Detection of Pneumonia clouds in Chest X-ray using Image processing approach

Abhishek Sharma
Institute for Plasma Research, HBNI
Bhat,Gandhinagar,India
abhishek211919@gmail.com

Daniel Raju
Institute for Plasma Research, HBNI
Bhat,Gandhinagar,India
raju@ipr.res.in

Sutapa Ranjan
Institute for Plasma Research
Bhat,Gandhinagar,India
sranjan@ipr.res.in

***Abstract*—Finding ways to automate diagnostics from medical images, has continuously been one of the most interesting areas of software development. This article presents a novel approach for detecting the presence of pneumonia clouds in chest X-rays (CXR) by using only Image processing techniques. For this, we have worked on 40 analog chest CXRs pertaining to Normal and Pneumonia infected patients. Indigenous algorithms have been developed for cropping and for extraction of the lung region from the images. To detect pneumonia clouds we have used Otsu thresholding which will segregate the healthy part of lung from the pneumonia infected cloudy regions. We are proposing to compute the ratio of area of healthy lung region to total lung region to establish a result. The task has been performed using Python and OpenCV as they are free, opensource tools and can be used by all, without any legality issues or cost implication.**

***Keywords*—Medical image processing, lung extraction, Pneumonia cloud detection, lung segmentation, automated diagnosis**

I. INTRODUCTION

Pneumonia is the leading cause of death among children in developing countries, including India. WHO estimates that one in three new born infant deaths is due to pneumonia[1]. About half of these deaths can be prevented as they are caused by the bacteria for which an effective vaccine is available. According to a report [2] which was released by International Vaccine Access Centre(IVAC), Johns Hopkins Bloomberg School of Public Health, India has the highest number of pneumonia and diarrhea deaths among children in the world with nearly three lakh children dying in 2016.

Chest X-ray (CXR), is an important tool for diagnosing pneumonia and many clinical decisions rely heavily on its radiological findings. Also it is relatively cheap compared to other imaging diagnostics and can be afforded by masses. Some work has been done on automated pneumonia detection through natural language processing and artificial neural network [3].However, such tools require elaborate hardware and software setup.

In order to develop a software that can be easily and freely ported to mobile devices, with no financial or license related obligations, so we have undertaken a purely an image processing approach as in [4], using free open source tools such as Python and its image processing libraries.

When interpreting chest X-rays for Pneumonia, the radiologist will look for white spots in the lungs called infiltrates that identify an infection. However, such cloudy patterns would also be observed in TB Pneumonia and severe cases of bronchitis too. For conclusive diagnosis, further investigations such as complete blood count (CBC), Sputum test, and Chest computed tomography (CT) scan etc. may be needed. Therefore, we are only attempting to detect possibility of pneumonia from Chest X-rays, by looking for cloudy region in the same. Conclusive detection will depend on pathological tests

II. DATA AND SOFTWARE TOOLS USED

For the purpose of coding, we have used Python version 2.7 with the OpenCV libraries. We have developed indigenous algorithms for cropping of images to remove abdomen region and for lung boundary identification. We have also looked for Japan Society of Radiological Technology (JSRT) image dataset as mentioned in [4], but we felt that image quality of that dataset is really very good compared to ground truth images. So we have used analog CXR images that matches rural scenario situations. As per availability, our dataset contains only 40 adult CXR.

Few samples of normal and Pneumonia infected CXR images from our dataset, are shown in *Fig. 1* and *Fig. 2*.

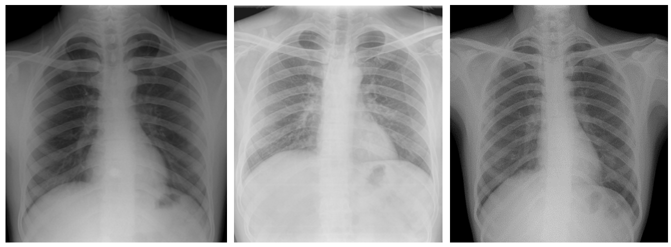

Fig. 1. Normal CXR images

978-1-5386-1747-2/17/$31.00 ©2017 IEEE

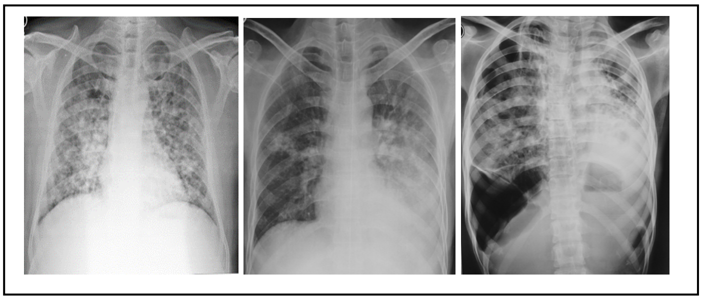

Fig. 2. Pneumonia affected CXR image

III. METHODOLOGY

The methodology used by us, is shown in form of block diagram in *Fig. 3*. We shall describe each section in detail, explaining the steps and the algorithm, specially the indigenous ones. We shall also mention why we decided on the threshold values used in our algorithms.

The task has been performed as follows. All images have been resized to optimal size for computational purposes. Then we have performed histogram equalisation on the same to enhance contrast of the images. Thereafter we have used our own algorithm to crop the abdomen area from CXR images. The lung boundary is then identified, to calculate the lung area. Thereafter, image thresholding has been performed to detect the non-cloudy region of the lung and its area is computed. The ratio of areas of this non-cloudy lung region to the total area of extracted lung region, would give an indication of the amount of cloud formation in the lungs. The following figure summarise or approach in form of the block diagram for detection of pneumonia clouds.

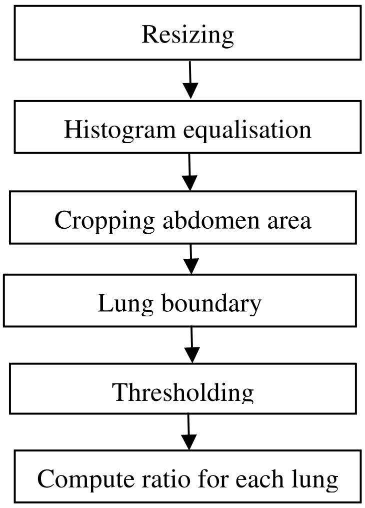

Fig. 3. Block diagram for detection of pneumonia clouds

A. Resizing

The original images were 2048x2048 pixels in size. We resized it to different values and compared the computational time and results observed. The comparison of data is shown in Table I.

TABLE I
Comparison Table for Different Sizes of Images

Size	**Computational time(s)**	**Result observed**
2048x2048	71	Good
1024x1024	9	Good
800x800	4.6	Good
500x500	3.2	Not satisfactory

For obtaining the above table, we have randomly selected the 5 CXR images from our dataset and performed our methodology on the same 5 CXR images by changing their image size. Then we have noted the thresholded images and ratio obtained in each case. The results are considered good, if they provide sufficient demarcation in the normal CXR images and pneumonia affected CXR images.

From the above *Table I*, we have selected an image size of 800x800 for further operations, as the computational time is lowest without compromising on the result quality. We resized all the images to 800x800, using the ***'resize'*** function of OpenCV.

B. Histogram Equalization

Histogram equalization is done to adjust image intensities so as to enhance contrast. The enhanced contrast aids in detection of the clouds. We used the ***'equalizeHist'*** function available in OpenCV.

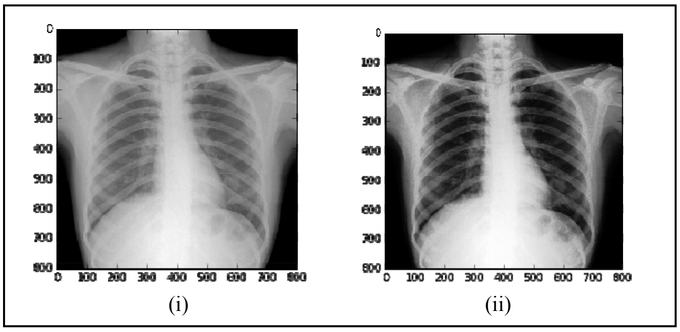

Fig. 4. (i) Image before histogram equalization (ii) Image after histogram equalization

C. Cropping of lower region

Lower region of the image has been cropped to eliminate the abdomen region, present in most of the CXR images as that region of CXR is insignificant in detection of pneumonia clouds. For this, we have started from the lower leftmost point of the image and traversed the image horizontally in batches of 5 pixels, till half of the image width. While traversing, we have compared the average intensities of two successive batches looking for a transition from white to black region with an intensity gradient of 100. We decided on the number 100, after multiple trials with various gradient values. After traversing the bottommost line, we move vertically upward in steps of 2 pixels to next pixel and continue the horizontal scanning till we locate the batches where the intensity gradient changes by 100. Now, the image is cropped horizontally from the pixel coordinates we have found while scanning.

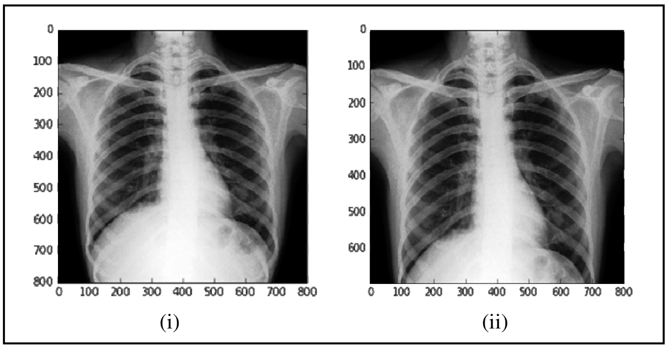

Fig. 5. (i) Image before cropping lower abdomen region (ii) Image after cropping lower abdomen region

D. Lung boundary identification

Lung boundary identification is important for calculating lung area. By finding the rib cage boundary or lung boundary we can actually get an idea of the actual area covered by lung in the given CXR image. For this purpose we have vertically split CXR into two halves, left and right. Then we have selected 100 equidistant horizontal lines in each lung and looked for a pixel intensity gradient value greater than some threshold value of 30, along each horizontal line. The number 30 was chosen after attempting lung boundary identification using multiple threshold values and selecting the one that produced the best results.

For values of pixel intensity an average of 5 consecutive pixels has been taken to nullify the effect of stray high or low intensity points. Wherever the desired gradient transition is noticed, the coordinate of the first pixel of the batch is taken as a boundary point to ensure that the lung region is not truncated.

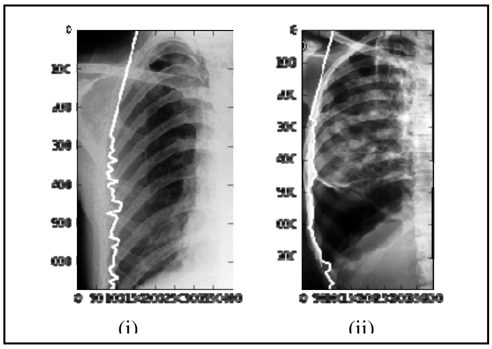

Fig. 6. Image after lung boundary detection (i) normal CXR (ii) Pneumonia affected CXR

E. Thresholding

Thresholding is done to segregate healthy part of lung region from the pneumonia affected part. For this purpose we have used Otsu thresholding. The image after thresholding gives us the binary image containing the dark region of image as black and remaining as white.

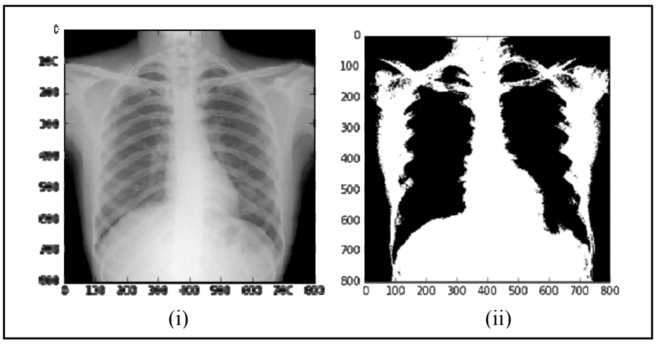

Fig. 7. (i) Image before thresholding (ii) Image after thresholding

Images of lung region after thresholding

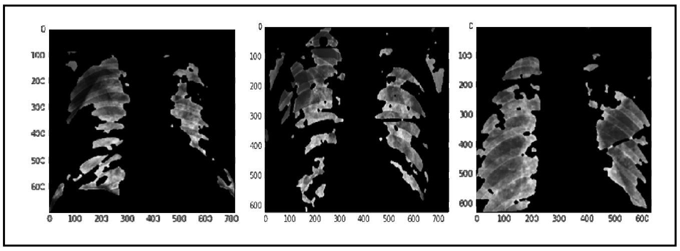

Fig. 8. Pneumonia affected lung region, after thresholding

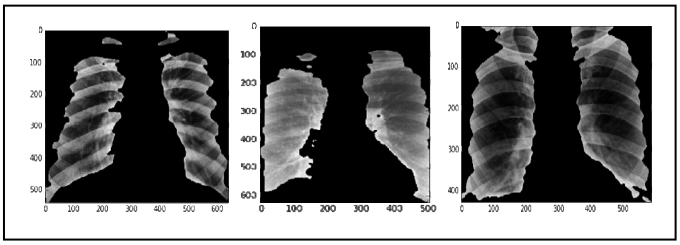

Fig. 9. Normal lung region, after thresholding

IV. CONCLUSION

Computer assisted detection of diseases from CXR are always very helpful at places where there is shortage of skilled radiologist. In countries like India where, we do not have experienced radiologists in rural areas, such tools can be of immense help by automatically screening people who need urgent medical care and further diagnosis.

At this stage of the project, we have identified the lung region by rib cage boundary identification and would be computing the area of this region. We have also used Otsu thresholding to segregate the pneumonia cloud from the healthy lung in the lung area, still we are working on other methods that can be adopted for thresholding the CXR images which can yield better results. Also, we are now looking to compute the ratio of the two areas - the area of the lung after Otsu threholding to the area of the total lung area. As pneumonia clouds are not visible in the image of the lung after Otsu thresholding, this ratio is expected to be much lower than when computed for healthy lungs without clouds.

REFERENCES

[1] WHO and Maternal and Child Epidemiology Estimation Group (MCEE) estimates 2015.

[2] International Vaccine Access Centre (IVAC) "The Pneumonia and Diarrhoea Progress Report for 2016".

[3] Oliveira, L.L.G., Silva, S.A.e., Ribeiro, L.H.V., de Oliveira, R.M., Coelho,C.J., S.Andrade, A.L.S. "Computer-aided diagnosis in chest radiography for detection of childhood pneumonia"(2008) *International Journal of Medical Informatics*, 77 (8), pp. 555-564.

[4] Mohd Nizam Saad, Noraidah Sahari Ashaari, Hamzaini Abdul Hamid, "*Image Segmentation for Lung Region in Chest X-ray Images using Edge Detection and Morphology*", 2014 IEEE International Conference on Control System Computing and Engineering, 28-30 November 2014.