ECE651 - Detection Theory
Computer Assignment 1, Due October 31st, 2017
Dr. M.R. Azimi

The purpose of this computer assignment is to apply both deterministic as well as random signal detection methods covered in class to a real-life detection problem. The objective of the assignment is to develop detection systems capable of discriminating calcified masses from normal tissue in mammogram imagery.

Data Description

The data set of this computer assignment consists of 10 grayscale images of mammograms taken from various test subjects. Each image has been cropped and zero-padded so that it is of size 1024×1024. Every image in the data set contains one or more masses that would be flagged by radiologists as areas of concern. The ground truth for these images is given at the end of the document. Displayed below is an example of one of these images where the calcified masses have been inscribed with circles.

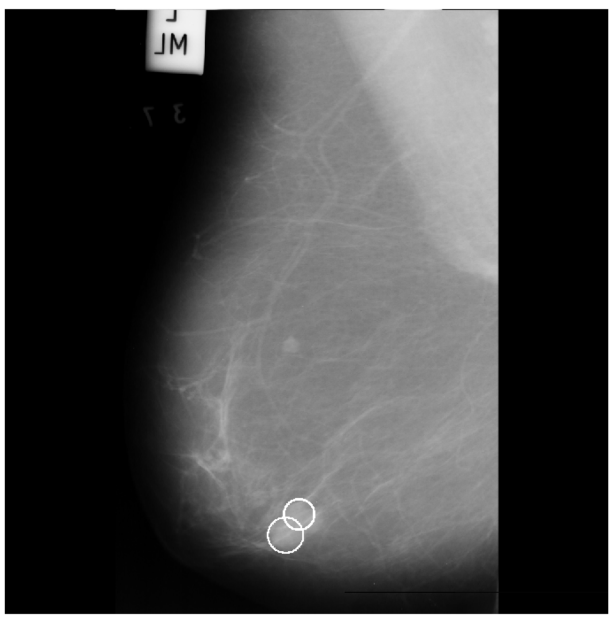

Figure 1: Example Mammogram Image.

The processing of these images to prepare them for detection and how one chooses to form realizations of a random vector is left up to the designer but be sure to fully describe and document your methods when writing the report. A simple example would be to partition the image into $N \times N$ blocks and "vectorize" each block into a random vector of size N^2.

Part 1: Multi-Hypothesis Matched Filter

1. From a subset of images extract M observations of calcified masses, $\{\mathbf{s}_i\}_{i=0}^{M-1}$, you believe would sufficiently represent other calcified masses within the rest of the data set and determine the mean ($\boldsymbol{\mu}$) and covariance matrix (R) of the noise distribution. If unsure of good choices in the mean and covariance one could again extract suitable realizations of normal tissue and estimate these two parameters.

2. Consider the following multi-hypothesis test

$$
\begin{aligned}
\mathcal{H}_0 &: \mathbf{x} = \mathbf{n} \\
\mathcal{H}_1 &: \mathbf{x} = \mathbf{s}_1 + \mathbf{n} \\
\mathcal{H}_2 &: \mathbf{x} = \mathbf{s}_2 + \mathbf{n} \\
&\vdots \\
\mathcal{H}_{M-1} &: \mathbf{x} = \mathbf{s}_{M-1} + \mathbf{n}
\end{aligned}
$$

   where $\mathbf{n} \sim \mathcal{N}(\boldsymbol{\mu}, R)$. Assuming that all hypotheses are equally probable, develop the minimum error probability detector for the detection problem given above. Give the empirical performance of your detector in terms of probability of detection (P_D) and false alarm rate (P_{FA}). Note that all hypotheses with the exception of $\mathcal{H}_0$ given in the problem above correspond to the presence of signal in your observation.

Part 2: Gaussian Detection

1. Now assume that the signal is random and consider the following hypothesis test

$$
\begin{aligned}
\mathcal{H}_1 &: \mathbf{x} = \mathbf{s} + \mathbf{n} \\
\mathcal{H}_0 &: \mathbf{x} = \mathbf{n}
\end{aligned}
$$

   where $\mathbf{n}$ is distributed according to the same model determined from Part 1 and $\mathbf{s} \sim \mathcal{N}(\boldsymbol{\mu}_s, R_s)$. Again if unsure of suitable values for $\boldsymbol{\mu}_s$ and R_s, extract a set of realizations of signal from a subset of images and estimate these parameters.

2. Develop the Neyman-Pearson detector for the detection problem given above. Give the empirical performance of the detector by plotting the ROC curve and giving the P_D and P_{FA} for a suitably chosen threshold.

3. Compare the performance of this detector to that developed in Part 1 and give the pros and cons of each.

Data Set Ground Truth

For each image in the data set, the following table gives the row and column indices of the center of each abnormality as well as the radius (in pixels) of a circle that encloses that abnormality.

Table 1: Data Set Ground Truth

Image Number	Row Index	Column Index	Radius
1	891	477	30
	856	500	26
2	599	525	33
3	566	471	40
4	343	538	29
5	618	462	44
6	875	432	20
7	530	680	20
8	265	470	29
9	388	347	26
10	671	337	45