Classification of Lung X-ray Images

Zixin Hao, 19th December 2023

1 Overview

This project aims to explore the effectiveness of identifying pneumonia and COVID-19 in the Edge Impulse platform based on the MobileNet model, using chest X-ray images as an analysis tool. Considering the global epidemic status caused by the COVID-19 virus and the problems of conventional detection methods that are labor-intensive, time-consuming and resource-limited, this study provides a new attempt and provides a reference for interested researchers to develop Tools to assist in diagnosing symptoms of lung disease. The technology supports rapid screening of large populations by reducing the burden of diagnosis by expert doctors [1]. The study applies transfer learning in deep learning technology and uses the MobileNet model to analyze X-ray images and identify specific patterns of disease tissue. Research results show that in identifying COVID-19, the model exhibits an accuracy of 60-70% under a specific data set, indicating that the model is expected to be an effective lightweight model that can be embedded in mobile phones. self-detection testing tool.

2 Project Background and Motivation

The preliminary diagnosis of pneumonia X-rays through machine learning is of great significance to improving the efficiency and accuracy of medical treatment. In X-rays, white represents high density and black represents low density. The following are two professional diagnostic reports for two images.

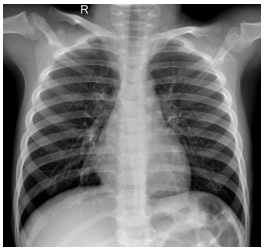

Figure 1: *A X-ray image for normal lung.*

Diagnostic report for Figure 1: There was no abnormal density shadow in the bilateral lung fields, the course of the lung markings was normal, the structure of the hilar shadow was normal, and there were no obvious abnormalities in the pleura, diaphragm, mediastinum and cardiac great vessels.

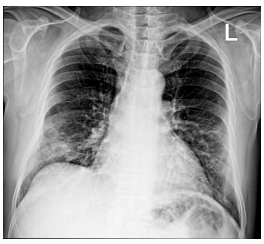

Figure 2: *A X-ray image for Covid-19 lung.*

Diagnostic report for Figure 2: The texture of both lungs is rough and blurred, and the hilum structure is disordered; patchy slightly high-density shadows can be seen in the middle and lower fields of both lungs near the subpleura with blurred edges; the permeability of the lung fields is increased, and local pleural hypertrophy and adhesions are observed; the trachea, bronchi and their The branches were unobstructed and their walls were partially thickened; the heart was active; other abnormalities were not found.

The doctor's initial diagnosis is: chronic bronchitis combined with emphysema; double pneumonia, please combine clinical and laboratory examinations

X-rays are relatively sensitive to bone imaging. Doctors cannot judge Covid-19 and other diseases using X-rays alone, and may lead to misdiagnosis due to interference from overlapping organs. In medicine, CT is used to diagnose Covid-19 more accurately. So, if X-rays can be used to initially determine the disease, redundant auxiliary medical steps (such as taking CT scans) can be reduced, and X-rays are relatively cheaper and faster, which can greatly improve the hospital's diagnostic efficiency. save costs.

3 Dataset

The public dataset can be found in roboflow [2]. It is resized into 416*416, and padded by reflecting method. There are three categories in total, namely: Covid-19, Normal and Pneumonia.

3.1 Data Analysis

Figure 3 lists all interference types. For these interferences, we have performed corresponding data processing.

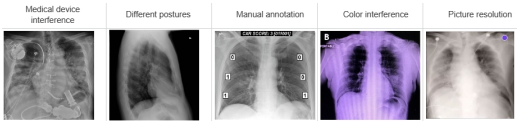

Figure 3: *Factors affecting accuracy in data sets*

First, we delete the pictures with large interference in the original data; then use Gaussian denoising to slightly denoise (the parameters used are k=3, sigma=1); and perform grayscale processing on the pictures. Through these preprocessing, we obtained relatively clean experimental data. Later experiments will mix and combine these data for comparative analysis.

4 Experiment

The platform the author used is the Edge Impulse. Edge Impulse is a leading development platform for machine learning on edge devices, enabling engineers to create intelligent device solutions with embedded Machine Learning. It provides an end-to-end workflow that includes data collection, processing, model training, and deployment. The platform supports a wide range of sensors and data types, making it versatile for applications in various industries. With its user-friendly interface, Edge Impulse allows both experts and novices to develop ML models efficiently. The models can run on low-power microcontrollers and are optimized for performance, making Edge Impulse a go-to platform for developing scalable ML solutions for the Internet of Things (IoT) ecosystem.

4.1 Baseline-Model Description

The baseline-model used in our experiments is a random model, that is, the labels are randomly assigned. So the accuracy rate is about 33%.

4.2 Model

The table details four experimental cases, each utilizing different neural network architectures and configurations for image processing:

Case 1 uses MobileNetV2 with a 96x96 grayscale input. The learning rate is set to 0.005, and the final layer of the network consists of 8 neurons with a dropout rate of 0.15. The data is split into 80% for training and 20% for validation, with 20 training cycles and a batch size of 32.

Case 2 employs EfficientNet B0, also with a grayscale image input. It has a lower learning rate of 0.001, with the same validation size, number of training cycles, and batch size as Case 1.

Case 3 repeats the use of MobileNetV2 architecture with grayscale images but differs from Case 1 by having 32 neurons in the final layer while maintaining the same dropout rate, learning rate, validation size, training cycles, and batch size.

Case 4 is similar to Case 3, utilizing MobileNetV2 with a grayscale image input and 8 neurons in the final layer but has a slightly lower dropout rate of 0.1.

These models are likely part of a study to compare the performance of different neural network configurations in processing grayscale images for a specific application, such as medical imaging. The variations in architecture, learning rate, and neuron count in the final layer are typical parameters adjusted to optimize model performance for the task at hand.

4.3 Result

The result of Case 3 is the best, as shown in Figure 4

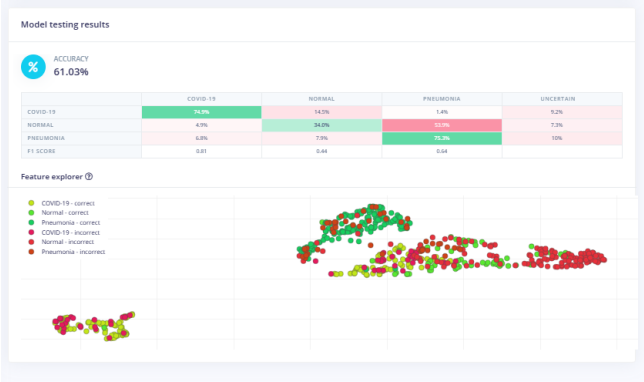

Figure 4: *Performance of case 3 in the original dataset*

In the Figure 4, the image presents the testing results of a model with an overall accuracy of 61.03%. It includes a confusion matrix and a feature explorer scatter plot. The confusion matrix shows the model's performance across four categories: COVID-19, Normal, Pneumonia, and Uncertain. The model is most effective at identifying Pneumonia with 73.5% accuracy, followed by COVID-19 at 74.9% accuracy for positive cases. However, it has a high misclassification rate for Normal cases as COVID-19 (34.0%) and a considerable uncertainty rate. The F1 Score is highest for COVID-19 (0.81), indicating better precision and recall, and lowest for Normal cases (0.44). The scatter plot below the matrix visualizes feature separability, with correct classifications marked in green and incorrect ones in red, suggesting moderate clustering but also overlap between classes, which correlates with the confusion matrix and supports the overall accuracy figure.

In the Figure 5, it depicts model testing results with an overall accuracy of 75.82%. The confusion matrix shows the model's classification performance for COVID-19, Normal, and Pneumonia cases, along with an Uncertain category. The model is highly accurate for Pneumonia detection at 82.5%, but less so for COVID-19, correctly identifying 45.8% of cases, with signif-

Cases	Pre-process	Architecture	Learning Rate	Validation Size	Training Cycles	Batch Size
1	image: grayscale	MobileNetV2 96*96 0.35 (final layer: 8 neurons, 0.15 dropout)	0.005	20%	20	32
2	image: grayscale	EfficientNet B0	0.001	20%	20	32
3	image: grayscale	MobileNetV2 96*96 0.35 (final layer: 32 neurons, 0.15 dropout)	0.005	20%	20	32
4	image: grayscale	MobileNetV2 96*96 0.35 (final layer: 8 neurons, 0.1 dropout)	0.005	20%	20	32

Table 1: *Experiment record, recording the adjusted hyperparameters*

Figure 5: *Performance of case 3 in the new dataset(pre-processed images)*

icant false negatives and misclassifications as Pneumonia (29.2%). Normal cases see better accuracy at 67.5%, with a low false positive rate for COVID-19 (0.2%). The F1 Scores reflect this distribution, being highest for Pneumonia (0.84) and lowest for COVID-19 (0.62). The feature explorer scatter plot illustrates the spread and overlap of the data points, with correct predictions in green and incorrect ones in red, highlighting the model's strengths and areas for improvement.

5 Challenges

In our study, we posit that the MobileNet model's proficiency in image classification tasks, particularly within the constraints of an embedded system, is commendable. We would like to delineate several prospective challenges and areas for future exploration:

Firstly, the inherent visual ambiguity in medical imagery poses a significant challenge due to the lack of distinct edges and textures, which complicates the differentiation between pathological and healthy tissues amidst extraneous background information.

Secondly, the necessity for uncertainty quantification in diagnostic models is paramount. Medical professionals require diagnostic tools that offer not only predictive insights but also confidence measures for each prediction, thus enabling focused analysis on areas of uncertainty.

Lastly, the limitations imposed by data-driven models are accentuated by the quality of the input data. An over-reliance on annotated data can lead to models that may overfit to non-essential features. The advent of semi-supervised learning approaches, which necessitate less annotated data, heralds a promising avenue for leveraging the growing volume of medical imagery, including unannotated datasets.

6 Conclusion

The report concludes that while the MobileNet model on the Edge Impulse platform shows promise in identifying pneumonia and COVID-19 from chest X-rays, with accuracy between 60-70%, there is significant room for improvement. The challenges highlighted involve dealing with the visual ambiguity inherent in medical imaging and the need for uncertainty quantification in diagnosis. The research underscores the potential of semi-supervised learning models to enhance future medical diagnostic processes, by learning from large datasets without extensive annotation. This direction could revolutionize the efficiency and cost-effectiveness of medical imaging diagnostics.

Acknowledgments

The authors would like to express their deepest gratitude to Professor Alex Rogers, Professor of Computer Science and Tutorial Fellow at St Anne's College, for his invaluable guidance and support throughout this research. His insights and expertise have been instrumental in the completion of this work. Furthermore, heartfelt thanks are extended to all the team members, whose relentless collaboration and dedication were pivotal in bringing this project to fruition.

7 Reference

[1] Civit-Masot, J., Luna-Perejón, F., Domínguez Morales, M., & Civit, A. (2020). Deep learning system for COVID-19 diagnosis aid using X-ray pulmonary images. Applied Sciences, 10(13), 4640.

[2] Roboflow. (n.d.). COVID-19 and Pneumonia scans. Retrieved [date you accessed the website], from https://public.roboflow.com/classification/covid-19-and-pneumonia-scans