Convolutional Neural Network for Chest X-ray Pneumonia Detection

Ilyas LAHSAINI
University of Tlemcen ,Algeria
Faculty of Technology
Biomedical Engineering Laboratory
ilies.lahsaini@univ-tlemcen.dz

Mostafa EL HABIB DAHO
University of Tlemcen ,Algeria
Faculty of Technology
Biomedical Engineering Laboratory
mostafa.elhabibdaho@univ-tlemcen.dz

Mohammed Amine CHIKH
University of Tlemcen ,Algeria
Faculty of Technology
Biomedical Engineering Laboratory
mea_chikh@univ-tlemcen.dz

ABSTRACT

Deep learning techniques are widely used to design robust classification models in several arias such as medical diagnosis tasks in which it achieves good performance. In recent years Pneumonia causes 15% of the total number of deaths in children under the age of 5. It can be caused by viruses, bacteria or fungi, which led the researchers to focus their studies on identifying pneumonia basing on ChestX-ray images, using deep learning techniques. In this paper, we propose a CNN model (Convolutional Neural Network) for the classification of Chest X-ray images. The proposed method is based on a non-complex CNN and without the use of transfer learning. Our proposed model uses fewer parameters and thus reduces the training time. In this study, and because of the low availability of data, we used the data augmentation method to eliminate overfitting and to improve the accuracy of the validation and classification of the model. The obtained results prove the efficiency of the proposed architecture compared with the state of the art methods.

KEYWORDS

Deep learning ,transfert learning ,Pneumonia ,Chest X-ray ,Convolutionel Neural Network CNN.

ACM Reference Format:
Ilyas LAHSAINI, Mostafa EL HABIB DAHO, and Mohammed Amine CHIKH. 2020. Convolutional Neural Network for Chest X-ray Pneumonia Detection. In *The international conference on Intelligent systems and Pattern recognition (ISPR '20), October 16–18, 2020, Virtual Event, Tunisia*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3432867.3432873

1 INTRODUCTION

Pneumonia is an infection of one or both lungs most often caused by bacteria or viruses and even by fungi. Pneumonia shows a wide variety of life-threatening infections.

Symptoms of pneumonia can cause muscle pain and headache. Detecting pneumonia is a tedious task and is only possible for expert radiologists. Chest pain that intensifies during cough and deep breaths, and it can create a sudden fever up to 41°C and significant chills.

The convolutional neural networks CNN (Convolutional Neural Network) is used to obtain more precision and to obtain fast results through their performances as training techniques for the classification of medical images, they are qualified to approximate complex non-linear functions to process large images. They are designed to automatically extract the features of the input images.

Deep Learning has been conducted using CNNs and has become the most used approach for the selection and recognition of shapes in an image over the past decade. For this reason, we used this architecture to help radiologists to detect pneumonia automatically from chest X-ray images.

-Convolutional neural network
The architecture of CNN (Figure 1) is composed of a set of processing layers. The first one the convolution layer, the purpose of this layer is to find the parameters in an image and locate them to keep the structure of the image with size compression. Then, it is important to use the ReLu correction layer to eliminate linearity in an image. The second step is to detect the features to minimize the spatial size of the image, thus reducing the number of parameters and the computing time in the network. It is therefore important to insert this pooling layer periodically between two successive convolutional layers to control overfitting. After several layers of convolution and pooling, we find the fully connected layer where the information will propagate until the artificial neural network is well trained. In the end, a prediction is made to categorize the image. The loss layer (Loss) consists of penalizing the difference between the output value and the real value, it is normally the last layer in the network. In practice, the outputs are not connected, so it is essential to use the "Sigmoid" or "Softmax" function which is used to extract a probability [1].

A CNN architecture is made up of the following processing layers:

- The convolution layer (CONV): consists of convolving the input image with a feature detector in order to extract the important characteristics

ISPR '20, October 16–18, 2020, Virtual Event, Tunisia

ACM ISBN 978-1-4503-7502-3/20/10. . . $15.00
https://doi.org/10.1145/3432867.3432873

- The pooling layer (POOL): is used to compress the dimensions of the images one more time while keeping these important characteristics.
- The correction layer (ReLU): allows the replacement of negative values in an image with zeros (Rectified Linear Unit).
- The fully connected (FC) layer: which is a perceptron type layer.

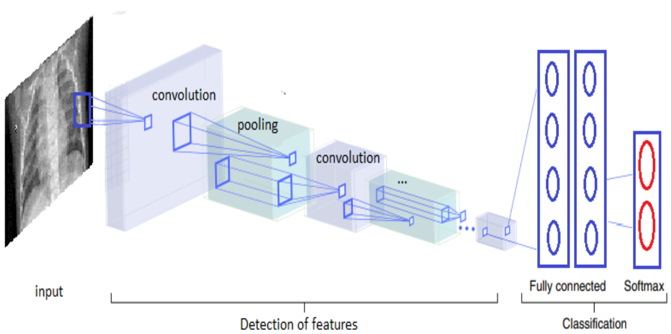

Figure 1: Convolutional neural network architecture.

2 RELATED WORKS

In the last few years, several automated machine learning approaches, to classify different classes of pneumonia, have been widely studied [5] - [6]. Fiszman et al [5] applied a natural language processing (NLP) tool to classify dangerous bacterial pneumonia-related diseases in chest X-ray (CXR). The results of this kind of application showed performance very comparable to that of the human expert.
Kallianos et al [17] published a state of art review explaining the importance and efficacity of AI in chest X-ray image classification and analysis. Wang et al.[7] treated this problem and prepared a new CXR8 database with 108,948 thoracic chest X-ray images of 32,717 unique patients. Each x-ray image could have various labels. They validated the results on this data using deep convolutional neural networks (CNN) and the results obtained were promising. In this study, they also cited that this chest X-Ray 8 database can be more extensive to include more classes of diseases and would be helpful for other research studies.
Rajpurkar et al. [4] proposed a convoluted neural network of 121 layers from CXR14 data set. This database is available to the public with over 100 thousand front view X-rays with 14 disease labels. They mentioned that their algorithm can predict 14 disease categories with great efficiency.
Irvin et al. [8] They mentioned that an increasing labeled dataset is the key to success for prediction and classification tasks. They proposed a huge dataset named CheXpert that consists of more than 0.2 million chest radiographic images of 65,240 patients. Then, they also used a CNN to assign labels to these patients based on the probability obtained by the model using the lateral and frontal radiographs. Furthermore, the use of a large data set is highly desirable as each object of the image is carefully detected and the segmentation of each instance is performed accurately. Therefore, a different method is needed to handle both instance segmentation and object detection. Such robust approaches are faster CNN (F-CNN) [9] and FCN (Fully convolutional network) [10].
Also, in the process of Faster R-CNN (FRCNN) we can extend the region by adding another branch, to predict the segmentation mask on each region of interest along with existing branches for the classification task. They have added to their approach a segmentation at the instance level by predicting the convolutional characteristics.
In the study of Kermany et al [14], they proposed a method that uses CNNs based on transfer learning. Transfer learning has shown to be effective when it comes to areas where data are limited. The main architecture was proposed for Glaucoma based on OCT images of the retina. The same architecture was also tested on chest X-rays images using transfer learning. In the case of retinal OCT, in a multi-class comparison between diabetic macular edema, choroidal neovascularization, drusen and normal, they obtained an accuracy of 96.6%, a sensitivity of 97.8%, and specificity of 97, 4%. In the comparison of chest X-rays with abnormal and normal cases, they obtained an accuracy of 92.8%, sensitivity of 93.2% and a specificity of 90.1%.

The works presented above shows the applicability of some CNN architectures that allow the selection and detection of features to improve as much as possible the performance of a deep learning algorithm. In this paper, we propose a CNN architecture applied to chest radiographic images containing manifestations of pneumonia.

3 MATERIALS AND METHODS

3.1 Materials

Our experiments relied on a Chest X-ray image data proposed in [4]. We have used the Keras open-source deep learning framework with TensorFlow backend to implement and train the convolutional neural network model. All experiments were run on an Nvidia GeForce GTX 1080 GPU card.

3.1.1 Description of the database. The dataset of Kermany et al [14] used in this work contains 5856 chestX-ray images of JPEG format for the training task divided on 3,883 classified as depicting pneumonia (2,538 bacterial and 1,345 viral) and 1,349 normal. The first remark is that the dataset used for this study is imbalanced Figure 2.

The model was then tested with 234 normal images and 390 pneumonia images (242 bacterial and 148 viral) from 624 patients. these radiographic images are chosen from children aged between 1 to 5 years old from the Medical Center in Guangzhou. Knowing that all these images were performed as part of clinical care, and they have undergone a pre-processing to remove all low-quality scans.

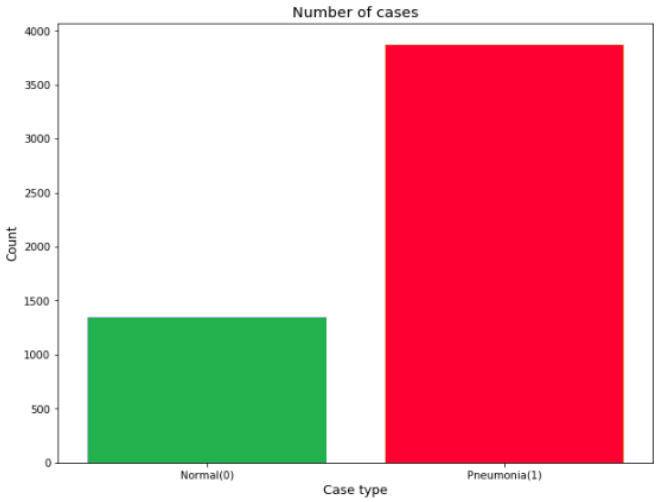

Figure 2: The datagram of the dataset used for the training phase which shows the imbalance of data.

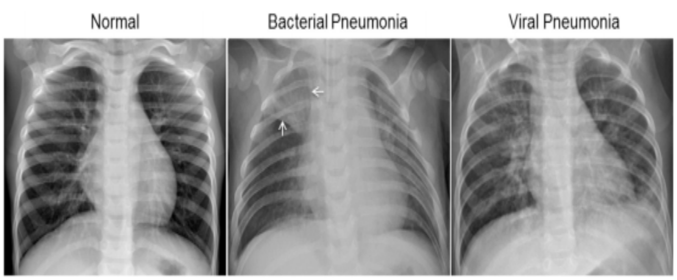

Figure 3: The normal chest X-ray (left panel) depicts clear lungs without any areas of abnormal opacification in the image. Bacterial pneumonia (middle) typically exhibits a focal lobar consolidation, in this case in the right upper lobe (white arrows), whereas viral pneumonia (right) manifests with a more diffuse interstitial pattern in both lungs[14].

3.2 Methods

3.2.1 Preprocessing and Data Augmentation: To solve the problem of overfitting and increase the model's ability to generalize during the training and to increase the size and quality of the data set, we used data augmentation methods that exist in the literature [18]. The settings deployed in data augmentation are exposed below in Table 1. The processing starts with a rescale to image (reduction or magnification during the augmentation process). After that, we do a rotation of the images which will be randomly turned during the training. Width Shift is the offset of the images horizontally and the height Shift is the offset of the images vertically in our case it is 10% for both cases. The zoom range is 20 percent randomly, and finally, the images were reversed horizontally.

methods	setting
Rotation range	30
Width shift	0.1
Height shift	0.1
Rescale	1/255
Shear range	0.2
Zoom range	0.2
Horizontal flip	true

Table 1: Settings for the image augmentation.

3.2.2 Proposed model: The proposed model in this work classifies normal and abnormal lung images showing various manifestations of pneumonia using a CNN architecture as shown in Figure 4. This architecture is composed of two large blocks, the first for the extraction of features where three input layers of 32 neurons were used, and three output layers of 64 neurons followed by layers of RELU activator to eliminate the negative values and replaced them with zeros, MaxPooling 2 x 2, Batch Normalization between them. Then we have 3 fully connected layers for the classification phase of 512, 256, 128 neurons and a two output using the Sigmoid function which allows us to get out a probability. Each neuron in the same layer is connected to all the neurons in the next block. After and to avoid the overfitting of our model, we took into consideration the Dropout layer defined at 50%. We used a dataset of 5840 images of chest X-rays (normal and abnormal classes) where 5216 images were used for the training phase (70 epochs) and the rest for the test phase. The output of the operations of convolution, max-pooling and Batch Normalizations are assembled into 2D planes called feature maps, and we acquired 248 248 32, 122 122 32, 59 59 32, 27 27 64, 11 11 64 and 3 3 64 sizes of feature maps, respectively, for the convolution operations and 124 124 32, 61 61 32, 29 29 32, 13 13 64, 5 5 64, and 1 1 64 sizes of feature maps from the pooling operations, respectively (See Table 2).

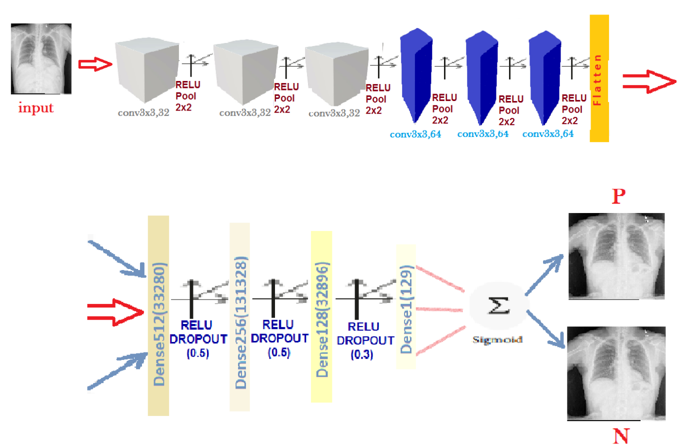

Figure 4: The proposed architecture.

Layer (type)	Output shape	Turtles
conv2d_1	*None, None, 248, 248, 32*	896
max_pooling2d_1	*None, 124, 124, 32*	0
batch_normalization_1	*None, 124, 124, 32*	128
conv2d_2	*None, 122, 122, 32*	9248
m_pooling2d_2	*None, 61, 61, 32)*	0
batch_normalization_2	*None, 61, 61, 32*	128
conv2d_3 (Conv2D)	*None, 59, 59, 32*	9248
max_pooling2d_3	*None, 29, 29, 32*	0
batch_normalization_3	*None, 29, 29, 32*	128
conv2d_4 (Conv2D)	*None, 27, 27, 64*	18496
max_pooling2d_4	*None, 13, 13, 64*	0
batch_normalization_4	*None, 13, 13, 64*	256
conv2d_5 (Conv2D)	*None, 11, 11, 64*	36928
max_pooling2d_5	*None, 5, 5, 64*	0
batch_normalization_5	*None, 5, 5, 64*	256
conv2d_6 (Conv2D)	*Lowest None, 3, 3, 64*	36928
max_pooling2d_6	*None, 1, 1, 64*	0
batch_normalization_6	*None, 1, 1, 64*	256
flatten_1	*None, 64*	0
dense_1	*None, 512*	33280
Dropout_1 (Dropout)	*None, 512*	0
dense_2	*None, 256*	131328
Dropout_2 (Dropout)	*None, 256*	0
dense_3 (Dense)	*None, 128*	32896
Dropout_2 (Dropout)	*None, 128*	0
dense_4 (Dense)	*None, 1*	129

Table 2: The output of the proposed network architecture

4 RESULTS AND DISCUSSION

As mentioned above, several methods such as data augmentation, learning rate variation, and annealing have been used to solve the problem of adapting a small data set into the deep convolutional neural network architecture [15]. The images sizes to be processed are fixed, so to evaluate the validation performance of our model we resized the radiographic images to 200x200x3, 250x250x3, 300x300x3, respectively, and we have trained them approximately four hours for each. The obtained results showed that the shape of the 250x250 image gave a better result than the others with an validation accuracy of 93.24% and training loss of 11,81% against of val-acc=90.71%, val-loss= 12.74% and val-acc=89.9%, val-loss=28.68% for 200 x 200,3; 300 x 300,3, respectively, as shown Table 3 and Figure 5 .

Data size	Training accuracy	Validation accuracy
200	95.72%	90.71%
250	**95.81%**	**93.24**
300	94.6%	89.9%

Table 3: Performance of the classication model on different data sizes.

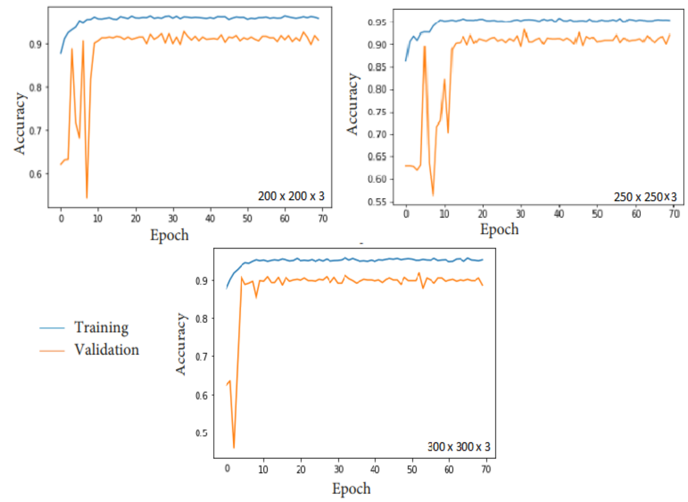

Figure 5: Performance of the classification model on varied data sizes.

Then to evaluate the results obtained for the dimensions of 250 x 250 x 3, four important metrics were considered, which are accuracy eq(1), sensitivity eq(2), specificity eq(3) and F-measure eq(4) . Our classification model has proven to be effective compared to the work of (Kermany et al., 2018). They obtained 92.8% of accuracy with a sensitivity of 93.2% and 90.1% of specificity and the current work with less parameters has an average accuracy of 93,24% with 95% of sensitivity, 91.4% of specificity , and F-Measure of 92.96% Table 4 .

$$\mathrm{ACC} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{P}+\mathrm{N}} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}} \quad (1)$$

$$\mathrm{S}N = \frac{\mathrm{T}P}{\mathrm{T}P+\mathrm{F}N} \quad (2)$$

$$\mathrm{S}P = 1 - \frac{\mathrm{F}P}{\mathrm{F}P+\mathrm{T}N} = \frac{\mathrm{T}N}{\mathrm{T}N+\mathrm{F}P} \quad (3)$$

$$\mathrm{F} - Meausure = \frac{2 \cdot precision \cdot recall}{precision + recall} \quad (4)$$

Authors	ACC	SN	SP	F-Measure	Amount of hyper-parameters
Kermany et al	92.8%	93.2%	90.1	–	23 million
Our model	**93.24%**	**95%**	**91.4%**	**92.96**	**0.31 million**

Table 4: Comparison of results

In this paper, we have proposed a model for the classification of Chest X-ray images of frontal views with promising validation accuracy. In the beginning, the algorithm minimizes the size of these chest images, then we used the data augmentation method to reduce the risk of overfitting. The next step is to apply this data to a convolutional neural network to extract the characteristics of the images for the identification and the classification. Due to the robustness of CNN, the accuracy of our model was encouraging and promising compared to other approaches as shown in Table 4 .

From this result we have demonstrated that to achieve good performance in this dataset, we don't need to use a complex architecture of CNN such as VGG16 [5], ResNet[13] or AlexNet[4], because this database is limited and deeper learning can cause overfitting and also it costs a lot of time for the training step.

The validation of the performances of the model trained on different radiographic image size is based on the variation of the sizes of the training and validation data set and we acquired relatively similar results. This will greatly contribute to improving the health of children by preventing and detecting pneumonia early.

5 CONCLUSION

The diagnosis of pathologies through chest X-rays is an independent practice in hospitals. This article presents one of the most recent techniques to accelerate the diagnosis of pneumonia, which is deep learning and more specifically convolutional neural networks (CNN) for image classification and recognition. The proposed model was developed taking into account the criteria of parameters and the time cost of the training step, which makes it different from other approaches that are based on transfer learning or which use a more complex architecture. One of the challenges encountered of using this is to find a suitable classification architecture, thus this work presented a method to generate and test different models of convolutional neural networks with different dimensions. The results were very satisfactory for this set. surpassing the accuracy, sensitivity, and specificity of the datasets author while showing a comparably lower necessity of computational power. This last point makes the model more used in sub-optimal situations.

In the future, we wish to improve the results obtained by using segmented images to reduce the computational effort and to facilitate the feature extraction process.

REFERENCES

[1] Esteva, A.; Robicquet, A.; Ramsundar, B.; Kuleshov, V.; DePristo, M.; Chou, K.; Cui, C.; Corrado, G.; Thrun, S. ; Dean, J. A guide to deep learning in healthcare Nature medicine, Nature Publishing Group, 2019, 25, 24

[2] He, K. ; Zhang, X. ; Ren, S. ; Sun, J. Deep residual learning for image recognition Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, 770-778

[3] Krizhevsky, A. ; Sutskever, I. ; Hinton, G. E. Imagenet classification with deep convolutional neural networks Advances in neural information processing systems, 2012, 1097-1105

[4] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, et al., Chexnet: radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv:1711.05225..

[5] M. Fiszman, W. W. Chapman, S. R. Evans, and P. J. Haug, Automatic identification of pneumonia related concepts on chest x-ray reports., in Proc. of the AMIA Symposium, p. 67, American Medical Informatics Association, 1999.

[6] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, et al., Chexnet: Radiologistlevel pneumonia detection on chest x-rays with deep learning, ArXiv preprint arXiv:1711.05225, 2017

[7] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R.M. Summers, Chestx-ray8: hospitalscale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 20972106.

[8] J. Irvin, P. Rajpurkar, M. Ko, Y. Yu, S. Ciurea-Ilcus, C. Chute, H. Marklund, B. Haghgoo, R. Ball, K. Shpanskaya, et al., Chexpert: a large chest radiograph dataset with uncertainty labels and expert comparison. arXiv:1901.07031..

[9] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: towards real-time object detection with region proposal networks, Adv. Neural Inf. Process. Syst. (2015) 9199.

[10] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 34313440

[11] K. He, G. Gkioxari, P. Dollr, R. Girshick, Mask r-cnn, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 29612969.

[12] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama, et al., Speed/accuracy trade-offs for modern convolutional object detectors, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 73107311.

[13] A. Shrivastava, R. Sukthankar, J. Malik, A. Gupta, Beyond skip connections: topdown modulation for object detection. arXiv:1612.06851..

[14] D. K. Kermany and M. Goldbaum, Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification, Mendeley Data, London, UK, 2018.

[15] Stephen, Okeke, et al. "An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare." Journal of healthcare engineering 2019 (2019).

[16] Simonyan, K. ; Zisserman, A. Very deep convolutional networks for large-scale image recognition arXiv preprint arXiv :1409.1556, 2014

[17] K. Kallianos, J. Mongan, S. Antani, T. Henry, A. Taylor, J. Abuya, M. Kohli, How far have we come? artificial intelligence for chest radiograph interpretation, Clin. Radiol..

[18] Roth, H.R., Lu, L., Seff, A., Cherry, K.M., Hoffman, J., Wang, S., Liu, J., Turkbey, E., Sum-mers, R.M.: A new 2.5 d representation for lymph node detection using random sets of deepconvolutional neural network observations. In: Medical Image Computing and Computer-Assisted InterventionMICCAI 2014, pp. 520527. Springer (2014).