Normal Vs. Abnormal classification using ROI identification

Jack Heuer

8/14/2023

Abstract:

The goal of this model was to be able to take in a wide range of chest x-rays and be able to classify them as normal or abnormal. This could aid a doctor who needs a preliminary screening for a second opinion. I used two machine learning models in my project, one for segmenting the lungs in order to create a ROI and another in order to diagnose the lungs. I was able to achieve a 98% accuracy in segmenting the lungs and a 95% accuracy in classifying between normal and abnormal lungs. In conclusion this model shouldn’t be used by itself or in place of a licensed doctor due to the accuracy not being high enough but it could be a useful tool in a first screening.

Introduction:

Around the world, a shortage of skilled radiologists has led to a critical issue: the misinterpretation of chest X-rays due to heavy workload and inadequate training. This problem holds dire consequences, with studies suggesting that over 30% of tuberculosis cases go undiagnosed on X-rays. This delay worsens patients' health and facilitates disease transmission.

To address this, focused efforts are essential. Training programs should be intensified, aiming to improve radiological proficiency among medical practitioners. For example, the ratio of radiologists to population in Africa is alarmingly low, for instance, only one radiologist per 1 million people in some regions. This scarcity calls for investment in medical infrastructure and increased recruitment. An automated classifier to help identify issues would be a huge benefit when there isn’t a radiologist around to recognise an issue

Method:

I used two datasets for training. Both datasets were composed of png images. The first dataset for which I used for Unet lung segmentation had 600 images and masks for training. For the second dataset I used 1250 normal and abnormal

The first priority was to create a bounding box around the lungs or an ROI(region of interest). First I had to change the pixel dimensions of each image to 112x112 due to memory constraints.I did this by detecting the lungs through a Unet model. The Unet architective was I was able to train my model to a dice score of around 97%. With more data I could have got the accuracy even higher, as seen in figure 1 the accuracy and loss flattens out around epoch 5 when the data has been overfitted. Places I lost accuracy were when the arms were visible in the x-ray, the x-ray was inverted in color, and when the x-ray was otherwise distorted. Other outlying data could be from x rays of babies or obesity. In order to create better lung masks I used CV’s erode and dilate functions in order to get rid of noise and not have any big floating islands.

With this x-ray mask I found the top and bottom of each lung and saved these points. Using these points I used CV’s convex hull in order to close up the lung area and capture all between these masks(Refer to figure 2 and then to figure 3). Using these chest masks I then selected which parts of the original images to keep (figure 4). Only letting my model look at the important parts of each xray would help to eliminate any confounding errors. A confounding error that may appear if I did not cut each x-ray image would be looking at the letter labeling of each x-ray. I believe this letter corresponds to the hospital at which the image was taken, however say all abnormal data was to come from one hospital and normal from another the machine might make its decisions based on the letters in the corners rather than the lung itself. A confounding error would be a serious problem because then on my testing data it would seem the model was working, but on outside data the model would suddenly stop working. Next I used this cut data and ran it through a tensorflow sequential model in order to get an accuracy of 94% detection of normal vs abnormal.

Results:

Training the Unet lung segmentation model on 566 training masks and images I was able to get a dice score of 97% and an AUROC of 95%. These numbers are before and post procession was done so the numbers might be a bit deflated. Once the images were processed an accuracy of 94% was achieved with the Normal Vs. Abnormal classification. The Model was much better at predicting on Abnormal data as it correctly predicted 96% of the time rather than 92% on Normal data. In 500 images tested 12 were incorrect on abnormal data and 16 were incorrect for normal data. In figure 5 you can see evidence of overfitting as I ran out of data.

Discussion:

With 5 out of 100 predictions being incorrect it is difficult to see this being used in the field, however it shows proof of concept and with more refinement of the photos selected for training I believe the model could have got up to an accuracy of 97%.

Ensuring global healthcare accessibility is imperative for equitable well-being. Universal access to healthcare can mitigate preventable deaths, reduce disease burdens, and enhance socio economic development. AI diagnostics can play a pivotal role in this endeavor by providing efficient, accurate, and scalable solutions, especially in regions with limited medical resources. By integrating AI-powered diagnostic tools, we can bridge gaps in healthcare delivery, enabling timely interventions and improved outcomes worldwide.

References:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3503514/>

<https://www.analyticsvidhya.com/blog/2022/10/image-segmentation-with-u-net/#:~:text=U%2DNet%20was%20developed%20in,to%20increase%20the%20model%20performance>.

<https://www.kaggle.com/datasets/nikhilpandey360/chest-xray-masks-and-labels>

<https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database>

Figure 1:

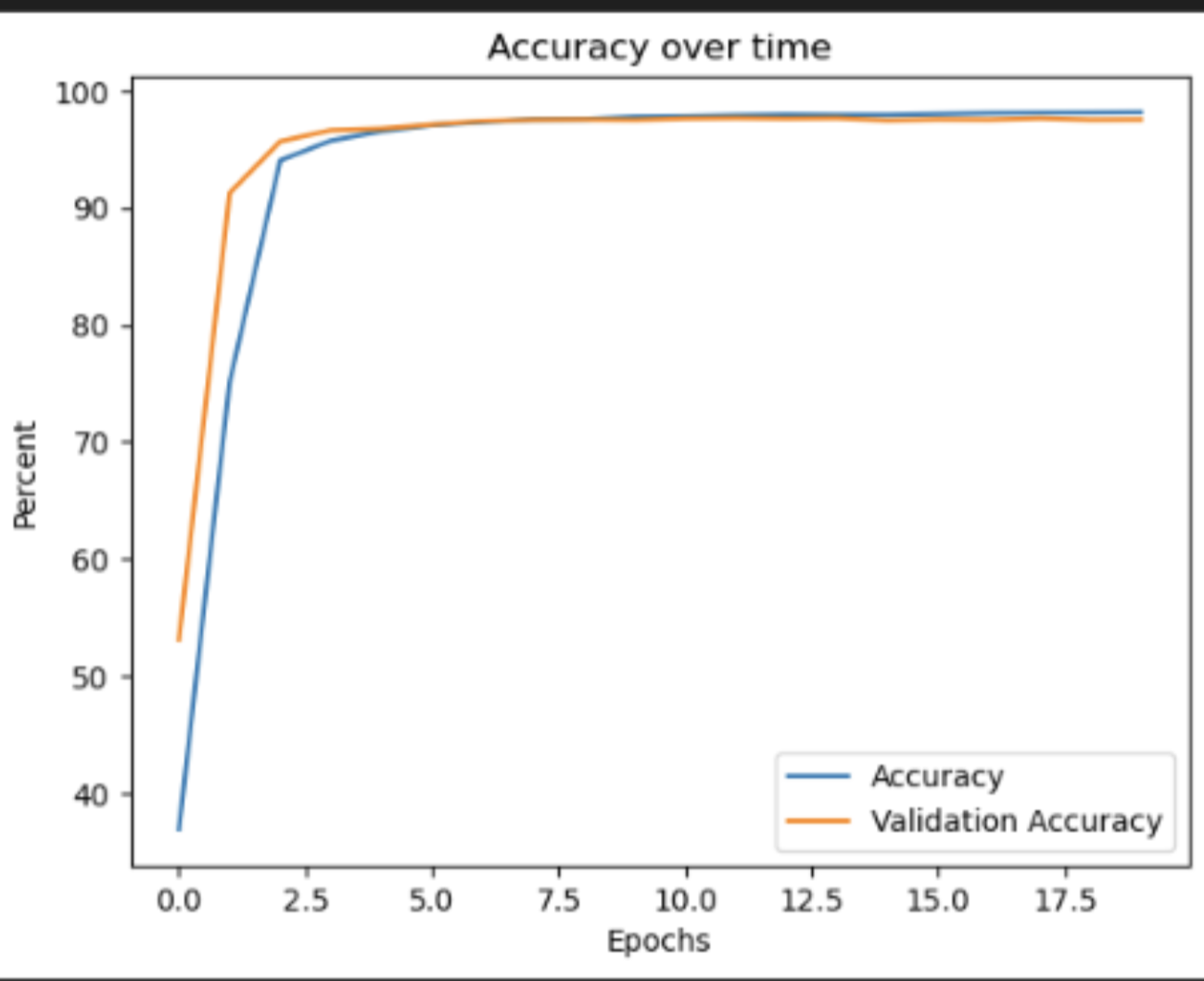


Figure 2:

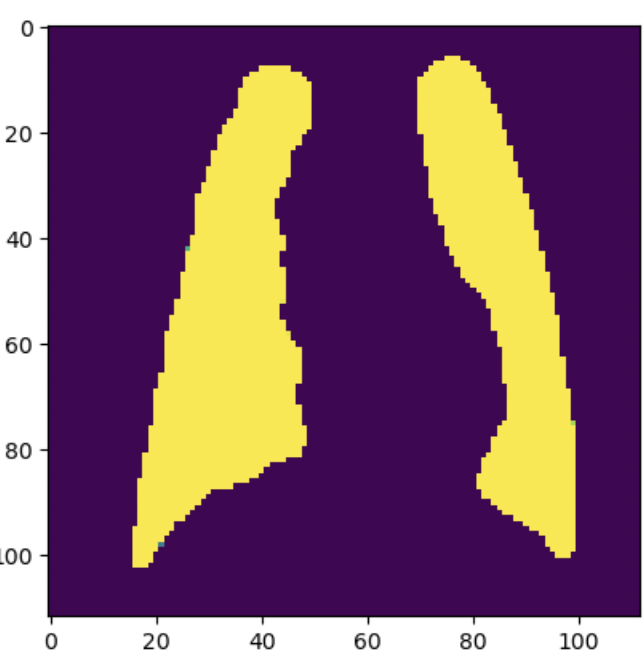


Figure 3:

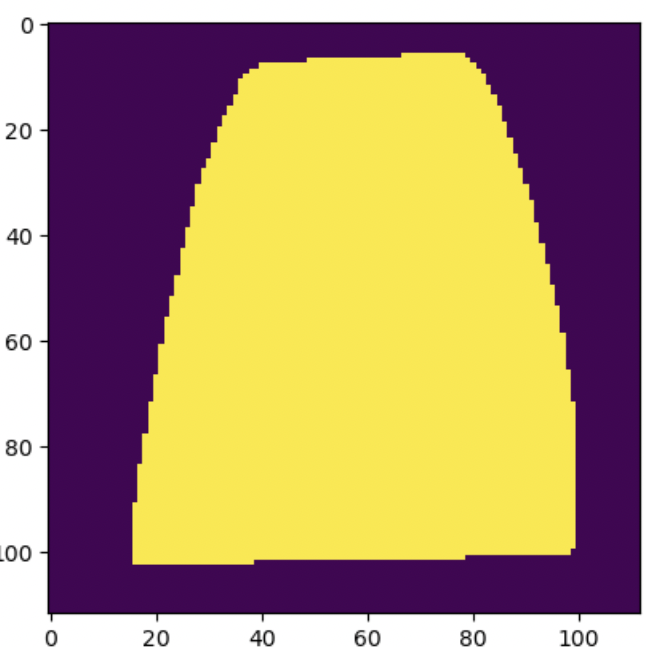


Figure 4:

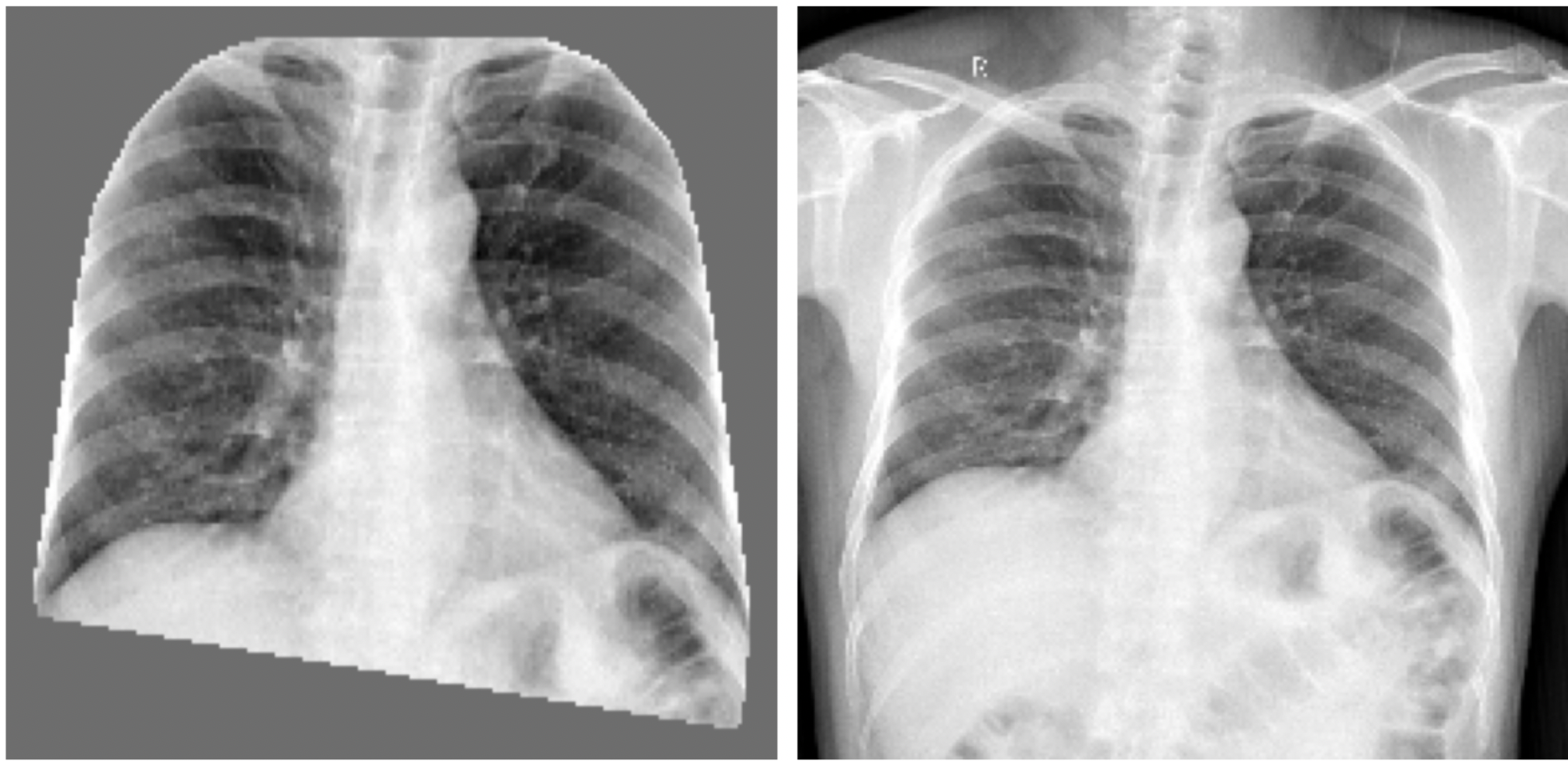


Figure 5:

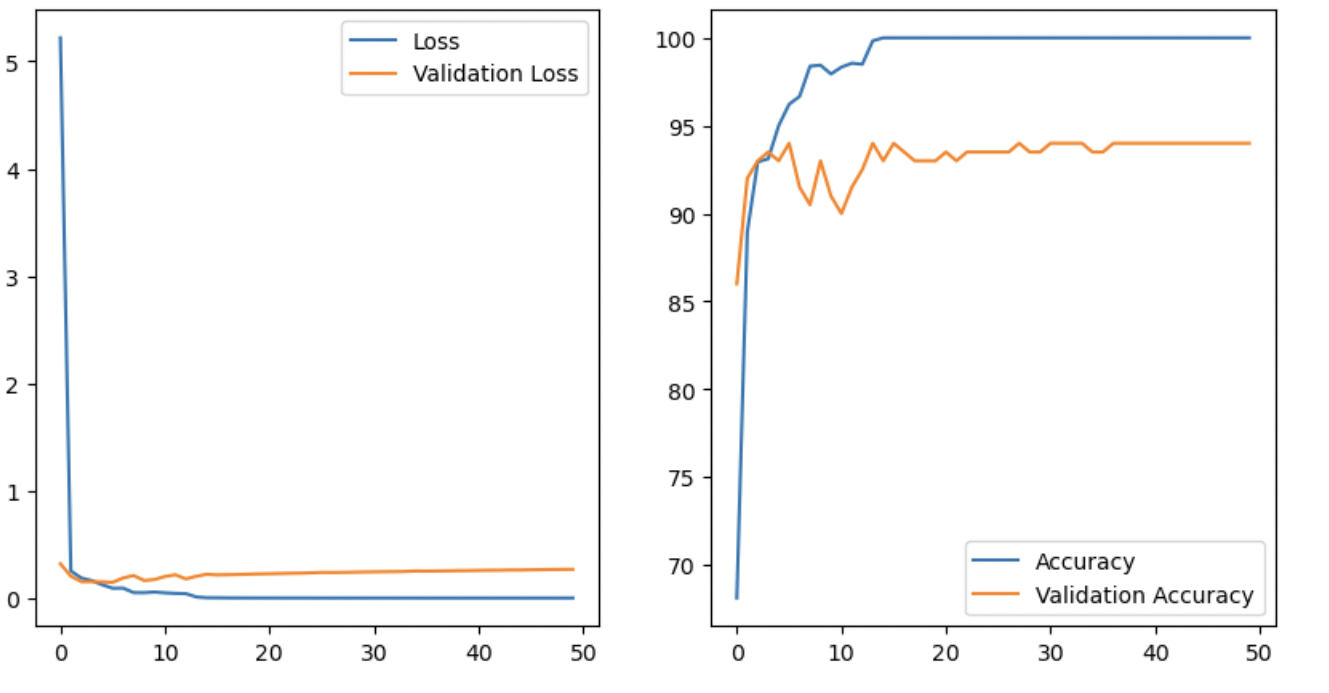


Figure 6:

