Expectation Maximization Testing

Contents

[Dataset: 1](#_Toc113272)

[Information about Testing Dataset: 1](#_Toc113273)

[EM-Based Testing: 1](#_Toc113274)

[EM-Based Testing Code: 1](#_Toc113275)

[Sampled Dataset: 1](#_Toc113276)

[Pre-processing: 2](#_Toc113277)

[Expectation Maximization: 2](#_Toc113278)

[Results: 2](#_Toc113279)

# Dataset:

## Information about Testing Dataset:

* + Publicly available dataset -> <https://www.kaggle.com/paultimothymooney/breast-histopathology-images>.
  + Dataset is Labelled as IDC in Breast Cancer.
  + Size of Patch -> 50x50x3.

# EM-Based Testing:

## EM-Based Testing Code:

<https://github.com/SalmanAhmed01911/cancer_reviewer.git>

## Sampled Dataset:

Randomly 12000 Patches where Cancer was not detected and 12000 Patches where cancer was detected.

## Pre-processing:

Each patch was divided by 255.0 to be normalized between 1 and 0.

## Expectation Maximization:

1. Trained Neural Network on labelled data.
2. Thresholding Scheme Applied on **probabilities of range of 0.3 – 0.7** (Thresholding Values and Model Epochs used from Research Paper “WHOLE SLIDE IMAGE CLASSIFICATION VIA ITERATIVE PATCH LABELLING.”).
3. Data within this range was considered as **Non-Discriminative** so neglected.
4. Data outside of range of thresholding scheme was further used for training again and considered as **Discriminative.**
5. Again **Step 1** and iteration continuous till **convergence (**Means There were no Non-Discriminative Patches**).**

## Results:

* **Converged in 11 Iterations.**
* **Non-Discriminative Patches for each iteration** 
  + **8610**
  + **175**
  + **42**
  + **43**
  + **34**
  + **35**
  + **6**
  + **6**
  + **250**
  + **3**
  + **3**
* **Updating Labels against each iteration**
  + **2305**
  + **73**
  + **35**
  + **26**
  + **27**
  + **3**
  + **5**
  + **5**
  + **9**
  + **4**
  + **0**
* **ACCURACY**
  + **Accuracy 84.82% -> Balanced Data Samples (24000)**
* **Remained Patches**
  + No Cancer **9724**
  + Cancer **5069**
* **Discriminative Patches Correctly Detected**
  + No Cancer **8489**
  + Cancer **6304**