Evaluating the Resnet-50 model on the COVID-QU-Ex Dataset

Le Mau Minh Phuc - BI12 351

March 19, 2024

Abstract

This study evaluates ResNet50's performance on the COVID-QU-Ex dataset, comprising 33,920 chest X-ray images categorized into COVID-19, non-COVID infections (Viral or Bacterial Pneumonia), and normal cases. Ground-truth lung segmentation masks accompany the dataset. We assess ResNet50's efficacy in multi-class classification and explore its performance with lung segmentation masks to enhance accuracy and interpretability. Evaluation metrics include accuracy, precision, recall, and F1-score, alongside qualitative analysis. Results demonstrate ResNet50's capability in identifying COVID-19 cases and emphasize the importance of lung segmentation masks in improving classification accuracy. This study contributes to the development of deep learning-based diagnostic tools for respiratory diseases, notably during the COVID-19 pandemic.

1 Introduction

The COVID-QU-Ex dataset is an incredible collection of 33,920 chest X-ray (CXR) images painstakingly put together by researchers from Qatar University. The objective of this dataset is to help advance research in the field of respiratory diseases, with a specific focus on COVID-19.

When it comes to medical diagnostics, chest X-rays are a commonly used tool to assess condition of lungs and provide valuable insights into respiratory infections like COVID-19. The COVID-QU-Ex dataset serves as a valuable resource for researchers, offering a vast array of chest X-ray images to develop and evaluate machine learning algorithms, computer-aided diagnostic systems, and other innovative solutions driven by artificial intelligence. By leveraging these tools, medical professionals can potentially improve accuracy and speed of COVID-19 and respiratory illness diagnoses, leading to better outcomes for patients and more effective public health interventions.
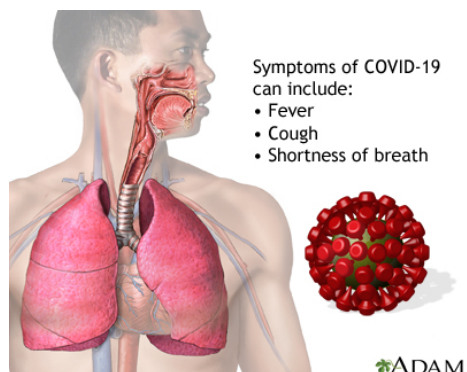

Figure 1: COVID-19

Having comprehensive datasets like COVID-QU-Ex is of utmost importance for researchers. Researchers can use these datasets to train and validate AI models, paving the way for creation of robust and reliable tools to diagnose and manage respiratory diseases. Moreover, this dataset holds potential to deepen understanding of radiological manifestations of COVID-19, helping researchers identify specific patterns or features associated with disease.

It's important to remember that while chest X-rays provide valuable information, a definitive diagnosis of COVID-19 typically requires additional tests such as polymerase chain reaction (PCR) tests or antigen tests, which directly detect presence of SARS-CoV-2 virus or its genetic material.

The COVID-QU-Ex dataset represents a significant contribution to scientific community's relentless efforts in combating COVID-19 pandemic. By fostering research and innovation in diagnosis and treatment of respiratory diseases, it plays a crucial role in collective fight against this global health crisis.

2 Dataset - COVID-QU-Ex

This annotation process guarantees good precision and dependability in examining lung pathology, representing a notable advancement in medical image segmentation research. The COVID-QU-Ex dataset, in particular, marks a significant milestone, being the most extensive collection of lung masks ever compiled.

Comprising 21,715, 5,417, and 6,788 images, respectively, the training, validation, and test datasets are divided into three main categories: COVID-19, non-COVID (which may encompass other infection types), and normal cases.

Table 1: Dataset Composition

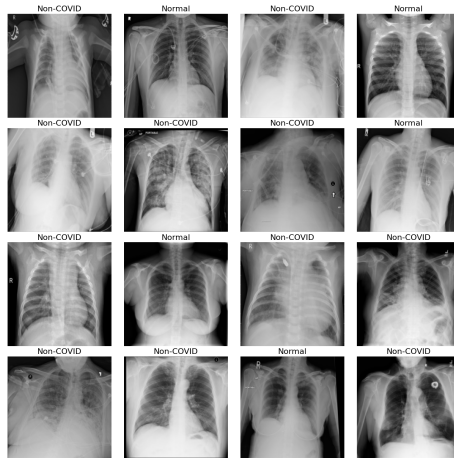|	**COVID-19**	**Non-COVID**	**Normal**
Training images	7658	7208	6849
Validation images	1903	1802	1712
Test images	2395	2253	2140

Figure 2: Parts of COVID-QU-Ex dataset

Notably, the COVID-QU-Ex dataset contains different cases of conditions, including 11,956 cases of COVID-19, 11,263 instances of non-COVID infections such as viral or bacterial pneumonia, and 10,701 normal cases. What truly sets this dataset apart is the inclusion of ground-truth lung segmentation masks for the entire repository, allowing for precise pixels of lung regions within each image. This dataset is an invaluable resource for medical professionals, data scientists, and researchers, providing critical insights into the wide range of lung pathologies.

3 ResNet-50

This report implements ResNet50, a powerful artificial neural network used for image recognition tasks. ResNet50 is able to accurately identify the contents of images by looking for patterns and recognizable objects.

What makes ResNet50 unique is its ability to train very deep neural networks effectively. Normally, making a neural network too deep can cause issues during training where the weights and biases start decaying and performance drops off. ResNet50 avoids this using "residual learning" which allows layers to bypass residual blocks as needed.

ResNet50 contains convolutional and pooling layers that process the raw image data into a form that the neural network can more easily understand. The core of the model is then a series of residual blocks that apply convolutional filters to detect lower-level features like edges and shapes, then builds up higher-level concepts through each successive residual block.

Figure 3: ResNet-50 Achitecture

After going through all the residual blocks, the final feature maps are averaged together using global average pooling. This averaging step helps make the model more robust. The flattened features are then passed through one final fully-connected layer that produces probability scores for each possible output category that could be depicted in the image.

For this implementation, ResNet50 was configured appropriately for the given image classification task, using pre-trained weights as a starting point before fine-tuning on the specific dataset. Standard practices were followed for the optimization algorithm, loss function, performance metrics, and other training settings.

The model summary confirms the architecture contains the expected residual blocks, convolutions, and connections. The total number of trainable and non-trainable parameters indicates it is a fairly large and capable model while still being practical to train on modern hardware.

4   Model Evaluation

The results of ResNet50 on the COVID-QU-Ex dataset indicate a high level of performance in classifying chest X-ray images. Specifically, the model excels at identifying COVID-19 cases with a high true positive rate as seen in the confusion matrix, with 1823 correctly identified cases out of 1903, reflecting a high recall in this class. Non-COVID infections are also well-identified with 1636 correct classifications out of 1802, suggesting a balanced precision and recall, which is important for medical diagnostics to minimize both false positives and negatives.

Table 2: Precision, Recall, F1-Score and Support

	Precision	Recall	F1-score	Support
COVID-19	0.99	0.96	0.97	1903
Non-COVID	0.90	0.91	0.90	1802
Normal	0.89	0.91	0.90	1712
Accuracy			0.93	5417
Macro avg	0.93	0.93	0.93	5417
Weighted avg	0.93	0.93	0.93	5417

For Normal cases, the model shows a slight decrease in precision but maintains a high recall, correctly identifying 1559 out of 1712 cases. This performance is critical, as it reduces the risk of falsely diagnosing healthy individuals with an infection The overall accuracy across the dataset stands at 93%, which is commendable given the challenges in differentiating between COVID-19 and other types of pneumonia from X-ray images.
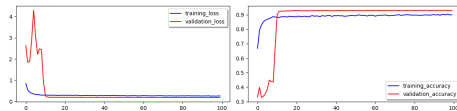

Figure 4: Training and Validation Loss

In Figure 4, the left plot demonstrates the progression of training and validation loss. Initially, there's a rapid decline in loss as the model trains, followed by a stabilization, indicating effective learning from the training data. Encouragingly, the validation loss mirrors this trend, decreasing alongside the training loss, suggesting the model's ability to generalize well to new data.

On the right side, the plot depicts the evolution of training and validation accuracy. Both metrics increase swiftly, with training accuracy reaching a high level and remaining steady thereafter. Similarly, validation accuracy rises rapidly and closely tracks training accuracy. This consistency between the two accuracy curves suggests minimal overfitting and the model's capacity to accurately classify both training and validation data.

The convergence of training and validation metrics in both plots, coupled with narrow gaps between them, signifies strong model performance and effective generalization to the validation dataset. However, the initial spikes in the graphs hint at some early training volatility, likely influenced by factors such as the chosen learning rate or the random initialization of model weights.
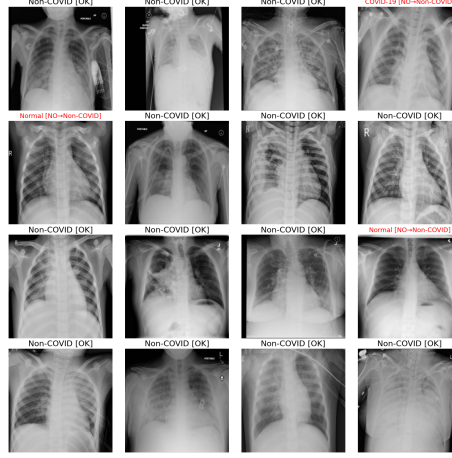

Figure 5: Result images

The majority of images in figure 5 are categorized as "Non-COVID [OK]," indicating that, according to the labeling, the individuals in the X-rays do not have COVID-19. A few of the X-rays are labeled as "Normal [NO→Non-COVID]," suggesting a normal finding not indicative of COVID-19 infection. There is one image labeled "COVID-19 [NO→Non-COVID]," possibly indicating a case initially suspected as COVID-19 but later determined to be non-COVID.

Chest X-rays serve as a valuable diagnostic tool, especially for respiratory illnesses like COVID-19, as they can reveal signs of pneumonia or other lung changes associated with the infection. However, it's essential to understand that diagnosing COVID-19 typically requires a comprehensive approach, including clinical evaluation, imaging, and testing (such as PCR or antigen tests), as various conditions besides COVID-19 can manifest lung changes visible on X-rays.

5 Conclusion

In summary, the COVID-QU-Ex dataset represents a pivotal advancement in medical imaging and artificial intelligence realms. Through its extensive array of chest X-ray images and precise lung segmentation masks, this dataset serves as a crucial resource for the development and assessment of machine learning algorithms aimed at diagnosing respiratory ailments, notably COVID-19.

Moreover, the assessment of the ResNet-50 model's performance on the COVID-QU-Ex dataset demonstrates encouraging outcomes, with notable precision, recall, and F1-scores across various categories. The model's adeptness in accurately identifying COVID-19 cases carries substantial implications for swift and efficient disease detection.

Collectively, the integration of the COVID-QU-Ex dataset with sophisticated deep learning models such as ResNet-50 contributes significantly to the global endeavor to combat the COVID-19 pandemic and bolster healthcare outcomes. By fostering collaboration among medical practitioners, data experts, and researchers, this dataset facilitates the exploration of inventive approaches to tackle respiratory illnesses, thereby enhancing patient care and enabling more impactful public health interventions.