

CS/EC 445/545

Sean Lesch, Hoi Dinh

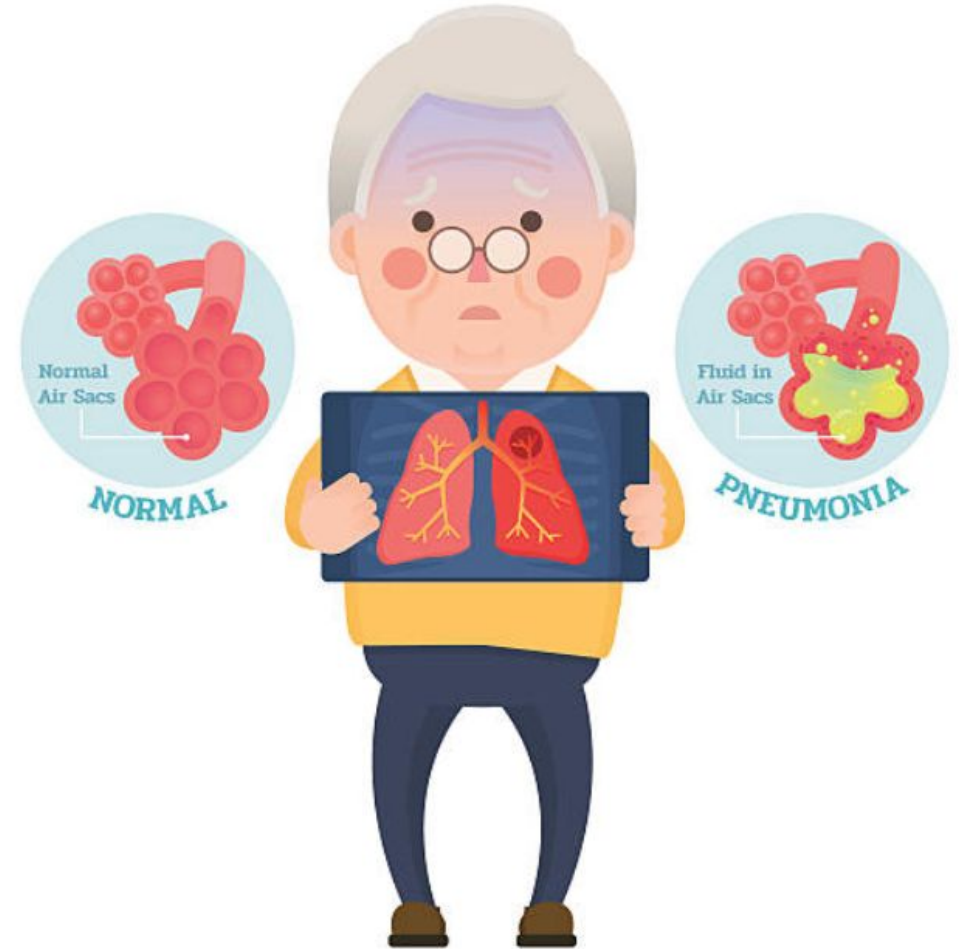
Wenjing Wu, Colleen Xu

Shalika Arora

Comparing Different Feature Extraction Methods: Pneumonia, X-ray images, and SVMs

Motivation

- Chest x-rays are a diagnostic tool to detect signs of pneumonia
- Overworked doctors, lack of radiologists in rural areas

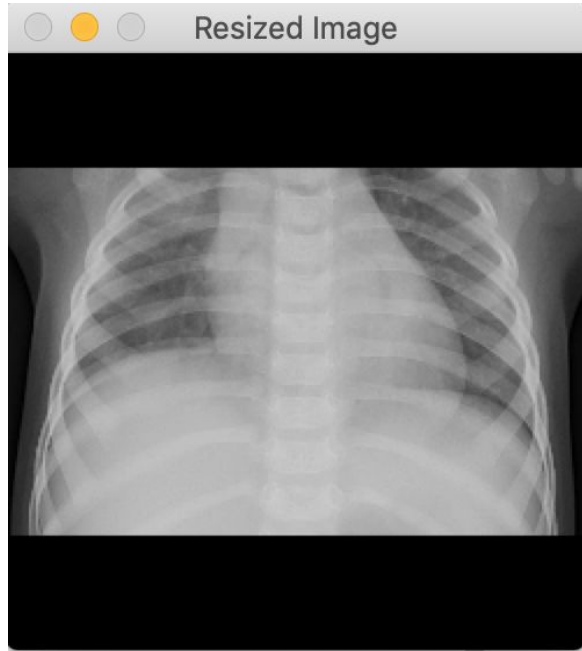


Chest x-rays

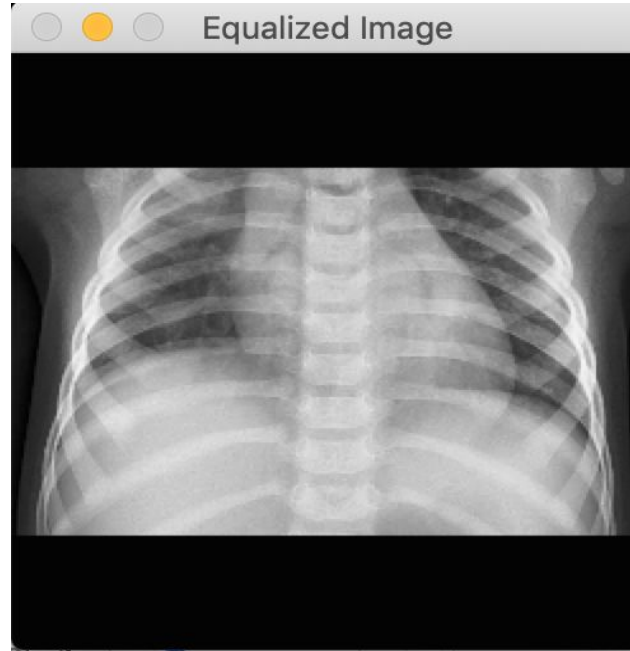
- Training images:
 - Normal: 1341 images (**Label : 0**)
 - Pneumonia: 1341 images (**label : 1**)
- Test images:
 - Normal: 234 images
 - Pneumonia: 390 images



Resize Image



Increase contrast (adaptive histogram equalization)



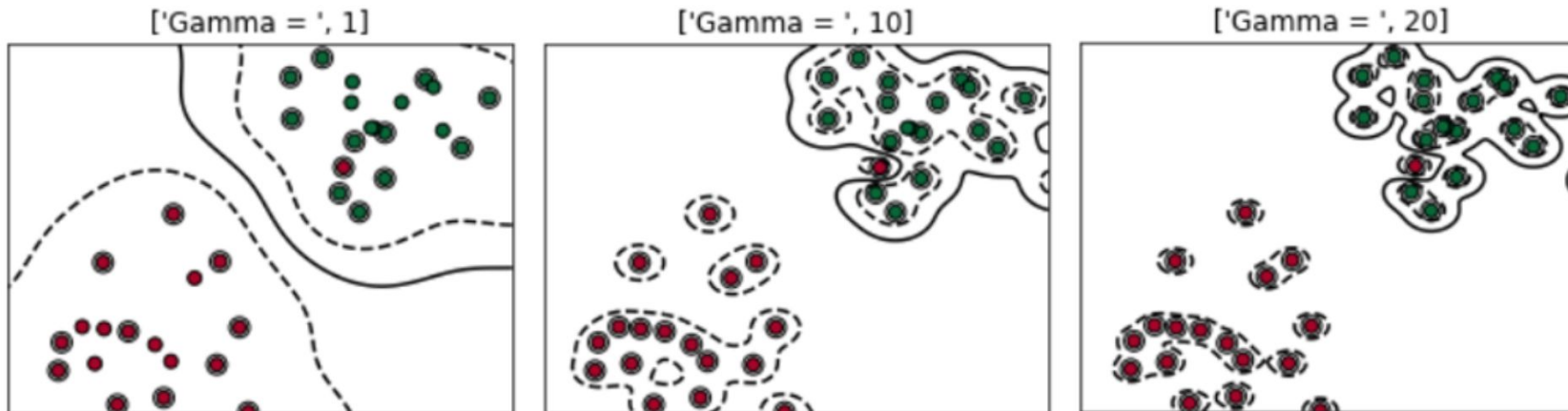
Feature Extraction

26 features/image for LBP
729 features/image for HOG
18 features/image for GLCM



SCALE

SVM



Local Binary Pattern

- LBPs compute a *local representation* of texture. Local representation is constructed by comparing each pixel with its surrounding neighborhood of pixels.
- Steps:
 - Convert to grayscale image.
 - Calculate the LBP mask.
 - Calculate the LBP Histogram and normalize it.

LBP uniformity

Uniform: if ***at***

most two *0-1* or *1-0* transitions
example, 00001000 (2 transitions)

Non-Uniform: 01010010

(six *0-1* or *1-0* transitions)

- $p + 1$ **uniform patterns**:
- $p + 2$ final dimensionality of the histogram ($p+1$ bins for uniform pattern, all non uniform patterns in one bin)

Top: $r=3$ $p=26$ Bottom: $r=1$

$n=8$

Local Binary Pattern

5	8	1
5	4	1
3	7	2



0	0	1
0		1
1	0	1

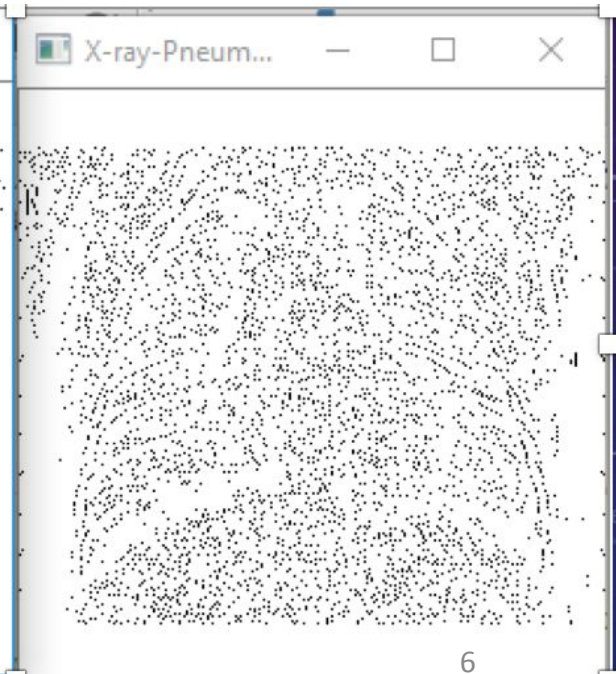
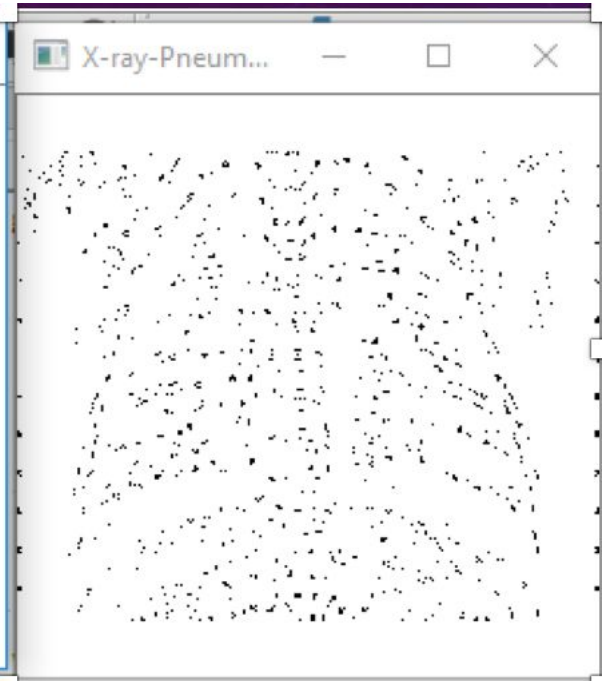
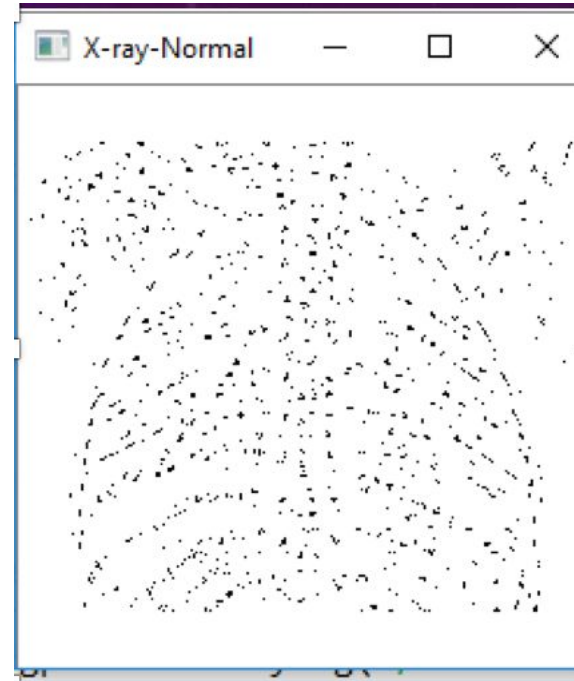
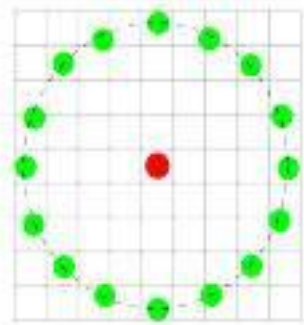
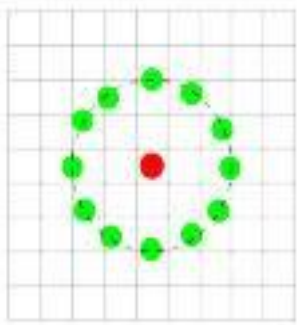
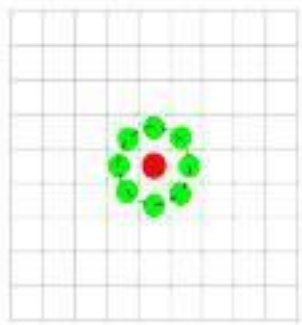
0	0	1
0		1
1	0	1



0	0	0	1	0	1	1	1
7	6	5	4	3	2	1	0

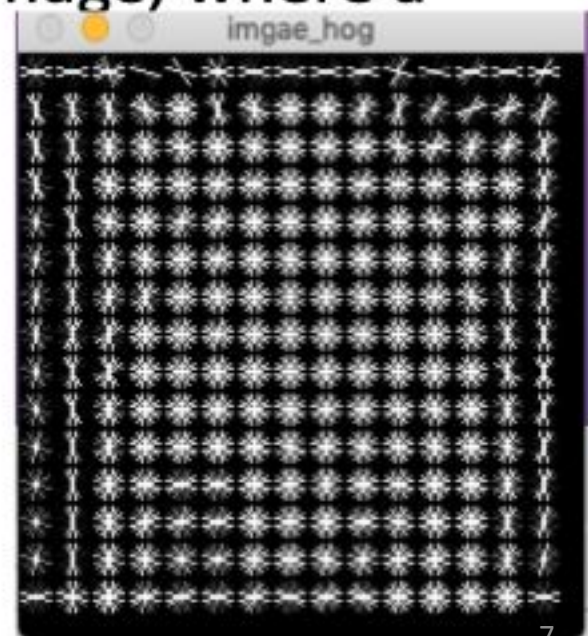
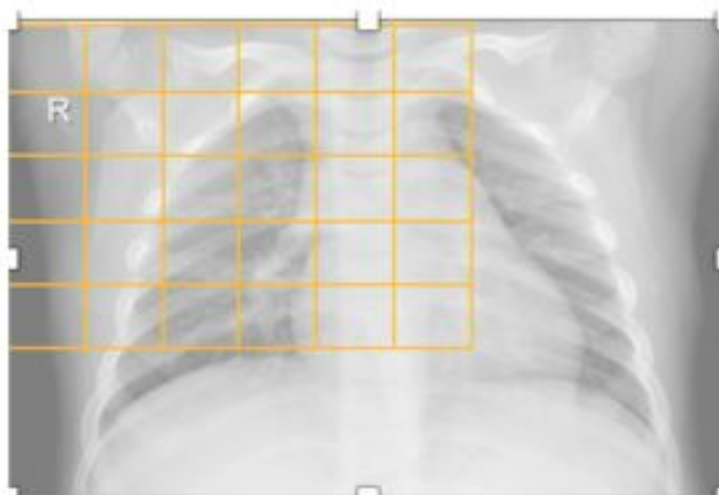
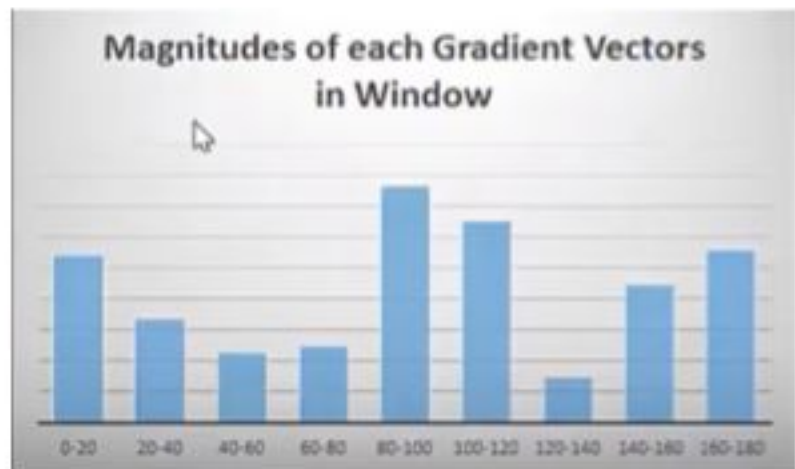
$$2^4 \quad 2^2 \quad 2^1 \quad 2^0$$

$$16 + 4 + 2 + 1 = 23$$



Histogram of Oriented Gradients (HOGs)

- HOGs are a **feature descriptor** that has been widely and successfully used for object detection
- It represents objects as a **single feature vectors** as opposed to a set of feature vectors where each represent a segment of the image.
- It's computed **by sliding window detector** over an image, where a HOG descriptor is a computed for each position.



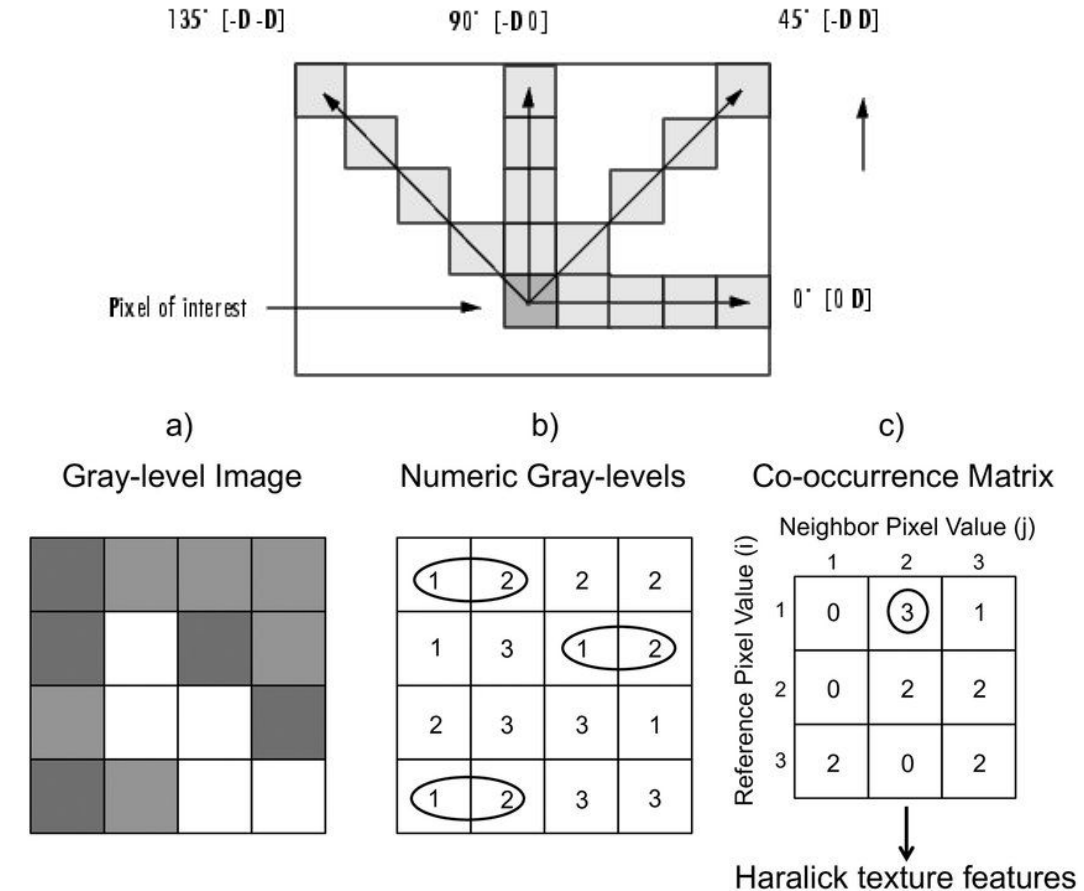
HOGs Step by Step

1. Using an 25×25 pixel detection or cell (yellow), we compute the gradient vector or edge orientations at each pixel
2. This generate 25×25 gradients vectors which are then represented as a histogram
3. Each cell is then split into angular bins, where each bins corresponds to a gradient direction.
4. As it stores gradients magnitudes, it relatively immune to deformations

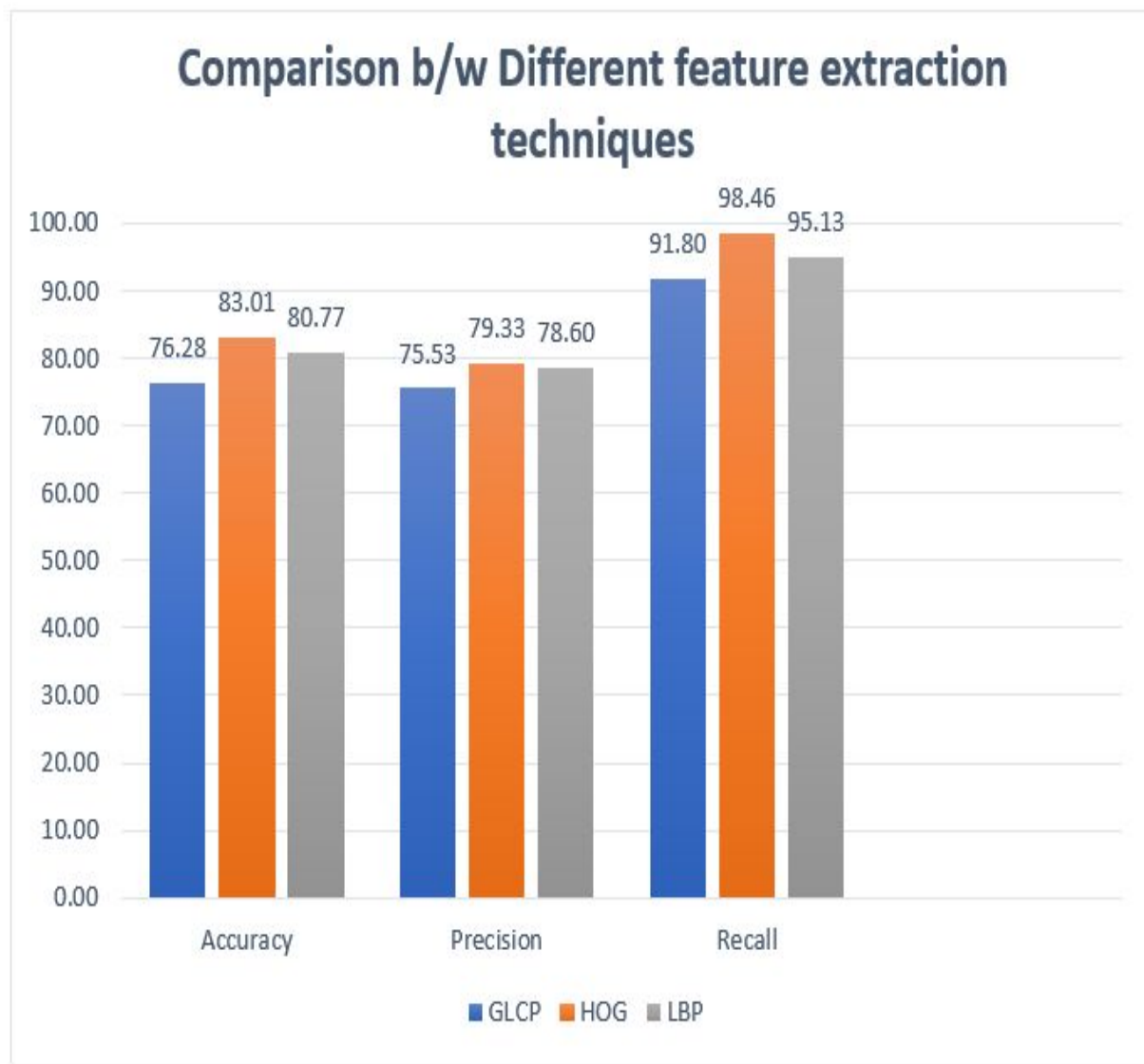


Gray Level Co-occurrence Matrix

- Tabulation of how often different combinations of pixel brightness values (gray levels) occur in an image.
- Result is a $256 \times 256 \times 1 \times 4$ matrix
levels x levels x number of distances
x number of angles.
- Properties of the resulting matrix are used for feature extraction after normalization.



Results



	0 (True Normal)	1 (True Pneumonia)
0 (False Normal)	134	6
1 (False Pneumonia)	100	384

References:

- Slide 2:
 - Image: <https://www.istockphoto.com/vector/senior-man-having-pneumonia-gm613235796-105823791>.
 - Statistics on Pneumonia (spoken):
<https://www.thoracic.org/patients/patient-resources/resources/top-pneumonia-facts.pdf>
- Slide 3: Image (Figure S6): [https://www.cell.com/cell/fulltext/S0092-8674\(18\)30154-5#fig6](https://www.cell.com/cell/fulltext/S0092-8674(18)30154-5#fig6)
- Slide 4:
 - The image showing RBF SVM with different gamma values is from
<https://towardsdatascience.com/support-vector-machine-simply-explained-fee28eba5496>
 - All other images were generated by us over the course of the workflow.
- Slide 6:
 - <http://hanzratech.in/2015/05/30/local-binary-patterns.html>
 - <https://www.pyimagesearch.com/2015/12/07/local-binary-patterns-with-python-opencv/>
- Slide 7: Left-most image:
<https://circuitdigest.com/tutorial/real-life-object-detection-using-opencv-python-detecting-objects-in-live-video>
 - All other images were generated by us over the course of the workflow.
- Slide 9: Top image: <https://www.mathworks.com/help/images/specify-offset-used-in-gldm-calculation.html>. Bottom image:
https://www.researchgate.net/figure/Haralick-texture-features-are-calculated-from-the-gray-level-co-occurrence-matrix-GLCM_fig1_330550795