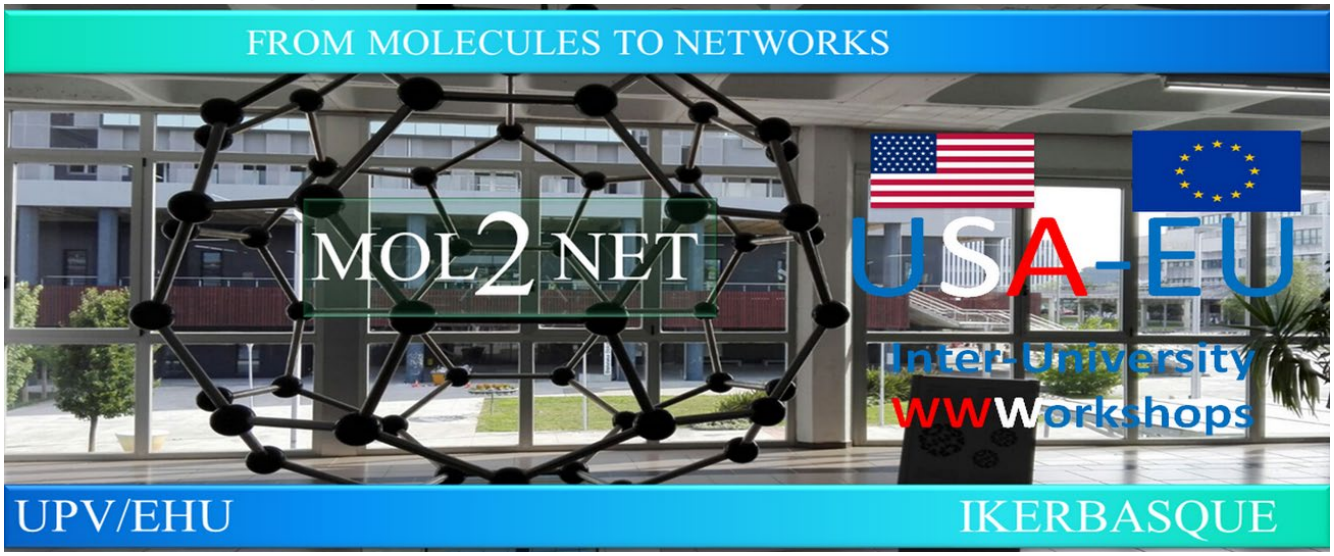
MDPI sciforum

MOL2NET, International Conference Series on Multidisciplinary Sciences

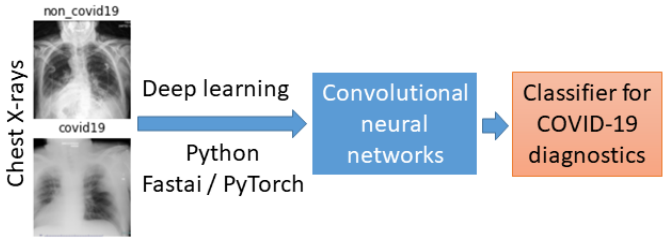


COVID-19 prediction using chest X-ray medical images and Deep Learning

Nuria Pereira Espasandin [a,], David Maseda Neira [a], Diana Marcela Noriega Cobo [a], Iago Iglesias Corrás [a], Alejandro Pazos [a,b,c], Julián Dorado [a,b], Cristian Robert Munteanu [a,b,c]*

[a] *Computer Science Faculty, University of Coruña, 15071 A Coruña, Spain*
[b] *Centro de Investigación en Tecnologías de la Información y las Comunicaciones (CITIC), Campus de Elviña s/n 15071 A Coruña, Spain*
[c] *Biomedical Research Institute of A Coruña (INIBIC), University Hospital Complex of A Coruña (CHUAC), 15006, A Coruña, Spain*

Graphical Abstract



Abstract

The current project is proposing a simple, fast and free solution to predict COVID-19 using a chest X-ray image as inputs and python scripts with deep learning from Fastai package. This is a prototype classifier to be improved and implemented as Web and mobile apps.

Introduction

The current evolution of COVID-19 [1,2] determines the need for fast and cheap solutions for the diagnostics. The majority of the patients with this disease suffers of severe lung problems and therefore the chest X-ray is the main medical test to check the disease evolution. Sometimes is difficult to say that a patient has COVID-19 using only the medical images because the images could be similar with other lung pathologies. Therefore, only an Artificial Intelligence method such as Deep Learning (DL) [3] could find complex patterns between the image pixels in order to make difference between COVID-19 and non-COVID-19 or other pathologies chest X-rays. The advantage of DL is that it is not needed to know which features in the input images are good for classification. Deep learning does the feature extraction for you (there is no need to have medical knowledge in the field). Several scientists published academic papers, free online datasets or scripts for CODIV-19 classifiers [4,5].

This project is proposing to find the simplest deep learning topology for this task by using python/Fastai. The future plan is to improve it with more medical images, to include it into free code repository and to implement it as Web and mobile free applications.

Materials and Methods

The medical images dataset was obtained from Minaee *et al.* [4]: 183 images for COVID-19 and 366 images for non-COVID-19 / other pathologies. The dataset was split into training and validation subsets (see Figure 1).

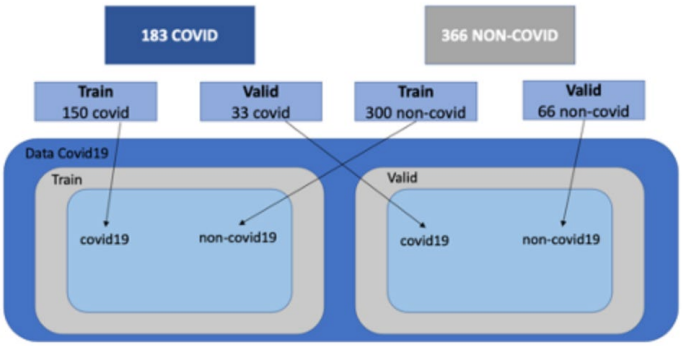


Figure 1. COVID-19 dataset composition.

In order to obtain the best classifier for COVID-19 prediction, Deep learning methodology has been used from python with the help of Fastai (based on pytorch). The dataset was normalized and extra data augmentation has been applied. Transfer learning techniques for training the classifiers have been used with different pre-trained networks. Thus, only the fully-connected part of the network is trained very fast. In addition, better classifiers have been trained with fine-tuning techniques where convolutional blocks are trained too. Five network topologies of Convolutional Neural Networks [6] have been trained: VGG19 BN [7], Resnet 34, Resnet152 [8], Squeezenet 1.1 [9], and Densenet 169 [10]. The metrics for training was based on accuracy and recall (random seed=2, batch size=16).

Results and Discussion

All DL topologies are able to build a DL classifier for COVID-19 / non-COVID-19 medical images by using pre-trained networks. The initial weights were obtained by training mentioned topologies with ImageNet [11]. The transfer learning method demonstrates its power by re-using learned convolutional

blocks. These blocks have been trained for other classes and represents basic image patterns that will be used in our classification by training only the fully-connected part of the networks for our particular task. Figure 2 shows for the models reads a sample of matrices of pixel inputs.

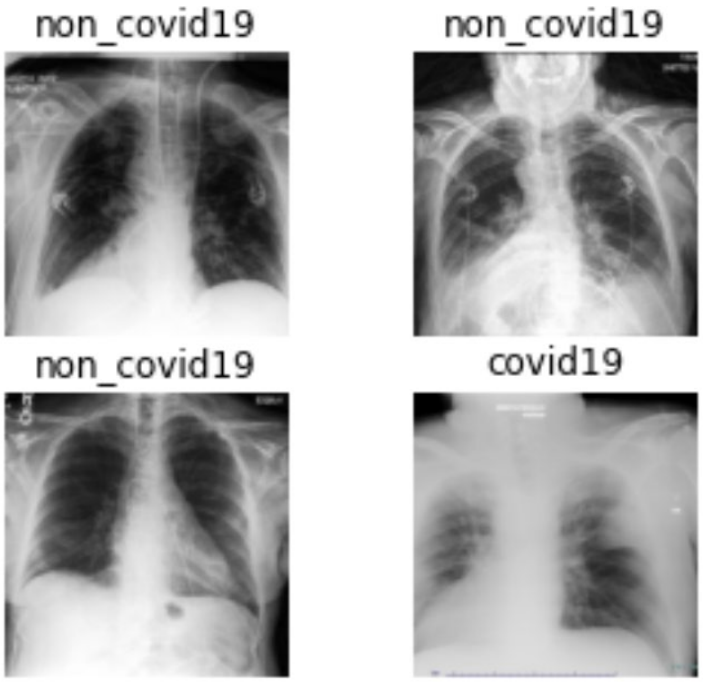


Figure 2. COVID-19 dataset samples read by deep learning algorithms.

In few epochs, all networks demonstrated the capacity to build classifiers for COVID-19 diagnostics in chest X-ray images with accuracy and recall of 1.00. Even with these perfect results, some topologies demonstrated more variability during the training and overfitting of the trained network. The best classifier was the one with the minimum variance during the training, the minimum number of batches to accuracy of 1.0, and the minimum number of parameters.

Thus, the best classifiers to predict COVID-19 in chest X-rays was provides by the Squeezenet 1.1.

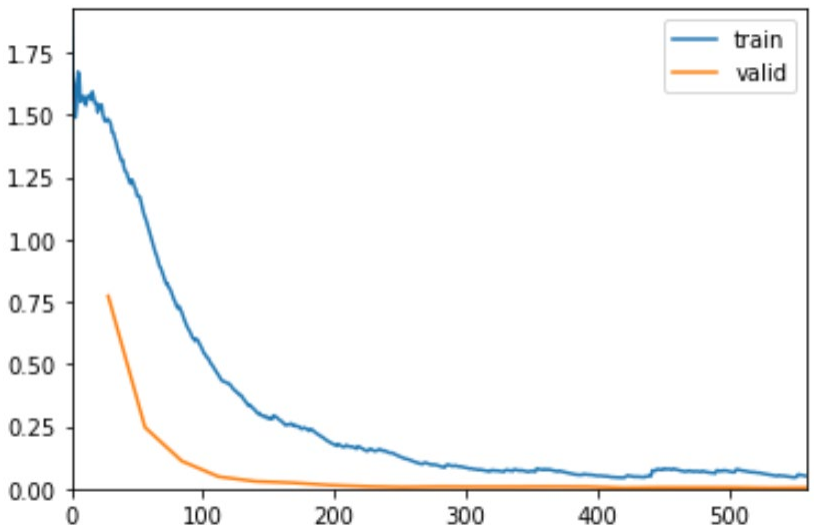


Figure 3. Error for training and validation subsets during the training of Squeezenet 1.1 with COVID-19 / non-COVID-19 chest X-rays

This network is extremely small compared with the other ones (number of trainable parameters) and it is very suitable for its use in small devices. With 50x fewer parameters compared with basic topologies such as AlexNet and less than 0.5 MB model size, Squeezenet is the perfect solution to create in the future a Web and mobile app that needs very low resources (RAM, CPU, GPU). Figure 3 presents the evolution of training and validation error during the training.

All the results for the rest of network topologies will be presented into an extended publication. The python scripts to build all the classifiers, the errors, all the plots and other statistics will be available Jupyter notebooks into a free GitHub repository. The Web and mobile apps will be available for free to be used in any device.

Conclusions

This project demonstrated that with a small public dataset and with only few lines of code in python/Fastai it is possible to obtain a very small and fast deep learning classifier as a model prototype for future Web and mobile applications. It is possible to obtain very accurate classifiers with small dataset by using data augmentation and pre-trained deep neural networks, without being expert in COVID-19 chest X-ray medical images - DL becomes the expert to find the differences of complex pixel patterns.

References

1. Rezaei N. Coronavirus Disease - COVID-19. Springer; 2021.
2. Riegelman R. COVID-19 Global Lessons Learned: Interactive Case Studies. Jones & Bartlett Learning; 2020.
3. Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press; 2016.
4. Minaee S, Kafieh R, Sonka M, Yazdani S, Jamalipour Soufi G. Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning. Med Image Anal. 2020;65: 101794.
5. Zhu Z. Detecting Coronavirus Disease 2019 Pneumonia in Chest X-Ray Images Using Deep Learning. 2020.
6. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521: 436–444.
7. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556.
8. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
9. Iandola F.N., Han S., Moskewicz M.W., Ashraf K., Dally W.J., Keutzer K. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <0.5 mb model size. arXiv 2016, arXiv:1602.07360.
10. Huang G., Liu Z., Van Der Maaten L., Weinberger K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
11. Deng J., Dong W., Socher R., Li L., Li Kai and Fei-Fei Li. ImageNet: A large-scale hierarchical image database, 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, pp. 248-255.