## 1. Analysis of Results reported by Mordang *et al*.

The results reported by Mordang *et al*. are reported in terms of sensitivity and false positive fraction. These results can be expressed in terms of sensitivity and FP per image based on the data given in Table 1 and 3 in their work.
The number of mammogram images used in the study is 1606 as per given in the Table 1. The distribution of positive and negative patches into training and testing set is reported in Table 3. The total number of negative patches used in testing set is 18,320,976.
The method reported sensitivities of 99.92%, 99.58%, 95.17%, and 74.63% at the false positive fractions of 0.1, 0.01, 0.001, and 0.0001.

The total number of false positives for the false positive fraction of 0.1 can be calculated as:
Total number of false positives = false positive fraction*number of negative samples

$$= 0.1*18,320,976$$
$$= 1,832,097.6 \approx 1,832,098$$

Thus, for 1,832,098 negative samples in 1606 mammogram images, the false positive per image can be calculated as 1,832,098/1606 = 1140.78. Thus, the technique achieves 99.92% sensitivity with 1140.78 false positive per image (corresponding to false positive fraction of 0.1).

Similarly, the false positive per image for the sensitivities of 99.58%, 95.17%, and 74.63% are calculated as 1140.78, 114.07, 11.47, and 1.14.

## 2. Sensitivity with standard deviation values at various FP per image of proposed method and competing techniques

**Table 1: Comparative results of proposed and competing methods in terms of sensitivity values with standard deviation at various FP per image for DDSM database**

| FP per image / Methods | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **mean multiscale 2D NEO** | 85±20 | 96±6.52 | 97±4.47 | 98±2.74 | 100±0 |
| **max multiscale 2D NEO** | 77±24.14 | 88±13.51 | 90±12.25 | 97±6.71 | 98±4.47 |
| **Karale et al.** | 63±20.18 | 77±17.89 | 86±7.42 | 90±6.12 | 91±5.48 |
| **Zhang et al.** | 16±6.52 | 34±20.74 | 47±20.49 | 54±18.51 | 58±16.05 |
| **El-Naqa et al.** | 24±16.36 | 43±24.39 | 64±25.1 | 74±17.82 | 82±13.04 |

**Table 2: Comparative results of proposed and competing methods in terms of sensitivity values with standard deviation at various FP per image for INbreast database**

| FP per image / Methods | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **mean multiscale 2D NEO** | 81.47±12.5 | 93.14±9.6 | 96±8.94 | 100±0 | 100±0 |
| **max multiscale 2D NEO** | 93.8±8.52 | 97.14±6.39 | 97.14±6.39 | 97.14±6.39 | 97.14±6.39 |
| **Karale et al.** | 83.62±15.7 | 90.47±14.68 | 93.8±8.52 | 93.8±8.52 | 93.8±8.52 |
| **Zhang et al.** | 64.28±41.65 | 82.14±20.82 | 88.8±10.96 | 88.8±10.96 | 97.14±6.39 |
| **El-Naqa et al.** | 65.14±19.23 | 79.14±14.82 | 84.14±17.12 | 92.14±11.41 | 92.14±11.41 |

**Table 3: Comparative results of proposed and competing methods in terms of sensitivity values with standard deviation at various FP per image for PGI database**

| FP per image / Methods | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **mean multiscale 2D NEO** | 94.73±11.77 | 100±0 | 100±0 | 100±0 | 100±0 |
| **max multiscale 2D NEO** | 85.14±24.75 | 98.94±2.35 | 98.94±2.35 | 98.94±2.35 | 98.94±2.35 |
| **Karale et al.** | 88.19±8.99 | 94.42±5.97 | 94.42±5.97 | 94.42±5.97 | 94.42±5.97 |
| **Zhang et al.** | 89.13±7.04 | 92.21±1.7 | 95.85±3.84 | 95.85±3.84 | 95.85±3.84 |
| **El-Naqa et al.** | 77.48±24.99 | 89.36±10.35 | 91.71±7.68 | 92.88±7.46 | 92.88±7.46 |

## 3. Estimated values of Gaussian width (σ) and soft margin constant (C) for SVM training

**Table 4: Following values of σ and C are selected for mean multiscale 2D NEO method in each fold of DDSM database:**

| | Fold1 | Fold2 | Fold3 | Fold4 | Fold5 |
|---|---|---|---|---|---|
| **σ** | 0.002 | 0.0078 | 0.0078 | 0.002 | 0.0313 |
| **C** | 512 | 64 | 64 | 512 | 64 |

**Table 5: Following values of σ and C are selected for max multiscale 2D NEO method in each fold of DDSM database:**

| | Fold1 | Fold2 | Fold3 | Fold4 | Fold5 |
|---|---|---|---|---|---|
| **σ** | 0.0156 | 0.002 | 0.0009 | 0.0009 | 0.002 |
| **C** | 64 | 64 | 128 | 64 | 512 |

**Table 6: Following values of σ and C are selected for mean multiscale 2D NEO method in each fold of INbreast database:**

| | Fold1 | Fold2 | Fold3 | Fold4 | Fold5 |
|---|---|---|---|---|---|
| **σ** | 0.002 | 0.0009 | 0.0009 | 0.0039 | 0.0009 |
| **C** | 64 | 64 | 1024 | 128 | 512 |

**Table 7: Following values of σ and C are selected for max multiscale 2D NEO method in each fold of INbreast database:**

| | Fold1 | Fold2 | Fold3 | Fold4 | Fold5 |
|---|---|---|---|---|---|
| **σ** | 0.0009 | 0.0009 | 0.0078 | 0.002 | 0.0039 |
| **C** | 2048 | 1024 | 2048 | 2048 | 2048 |

**Table 8: Following values of σ and C are selected for mean multiscale 2D NEO method in each fold of PGIMER-IITKGP database:**

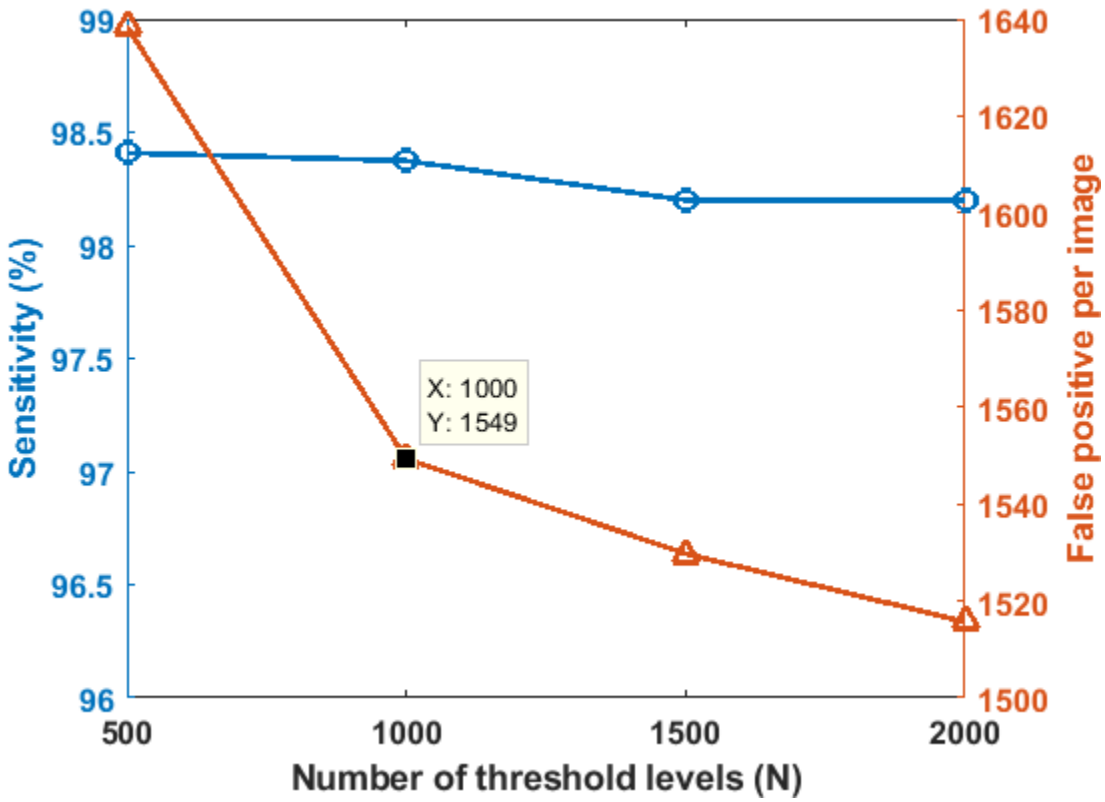
| | Fold1 | Fold2 | Fold3 | Fold4 | Fold5 |
|---|---|---|---|---|---|
| **σ** | 0.0625 | 0.0009 | 0.0625 | 0.0009 | 0.0009 |
| **C** | 128 | 64 | 2048 | 128 | 64 |

**Table 9: Following values of σ and C are selected for max multiscale 2D NEO method in each fold of PGIMER-IITKGP database:**

| | Fold1 | Fold2 | Fold3 | Fold4 | Fold5 |
|---|---|---|---|---|---|
| **σ** | 0.002 | 0.0313 | 0.0156 | 0.0156 | 0.0009 |
| **C** | 128 | 64 | 2048 | 256 | 128 |

## 4. Selection of number of threshold levels (N) in thresholding step

For the purpose of analyzing the effect of number of threshold levels (N), we have analyzed the training data from one of the folds of 5-fold validation on DDSM database. The effect of variation of N on the senstivity and false positive per image of individual microcalcification detection is shown in figure below.




As shown in the figure, increasing the step size from 500 to 2000 does not change the sensitivity significantly. The overall decrease in sensitivity is from 98.41% to 98.2%. The false positive per image decreases by 89 when the N is changed from 500 to 1000, but false positive per image decreases by only 33 when N is changed from 1000 to 2000. On the other hand, increasing the value of N increases the computational burden. Thus, the value of N is chosen as 1000 since the decrease in false positive per image is highest without significant change in sensitivity. Similar observations were made for other folds of all databases.