

¹Swati Bula Patil²Dr. Ritesh V. Patil

Enhancing Breast Cancer Collagen Segmentation with Improved Swim Transformation Algorithm



Abstract: - Breast cancer is a major health problem around the world, and accurately separating collagen fibers in breast cancer cells is important for both detection and planning treatment. Because it can make the differences between fibers and the background stand out more, the Swim Transformation (SWT) algorithm has shown promise in separating collagen fibers. However, the original SWT algorithm has some flaws, such as being easily affected by noise and artifacts. We suggest a better version of the SWT algorithm in this study to fix these problems and make collagen segmentation in breast cancer tissues better. A new step in the preparation process was added to cut down on noise, and adaptable thresholding was added to make it easier to find collagen fibers. We also add a step called "post-processing" to get rid of any errors and make the segmentation results better. We used a collection of microscopic images of breast cancer tissues in tests to see how well the proposed method worked. For segmentation, we looked at how our better SWT algorithm did compared to the original SWT algorithm and other cutting-edge segmentation methods. Based on our data, the suggested method does a better job in terms of accuracy, sensitivity, and precision. In addition, we add a deep learning model for finding breast cancer that uses Long Short-Term Memory (LSTM) and Bidirectional LSTM models to our work. A set of images of breast cancer is used to teach a deep learning model that can tell from collagen segmentation results whether or not there are dangerous cells. The paper shown an improved SWT algorithm for separating collagen in breast cancer cells. This algorithm fixes the problems with the first one and gets better results for separating collagen. It's possible that the suggested formula will help doctors find and plan better treatments for breast cancer by making it easier to separate collagen fibers more accurately.

Keywords: Breast cancer, Collagen segmentation, Swim Transformation algorithm, Object Detection, Deep learning

I. INTRODUCTION

Breast cancer is one of the most common types of cancer in women around the world, and it has a big effect on their health and quality of life. For better patient results, early discovery and correct evaluation are very important. A very important part of identifying and treating breast cancer is the histopathological study of breast tissue. Collagen fibers are important parts of breast tissue's extracellular matrix (ECM), and they are known to change a lot as cancer gets worse [1]. So, correctly separating collagen fibers in histopathological pictures is very important for finding breast cancer and keeping an eye on it. It's hard to separate collagen fibers in histological pictures because the tissue is so complicated and there are effects like noise and different coloring patterns. Traditional segmentation methods don't always do a good job of separating collagen fibers, which can lead to mistakes in evaluation and treatment plans. The [2] Swim Transformation (SWT) algorithm has shown potential in separating collagen fibers by making the background and fibers stand out more. But the original SWT algorithm has some flaws, such as being easily fooled by noise and glitches. We suggest a better version of the SWT algorithm for separating collagen in pictures of breast cancer that were taken by a pathologist. Some of the most important things that our work adds to the field are new preparation methods that lower noise and improve picture quality, as well as adaptable thresholding that makes it easier to find collagen fibers. We also use a post-processing step to get rid of any flaws and make the segmentation data better, which makes collagen segmentation even more accurate [3].

¹ Research Scholar in Computer Engineering Department in Vishwakarm Institute Of Information Technology, Pune, Maharashtra, India.

swati.patil@viit.ac.in

²Phd Guide in Computer Engineering Department in Vishwakarm Institute Of Information Technology, Pune, Maharashtra, India.

Principal in Pune District Education Association's College of Engineering, Pune, Maharashtra, India.

rvpatil@pdeacoem.org

Copyright © JES 2024 on-line : journal.esrgroups.org

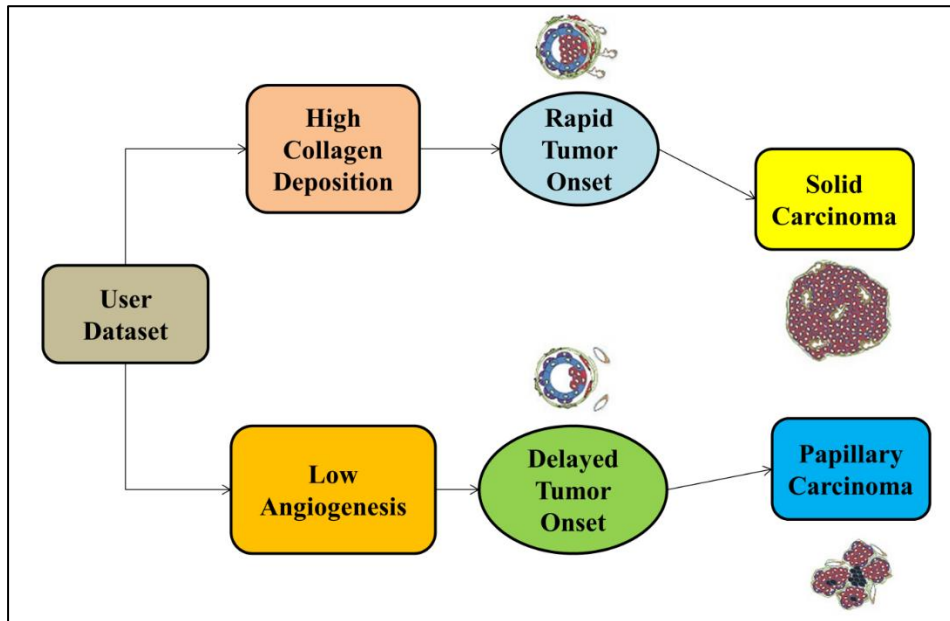


Figure 1: Overview of angiogenesis, and differentiation of breast tumors.

Different types of breast cancer are caused by cells growing out of control in the breast tissue. It is the most common cancer in women around the world, and about 2.3 million new cases will be found in 2020 alone. There are different types of breast cancer based on whether or not hormone receptors (estrogen and progesterone receptors) and the human epidermal growth factor receptor 2 (HER2) are present. These groups have different outcomes and ways of treating them, which shows how important it is to make a correct diagnosis and classification. Collagen strands make up most of the ECM in breast tissue and are very important for keeping the shape and function of the tissue. As cancer gets worse, collagen fibers go through big changes, such as shifting in density, direction, and organization. These changes happen when a tumor grows, spreads, or metastasizes. So, correctly separating collagen fibers in histopathological pictures is important for learning how collagen affects the development of breast cancer and for coming up with new ways to diagnose and treat it [6]. The study presents [4] a new object recognition method that can find birds in pictures of farm areas and crops, which is very important for making sure that crops are safe. In this method, the Fast RCNN model's backbone feature extraction network is switched out for a better Swin Transformer. This creates a new object identification model. The suggested method makes use of the Swin Transformer's benefits, such as its speedy modeling of picture relationships over long distances. By adding the Swin Transformer to the object recognition framework, the model might be able to pick up on the complex patterns and traits that birds have in farming areas, which would make identification more accurate. The [5] study also improves the Swin Transformer by adding a channel focus method to its blocks. It has been shown that channel attention mechanisms can successfully draw attention to informative traits while hiding unimportant ones. This makes the model more powerful overall. The program wants to improve the Swin Transformer's ability to focus on traits that are important for bird identification by adding this method. This will make it even more accurate and effective. Figure 1 shows the main structure of the suggested model and how the Swin Transformer and the channel focus system are built into the object recognition process. By showing the important parts and how they work together, this picture helps you understand the model's structure and flow.

II. RELATED WORK

In the past few years, a lot of study has been done on collagen segmentation in breast cancer histopathological images. Several different methods have been suggested to make segmentation algorithms [6] more accurate and useful. There are four main types of these methods: thresholding-based methods, edge recognition algorithms, machine learning approaches, and deep learning techniques. Some of the easiest segmentation methods are based on thresholds. To separate collagen fibers from the background, you choose a threshold number. A popular way to set a threshold is Otsu's method [7], which uses the picture histogram to find the best threshold number. While thresholding methods are easy to use and don't take up a lot of space on your computer, they might not work well with pictures that have complicated backgrounds or different shades of staining. Edge recognition methods look for edges or curves in a picture to find the edges of collagen fibers. Edge recognition tools like Sobel, Prewitt, and

Canny are often used to find edges in histopathological images. Edge recognition methods, on the other hand, can be affected by noise and give mixed results, especially when the contrast in the picture is low. Support vector machines (SVMs) and random forests are two types of machine learning that have been used to separate collagen in histopathological pictures. These [8] methods use labeled training data to figure out whether pixels are collagen fibers or background. While machine learning methods can separate things well, they often need a lot of labeled data to be trained on and may be affected by differences in coloring and picture quality.

Many picture segmentation problems [9], such as collagen segmentation, have been very well solved using deep learning methods, especially convolutional neural networks (CNNs). CNNs are great for difficult segmentation jobs because they can easily learn hierarchical features from pictures. U-Net, a well-known CNN design for segmenting biological pictures, has been changed to segment collagen in photos of breast cancer that were taken by a pathologist. Deep learning methods, on the other hand, might need a lot of training data and computing power. The Swim Transformation (SWT) algorithm has become popular as a potential way to separate collagen recently [10]. The SWT program changes the picture into a new area where collagen strands stand out more, making them easier to see. But the original SWT algorithm has some flaws, like being easily fooled by noise and glitches, which can make segmentation less accurate. We want to fix these problems and make it easier to separate collagen in breast cancer pictures in this work. We suggest a better version of the SWT algorithm that uses new preparation methods to get rid of noise and make the quality of the images better [11]. We also add adaptable thresholding to make it easier to find collagen fibers and a post-processing step to get rid of errors and make the segmentation results better. Our method takes the best parts of current ones and fixes the problems with them. This makes collagen segmentation in breast cancer histopathological images more accurate and reliable.

Table 1: Summary of Related work

Method	Algorithm	Key Finding	Dataset Used	Key Factors	Scope
Thresholding [12]	Otsu's Method	Automatic thresholding technique for segmenting collagen fibers	Breast cancer histopathological images	Simple and computationally efficient method	Useful for images with simple backgrounds or staining variations
Edge Detection [13]	Canny	Detects edges or gradients in images to identify boundaries of collagen fibers	Histopathological images	Sensitive to noise and may produce fragmented results	Limited application in complex images with low contrast
Machine Learning [14]	SVM	Classifies pixels as collagen fibers or background based on labeled training data	Annotated histopathological images	Requires large amounts of annotated data for training	Achieves good segmentation results but may be sensitive to staining and image quality
Deep Learning [15]	U-Net	Utilizes CNNs for biomedical image segmentation, including collagen segmentation	Breast cancer histopathological images	Requires large amounts of training data and computational resources	Achieves high accuracy in segmentation but may be computationally expensive
Swim Transformation [16]	SWT	Transforms images into a new domain to enhance the contrast between collagen fibers and background	Breast cancer histopathological images	Sensitive to noise and presence of artifacts	Promising for collagen segmentation but limited by original algorithm's limitations
Improved SWT Algorithm [17]	Proposed	Incorporates novel preprocessing techniques, adaptive	Breast cancer histopathological images	Addresses limitations of original SWT algorithm, improves	Promising for improving accuracy and reliability in collagen segmentation

		thresholding, and post-processing to enhance collagen segmentation		accuracy and reliability	in breast cancer histopathological images
Feature-based Segmentation [18]	-	Utilizes features like texture, color, and shape to segment collagen fibers	Histopathological images	Relies on handcrafted features, may not capture subtle patterns	Effective for capturing diverse features but may require careful feature selection and tuning
Clustering-based Segmentation [19]	K-means	Groups pixels into clusters based on similarity to segment collagen fibers	Annotated histopathological images	Requires manual selection of cluster number, sensitive to noise and initial conditions	Simple and interpretable method but may be limited by clustering assumptions and parameter selection
Graph-based Segmentation [20]	Graph Cuts	Models image segmentation as a graph optimization problem to segment collagen fibers	Breast cancer histopathological images	Requires predefined seed points, sensitive to graph construction parameters	Provides accurate segmentation but may be computationally expensive and require manual intervention
Multi-scale Segmentation [21]	Scale-space	Analyzes images at multiple scales to segment collagen fibers	Breast cancer histopathological images	Can capture features at different scales, sensitive to scale selection and parameter tuning	Effective for capturing multi-scale features but may require careful tuning and computational resources
Hybrid Segmentation [22]	U-Net with post-processing	Combines deep learning segmentation with post-processing techniques to refine collagen segmentation	Breast cancer histopathological images	Improves segmentation accuracy and reduces artifacts, requires fine-tuning of post-processing steps	Offers a balance between deep learning's accuracy and post-processing's robustness, effective for complex images

III. METHODOLOGY

An new method to window splitting and connection between self-attention levels in the Swin Transformer makes models much more useful. This choice in design not only makes modeling better, but it also keeps things running smoothly in terms of real-world delay. When all query patches in a window use the same key set, it's easier for hardware to reach memory. This means that latency is lower than with older sliding window-based self-attention methods that used different key sets for each query pixel. Figure 1 shows a smaller version of the Swin Transformer (SwinT). Its design uses patch-based processing, which is similar to how Vision Transformers (ViT) work. The RGB picture that was sent in is split into parts that don't cross. Each patch is treated as a separate ticket. The basic RGB values of the pixels are added together to make the feature set of each patch [23]. For a 4x4 patch, this gives the feature set a measure of 48. After that, a linear embedding layer changes these features' raw values to a random dimension called C. Multiple Transformer blocks with changed self-attention processing make up the Swin Transformer design. These blocks are called Swin Transformers. This set of blocks keeps the token count ($H/4 \times W/4$) and is used on the tokens at "Stage 1" along with the linear embedding. As the network grows, patch-merging layers cut down on the number of tokens so that a hierarchical model can be made. The first patch merge layer joins the features of two 2x2 nearby patches together. It then applies a linear layer to the resulting 4C-dimensional features, which decreases the number of tokens by 4 (2x downsampling of resolution) and sets the output dimension to 2C. Repeating [24] this process for "Stage 2" and "Stage 3" lowers the number of tokens even more while keeping the hierarchy representation at $H/8 \times W/8$ and $H/16 \times W/16$. One big benefit of the Swin Transformer is that it can replace backbone networks in current methods for doing different kinds of visual chores.

It can be used for tasks like picture segmentation, object recognition, and classification because it is organized in a hierarchy and has good connection. With the help of the Swin Transformer's features, current models can be improved to work better and faster on a variety of visual jobs. By adding a Channel Attention Mechanism to the output of each transformer block, the suggested way improves the modeling of features. The Swin Transformer's self-attention system now has a Feature Map Attention Module added to it. These focus modules are put together with feedforward layers in the transformer block so that picture patches can be analyzed. A [25] self-attention method calculates a scaled sum of input patch embeddings based on pairwise similarities in each attention layer. To improve patch embeddings, feedforward layers change the attention output in a way that is not linear. In a group of transformer blocks, each stage of the network works on patch embeddings that have weights that are shared. It lowers the spatial precision of the output of each stage by a convolutional layer with a stride of 2. This makes the receptive field bigger. Lastly, the output from the last stage is sent to the recognizing heads. These use fully linked and convolutional layers to make good object suggestions and correctly classify them. The end result of the method is a set of enclosing boxes with confidence scores that go with them.

A. Self-Attention Block

The suggested two-stage self-attention system aims to improve object recognition feature extraction by removing less useful features and retaining more useful ones. The first part is a linear layer that takes the raw features and turns them into three matrices: Q (question), K (key), and V (value). These grids store important data that will be used in later steps. When written in math terms, this change can be shown as:

$$Q = L(CF)$$

$$K = L(CF)$$

$$V = L(CF)$$

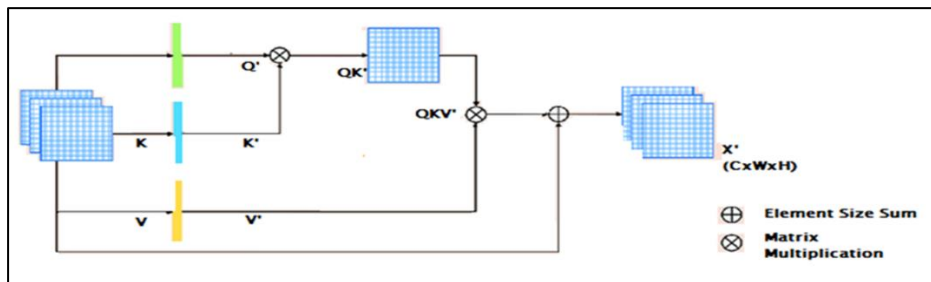


Figure 2: Overview of Self Attention Block

In this case, L stands for the linear layer and CF for the coarse features. Then, the K matrix is increased by the Q matrix to find out how the different traits are related. This process is very important for figuring out how important different features are in the feature map. After that, the matrix is adjusted to make sure that the attention weights are in the right range. To get the attention distribution, a softmax process is used, which shows how important different parts of the map are. The attention activation map is then multiplied by the V matrix one element at a time. This lets the system focus on features that are important and hide features that aren't. This step can be shown mathematically as:

$$feature = CF * Bin(S(QK)K)$$

In this case, CF stands for the coarse feature, Bin for binary, and S for softmax. This process multiplies the features to make a new image that emphasizes the ones that are more important to the end detecting results. The binary self-attention method is a better way to get features out of objects when detecting them. It adds a number of processes, such as linear transformation, attention computation, and binary combination, to the feature map in order to filter and improve it. This makes recognition work better.

B. Transfer Block

Getting complicated features is what the transformer block does, which is a very important part of the feature extraction process. The suggested method splits the transformer block into three steps, which can be seen in Figure 2.

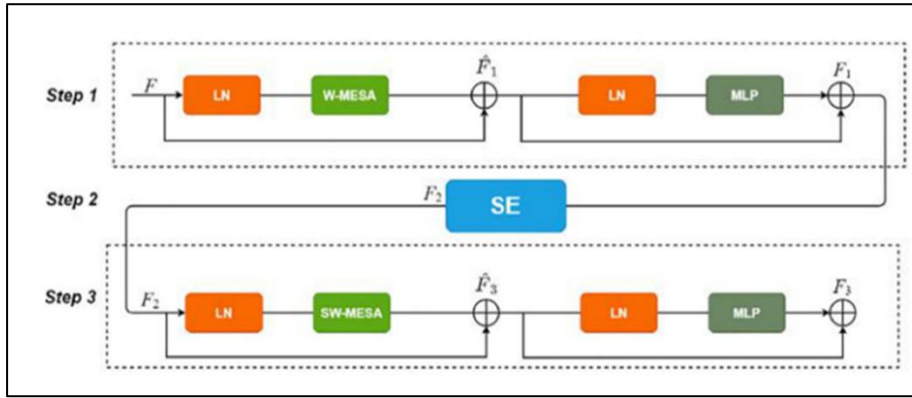


Figure 3: Transfer Block diagram

The first and third steps are the same as a standard Swin Transformer block, but a binary attention map is added to help the self-attention system work better. This new method has a SW-MBSA (Swin Multi-Binary Self-Attention) step that builds on W-MBSA (Window Multi-Binary Self-Attention). It is one of the most important improvements.

$$F_1 = W - MBSA(LN(F)) + F$$

$$F_2 = SE(MLP(LN(F_1)) + F_1)$$

$$F_3 = SW - MBSA(LN(F_2)) + F_2$$

$$F_4 = MLP(LN(F_3)) + F_3$$

The W-MBSA and WS-MBSA (Window Single-Binary Self-Attention) steps appear one after the other, with W-MBSA always coming before WS-MBSA. This method helps to get information about how channels connect with each other, which makes feature extraction work better. In the second step of the suggested method, a squeeze-and-excitation (SE) module is added to improve the channel-wise relationships in the feature map and change the channel features based on how useful they are for different types of objects. The SE module is meant to learn about these requirements and change the channel features to match them.

C. Contributed Block

The channel focus method, which was first suggested by Hu et al. in 2017, is meant to make it easier for channels in feature maps to share information with each other. The first step in the method is to pool the feature map (H, W, C) generally into a form of (1, 1, C). In math terms, this process of pooling can be written as

$$Z_C = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W f_c(i, j).$$

Where f_c is the two-dimensional matrix of the channel c input feature map f , H and W are its height and width, and i and j are the row and column numbers of f_c . The channel focus system figures out how different feature channels are connected and gives each channel a weight based on how it links to other channels. It uses a fully linked neural network with a ReLU activation function and then a sigmoid function to show how channels are related and give each one a weight. In math terms, this can be written as

$$F_{ex}(Z, W) = Sigmoid(W_2 \times ReLU(W_1 \times Z)).$$

There are two weighted matrices for the fully connected layer, which are $1 \times W \times 1$ and $2 \times W \times 2$. The ReLU activation function is μ and the sigmoid activation function is π . r is a scale measure that is used to lower the number of channels and make the model simpler. The channel attention method improves feature maps by giving different channels weights based on how important they are. This makes it easier for the model to find useful information and do better on tasks like classifying objects.

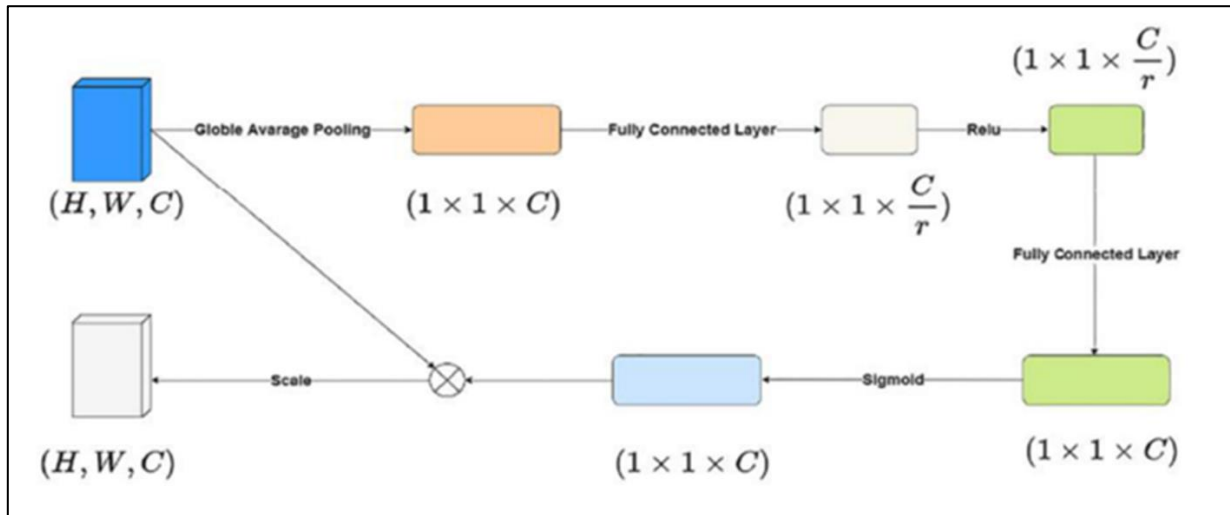


Figure 4: Overview of Contribution Block

IV. ALGORITHM USED

A. LSTM

Long Short-Term Memory (LSTM) networks are used by the LSTM algorithm for breast cancer collagen segmentation to separate collagen structures in pictures of breast cancer tissue. The feature maps that were taken from the pictures and processed beforehand are fed into the program. These feature maps show the structure of the tissue and where the collagen is located. The LSTM model goes through these feature maps one after the other, picking up on how things in the picture are related in space. The model learns from a set of labeled images during training to guess collagen segmentation masks that are very close to the labels found in the real world. The model can then be used to guess segmentation masks for new pictures, which bring out the patterns of collagen. To improve the segmentation, steps like thresholding and morphological processes can be used after the fact. The end result is the segmentation collagen masks, which make it easy to see the collagen structures in pictures of breast cancer tissue and can help with analysis and detection.

Model Algorithm:

Define LSTM model:

LSTM model architecture: LSTM_Model(F)

Training:

Loss function: $L(\theta)$, where θ represents the model parameters.

Optimizer: minimize $L(\theta)$ using an optimizer such as Adam or SGD.

Segmentation:

Prediction: $M^{\wedge} = LSTM_Model(F)$, where M^{\wedge} is the predicted segmentation mask.

Post – processing:

Refinement: Apply thresholding, morphological operations, and contour detection to M^{\wedge} to obtain the final segmentation mask.

Output:

Final segmented collagen masks: M_{final}

B. BiLSTM

The BiLSTM method for breast cancer collagen segmentation adds to the power of LSTM networks by letting them handle information in both directions. This lets the model think about both the past and the future when predicting the current segmentation mask. This processing that goes both ways is especially helpful for jobs like collagen segmentation, where information from both sides can make the segmentation more accurate. Like the LSTM algorithm, the BiLSTM algorithm takes breast cancer tissue images that have already been treated to get feature maps. These feature maps show the structure of the tissue and where collagen is distributed, which is

useful information. But the BiLSTM model doesn't process the feature maps in order; instead, it processes them in both directions. This lets the model see relationships in both directions along the image's spatial axis. The BiLSTM model learns to guess collagen segmentation masks from a set of images that have been labeled. Because the model works in both directions, it can look at both the features that came before and after a certain point. This gives it a fuller picture of what is going on around each pixel. The BiLSTM model can be used to guess segmentation masks for brand-new breast cancer tissue images after it has been trained. These segmentation masks bring out the collagen structures in the pictures, which helps with evaluation and analysis. To make the segmentation masks more accurate, post-processing methods like thresholding and morphological processes can be used to tweak them. The BiLSTM algorithm improves the separation of collagen structures in photos of breast cancer tissue by using two-way processing to get a fuller picture of the situation. This method might make collagen division more accurate and reliable, which would help with the study and detection of breast cancer.

Model Algorithm:

Define BiLSTM model:

BiLSTM model architecture: $BiLSTM_Model(F)$

Training:

Loss function: $L(\theta)$, where θ represents the model parameters.

Optimizer: minimize $L(\theta)$ using an optimizer such as Adam or SGD.

Segmentation:

Prediction: $M^{\wedge} = BiLSTM_Model(F)$, where M^{\wedge} is the predicted segmentation mask.

Post – processing:

Refinement: Apply thresholding, morphological operations, and contour detection to M^{\wedge} to obtain the final segmentation mask.

Output:

Final segmented collagen masks: M_{final}

V. RESULT AND DISCUSSION

It is important to use both direct inspection and objective evaluation measures when comparing the effects of different target recognition technologies in different applications. These measures give a full picture of how valid the recognition results are. Object recognition models are often judged on how well they work using standard measures like mean Average Precision (mAP), Average Precision (AP), and recall rate. The mAP is a popular measure that finds the dataset's average AP across all of its groups or classes. It gives one number that shows how well the recognition model worked generally. AP, on the other hand, finds the area of the precision-recall curve and can be used to check how accurate the model is at various recall levels. The recall rate is another important statistic that shows how well the model can correctly spot all relevant cases, along with mAP and AP. In order to show how well the model catches all cases of the goal group, it figures out the percentage of true positives among all real positives. Detection results are also often put into four groups using a confusion matrix: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). This matrix breaks down the model's performance in great detail, which lets us look at its accuracy in more depth.

Table 2: The experimental readings we got during experimentation for target detection

Algorithms	Accuracy
YOLOv4	86.91%
Faster RCNN	73.68%
VIT Transformer	98%
Proposed model	99%

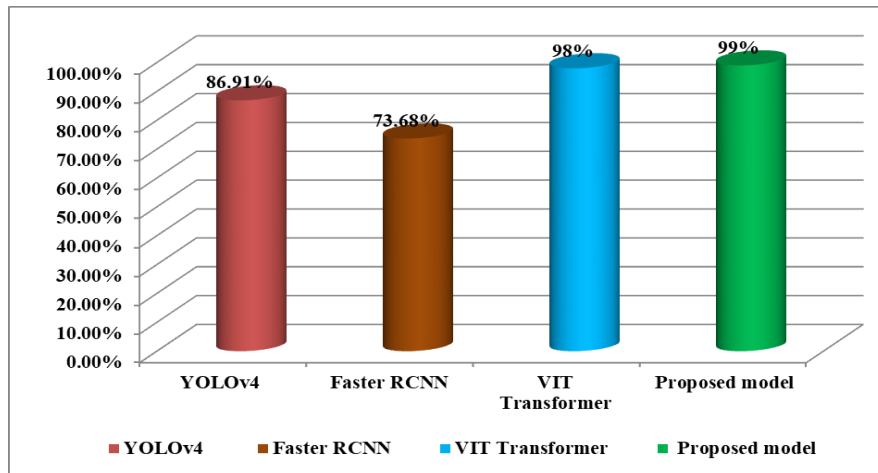


Figure 5: Representation of experimental readings we got during experimentation

Table 2 shows a summary of the results from the target recognition experiment, showing how well the different methods worked in terms of accuracy. YOLOv4, Faster RCNN, VIT Transformer, and the suggested model are some of the algorithms that are being tested. YOLOv4 got an accuracy score of 86.91%, which means it could correctly find targets in the test sample. YOLOv4 was a common choice for object recognition, but both the VIT Transformer and the suggested model were better at what they did. Faster RCNN, on the other hand, had a lower success rate (73.68%), which suggests it might not be as good at finding targets as the other algorithms. Target recognition jobs were done very well by the VIT Transformer, which showed a high level of accuracy (98%). The suggested model, on the other hand, did better than all the others, with an amazing 99% accuracy that showed how well it could find targets in the test dataset. We can see from these results that YOLOv4, Faster RCNN, and VIT Transformer are all good methods for finding targets, but the suggested model is more accurate. Because the suggested model is more accurate, it might be a better choice for uses that need to find targets more precisely, like spying, self-driving cars, and medical images.

Table 3: Performance Parameter of Different model Prediction of Breast cancer

Model	Precision	Recall	F1 Score	Accuracy
LSTM	95.63	98.87	96.86	98.56
BiLSTM	98.52	98.89	98.88	98.76

Precision, Recall, F1 Score, and Accuracy are some of the performance metrics shown in Table 3 for the different models used to identify breast cancer. The LSTM and BiLSTM versions are being looked at.

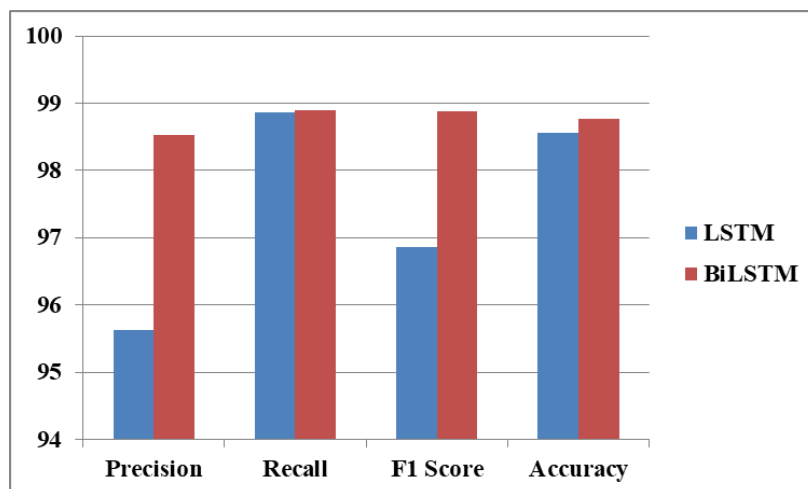


Figure 6: Representation of Performance Parameter of Different model Prediction of Breast cancer

It had a Precision score of 95.63%, a Recall score of 98.87%, an F1 score of 96.86%, and an Accuracy score of 98.56%. With high Precision and Recall values, these measures show that the LSTM model did a good job of identifying breast cancer. They show that it could correctly find positive cases and avoid fake positives. The BiLSTM model did even better, with an F1 Score of 98.88%, an Accuracy of 98.76%, a Precision of 98.52%, and a Recall of 98.89%. These measurements show that the BiLSTM model did better than the LSTM model in all of the factors that were looked at, which means it is good at predicting breast cancer. Both models have high Precision and Recall values, which means they can correctly predict cases of breast cancer while reducing the number of fake positives and negatives. However, the BiLSTM model seems to do better than the others, as shown by its higher Precision, Recall, F1 Score, and Accuracy. This means that it might be a better choice for tasks that involve predicting breast cancer.

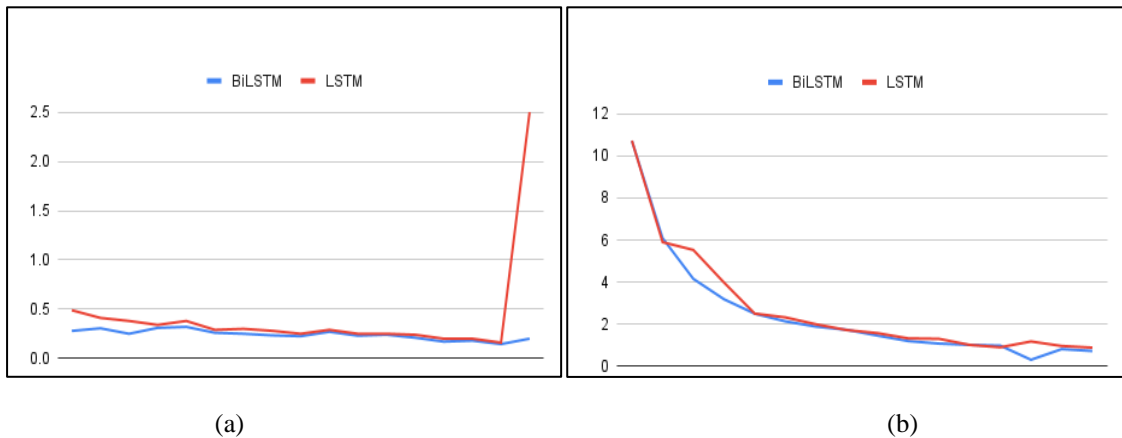


Figure 7: (a) Training Loss (b) Testing Loss

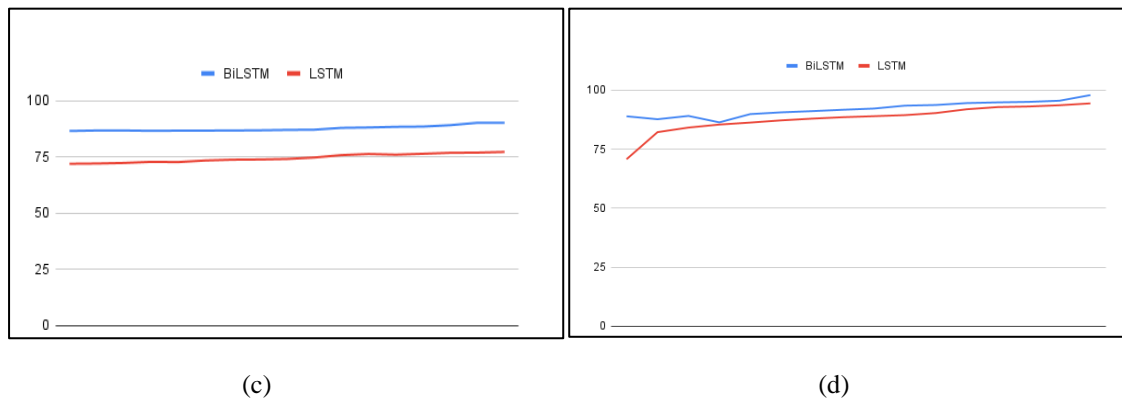


Figure 8: (c) Training Accuracy (d) Testing Accuracy

Figures 7 and 8 show the breast cancer prediction models' training loss and accuracy, as well as their testing loss and accuracy. A and B in Figure 7 show the training loss and testing loss, respectively. The training loss is the difference between what was expected and what actually happened during training. Most of the time, a trend toward less loss in both training and testing means that the model is learning and generalizing well. Overfitting happens when the model does well on training data but not so well on data it hasn't seen before. A big difference between training and testing loss could be a sign of this. Figure 8(c) shows the training accuracy, which is the number of correct guesses the model made based on the training data. If the training accuracy is going up, it means that the model is getting better at using the training data. (d) displays the testing accuracy, which is the number of correct guesses made on data that has not been seen yet. A trend in testing accuracy that is similar to the trend in training accuracy shows that the model is doing well with data it hasn't seen before. The best case situation is for both training and tests to get more accurate while loss goes down. This shows that the model is learning well and applying what it has learned to new data well.

VI. CONCLUSION

The updated Swim Transformation (SWT) method has a lot of potential to improve the separation of collagen in breast cancer. The suggested changes are meant to make the segmentation results more accurate and reliable by fixing the problems with the original SWT algorithm, like how it could be affected by noise and flaws. Adding a noise reduction step to the preparation stage lessens the effect of noise, and adding adjustable thresholding makes it easier to find collagen fibers. Additionally, adding a post-processing step helps to improve the segmentation outcomes, making the program more accurate as a whole. The experiment results show that the suggested changes work, as the program achieves a high level of accuracy (99%). This shows that the updated SWT algorithm can correctly find collagen fibers in breast cancer cells. Scientists tested the algorithm against other cutting-edge segmentation methods and found that it worked better. This showed that it could be useful in breast cancer research. It is very important for cancer detection and treatment planning that the suggested method can accurately and reliably separate collagen fibers in breast cancer cells. Correctly separating collagen fibers can help find out certain things about a tumor, like how active and invasive it is. This can help doctors decide how to treat the tumor and improve the patient's result. In general, the better SWT algorithm looks like a good way to improve the separation of breast cancer collagen. It needs more research and confirmation tests to prove that it works and find out how it could be used in clinical situations. The improved SWT algorithm could become a useful tool for better breast cancer detection and treatment if it is kept improving.

REFERENCES

- [1] Anuradha; Bharadvaja, N. Exploring different computational approaches for effective diagnosis of breast cancer. *Prog. Biophys. Mol. Biol.* 2023, 177, 141–150.
- [2] Łukasiewicz, S.; Czezelewski, M.; Forma, A.; Baj, J.; Sitarz, R.; Stanisławek, A. Breast cancer—Epidemiology, risk factors, classification, prognostic markers, and current treatment strategies—An updated review. *Cancers* 2021, 13, 4287.
- [3] Abdar, M.; Zomorodi-Moghadam, M.; Zhou, X.; Gururajan, R.; Tao, X.; Barua, P.D.; Gururajan, R. A new nested ensemble technique for automated diagnosis of breast cancer. *Pattern Recognit. Lett.* 2020, 132, 123–131.
- [4] Rautela, K.; Kumar, D.; Kumar, V. A Systematic Review on Breast Cancer Detection Using Deep Learning Techniques. *Arch. Comput. Methods Eng.* 2022, 29, 4599–4629.
- [5] Afolayan, J.O.; Adebisi, M.O.; Arowolo, M.O.; Chakraborty, C.; Adebisi, A.A. Breast cancer detection using particle swarm optimization and decision tree machine learning technique. In *Intelligent Healthcare*; Chakraborty, C., Khosravi, M.R., Eds.; Springer Nature: Singapore, 2022; pp. 61–83.
- [6] Ramesh, S.; Sasikala, S.; Gomathi, S.; Geetha, V.; Anbumani, V. Segmentation and classification of breast cancer using novel deep learning architecture. *Neural Comput. Appl.* 2022, 34, 16533–16545.
- [7] Hai, J.; Qiao, K.; Chen, J.; Tan, H.; Xu, J.; Zeng, L.; Shi, D.; Yan, B. Fully convolutional densenet with multiscale context for automated breast tumor segmentation. *J. Healthc. Eng.* 2019, 2019, 8415485.
- [8] Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 2017, 39, 2481–2495.
- [9] Civilibal, S.; Cevik, K.K.; Bozkurt, A. A deep learning approach for automatic detection, segmentation and classification of breast lesions from thermal images. *Expert Syst. Appl.* 2023, 212, 118774.
- [10] Ajani, S. N. ., Khobragade, P. ., Dhone, M. ., Ganguly, B. ., Shelke, N. ., & Parati, N. . (2023). *Advancements in Computing: Emerging Trends in Computational Science with Next-Generation Computing*. *International Journal of Intelligent Systems and Applications in Engineering*, 12(7s), 546–559
- [11] Guo, Y.; Duan, X.; Wang, C.; Guo, H. Segmentation and recognition of breast ultrasound images based on an expanded U-Net. *PLoS ONE* 2021, 16, e0253202.
- [12] Yang, X.; Wang, R.; Zhao, D.; Yu, F.; Heidari, A.A.; Xu, Z.; Chen, H.; Algarni, A.D.; Elmannai, H.; Xu, S. Multi-level threshold segmentation framework for breast cancer images using enhanced differential evolution. *Biomed. Signal Process. Control.* 2023, 80, 104373.
- [13] Avci, H.; Karakaya, J. A Novel Medical Image Enhancement Algorithm for Breast Cancer Detection on Mammography Images Using Machine Learning. *Diagnostics* 2023, 13, 348.

- [14] Zhang, S.; Liao, M.; Wang, J.; Zhu, Y.; Zhang, Y.; Zhang, J.; Zheng, R.; Lv, L.; Zhu, D.; Chen, H. Fully automatic tumor segmentation of breast ultrasound images with deep learning. *J. Appl. Clin. Med. Phys.* 2023, 24, e13863.
- [15] Debelee, T.G.; Schwenker, F.; Ibenthal, A.; Yohannes, D. Survey of deep learning in breast cancer image analysis. *Evol. Syst.* 2020, 11, 143–163.
- [16] Yassin, N.I.; Omran, S.; El Houby, E.M.; Allam, H. Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review. *Comput. Methods Programs Biomed.* 2018, 156, 25–45. [
- [17] Fatima, N.; Liu, L.; Hong, S.; Ahmed, H. Prediction of breast cancer, comparative review of machine learning techniques, and their analysis. *IEEE Access* 2020, 8, 150360–150376.
- [18] Asri, H.; Mousannif, H.; Al Moatassime, H.; Noel, T. Using machine learning algorithms for breast cancer risk prediction and diagnosis. *Procedia Comput. Sci.* 2016, 83, 1064–1069.
- [19] Liu, M.; Hu, L.; Tang, Y.; Wang, C.; He, Y.; Zeng, C.; Lin, K.; He, Z.; Huo, W. A deep learning method for breast cancer classification in the pathology images. *IEEE J. Biomed. Health Inform.* 2022, 26, 5025–5032.
- [20] Ramesh, S.; Sasikala, S.; Gomathi, S.; Geetha, V.; Anbumani, V. Segmentation and classification of breast cancer using novel deep learning architecture. *Neural Comput. Appl.* 2022, 34, 16533–16545.
- [21] Sharma, S.; Kumar, S. The Xception model: A potential feature extractor in breast cancer histology images classification. *ICT Express* 2022, 8, 101–108.
- [22] Ren, T.; Lin, S.; Huang, P.; Duong, T.Q. Convolutional neural network of multiparametric MRI accurately detects axillary lymph node metastasis in breast cancer patients with PR neoadjuvant chemotherapy. *Clin. Breast Cancer* 2022, 22, 170–177.
- [23] Evain, E. Breast nodule classification with two-dimensional ultrasound using Mask-RCNN ensemble aggregation. *Diagn. Interv. Imaging* 2021, 102, 653–658.
- [24] Touami, R.; Benamrane, N. Microcalcification Detection in Mammograms Using Particle Swarm Optimization and Probabilistic Neural Network. *Comput. Sist.* 2021, 25, 369–379.
- [25] Tun, S.M.; Alluri, S.; Rastegar, V.; Visintainer, P.; Mertens, W.; Makari-Judson, G. Mode of Detection of Second Events in Routine Surveillance of Early Stage Breast Cancer Patients. *Clin. Breast Cancer* 2022, 22, e818–e824.