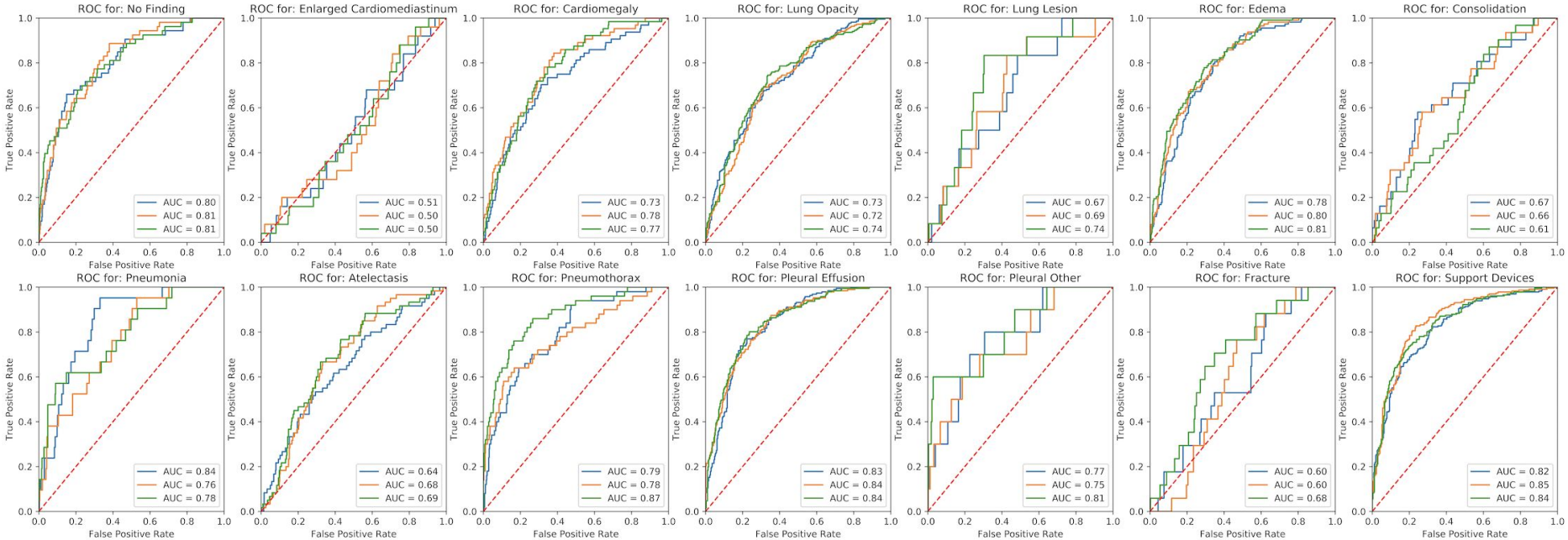
Loss: BCE Loss

Optimization: Adam optimizer



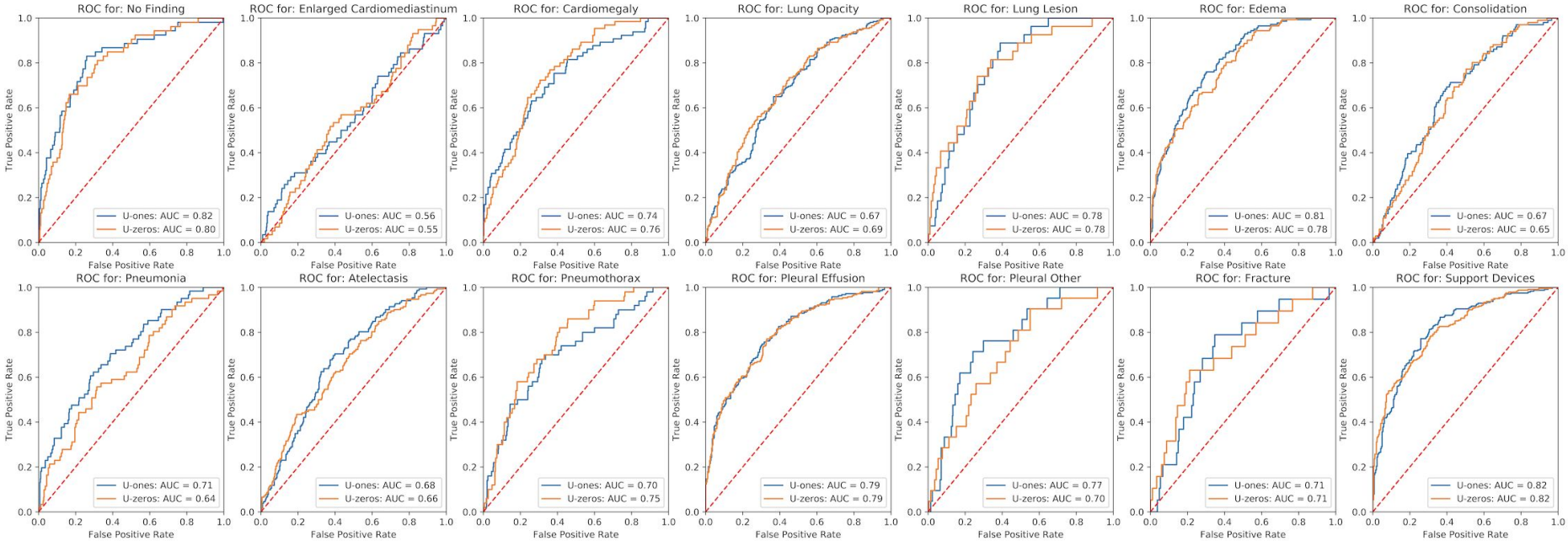
Courbes ROC pour différentes epoch avec DenseNet

AUROC means: epoch1 (0.72) < epoch2 (0.74) < epoch3(0.76)



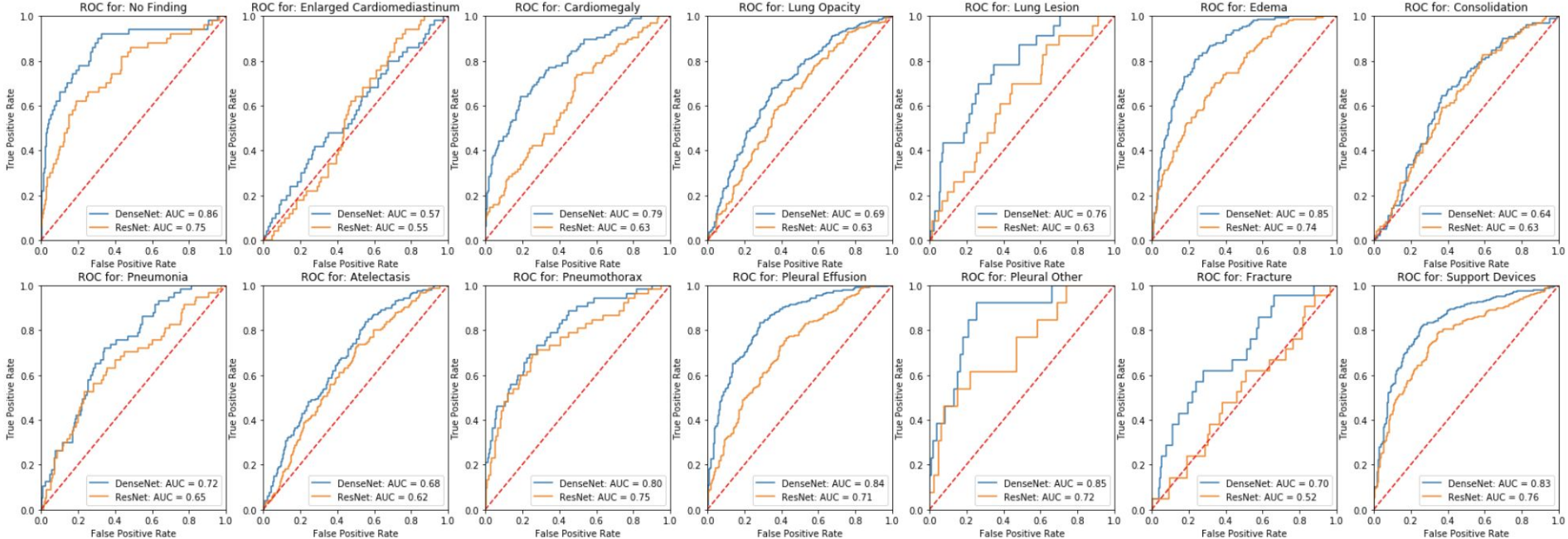
Courbes ROC pour deux *pre-processing* différents avec DenseNet

AUROC means: U-zeros (0.71) < U-ones (0.74)



Courbes ROC pour deux réseaux différents : ResNet et DenseNet

AUROC means: DenseNet (0.76) > ResNet(0.66)



III. Résultats et comparaison avec les radiologues

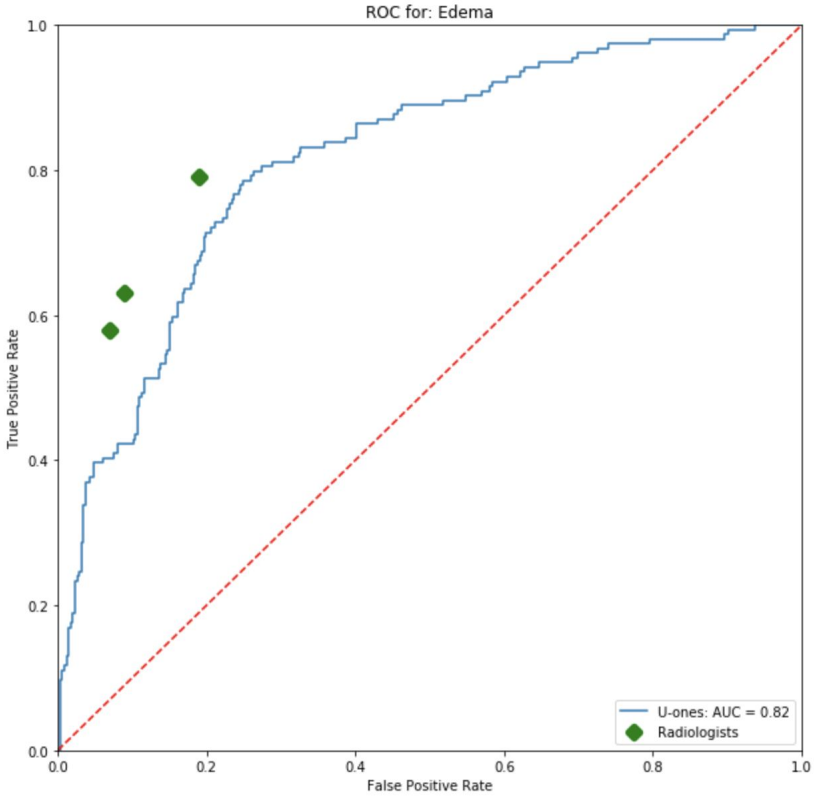
Comparaison avec le projet CheXpert: https://stanfordmlgroup.github.io/competitions/chexpert/

Attention: pas le même test set…

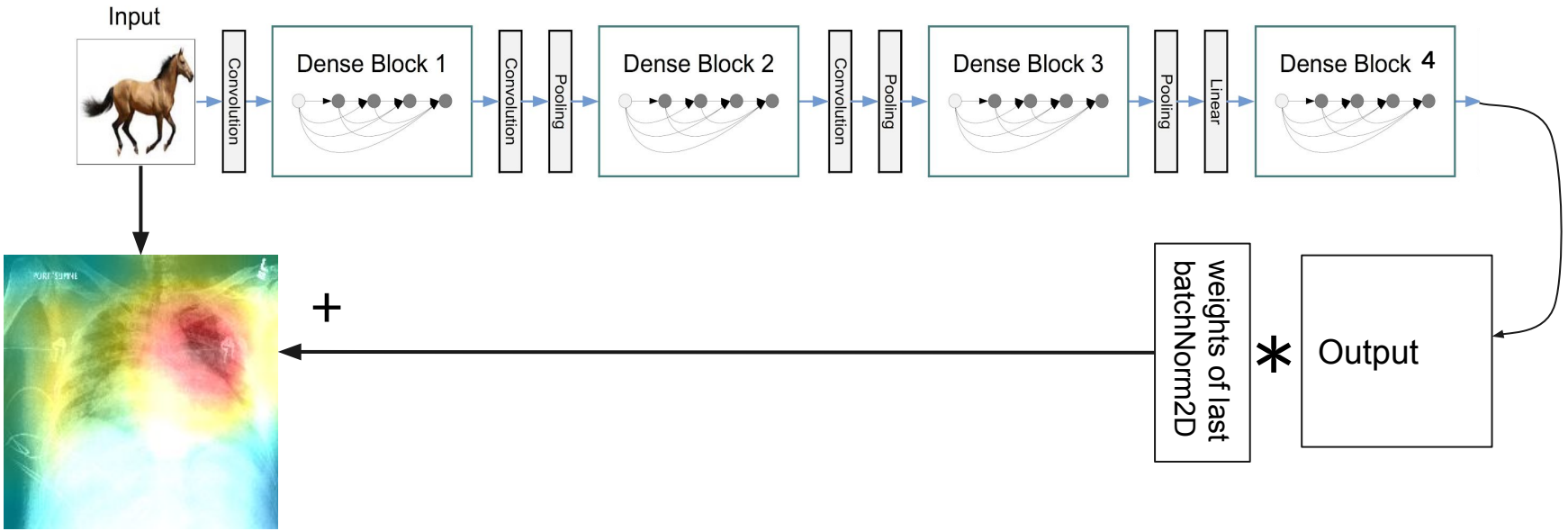
	AUROC Stanford	AUROC
Atelectasis	0,85	0,69
Cardiomelagy	0,9	0,77
Consolidation	0,9	0,67
Edema	0,92	0,81
Pleural Effusion	0,97	0,84

III. Résultats et comparaison avec les radiologues

Comparaison avec les radiologues



IV. Heatmap


Input
Convolution
Dense Block 1
Convolution
Pooling
Dense Block 2
Convolution
Pooling
Dense Block 3
Pooling
Linear
Dense Block 4
+
weights of last batchNorm2D
*
Output