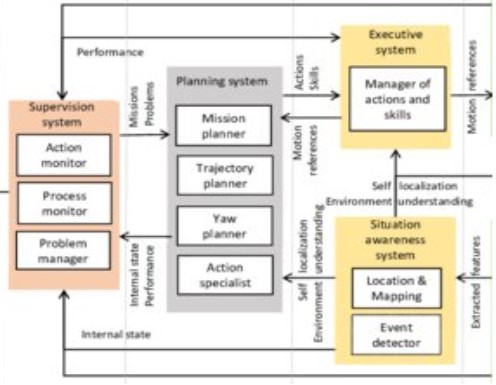
Project Design Phase-I

Solution Architecture

|  |  |
| --- | --- |
| Date | 24-05-2023 |
| Team ID | NM2023TMID22553 |
| Project Name | **Advanced Breast Cancer Prediction With Deep Learning** |
| Maximum Marks | 4 Marks |



The aim of the tissue level is to get the tumor region according to the labeled mask so as to assign the TNM stage of WSI. As shown in Figure [4](https://www.hindawi.com/journals/bmri/2021/2567202/fig4/), the tissue level and cell level share the same basic network. They are all based on DeconvNet which consists of the stacked convolution layers, max-pooling layers, and deconvolution layers. In the front three blocks, max-pooling layers are followed with convolution layers, and ReLU activation layer and convolution layer which form deconvolution layers are followed with upsampling layers which are followed by the last three blocks. Finally, a convolution layer was used to replace fully convolutional layers in the end of the network in order to obtain a density map. Totally speaking,  kernel size, ReLU activation, Mean Squared Error (MSE) loss function, and Adam optimizer were used in all network structures. In order to handle different tasks, the basic network was fine-tuned. As shown in Figure [2](https://www.hindawi.com/journals/bmri/2021/2567202/fig2/), the fully connected layer was connected to the last convolution layer to get confidence result of patch binary classification. With characteristic of DeconvNet, images of any size can be input. However, in actual application, images with the dimensions reduction were used to avoid cutting the edge of the output density map.

For the high-resolution WSI images, the content is usually up to 2 GB for a single image, and the tumor region locates at somewhere in the whole slide. For a large image, the network has to be trained on the image patches generated from the ground scale. In order to keep the patch sequencing, we proposed a LSTM-CFCN-based model for the segmentation task. In this model, the stacked channel FCN is for the patch feature encoding and the LSTM model is used to merge these patches into a large image from these FCN encoded patches; more details are illustrated in Figure [5](https://www.hindawi.com/journals/bmri/2021/2567202/fig5/). In this part, the cancer region in the training data set is labelled by using a mask; the channel FCN is used to estimate the density map of a tumor region in each sequential patch; and LSTM block is used to combine the detected result into an integrated figure. From this point of view, the tumor density map is predicted by the deconvolution part, and a Euclidean distance is used for measuring the difference between the generated density map and the ground truth.

In this segmentation stage, only correctly detected and classified masses by YOLO model were considered and the false predictions were discarded as similarly highlighted in previous works[7](https://www.nature.com/articles/s41523-021-00358-x#ref-CR7),[23](https://www.nature.com/articles/s41523-021-00358-x#ref-CR23). Some cases of mammograms have more than one detected mass lesion, therefore, a total of 1467, 112, and 638 masses were, respectively, considered from the CBIS-DDSM, INbreast, and the private dataset. Our network is applied on single detected ROIs and therefore our intention was to consider mammograms with multiple lesions at the detection stage and treat them separately as single ROIs of mass lesions. The predicted ROI masses were next resized into 256 × 256 using a bi-cubic interpolation in case the original size is small, or using an inter-area resampling interpolation in case it is large. All images were preprocessedto remove additional noise and degradation caused by the scanning technique of digital X-ray mammography[28](https://www.nature.com/articles/s41523-021-00358-x#ref-CR28),[29](https://www.nature.com/articles/s41523-021-00358-x#ref-CR29). Thus, we applied a histogram equalization to enhance the compressed regions and smooth the distribution of the pixels that helps the pixel segmentation. All images were normalized to a range of [0, 1].

To train the proposed segmentation deep learning models, a large amount of annotated samples should be prepared to generalize the learning curve of the models. Due to the limited amount of ROI masses in each dataset, we have augmented the original ROIs four times by rotating them with the angles Δθ = {0°, 90°, 180°, 270°}. We have also transformed them twice differently using the Contrast Limited Adaptive Histogram Equalization (CLAHE) method. Consequently, raw data of single ROI images were augmented six times into a total of 8802, 672, and 3828 of ROI masses were, respectively, prepared from the CBIS-DDSM, INbreast, and private set to train and propse the architecture.