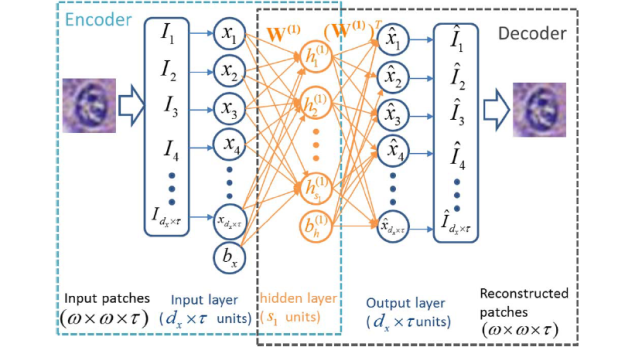
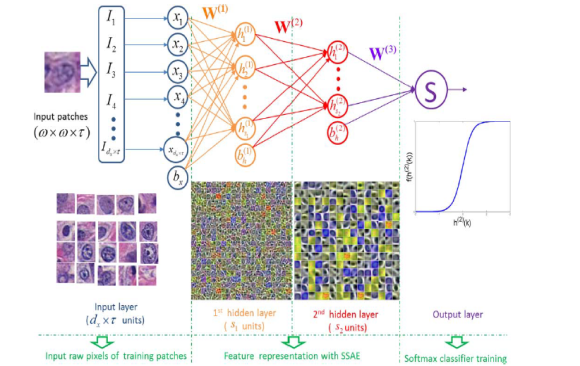
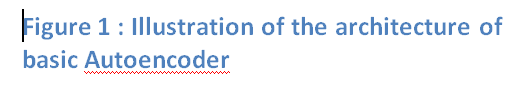
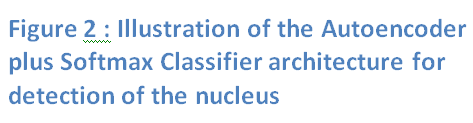
***Deep Autoencoder Network for automated nuclei detection and segmentation in Breast Cancer Histopathology Images***

Autoencoders are Neural Networks for unsupervised learning of efficient encoding. The aim of an autoencoder is to learn a representation (encoding) for a set of data, typically for the purpose of dimensionality reduction. Recently, the autoencoder concept has become more widely used for learning generative models of data. For the purpose of automated nuclei detection in Breast Cancer tissues we use a 2 layer deep autoencoder followed by a Softmax Classifier. The main challenge for unsupersived learning is lack of properly labelled data. In our case, we have 6000 colored images of background and 6000 images of cells. The Autoencoder learns high-level features from just pixel intensities alone in order to identify distinguishing features of nuclei. Autoencoder is an encoder-decoder architecture where the “encoder” network represents pixel intensities modeled via lower dimensional attributes, while the “decoder” network reconstructs the original pixel intensities using the low dimensional features. For our application, the size of nuclear and non-nuclear patches was set to 61 X 61 pixels. Additionally, each image patch may contain up to a single object (in this case a nucleus) that would be appropriate for construction of a full connection model.

Deep Autoencoders have several layers of the basic Autoencoder that learning the parameters for higher level representation while simultaneously minimizing the discrepancy between the input image and its reconstruction. As a result, each patch is represented by a higher level representation of nuclei and non-nuclei patches in the subsequent layers. Our 2 layer deep autoencoder had 400 and 255 hidden units in the first and the second layer respectively. During training the autoencoder, the pixel intensities of the input patch is fed to the first layer and it is allowed to reconstruct the original image while adjusting the weights. After the first layer of the autoencoder is trained, this primary representation is then fed to the second layer as the input which learns the secondary representation by adjusting its weights. Note that in the autoencoder learning procedure, the label information is not used. Therefore, autoencoder learning is an unsupervised learning scheme.

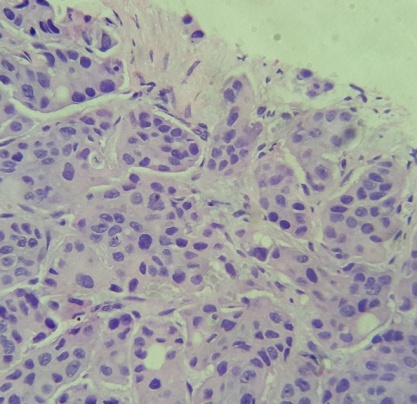
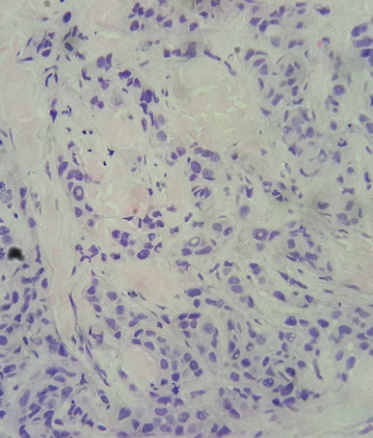
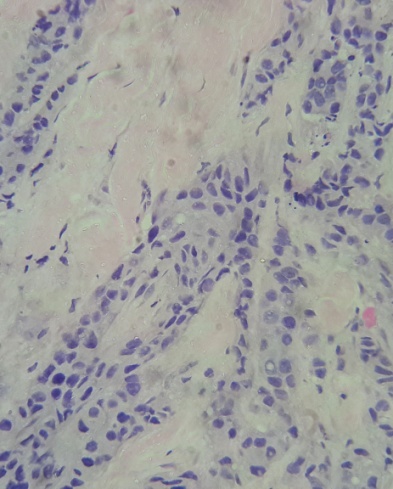
Softmax Classifier is a supervised learning model which generalizes Logistic Regression. The input of the softmax classifier is the higher level feature representation learned by the autoencoder. By minimizing the cost fuction via Gradient Descent, the parameters of the softmax Classifier is learnt. During detection process, each image patch detected by a sliding window is first represented by high-level feature by the autoencoder network. This is then fed to the SMC that obtains the probability of the patch containing the nuleus or not.

***Figure 3 : Batch loss curve while training layer 1 of the autoencoder***

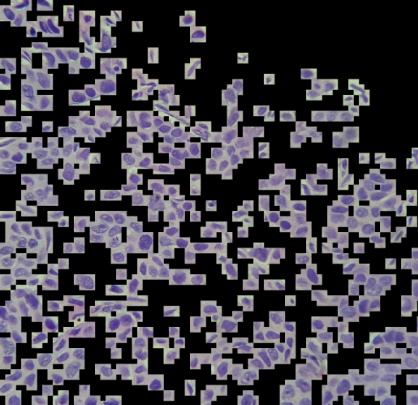
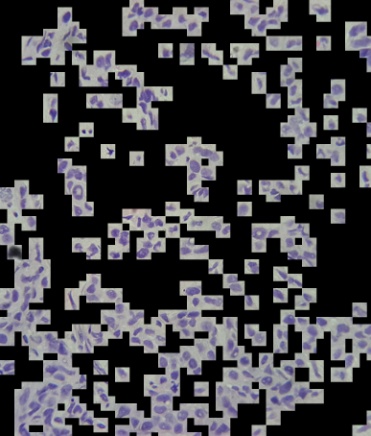
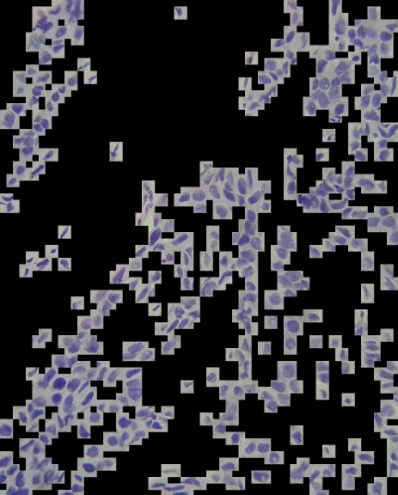
***Figure 4 : Batch loss curve while training layer 2 of the autoencoder***

***Figure 5 : Batch Loss curve while training softmax classifier***

Our strategy involves identifying the presence or absence of a nucleus within every individual image patch in a histopathologic image. A sliding window scheme is used to select candidate patches. Since the sliding window detector will typically evoke multiple responses around target nuclei, non-maxima suppression is applied to only retained those evoked responses above a pre-defined threshold. The threshold and overlapping rate for the non-maxima suppression algorithm are empirically defined as 0.8 and 50%, respectively.

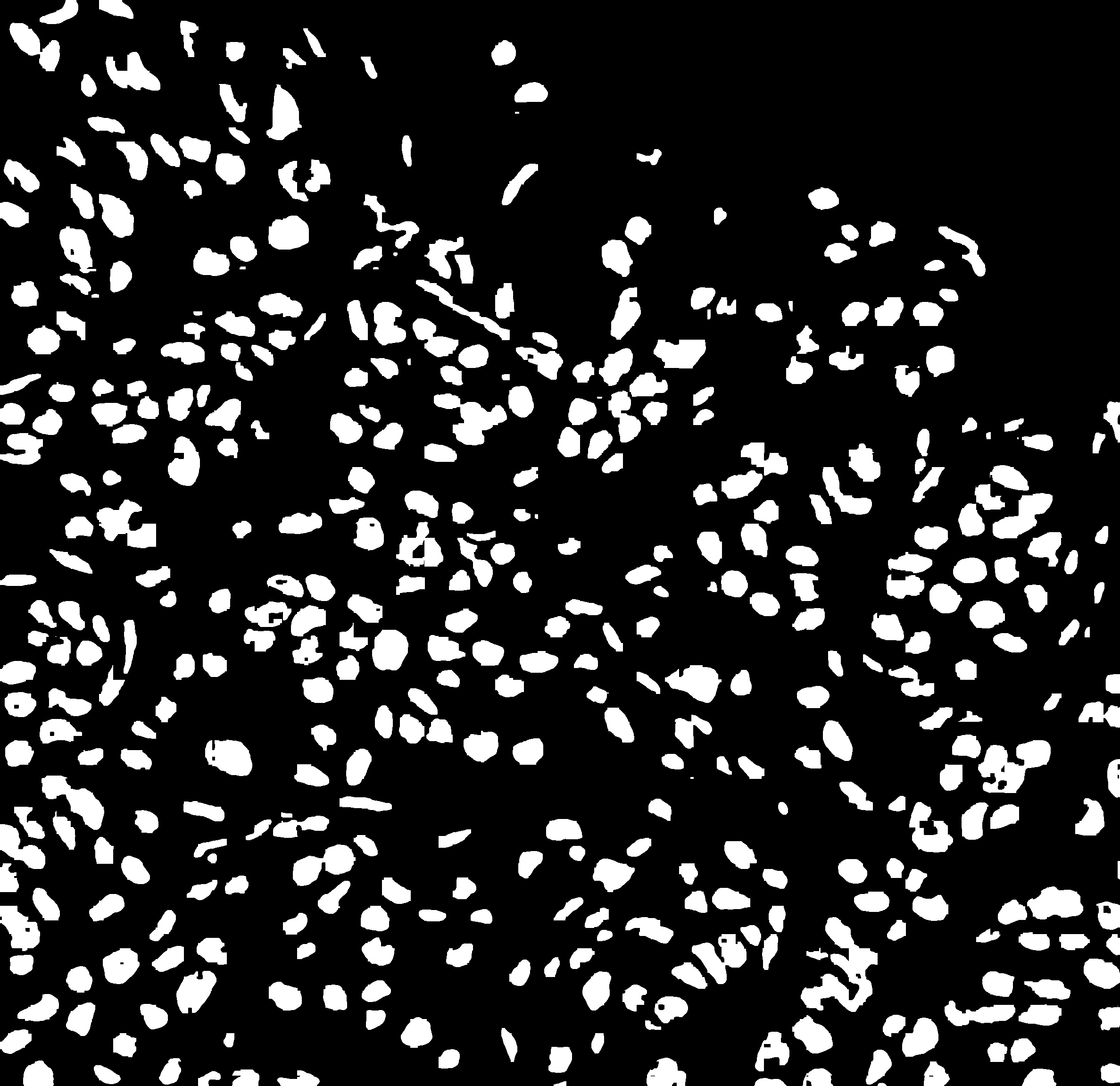
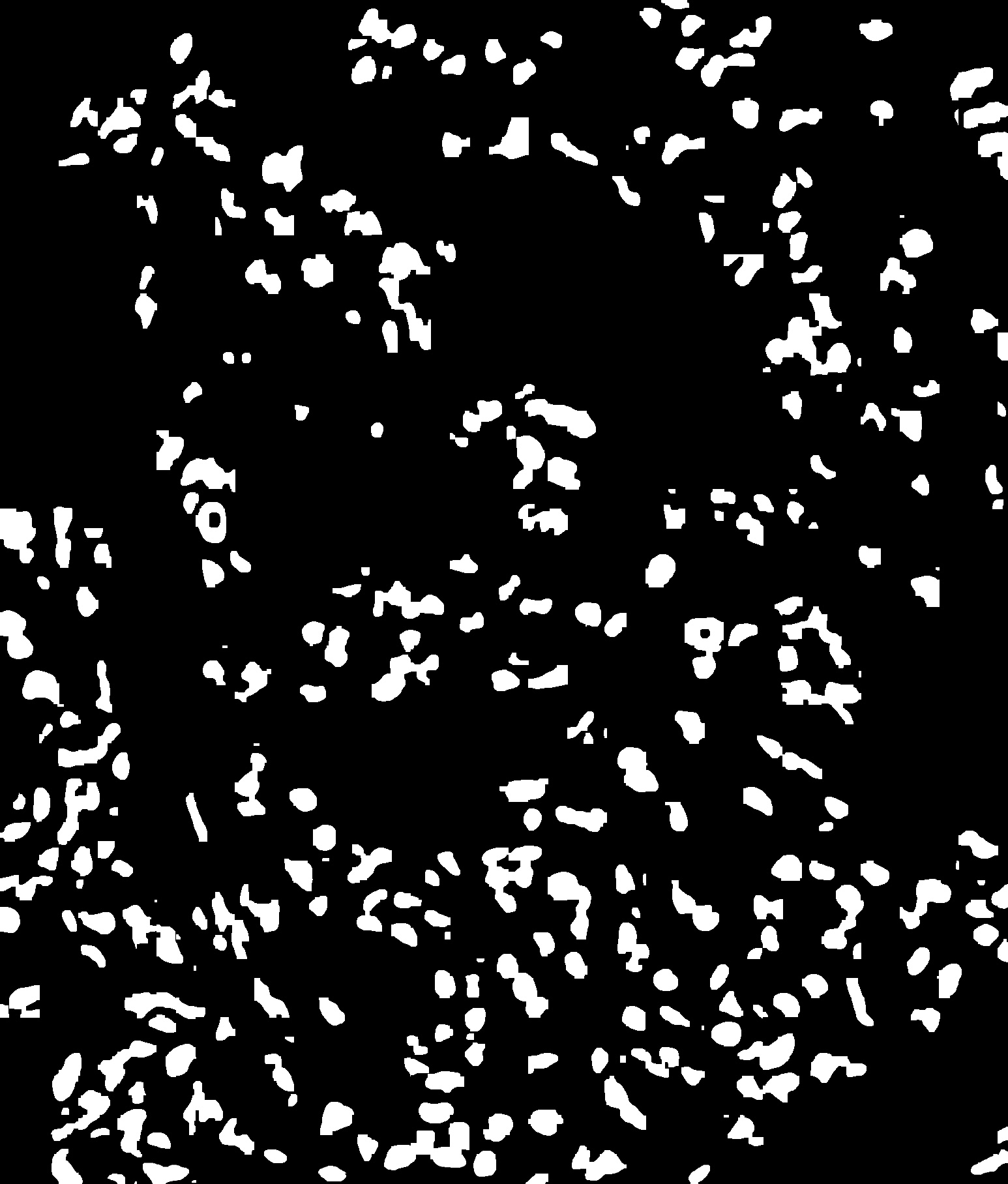
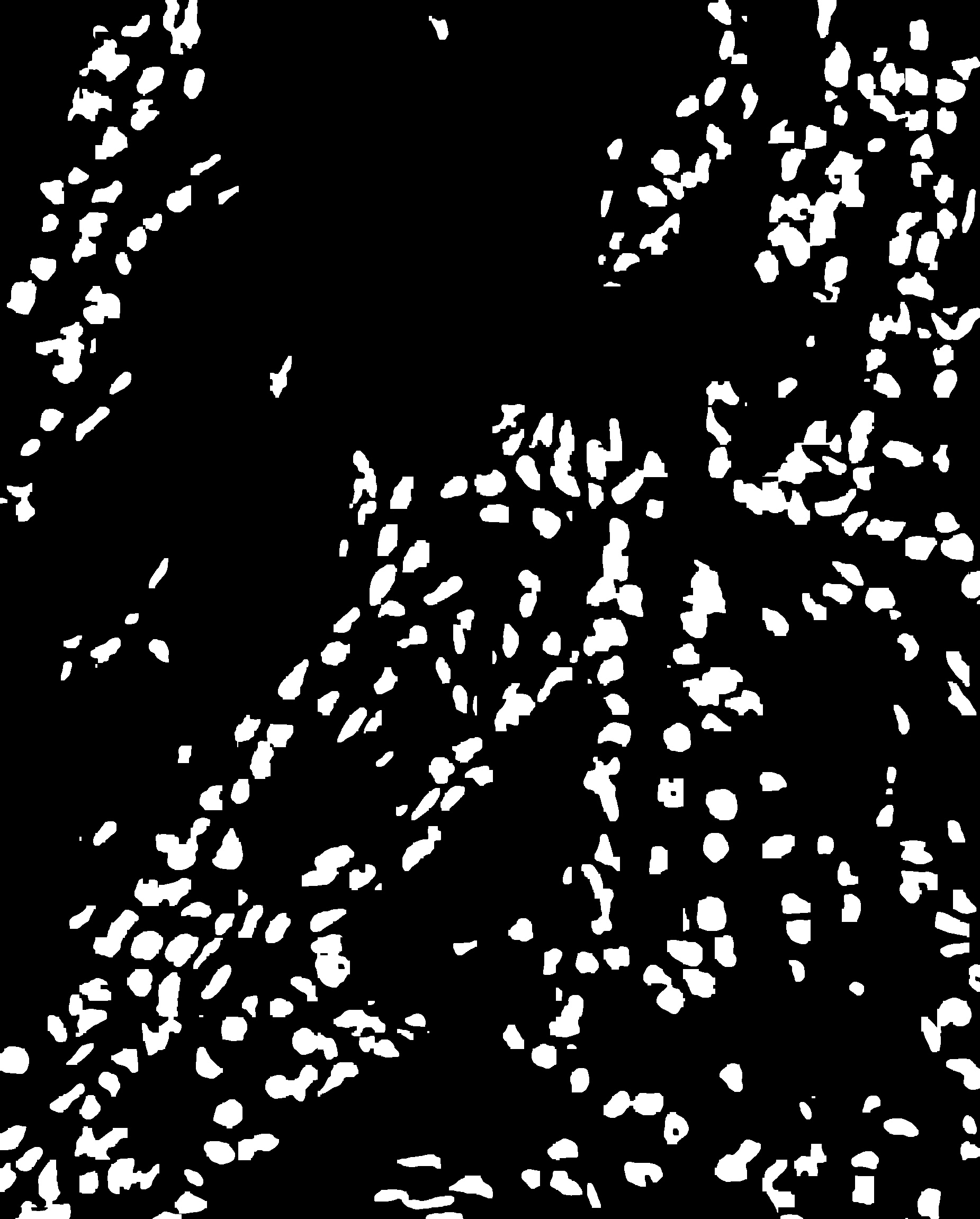
  

***Figure 6 : Original Breast Cancer Histopathology Images***

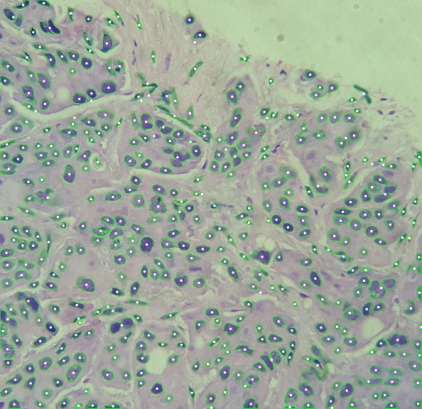
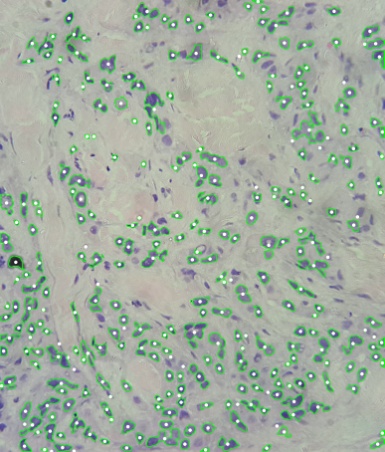
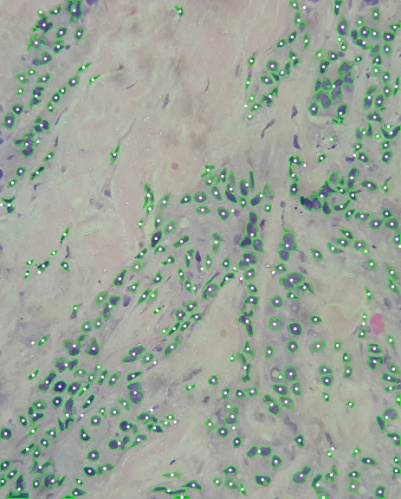
***Figure 7 : Output of the trained autoencoder plus softmax classifier architecture using sliding window***

As we can see that the autoencoder and the softmax classifier network accurately detects the regions where the nuclei as present. Further to remove the white spaces around the nuclei and make the segmentation map more accurate, we apply otsu filter to each of the window that has been detected to contain a nuclei by the autoencoder and the softmax classifier network. To further improve the model by making the output mask noiseless, we apply Gaussian Blur having a kernel size of 7 X 7 and opening and closing morphological transforms before the otsu filter. Figure 5 shows the output mask of the model as described above.

***Figure 8 : Mask generated when otsu filter is applied to each sliding window containing the nuclei.***

Lastly, the goal of the model is to generate the final output image that correctly identify the border of the nuclei and the center. This is obtained by locating and drawing the contours in the mask generated when otsu filter as applied to each of the window which is detected to contain a nuclei by the autoencoder and the softmax classifier network.

***Figure 9 : The final output.***

Hence, we see that with the help of an unsupervised learning based autoencoder network, softmax classifier and otsu filter, we can design a model which can efficiently detect and segment the nuclei from breast cancer histopathology images. This deep learning model is robust to noise and the shape of the nuclei. It can work with high rsolution images and is capable of detecting the shape and the center of the nuclei irrespective of the location of the nuclei in the image.