Gabor Wavelet-based image classification using CNNs

Rajdeep Pathak

Department of Mathematics,

IIT Hyderabad

May 6, 2025

Motivation and Content

- Image classification - pneumonia detection, skin lesion detection, tumour detection, other medical diagnoses
- Gabor wavelet and Gabor filters
- Gabor representation of an image
- Training multiple CNNs with original image and Gabor-transformed images
- **Goal:** Improve classification accuracy by using models trained with Gabor representations of the images

Gabor wavelet

- The equation of a 1D Gabor wavelet is a Gaussian modulated by a complex exponential:

$$f(x) = e^{\frac{-(x-x_0)^2}{a^2}} e^{-ik_0(x-x_0)}$$

- Here, x_0 denotes the centre, a denotes the standard deviation of the Gaussian envelope/the spread, and k_0 controls the modulation rate.

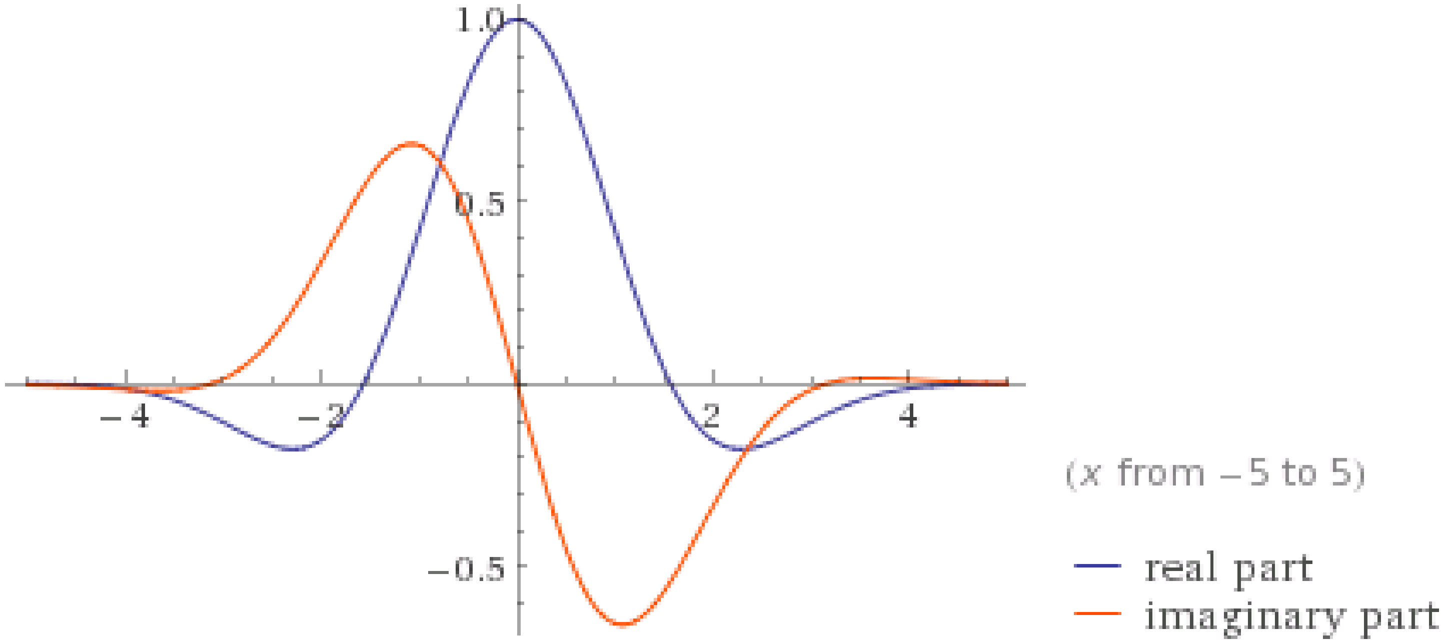


Figure 1. A Gabor wavelet with a = 2, x_0 = 0, and k_0 = 1. (Generated from Wolfram Alpha)

Properties of the Gabor Wavelet

- Gabor wavelets are localized - as the distance from the centre increases, the value of the function becomes exponentially suppressed.

- Fourier transform of the Gabor wavelet is given by:

$$\mathcal{F}(k) = ae^{-(k-k_0)^2a^2}e^{-ix_0(k-k_0)}$$

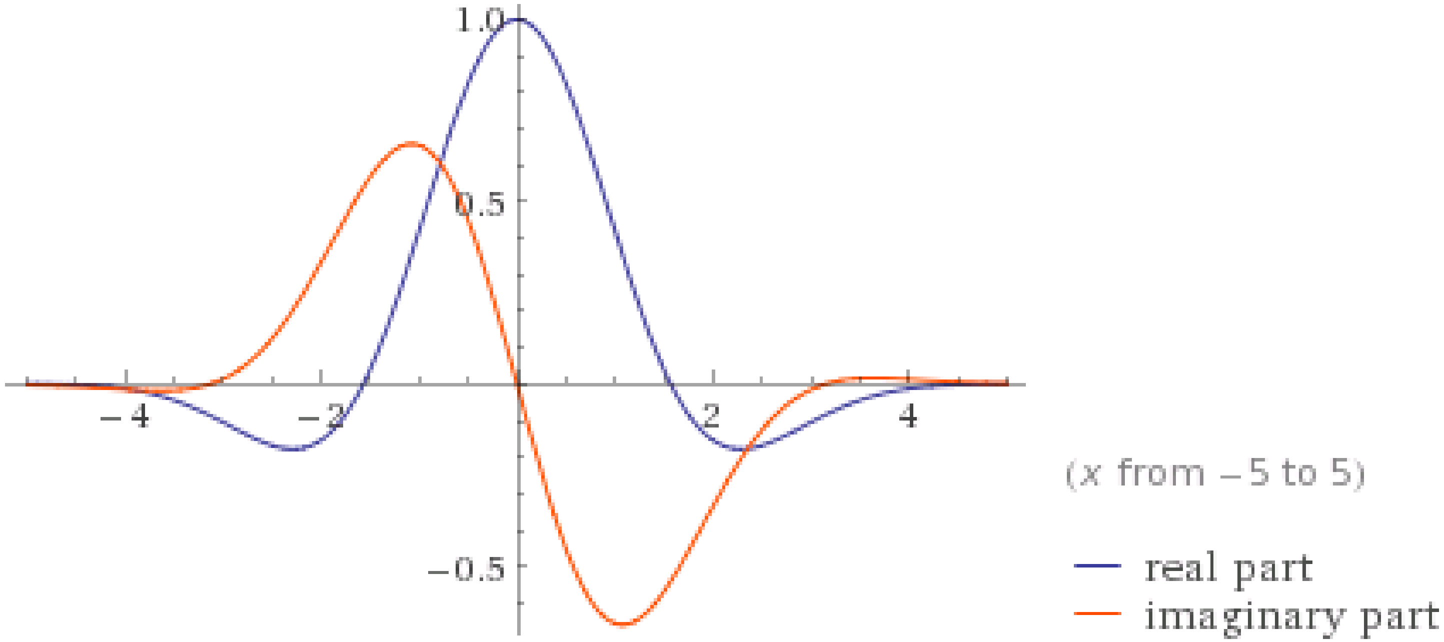


Figure 1. A Gabor wavelet with $a = 2$, $x_0 = 0$, and $k_0 = 1$.

Gabor Filter

- Gabor Filter was first proposed as a 1D filter by Dennis Gabor.
- It was first generalized to 2D by Gösta Granlund [1] in 1978.
- It is defined by a sinusoidal wave (a plane wave for 2D Gabor filters) multiplied by a Gaussian function.
- The filter has a real and an imaginary component representing orthogonal directions:

Complex

$$g(x,y;\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\exp\left(i\left(2\pi\frac{x'}{\lambda}+\psi\right)\right)$$

Real

$$g(x,y;\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\psi\right)$$

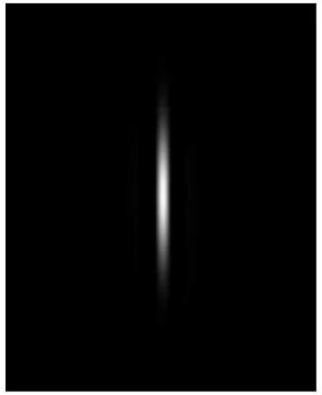
Imaginary

$$g(x,y;\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\sin\left(2\pi\frac{x'}{\lambda}+\psi\right)$$

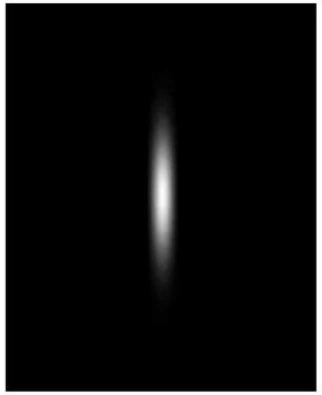
where $x' = x\cos\theta + y\sin\theta$ and $y' = -x\sin\theta + y\cos\theta$.

where,
λ = wavelength of the sinusoidal factor
σ = standard deviation of Gaussian envelope
γ = spatial aspect ratio (specifies ellipticity of support of the Gabor function)
θ = orientation
ψ = phase offset

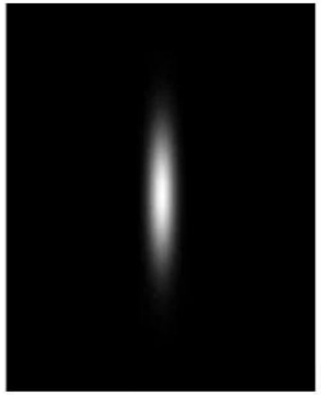
Effects of the parameters on Gabor Filters



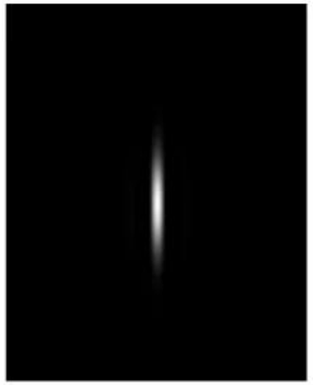
Lambda (λ) = 30



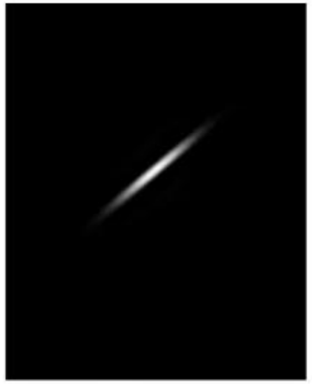
Lambda (λ) = 60



Lambda (λ) = 100



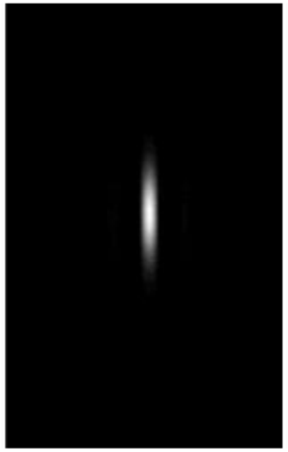
Theta (Θ) = 0



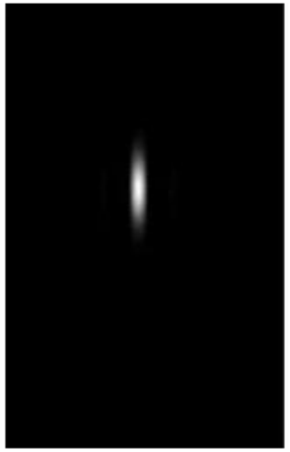
Theta (Θ) = 45

Theta (Θ) = 90

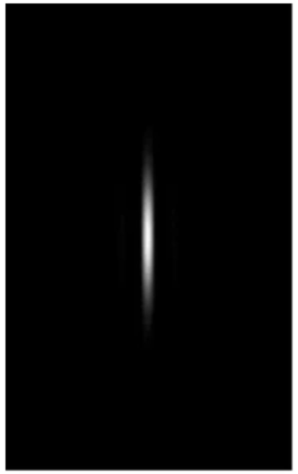
Gamma (γ) = 0.25



Gamma (γ) = 0.5



Gamma (γ) = 0.75



Sigma (σ) = 10

Sigma (σ) = 30

Sigma (σ) = 45

Gabor Filters and Gabor Wavelets

- Gabor filters are directly related to Gabor wavelets, since they can be designed for a number of dilations and rotations.
- Usually, a filter bank consisting of Gabor filters with various scales and rotations is created.
- The filters are convolved with the signal (or image), resulting in a so-called Gabor space.
- The Gabor filter is used for texture analysis and edge detection - it analyzes whether there is any specific frequency content in the image in specific directions in a localized region around the point or region of analysis.

Building a Gabor Kernel and Transforming an Image

- A Gabor kernel/filter is a 2D matrix.
- We use OpenCV's getGaborKernel() function to build Gabor kernels.
- Specify the kernel size: in our case, we choose ksize = 31.
- Specify the other parameters: lambda, sigma, gamma, theta, and psi.
- getGaborKernel((ksize,ksize), sigma, theta, lambda, gamma, psi) returns a Gabor kernel G, a matrix of size 31 x 31.
- Now, we want to apply the Gabor filter on an image K.
- The image K is another matrix of size, say, 128 x 128.
- The image is transformed by convolving the image with the Gabor kernel: K*G.

Example

- Set sigma = 1.3, lambda = 5.7, gamma = 0.5, psi = 0, and vary theta (this captures localized properties at different orientations).
- Make three Gabor filters using theta = [0^o, 22.5^o, 90^o]
- Convolve each with the original image. The result is a Gabor-transformed image at that particular orientation.
- Visually, does Gabor-transforming help in detecting pneumonia from chest X-rays?

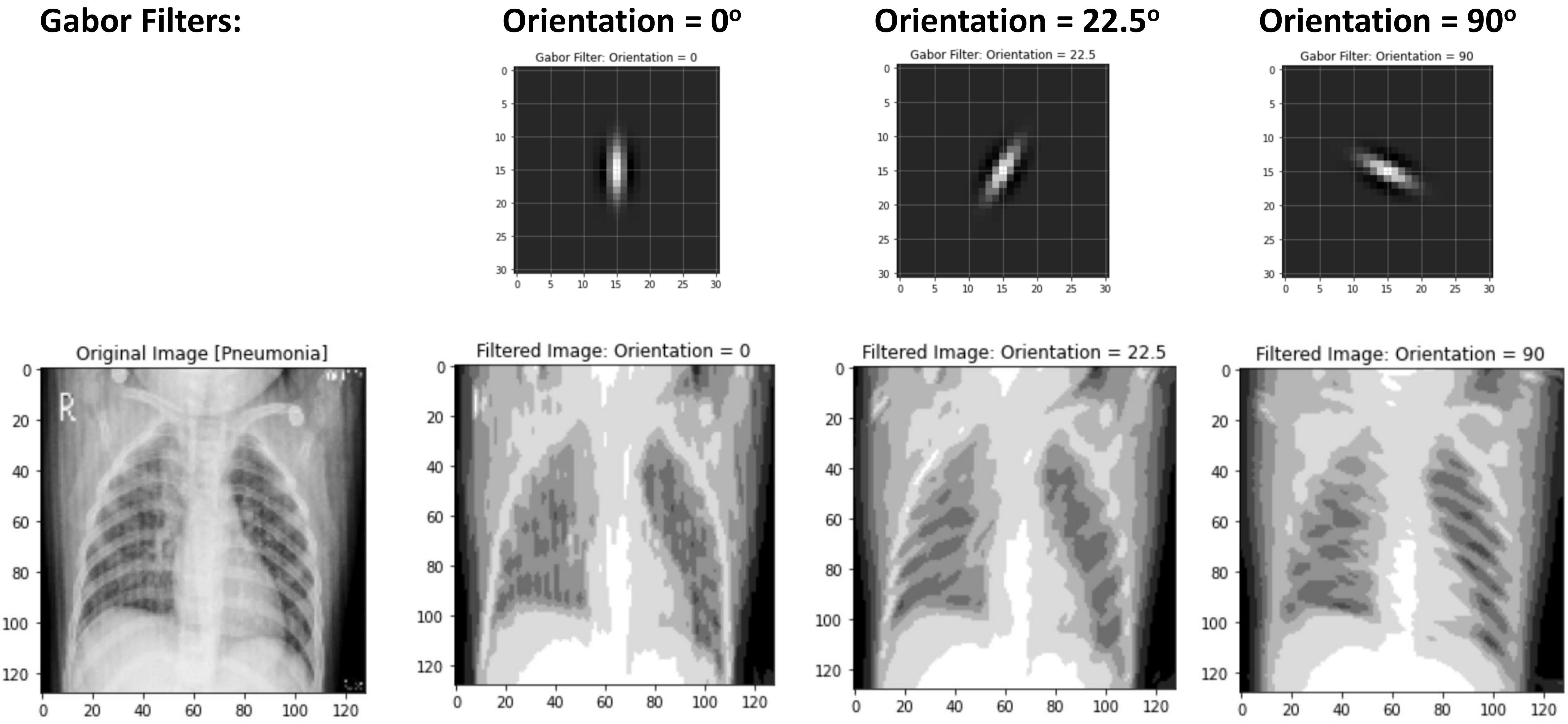


Figure 2. Original image (Pneumonia) and its Gabor-transformed images

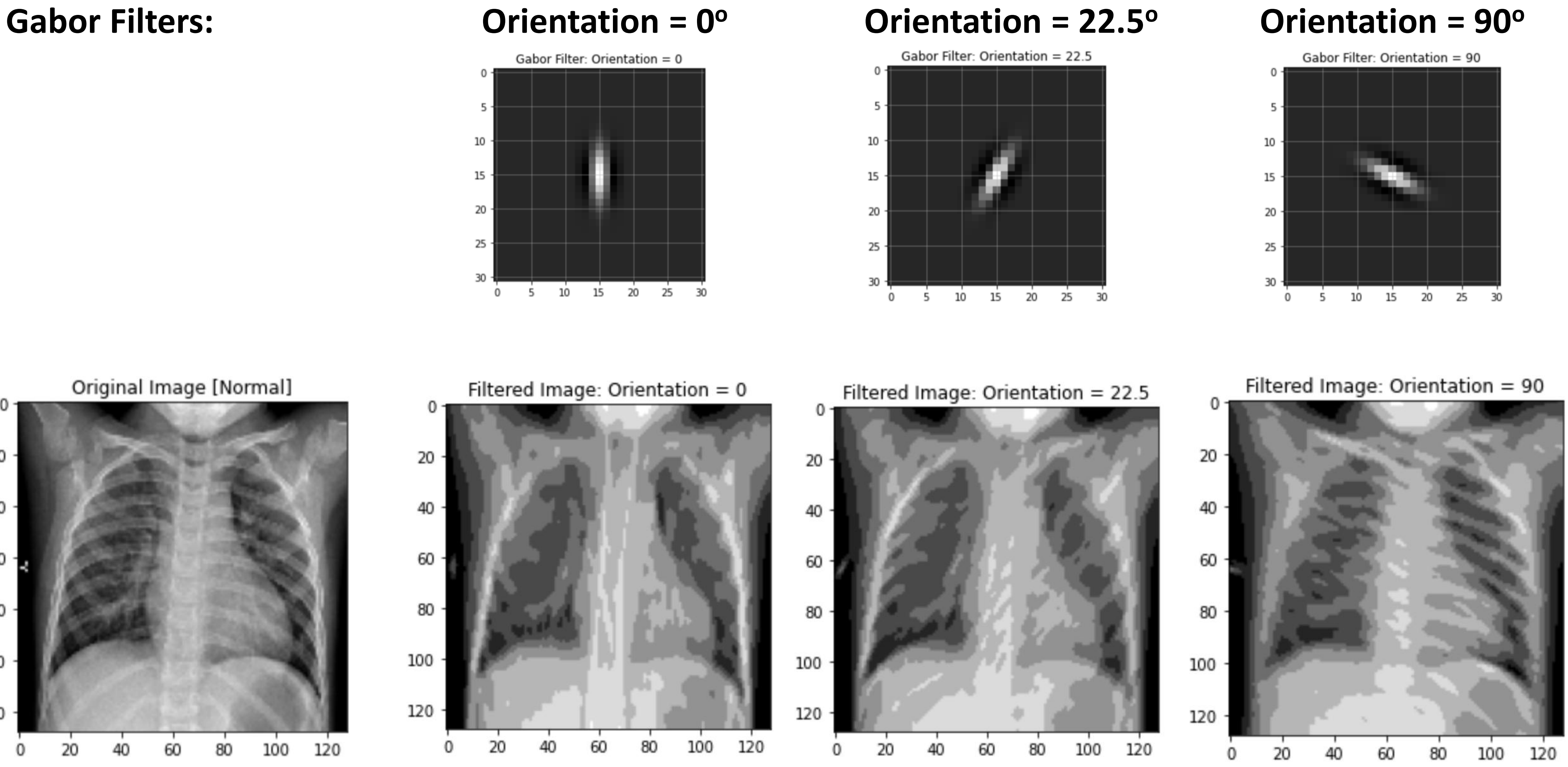
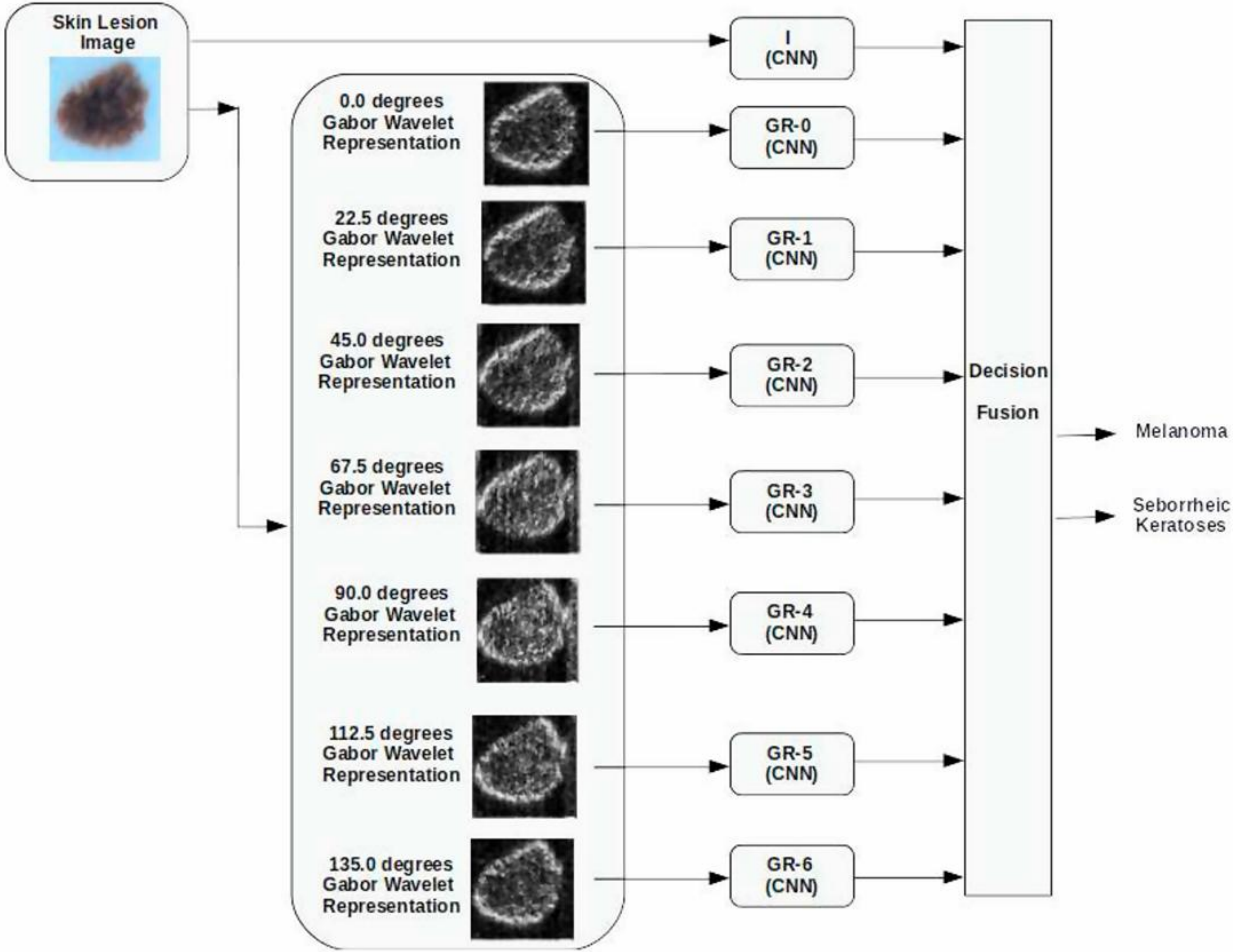


Figure 3. Original image (Normal) and its Gabor-transformed images

Gabor Wavelet-based CNN Classifier

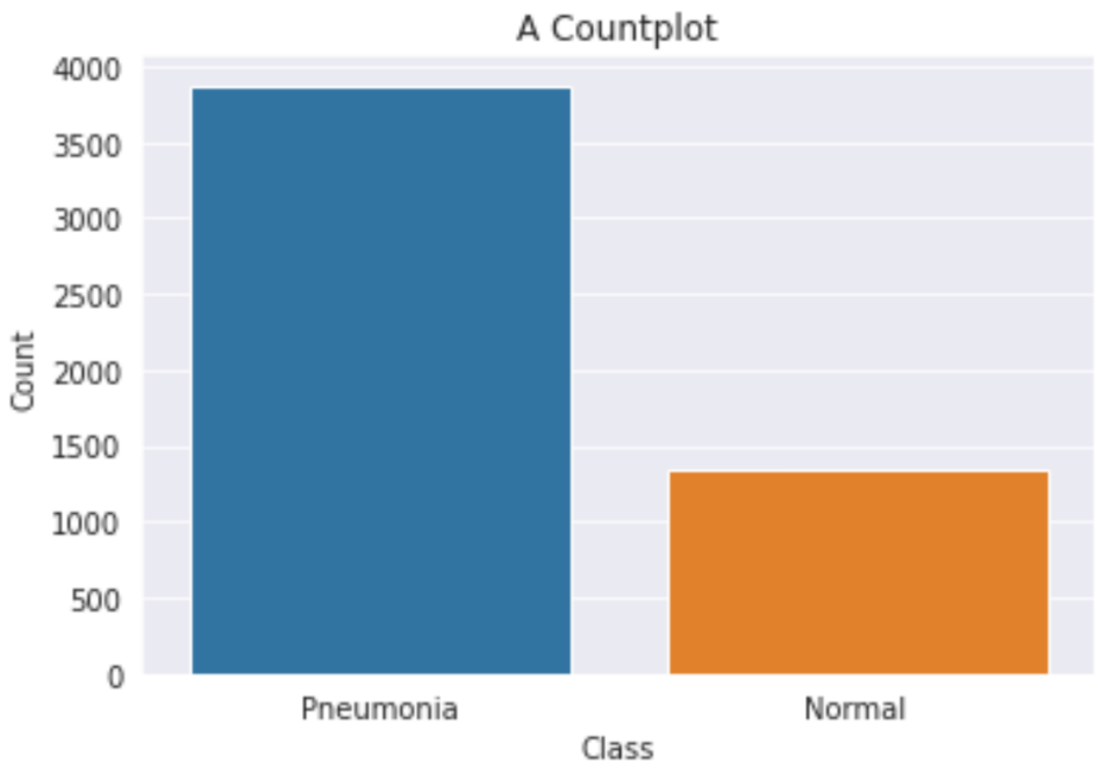
- Normalize the images (divide each element of the array by 255).
- Decompose the given image into 8 directional sub-bands (based on different orientations of the Gabor filter).
- Use these eight sub-band images and the original input image to train 9 parallel CNNs.
- For a test image, decompose it into sub-bands and pass it along with the input image to the trained CNNs.
- The CNNs generate 9 probabilistic predictions.
- Perform decision fusion using the sum rule for the final decision.
- We implement this idea on chest X-ray images for pneumonia detection.



Skin Lesion Image
I (CNN)
0.0 degrees Gabor Wavelet Representation
GR-0 (CNN)
22.5 degrees Gabor Wavelet Representation
GR-1 (CNN)
45.0 degrees Gabor Wavelet Representation
GR-2 (CNN)
67.5 degrees Gabor Wavelet Representation
GR-3 (CNN)
90.0 degrees Gabor Wavelet Representation
GR-4 (CNN)
112.5 degrees Gabor Wavelet Representation
GR-5 (CNN)
135.0 degrees Gabor Wavelet Representation
GR-6 (CNN)
Decision Fusion
Melanoma
Seborrheic Keratoses

The Data

- There are 5,856 X-Ray images (JPEG) and 2 categories (Pneumonia/Normal).
- Chest X-ray images (anterior-posterior) were selected from retrospective cohorts of pediatric patients of one to five years old from Guangzhou Women and Children's Medical Center, Guangzhou.
- Number of
  - Training images = 5216
  - Validation images = 16
  - Test images = 624



Decision Fusion and Prediction

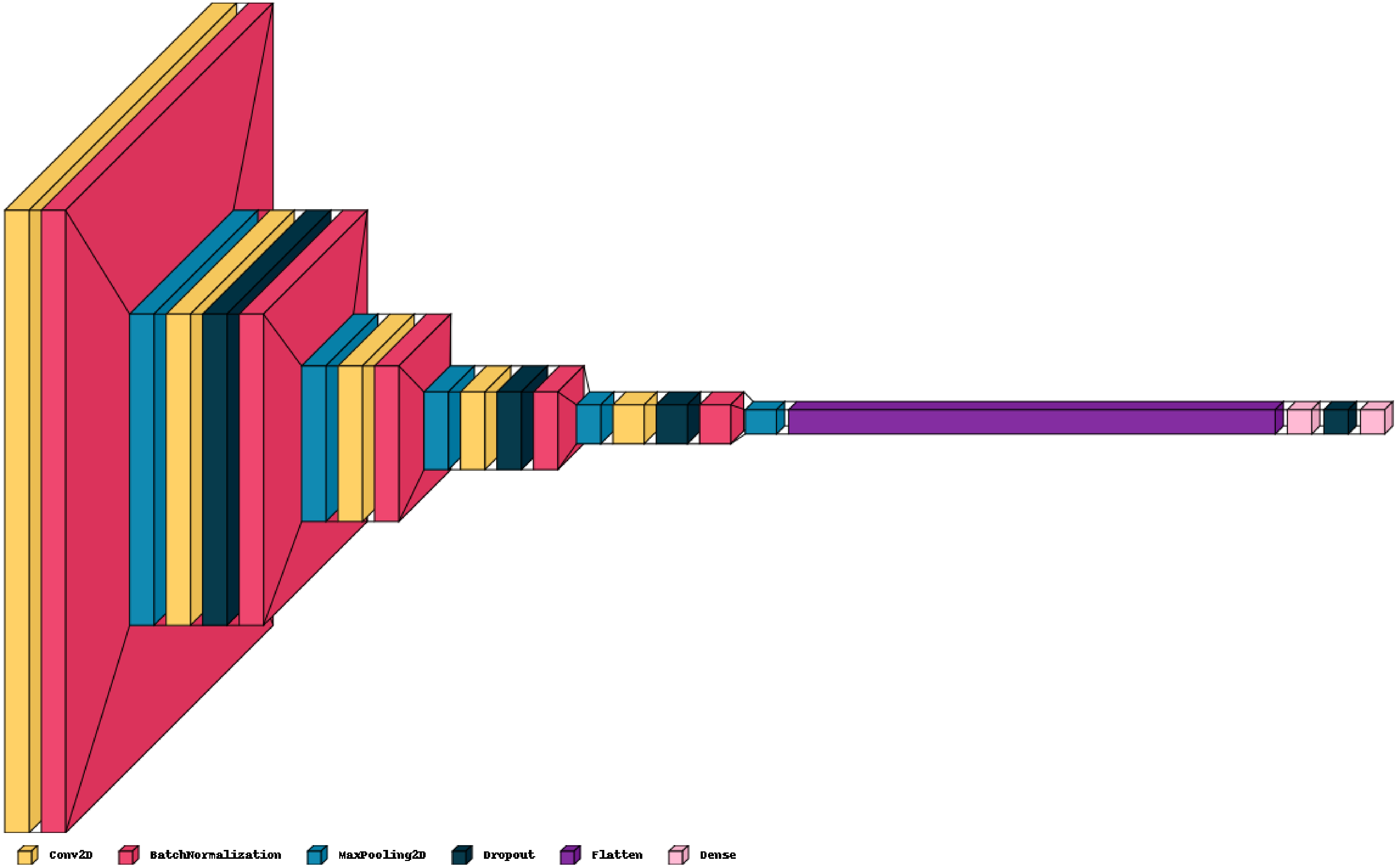
- Final decision probability is obtained using the Sum of Probabilities (SMP) rule.
- Note that we have 2 classes (Pneumonia/Normal) and 9 CNNs (one trained with original input images and 8 others with 8 different Gabor representations)
- Suppose p_{ij} is the probability for class i given by model CNN_j

  (i = {0,1} and j = [0:8])
- The final probability for class i is given by

$$p_i = \frac{\sum_{j=0}^{8} p_{ij}}{9}$$

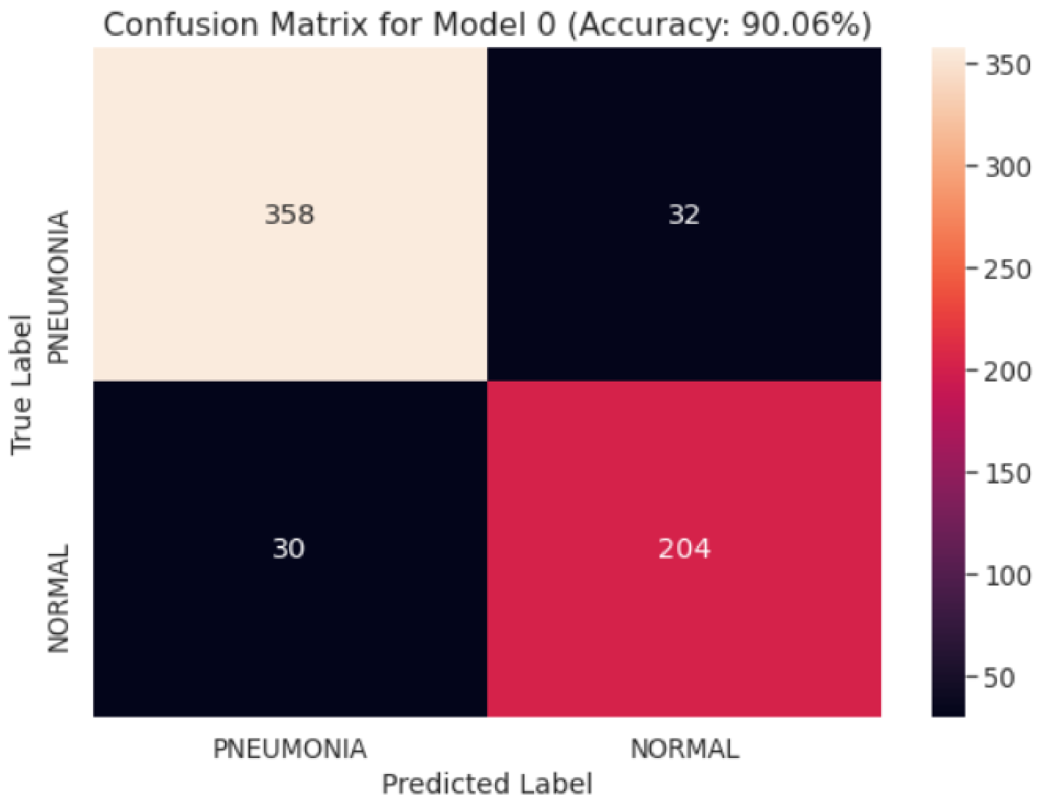
- We use a decision threshold of p = 0.4 (classify as Normal if probability of being normal is > 40%, otherwise pneumonia) because it gives the best TPR-FPR tradeoff.

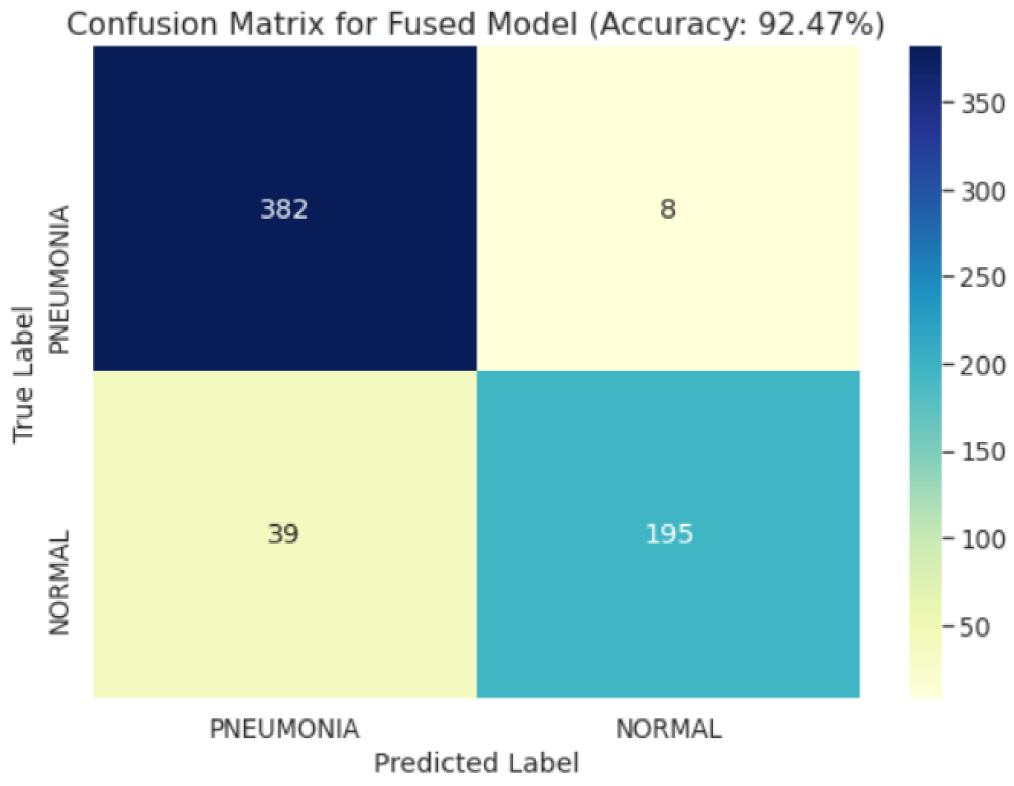
CNN Architecture

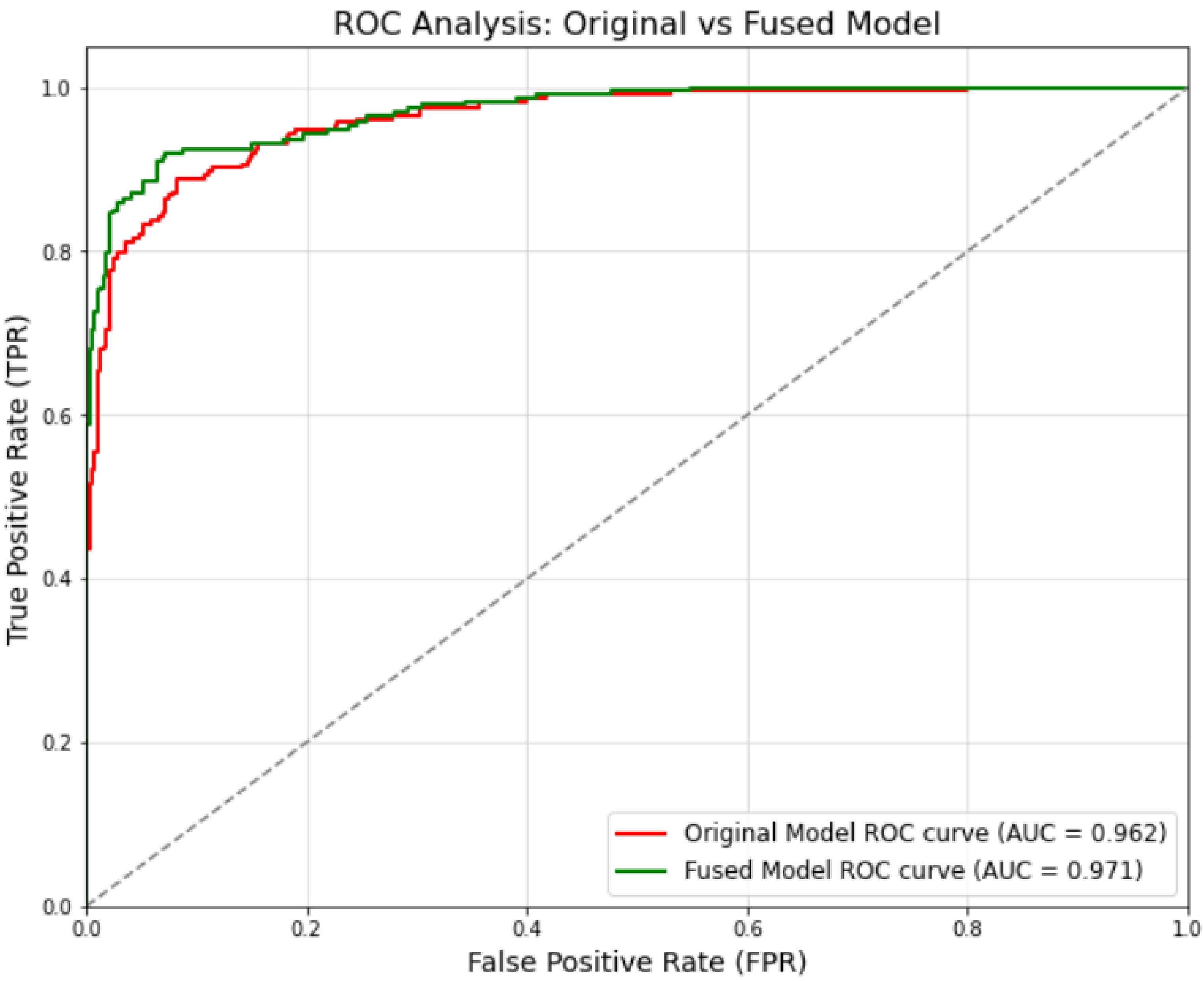


Results

	Accuracy	Precision	Recall	F1-score
Model 0	90.06%	92.27%	91.79%	0.92
Fused Ensemble Model	**92.47%**	90.73%	**97.94%**	**0.94**







Threshold	FPR	TPR
0.3	0.041	0.863
0.4	0.021	0.799
0.5	0.013	0.752
0.6	0.003	0.679
0.7	0.000	0.590

TPR and FPR at different thresholds for the fused model.

Future directions

- Trying using other Wavelets for representing the images
- Trying using other models for image classification
- Trying **Wavelet Packets**

References

[1] Granlund G. H. (1978). "In Search of a General Picture Processing Operator". Computer Graphics and Image Processing. 8 (2): 155–173. doi:10.1016/0146-664X(78)90047-3. ISSN 0146-664X

[2] Sertan Serte, Hasan Demirel, Gabor wavelet-based deep learning for skin lesion classification, Computers in Biology and Medicine, Volume 113, 2019, 103423, ISSN 0010-4825, https://doi.org/10.1016/j.compbiomed.2019.103423.