25 Maggio 2022

Pneumonia Detection On X-Ray Images

An Artificial Intelligence research project by

Alberto Zerbinati, Stefano Binotto, Simone Mosco

Useful material

- Our code is on Github: https://github.com/AlbertoZerbinati/pneumonia-xray-detection
- We run our code on Colab for exploiting GPU acceleration.
- Kaggle dataset (5,863 images, normal vs. pneumonia): https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia.
- ChestX-ray14 dataset (112,120 images, 14 common diseases): https://paperswithcode.com/dataset/chestx-ray14
- Sample implementation guide: https://www.projectpro.io/article/deep-learning-for-image-classification-in-python-with-cnn/418.
- Video paper: AI Beats Radiologists at Pneumonia Detection | Two Minute Papers #2...
.
  - (Related paper: https://stanfordmlgroup.github.io/projects/chexnet/).
  - (Paper commentary: https://laurenoakdenrayner.com/2017/11/18/quick-thoughts-on-chestxray14-performance-claims-and-clinical-tasks/)
- Different paper: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7345724/
- Great info about the disease: https://www.nhs.uk/conditions/pneumonia/

- For the report, also use this paper, which is very well-written: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0256630#pone-0256630-g002

Notes from the first paper

- 121 layer CNN → ~7mln parameters.
- Not only pneumonia but also 13 other similar diseases.
- Pneumonia = inflammatory lung condition responsible for 1mln hospitalizations and 50k in the US. Serious illness.
- Training set: contains annotation on WHERE the pneumonia is detected, not only *NORMAL vs. PNEUMONIA*. >100k images for training.
- Performed better than radiologists in the Test set (on average). SUPERHUMAN performance.
- Performance measured via:
    - Sensitivity: true-positive rate
    - Specificity: true-negative rate
    - → I think this means they used AUROC
- Pneumonia performance: 0.7680 AUROC
- 2017.
- So this is very interesting because it can spot various diseases and indicate the location, and the reason for the output (explainability++).
- Limitations: no lateral view, but only frontal, no supplementary information on the patient (which would be used in general by doctors)
- Very cool-looking heatmaps are obtained by making a weighted sum of the feature maps of the last convolutional layer with the weight they are connected to the fully connected part.

- Better explanation of pneumonia detection: *Pneumonia is an infection in the lungs. This means a pathogen, like bacteria or a virus, has set up shop and is causing trouble. In most* *cases, trouble means swelling and pus. On a chest xray, this usually means you get fluid in the airspace (ignoring interstitial change for the moment). When the air (low density) is replaced by fluid (higher density), you get white clouds on the x-ray, otherwise known as consolidation.*

- Clinical relevance yes, but not really

Slides ideas

- A possible start line:

  "Pneumonia causes the death of around 700,000 children every year and affects 7% of the global population." *[https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7345724/#:~:text=Currently%2C%20chest%20X%2Drays%20are,available%20in%20many%20underdeveloped%20regions.]*

  "Pneumonia is [...] a very common disease. It may have various sources. The Kaggle dataset we used consists of pneumonia samples for bacterial and viral cases. It also contains fairly few samples and it is unbalanced (# desease >> # normal). The aim of this project is to develop a robust deep learning model from scratch on this limited amount of data. We all know that deep learning models are data-hungry but if you know how things work, you can build good models even in limited scenarios."

  Source:https://www.kaggle.com/code/aakashnain/beating-everything-with-depthwise-convolution

  Also in the start something about deep CNNs on images

  Stress on social benefits from

- *"[show NORMAL vs. PNEUMONIA] images"*
- *"X-Ray analysis is as of today one of the best tools available to doctors for diagnosing such a disease timely."*

- "We found quite a bit of papers on the matter, the oldest dating back to 2017, but some are very recent."
- *I will take care of data analysis and visualization at the end: Accuracy, AUROC, Confusion Matrix, precision, recall,. etc. Include images.*

Slides organization

- First part [Alberto]: introduce pneumonia, and explain why we did this research project. Show some data and stress on social impact. Mention related work etc.
- Second part [Stefano]: our model
- Third part [Simone]: results analysis and comparisons. Future work and summary.

Deliver

25/05/2022: Slides before the presentation on Moodle.

04/06/2022: Doc decribing the project on Moodle. Link the GitHub repo

Presentation

Alberto:

Good afternoon everyone. Today we are presenting our AI project which is about Pneumonia Detection on X-Ray Images.

After a quick introduction, we'll present our deep learning approach and finally the results that we got.

So, what is pneumonia? Well, pneumonia is an inflammatory condition affecting humans in their lungs' air sacs. These get filled up with fluid or pus, causing symptoms like cough, fever, or trouble breathing. Pneumonia is generally caused by a bacteria or virus infection and it can get very serious really fast.

So, why do we care so much about this particular condition? Well, as for 2019 alone, pneumonia caused the death of 2 point 5 million people around the world, among which 6 hundred thousands were children under the age of 5. So this is the most recent and accurate data we could find, but even though the trend seems negative over the years, unfortunately, COVID-19 may cause pneumonia as a side effect and so we can imagine the curves to level up in the more recent years.

With that said, how can we detect pneumonia? The best way to do it is by inspecting chest x-rays. So this is a sample of the input we provided to train our model. Here on the left we can see a normal patient, while on the right the lungs of an affected person. They might look very similar to our untrained eye, so the challenge was to build an intelligent model that could help classifying pneumonia cases.

So I'm leaving it to Stefano to tell us more about how we used deep learning to tackle our mission.

Report info

The project report should be named as follows:

Project – Zerbinati, Binotto, Mosco – Pneumonia Detection On X-Ray Images – 25/05/2022.pdf

It should be uploaded in the assignment activity called Project delivery. The maximum length of the report should be 2 pages.

The report should describe: 1) The idea of the project 2) How the project could help people 3) AI tools used in the project 4) Links to topics seen in other courses if any (or links to other papers).

Report

Introduction

Pneumonia is a respiratory disease that causes inflammation of the air sacs and pleural effusion, a condition in which fluids fill the lungs, hindering breathing. Pneumonia is primarily driven by bacteria or viruses and less commonly by fungi or parasites.

Pneumonia affects many individuals, especially in underdeveloped countries, where poor hygienic conditions facilitate the spread of the disease, and the scarcity of medical resources inhibits detection and cure.

Whenever not timely diagnosed, pneumonia can be deadly. World Health Organization data [https://www.who.int/news-room/fact-sheets/detail/pneumonia] show that in 2019 alone, pneumonia accounted for more than 15% of deaths in children under the age of five years and caused the death of more than 2.5 million people around the world.

Therefore, early diagnosis is essential for preventing the disease from becoming fatal, and any effort towards this goal can save lives.

Radiological examination of the lungs through X-rays is one of the best-accounted diagnosis techniques. This is mainly because it constitutes a non-invasive and relatively inexpensive exam that gives medical professionals thorough insights into the patient’s lungs. However, chest-X-ray examinations for pneumonia detection are prone to subjective biases, and human read mistakes.

With this work, we propose a new automated system leveraging deep learning for automatically detecting pneumonia infections. The proposed system is based on X-ray lung image inspection, and focuses on high accuracy prediction and granting of output explainability.

Deep learning for pneumonia detection is a well-established field of research, so the actual aim of this research project was…

“We used the Kaggle dataset of pneumonia samples. It contains few samples and is unbalanced (# desease >> # normal). This project aims to develop a robust deep learning model from scratch on this limited amount of data. We all know that deep learning models are data-hungry, but you can build good models even in limited scenarios if you know how things work.”

For our project we implemented a Convolutional Neural Network (CNN), which is a particular kind of network based on the visual cortex of human beings. In fact it’s mainly used for image processing tasks, for example image classification or object detection. The CNNs need a huge amount of data for the training process. In our case we found a Pneumonia dataset available on kaggle.com, with 2 classes, Pneumonia or Normal. This dataset has 2 main problems. The first problem is the small amount of images for the training process. In fact we only have 5000 images, which is not enough for a proper training. So we used a data augmentation method in order to create artificial images by applying random transformations to the original ones. This is the list of transformations we applied:

- rotation range 20%
- zoom range 20%
- width shift range 10%
- height shift range 10%
- horizontal flip
- vertical flip

The second main problem is the imbalance of the dataset. In fact, among the 5000 original images, only 1000 were labeled as ‘Normal’. In this way the model learns to identify only the most represented class. In order to overcome this problem we implemented a Weighted Loss Function. We used the binary cross-entropy loss function, which is one of the most common functions for binary image classification, and we applied a weight W(i) for every class i. This weight is higher than 1 for the least represented class, and lower than 1 for the most represented class. Basically we tell the classifier to focus the learning process on the least represented class.

Pneumonia is a respiratory infection caused by bacteria or viruses; it affects many individuals, especially in developing and underdeveloped nations, where high levels of pollution, unhygienic living conditions, and overcrowding are relatively common and inadequate medical infrastructure. Pneumonia causes pleural effusion, a condition in which fluids fill the lung, causing respiratory difficulty. Early diagnosis of pneumonia is crucial to ensure curative treatment and increase survival rates. Chest X-ray imaging is the most frequently used method for diagnosing pneumonia. However, the examination of chest X-rays is a challenging task and is prone to subjective variability. In this study, we developed a computer-aided diagnosis system for automatic pneumonia detection using chest X-ray images. We employed deep transfer learning to handle the scarcity of available data and designed an ensemble of three convolutional neural network models: GoogLeNet, ResNet-18, and DenseNet-121. A weighted average ensemble technique was adopted, wherein the weights assigned to the base learners were determined using a novel approach. The scores of four standard evaluation metrics, precision, recall, f1-score, and the area under the curve, are fused to form the weight vector, which in studies in the literature was frequently set experimentally, a method that is prone to error. The proposed approach was evaluated on two publicly available pneumonia X-ray datasets, provided by Kermany et al. and the Radiological Society of North America (RSNA), respectively, using a five-fold cross-validation scheme. The proposed method achieved accuracy rates of 98.81% and 86.85% and sensitivity rates of 98.80% and 87.02% on the Kermany and RSNA datasets, respectively. The results were superior to those of state-of-the-art methods and our method performed better than the widely used ensemble techniques. Statistical analyses on the datasets using McNemar's and ANOVA tests showed the robustness of the approach. The codes for the proposed work are available at https://github.com/Rohit-Kundu/Ensemble-Pneumonia-Detection.

Final Part

In the evaluation process of the model, the Confusion Matrix is a useful tool that allows us to visualize the performances. Indeed, from it we extracted the values of precision (93%), recall (87%) metrics and most important, the accuracy which was set to 88%.

Moreover, the ROC Curve plot indicates how precise the network is in distinguishing pneumonia X-ray images from the normal one. The area below the curve is close to the value 1, the perfect binary classification.

A better visualization of what the model has learned is provided by the heatmap structure, generated by the last convolutional layers' weights. The information represented, according to the red and yellow pixels, seems to describe more cloudy and less clean areas of the lungs when dealing with pneumonia images, compared to the normal ones.

Finally, an overview on the development process and a comparison with other models took place. As expected, an increasing complexity and a larger number of layers in the network, led to better performances and an greater accuracy value, starting from 69%, going to 88%. Even a custom implementation of the AlexNet model [https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf] for the binary classification task could not reach the performances of our network.

Actual Report

Introduction

Pneumonia is a respiratory disease that causes inflammation of the air sacs and pleural effusion, a condition in which fluids fill the lungs, hindering breathing. It is primarily driven by bacteria or viruses and less commonly by fungi or parasites. Pneumonia affects many individuals, especially in underdeveloped countries, where poor hygienic conditions facilitate the spread of the disease, and the scarcity of medical resources inhibits detection and cure.

Whenever not timely diagnosed, pneumonia can be deadly. World Health Organization data *[https://www.who.int/news-room/fact-sheets/detail/pneumonia]* show that in 2019 alone, pneumonia accounted for more than 15% of deaths in children under the age of five years and caused the death of more than 2.5 million people around the world.

Therefore, early diagnosis is essential for preventing the disease from becoming fatal, and any effort towards this goal can save lives.

Radiological examination of the lungs through X-rays is one of the best-accounted diagnosis techniques. This is mainly because it constitutes a non-invasive and relatively inexpensive exam that gives medical professionals thorough insights into the patient’s lungs. However, chest-X-ray examinations for pneumonia detection are prone to subjective biases, and human read mistakes.

With this work, we propose a new automated system leveraging deep learning for automatically detecting pneumonia infections. The proposed system is based on X-ray lung image inspection, and focuses on high accuracy prediction and granting of output explainability.

Why deep learning

Having to deal with an image classification task, we decided to implement a Convolutional Neural Network (CNN). This particular type of classifier is inspired by the human visual cortex and was frequently used in the literature for image-related tasks. CNNs can deliver impressive performances, but they need a massive amount of data for the training process. This is to the first challenge we faced: finding a suitable dataset from which we could extract knowledge.

Dataset

We adopted a pneumonia dataset available on the Kaggle platform [https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia]. This dataset contains images labeled with two classes: `PNEUMONIA` and `NORMAL`. We evaluated the dataset as having a good quality and good integration with our tools, mainly Google Colablоratory [ ].

This dataset presented two main problems: the small number of images and the imbalanced classes.

The dataset contains only 5000 images, which is insufficient for properly training any deep CNN model we designed, considering that we wanted to split the images in a train-test-validation fashion. On top of that, among the 5000 original images, only 1000 were labeled as `NORMAL`. In this way, the model tended to identify only the most represented class.

Data Preprocessing

To overcome the issues above, we applied two well-known preprocessing techniques: we used data augmentation to create artificial images by applying random transformations to the original ones, for example, by rotating, zooming, or flipping the images; also, we provided a weighted loss function to our learning algorithm, to weigh more accurate and generalized learning of the least represented class.

Proposed CNN architecture

The final proposed CNN architecture uses 19 layers organized classically. A fully connected two-layer network follows a series of convolutional rounds. The code at this link can provide further details.

Performance analysis

To further test the robustness of the proposed model, we obtained a confusion matrix and calculated the accuracy, precision, recall, and area under the ROC curve (AUROC) score. The proposed model reaches 88% accuracy, 93% precision, and 87% recall.

Finally, we compared some of our CNN architecture trials. As expected, increasing complexity and more layers in the network led to better performance scores. Still, this was possible only thanks to the described preprocessing stage. We also compared the accuracy score with some well-established ones in the literature, for example, the general-purpose AlexNet [ ] or the pneumonia finely-tuned CheXNet [ ], which the proposed model can outperform in some instances.

Model explainability

Too often, deep learning leads to black-box agents. These cannot justify their output, which prevents their applicability, especially in the medical field. Considering the importance for a patient of the output of the proposed predictor, we decided to understand better what the model is considering when making a diagnosis. To do so, we implemented a visualization of the heatmap structure generated by the last convolutional layers' weights. The red and yellow areas in these activation maps seem to describe the portions of the lungs where possible pneumonia effects, like pleural effusion, are visible via X-rays.

References