國立臺灣大學電機資訊學院電信工程學研究所

碩士論文

Graduate Institute of Communication Engineering

College of Electrical Engineering and Computer Science

National Taiwan University

Master's Thesis

# Filters and Learning-based Methods for Tumor Detection from Ultrasonic Images

Daisuke Osako

指導教授：丁建均 博士

Advisor: Jian-Jiun Ding, Ph.D.

中華民國 114 年 1 月

Jan 2025

doi:10.6342/NTU202500232

# ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor, Prof. Jian-Jiun Ding, for his unwavering support and guidance throughout my academic journey. Coming from a background in economics with no prior knowledge of image processing, I was warmly welcomed into his research lab and given the opportunity to learn from the very basics. Prof. Ding's patient and thorough teaching provided me with the foundation to tackle challenging research topics, while his insightful advice continually guided me in the right direction.

Beyond academics, Prof. Ding has also been a source of immense support during my time at the university, helping me navigate various aspects of university life and ensuring that I could focus on my studies and personal growth. His encouragement and belief in my abilities have been instrumental in my development as a researcher and as an individual. I am sincerely grateful for the mentorship and care he has shown me, which has made this journey not only successful but also deeply rewarding.



doi:10.6342/NTU202500232

# ABSTRACT

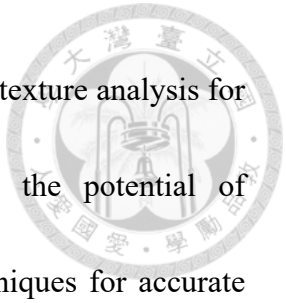
Tumor segmentation in medical imaging plays a critical role in the accurate diagnosis and treatment planning of cancer. This study proposes a hybrid framework that combines complementary convolutional neural network (CNN) models and advanced post-processing techniques to achieve robust and accurate tumor segmentation. The initial model (Model 1) employs CLAHE preprocessing, CNN-based predictions, and active contour refinement to provide a baseline segmentation. However, its performance is limited by difficulties in capturing complex tumor boundaries. To address these challenges, a second model (Model 2) incorporates noise-augmented preprocessing and iterative detection, enhancing the segmentation of subtle and irregular tumor regions.

The outputs of both models are merged using logical operations and refined further with edge correction and size filtering. Additionally, an enhanced merging model integrates a Spatial Intensity Metric (SIM) expansion, which leverages spatial and intensity relationships to refine and expand tumor regions, particularly addressing under-segmented areas. This enhancement results in significant improvements, as demonstrated by higher F1 and IoU scores compared to earlier models.

The study also highlights the limitations of the grid-based $16 \times 16$ classification approach, especially for large tumors, and suggests future directions such as adaptive grid



doi:10.6342/NTU202500232



sizes, more detailed labeling schemes, and the incorporation of local texture analysis for malignancy assessment. The proposed framework demonstrates the potential of integrating machine learning and traditional image processing techniques for accurate tumor segmentation, paving the way for more reliable and clinically valuable diagnostic tools.



doi:10.6342/NTU202500232

# CONTENT





doi:10.6342/NTU202500232

doi:10.6342/NTU202500232

# LIST OF FIGURES





doi:10.6342/NTU202500232

# LIST OF TABLES





doi:10.6342/NTU202500232

# Chapter 1 Introduction

## 1.1 Motivation

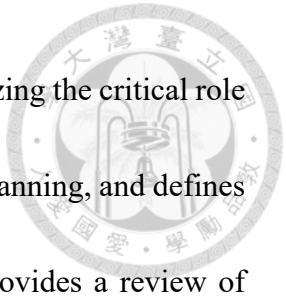
The motivation for this research stems from both personal and academic influences. Growing up with a mother working as a healthcare professional, I developed a deep appreciation for the challenges and intricacies of the medical field. Witnessing her dedication to patient care and the critical role of accurate diagnoses inspired me to contribute to this domain through technological innovation. During my academic journey, I became increasingly interested in the intersection of computer science and medicine, particularly in how advanced computational methods could address real-world healthcare challenges. Tumor segmentation in medical imaging emerged as a compelling focus, as it directly impacts the accuracy of cancer diagnosis and treatment planning—a field where even incremental improvements can significantly enhance patient outcomes. This research represents a synthesis of my personal connection to the medical field and my technical expertise, aiming to contribute to the development of more reliable and efficient diagnostic tools in clinical practice.

## 1.2 Thesis Organization

This thesis is structured to systematically present the development and refinement of tumor segmentation techniques in medical imaging. The study begins with Chapter 1:



doi:10.6342/NTU202500232



Introduction, which outlines the motivation for the research, emphasizing the critical role of accurate tumor segmentation in medical diagnosis and treatment planning, and defines the scope and objectives of the work. Chapter 2: Related Works provides a review of existing methods, encompassing traditional image processing and modern deep learning approaches, highlighting their strengths, limitations, and the evaluation metrics commonly used for segmentation tasks. Chapter 3: Grid-Based CNN Model describes the design and implementation of the proposed grid-based framework, detailing preprocessing methods like CLAHE and noise augmentation, and discusses the challenges faced in segmenting large tumors due to grid constraints. In Chapter 4: Post-Processing Techniques, various refinement methods are introduced, including active contour modeling, size filtering, and the novel Spatial Intensity Metric (SIM) expansion, which enhances boundary accuracy and addresses under-segmentation. Chapter 5: Experiments and Results presents comprehensive evaluations of the segmentation models, demonstrating the improvements achieved through iterative refinements and SIM expansion, supported by quantitative metrics and comparative analyses. Finally, Chapter 6: Conclusion and Future Work summarizes the findings, emphasizing the contributions of the proposed hybrid framework while discussing its limitations and potential extensions, such as adaptive grid sizes, detailed labeling strategies, and texture-based malignancy classification. This organization provides a logical and cohesive flow,



doi:10.6342/NTU202500232

guiding the reader through the progression of the research from foundational concepts to its practical implications and future directions.



doi:10.6342/NTU202500232

# Chapter 2 Related Works

## 2.1. General Image Processing Method

### 2.1.1 Morphological processing

Morphological processing is a set of image processing techniques used to analyze and manipulate the structure of objects within binary or grayscale images. Rooted in mathematical morphology, these operations focus on the shape and structure of regions in an image, making them particularly useful in tasks such as noise reduction, shape extraction, and image segmentation.

Morphological operations typically rely on a structuring element, a small binary matrix that defines the neighborhood for the operation. Common morphological operations include:

**Erosion**: Erosion shrinks objects in the image by removing pixels along the boundaries. It removes small, isolated noise and separates connected objects. Formally, a pixel is retained only if all pixels under the structuring element are foreground pixels.

**Dilation**: Dilation expands objects in the image by adding pixels to their boundaries. It helps bridge small gaps, enhance connectivity, or fill holes within objects. A pixel is added if at least one pixel under the structuring element is a foreground pixel.



doi:10.6342/NTU202500232

### 2.1.2 Gaussian filter

The Gaussian filter is a widely used image processing technique for smoothing and noise reduction. It applies a Gaussian function to the image, effectively averaging pixel intensities in a manner that preserves edges better than simple averaging filters. The filter is named after the Gaussian distribution, characterized by its bell-shaped curve, and is defined as:

$$G(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (1)$$

Where x and y are spatial coordinates, $\sigma$ is the standard deviation, controlling the filter's spread and smoothing strength.

### 2.1.3 Contrast Limited Adaptive Histogram Equalization (CLAHE)

Contrast Limited Adaptive Histogram Equalization (CLAHE) is an advanced image processing technique designed to improve the contrast of images while minimizing noise amplification. Unlike standard histogram equalization, which applies a uniform adjustment across the entire image, CLAHE operates locally, dividing the image into small, non-overlapping regions called tiles. For each tile, it redistributes pixel intensity values based on the local histogram to enhance contrast. To prevent over-amplification of noise in homogeneous areas, CLAHE imposes a contrast limit by clipping the histogram at a predefined threshold and redistributing the clipped values uniformly. This process ensures that the contrast enhancement is adaptive and preserves details in both bright and



doi:10.6342/NTU202500232

dark regions of the image. Once the tiles are processed, CLAHE smoothly combines them using bilinear interpolation to avoid visible boundary artifacts between adjacent tiles. CLAHE is particularly effective in enhancing medical images, low-light photographs, and other applications where global contrast adjustment may not adequately reveal local details.

### 2.1.4 Active Contour Model

Active Contour Model, also known as Snake model, is a computational technique widely used for image segmentation and object boundary detection. The Snake model represents a deformable curve that iteratively adjusts its position to fit the contours of an object within an image. This curve, defined as a parametric spline, is influenced by three forces: internal forces, image forces, and external constraint forces. Internal forces ensure smoothness and continuity of the curve by penalizing excessive bending or stretching. Image forces, derived from the image's intensity gradients, attract the curve toward edges or regions with high contrast, effectively guiding it to the object's boundaries. External constraint forces, often user-defined or application-specific, help the curve overcome local minima or focus on specific regions of interest. The Snake model evolves by minimizing an energy function that combines these forces, achieving an optimal balance between fitting the object's boundaries and maintaining the curve's shape. Typically solved using numerical optimization techniques, the Snake model is versatile and can



doi:10.6342/NTU202500232

adapt to various applications such as medical image analysis, object tracking, and shape recognition.

## 2.2. Deep Learning Method

### 2.2.1 Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNNs) are a class of deep learning models specifically designed for processing data with a grid-like structure, such as images or time series. They are widely used in tasks such as image classification, object detection, and semantic segmentation due to their ability to automatically and hierarchically learn spatial features from raw data.

Key attributes of CNNs include weight sharing (reducing the number of parameters) and local connectivity (capturing spatial hierarchies). These properties make CNNs computationally efficient and highly effective for handling high-dimensional data like images.

CNNs have been instrumental in advancing computer vision tasks, outperforming traditional machine learning approaches in accuracy and scalability. Modern architectures, such as ResNet and EfficientNet, have further enhanced CNN capabilities by introducing innovations like residual connections and efficient scaling.



doi:10.6342/NTU202500232

## 2.3. Evaluation Method

### 2.3.1 Intersection over Union (IoU)

The Intersection over Union (IoU) is a commonly used evaluation metric for object detection, segmentation, and other tasks that involve spatial predictions. IoU quantifies the accuracy of a predicted region compared to the ground truth by measuring the overlap between the two. It is defined as the ratio of the area of intersection to the area of union between the predicted region $A_p$ and the ground truth region $A_g$, as follows:

$$IoU = \frac{|A_p \cap A_g|}{|A_p \cup A_g|} \tag{2}$$

Where $|A_p \cap A_g|$ is the area of overlap between the predicted and ground truth regions, and $|A_p \cup A_g|$ is the total area covered by both regions. The IoU value ranges from 0 to 1, where a higher IoU indicates better alignment between the prediction and the ground truth. An IoU of 1 signifies perfect overlap, while an IoU of 0 indicates no overlap.

### 2.3.2 F1 Score

The F1 Score is a widely used performance metric for classification tasks, particularly in imbalanced datasets. It is the harmonic mean of precision and recall, providing a single value that balances the trade-off between these two measures. The F1 Score is calculated as follows:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{3}$$



doi:10.6342/NTU202500232

Precision measures the proportion of correctly identified positive predictions out of all positive predictions, calculated as:

$$Precision = \frac{TruePositive(TP)}{TruePositive(TP) + FalsePositive(FP)} \quad (4)$$

Recall measures the proportion of correctly identified positive predictions out of all actual positive instances, calculated as:

$$Recall = \frac{TruePositive(TP)}{TruePositive(TP) + FalseNegative(FN)} \quad (5)$$

The F1 Score ranges from 0 to 1, where a value closer to 1 indicates better performance. It is particularly useful when the cost of false positives and false negatives are both significant, as it balances the need for precision and recall. Unlike accuracy, the F1 Score provides a more nuanced evaluation of performance in cases where class distributions are skewed.



doi:10.6342/NTU202500232

# Chapter 3 Grid-based CNN Model Part

## 3.1. Preprocessing

### 3.1.1 Model 1

This preprocessing method incorporates noise reduction, brightness adjustment, and contrast enhancement using CLAHE (Contrast Limited Adaptive Histogram Equalization). The steps include:

**Gaussian Blur**: A Gaussian filter with a kernel size of $5 \times 5$ and a sigma value of 1.5 is applied to smooth the images and reduce high-frequency noise.

**Brightness Adjustment**: A custom function adjusts the brightness of the image to a target value of 120, ensuring consistent lighting conditions across the dataset.

**CLAHE**: A CLAHE operation with a clip limit of 4.0 and a tile grid size of $8 \times 8$ enhances local contrast and reveals finer image details.

**Grid Division**: The processed images are divided into $16 \times 16$ grids. Each grid's content is labeled as "Tumor" or "Normal Tissue" based on the percentage of tumor pixels, enabling localized analysis.



doi:10.6342/NTU202500232

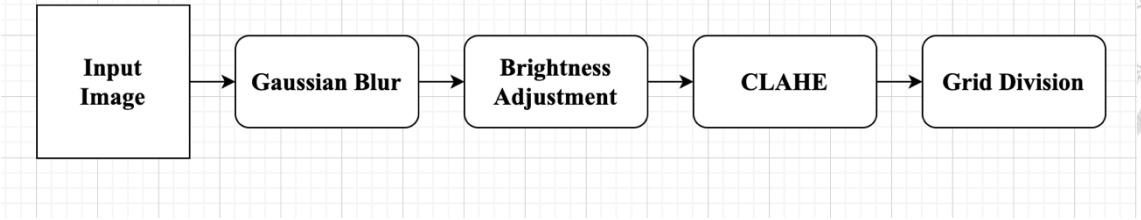



Fig 1 Model1 Preprocessing Flow Chart

### 3.1.2 Model 2

This preprocessing method focuses on applying noise addition, brightness adjustment, and Gaussian blurring to the images. The steps involve:

**Noise Addition**: Gaussian noise with a mean of 0 and a standard deviation of 30 is added to the image to simulate variability and enhance the robustness of the model.

**Brightness Adjustment**: A custom function adjusts the brightness of the image to a target value of 120, ensuring consistent lighting conditions across the dataset.

**Gaussian Blur**: A Gaussian filter with a kernel size of $5 \times 5$ and a sigma value of 1.5 is applied to smooth the images and reduce high-frequency noise.

**Grid Division**: The processed images are divided into $16 \times 16$ grids. Each grid's content is labeled as "Tumor" or "Normal Tissue" based on the percentage of tumor pixels, enabling localized analysis.

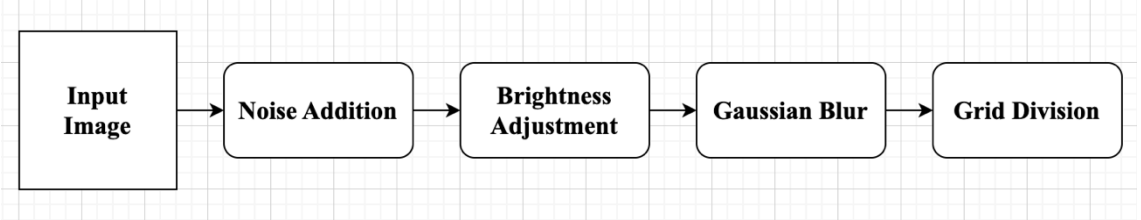



Fig 2 Model2 Preprocessing Flow Chart



doi:10.6342/NTU202500232

## 3.2. Model Structure

Fig3 below shows the common layer structure, and Table 1 below indicates the detailed values for each layer.

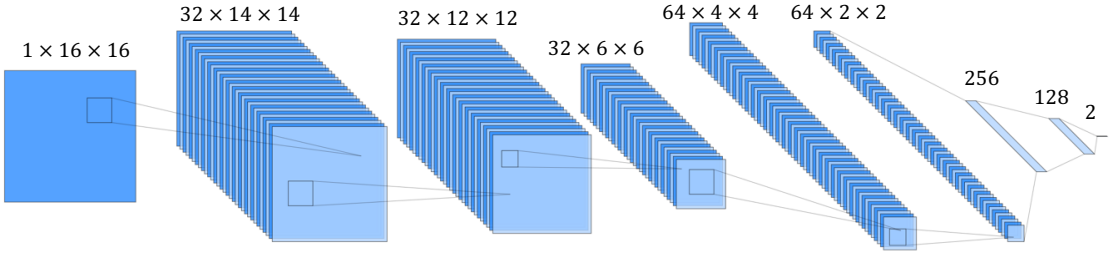



Fig 3 Overall layer structure

| Layer Type | Output Shape | Parameters |
|---|---|---|
| Convolutional Layer (32 filters, 3 × 3, ReLU activation) | 14 × 14 × 32 | 320 |
| Convolutional Layer (32 filters, 3 × 3, ReLU activation) | 12 × 12 × 32 | 9248 |
| Max Pooling Layer (2× 2) | 6 × 6 × 32 | 0 |
| Convolutional Layer (64 filters, 3 × 3, ReLU activation) | 4 × 4 × 64 | 18496 |
| Max Pooling Layer (2× 2) | 2 × 2 × 64 | 0 |
| Flatten Layer | 256 | 0 |
| Fully Connected Layer (128 units, ReLU activation) | 128 | 32896 |



doi:10.6342/NTU202500232

| | | |
|---|---|---|
| Fully Connected Layer (2 units, ReLU activation) | 2 | 258 |

Table 1 Layer Structure Details

## 3.3. Performance

The performance of the two preprocessing methods was evaluated based on their ability to predict whether each $16 \times 16$ grid corresponds to a tumor or normal tissue, with the test accuracy shown below:

| Model | Test Accuracy |
|---|---|
| Model 1 | 0.955 |
| Model 2 | 0.9229 |

Table 2 Test accuracy of CNN models

## 3.4. Grid Prediction

### 3.4.1 Model Prediction

The grayscale image is divided into non-overlapping grids of size 16×16 pixels. Each grid serves as an individual test input for the CNN model.

Each segmented grid is passed to the CNN model for classification. The model outputs a predicted probability for each grid, indicating whether the grid belongs to a tumor or non-tumor region. If the tumor probability > 0.3, the grid is likely to contain tumor tissue and is considered a potential tumor grid.



doi:10.6342/NTU202500232

### 3.4.2 Grid Connection

Grids classified as potential tumor regions are connected to form continuous regions. Tumor grids in close proximity are linked horizontally, vertically, and diagonally to represent cohesive potential tumor regions.

Among these connected potential tumor regions, those that are in contact with four or more edges are excluded. The remaining potential tumor regions after this processing serve as the basis for subsequent tumor detection.



doi:10.6342/NTU202500232

# Chapter 4 Post Processing Part

## 4.1 Region Expansion Method

### 4.1.1 Refinement of Predicted Regions Using Active Contour Model

To further refine the potential tumor regions predicted by CNN model, we utilized an active contour model to segment and enhance the predicted regions.

The corresponding grayscale image is normalized to the range [0, 1], ensuring compatibility with the active contour model.

If there are multiple potential tumor regions, each region is calculated independently using the active contour model.

Key parameters used include:

$\mu = 0.15$: Regularization parameter for smoothness.

$\lambda_1 = 2.6, \lambda_2 = 1.0$: Weights for the data fidelity terms inside and outside the contour.

Maximum iterations of 100 and a time step $(dt)$ of 2.5.

The algorithm refines the region boundaries by iteratively evolving the contour to minimize energy functions.

Only the subregions generated by the active contour model that overlap with the potential tumor regions mask are retained as valid tumor regions, which ensures that only contours



doi:10.6342/NTU202500232

closely related to the original prediction are retained, reducing the chance of false positives.

### 4.1.2 Region Expansion Using Spatial Intensity Metric (SIM)

The refinement of segmentation masks is achieved through a novel expansion technique based on the Spatial Intensity Metric (SIM), which combines spatial and intensity-based constraints. This method begins with the extraction of contours from the binary segmentation mask. Using OpenCV's contour detection algorithm, the outermost boundary of the segmented region is identified. From the detected contour, up to 8 key points are selected to serve as seed points for the expansion process. If fewer than 8 points exist, all contour points are used; otherwise, the contour is sampled at evenly spaced intervals to ensure representative key points across the boundary.

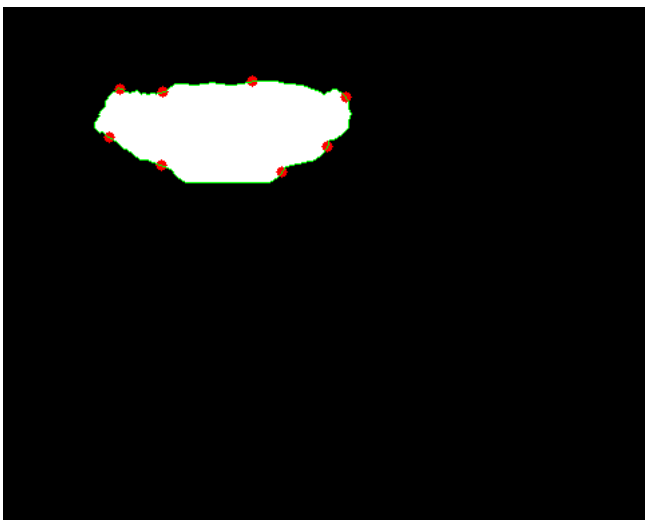


Fig 4 Selected Points Example

For each key point $(x, y)$, an intensity-based expansion is performed. The local intensity at the key point in the original grayscale image serves as a reference, denoted as $I(x, y)$.



doi:10.6342/NTU202500232

A rectangular search area is defined around the key point with a width and height proportional to the expansion threshold $\tau$. Within this search area, each pixel $(i, j)$ is evaluated based on its spatial distance $d_s$ from the key point and its intensity difference $d_I$ relative to the reference intensity. These terms are defined as:

$$d_s = \sqrt{(x-i)^2 + (y-j)^2} \tag{6}$$

$$d_I = |I(i,j) - I(x,y)| \tag{7}$$

The Spatial Intensity Metric (SIM) is computed as the sum of these two terms:

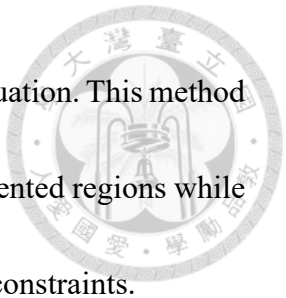
$$SIM(i,j) = d_s + d_I \tag{8}$$

If the SIM value is less than the predefined threshold $\tau$, the pixel $(i, j)$ is included in the expanded mask. This ensures that only pixels within an acceptable range of spatial and intensity similarity are added, preserving the region's coherence.

After evaluating all pixels within the search area for each key point, the resulting expanded regions are merged with the original segmentation mask. The combined mask is then processed to identify connected components. From these, the largest connected region is selected, as it most likely represents the primary object of interest. To further enhance the mask, any internal holes are filled using morphological operations, resulting in a refined and contiguous segmentation.



doi:10.6342/NTU202500232



Finally, the expanded mask is compared with the ground truth for evaluation. This method demonstrates a robust ability to capture subtle extensions of the segmented regions while maintaining accuracy by leveraging both spatial and intensity-based constraints.

## 4.2 Region Exclusion Method

### 4.2.1 Region Exclusion Based on Geometric Feature

The exclusion of unwanted regions is performed using a geometric feature-based approach, focusing on the shape and size properties of connected regions. This method ensures that only regions with meaningful geometric characteristics are retained for further analysis.

The process begins by labeling the connected components in the binary mask. If the number of connected components is fewer than two, the input mask is returned unchanged, as no additional filtering is required. For cases with multiple regions, each labeled region is processed individually to evaluate its geometric properties.

For each labeled region, contours are extracted to identify the outer boundary. The contour with the largest area is selected as the primary shape descriptor for the region. If the contour contains fewer than five points, it is skipped, as it cannot form a valid ellipse.



doi:10.6342/NTU202500232

An ellipse is then fitted to the largest contour using the least-squares method, providing key parameters including:

Center of the ellipse: $(c_x, c_y)$

Semi-major axis: $a$

Semi-minor axis: $b$

Rotation angle: $\theta$

These parameters are used to construct an ellipse mask, $E(x, y)$, where a pixel $(x, y)$ lies within the ellipse if:

$$\frac{\left((x - c_x)\cos\theta + (y - c_y)\sin\theta\right)^2}{a^2} + \frac{\left((x - c_x)\sin\theta + (y - c_y)\cos\theta\right)^2}{b^2} \leq 1 \quad (9)$$

Next, the geometric proportions of the ellipse are analyzed. The major $(a)$ and minor $(b)$ axes are compared, and the longer and shorter axes are assigned as:

$$long = Max(a, b) \quad (10)$$

$$short = Min(a, b) \quad (11)$$

If the ratio of the longer axis to the shorter axis exceeds a predefined threshold, the region is considered too elongated and is excluded from the final mask. This ensures that only regions with balanced proportions are retained, filtering out artifacts such as linear streaks or noise.



doi:10.6342/NTU202500232

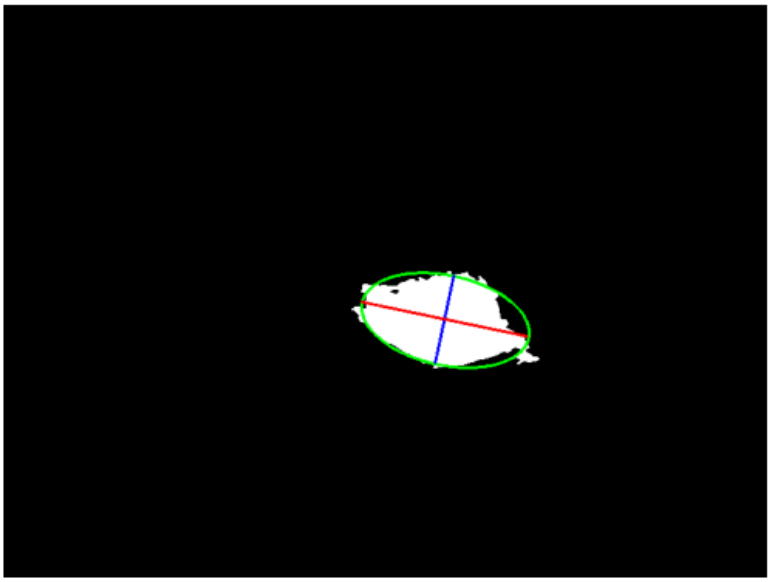


Fig 5 Geometric Feature

To further refine the selection, the intersection area between the region mask and the ellipse mask is calculated, along with the total area of the ellipse:

$$Intersection\ Area = \sum_{x,y} \big(R(x,y) \wedge E(x,y)\big) \tag{12}$$

$$Ellipse\ Area = \sum_{x,y} E(x,y) \tag{13}$$

The overlap ratio is computed as:

$$Overlap\ Ratio = \frac{Intersection\ Area}{Ellipse\ Area} \tag{14}$$

If the overlap ratio exceeds a threshold, the region is considered valid and included in the final mask. Otherwise, it is excluded as an unwanted artifact.

After processing all regions, the valid regions are combined into a unified binary mask. If no valid regions remain, the original mask is returned to preserve input integrity. This



doi:10.6342/NTU202500232

method ensures that retained regions exhibit consistent geometric features, such as ellipticity.

### 4.2.2 Region Exclusion Based on Positional Constraints

In addition to geometric features, regions are filtered based on their positional characteristics within the image. This approach ensures that only regions entirely contained within the image boundaries are considered valid, avoiding potential artifacts introduced by partial regions at the edges.

The algorithm uses a helper function, *is_white_area_on_edge(img)*, to detect whether a region intersects with any edge of the image. This function evaluates the binary mask to check if any white pixels (255) are present in the top row, bottom row, leftmost column, or rightmost column of the mask. If such pixels are detected, the region is classified as an edge-intersecting region. Mathematically, the edge condition is defined as:

$$\exists(x, y) = R \; such \; that \; x = 0 \vee x = H - 1 \vee y = 0 \vee y = W - 1 \qquad (15)$$

where $R$ represents the region, $H$ is the image height, and $W$ is the image width.

Regions that meet this criterion are excluded from further analysis. This exclusion logic is integrated into multiple processing steps within the algorithm, ensuring consistent application across the entire pipeline. Specifically, regions that do not intersect with the



doi:10.6342/NTU202500232

edges (*if not is_white_area_on_edge(img)*) are identified as isolated regions within the image and are subjected to additional filtering based on other criteria, such as size or shape.

By incorporating this edge-based exclusion mechanism, the algorithm effectively removes partial or incomplete regions, which are often noise or artifacts resulting from boundary effects. This refinement improves the reliability of segmentation results, ensuring that only meaningful and complete regions are retained for subsequent analysis.

### 4.2.3 Region Exclusion Based on Size Constraints

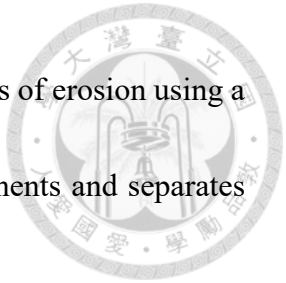
To improve the accuracy of tumor region detection, the algorithm incorporates a size-based filtering mechanism. This step addresses cases where detected tumor regions are disproportionately large, potentially due to the inclusion of irrelevant shadows or artifacts.

The size threshold for exclusion is set to 17% of the total image area. If the detected tumor region exceeds this threshold, it is flagged as a candidate for refinement.

When this threshold is exceeded, *split_largearea* function is applied to refine the segmentation. The process is as follows:



doi:10.6342/NTU202500232



**Iterative Erosion for Separation**: The mask undergoes five iterations of erosion using a cross-shaped kernel. This step reduces the size of connected components and separates overlapping regions, making it easier to isolate distinct subregions.

**Region Labeling and Filtering**: The eroded mask is labeled to identify connected components, which are then evaluated to ensure they meet specific criteria for inclusion. A region is considered valid if it does not intersect the image boundaries, as determined by the *is_white_area_on_edge* function, and if its area exceeds 0.1% of the total image size.

**Iterative Dilation for Restoration**: The filtered regions are dilated five times using the same kernel to restore their original size and shape, ensuring continuity and completeness in the refined mask.

By incorporating the *split_largearea* function, the algorithm effectively handles scenarios where initial segmentation results contain large, conflated regions.



doi:10.6342/NTU202500232

# Chapter 5 Experiments

## 5.1 Tumor Segmentation Using Model 1

### 5.1.1 Experimental Methodology 1

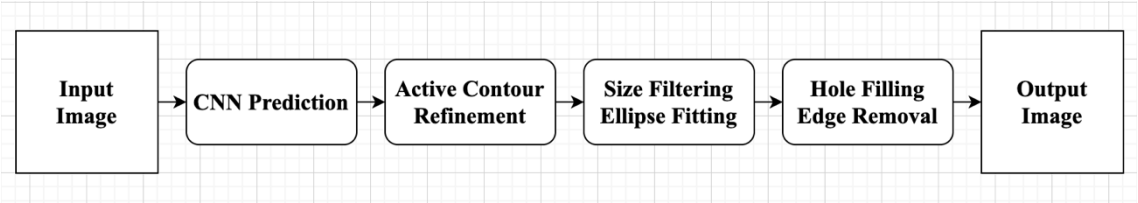



Fig 6 Flow Chart of Tumor Segmentation Using Model1

The following is the flow of the tumor detection model using model 1.

- Tumor Prediction Using a CNN Model1

  The input grayscale medical images are preprocessed to enhance contrast and suppress noise. Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to improve local contrast, followed by Gaussian Blur to reduce noise. The preprocessed images are then divided into non-overlapping grids of $16 \times 16$ pixels and classified using a pre-trained CNN model.

  Each grid is evaluated for the presence of tumor or non-tumor regions. Grids classified as tumor regions (confidence > 0.3) are marked in a tumor prediction mask. To ensure spatial continuity, tumor grids are further analyzed: if adjacent tumor grids



doi:10.6342/NTU202500232

are detected within two grids in the vertical or horizontal directions, they are connected to form continuous tumor regions.

- Active Contour Refinement

  To accurately detect tumor boundaries based on the CNN-predicted tumor grids, the algorithm employs the Chan-Vese Active Contour Model. Each connected component in the tumor mask is extracted and used as the initial contour for the active contour model.

  The Chan-Vese model iteratively evolves these contours, optimizing their alignment with the actual tumor boundaries by minimizing an energy function based on region homogeneity. Only regions that overlap significantly with the original tumor mask are retained, ensuring consistency between the CNN predictions and the refined boundaries.

- Size Filtering

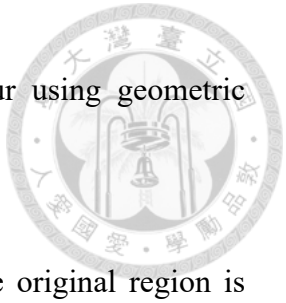
  Regions smaller than a threshold (e.g., 0.1% of the image area) are removed to avoid noise.

  Large regions exceeding predefined size limits (e.g., 17% of the image area) are split into smaller components using erosion and morphological operations.

- Ellipse Fitting



doi:10.6342/NTU202500232



For each segmented region, an ellipse is fitted to its contour using geometric properties.

The ratio of the intersection between the fitted ellipse and the original region is computed. Regions with a significant intersection ratio ($\geq 60\%$) are retained, while others are discarded.

- Hole Filling

  Internal holes within the tumor regions are filled using a binary hole-filling algorithm, ensuring complete and continuous masks.

- Edge Removal

  Remaining small regions touching the edges are carefully excluded, focusing the final output on central tumor areas.

### 5.1.2 Results and Analysis 1

The experimental results for tumor segmentation using Model 1 were evaluated based on two metrics: F1 Score and Intersection over Union (IoU). The detailed results and their distribution are as follows:

| | |
|---|---|
| Average F1 Score | 0.3831 |
| Average IoU Score | 0.2854 |

Table 3 Average F1 and IoU Score using Model1



doi:10.6342/NTU202500232

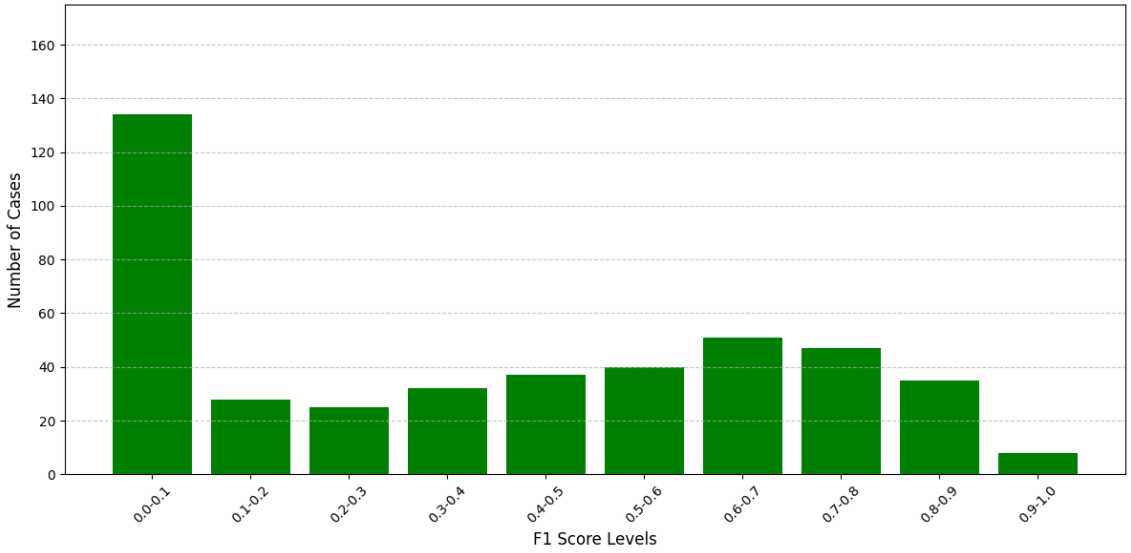



Fig 7 F1 score distribution using Model 1

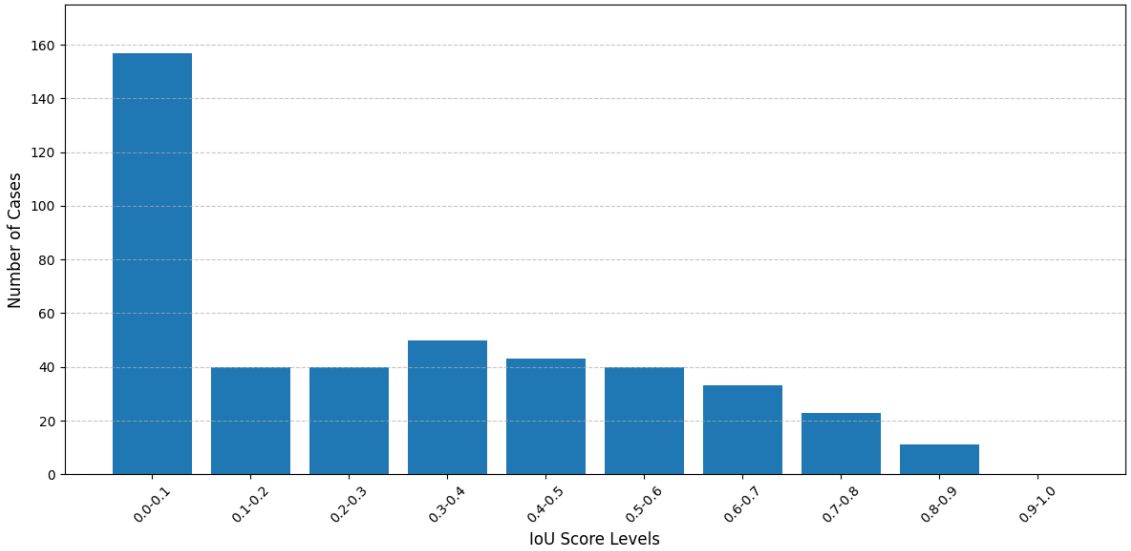



Fig 8 IoU score distribution using Model 1

The performance metrics (F1 and IoU scores) exhibit significant variability across the dataset. While most predictions indicate low overlap and precision, certain cases demonstrate the model's capability to achieve reasonable segmentation accuracy.

Fig 9 illustrates the transition of images in successful segmentation cases from this experiment.



doi:10.6342/NTU202500232

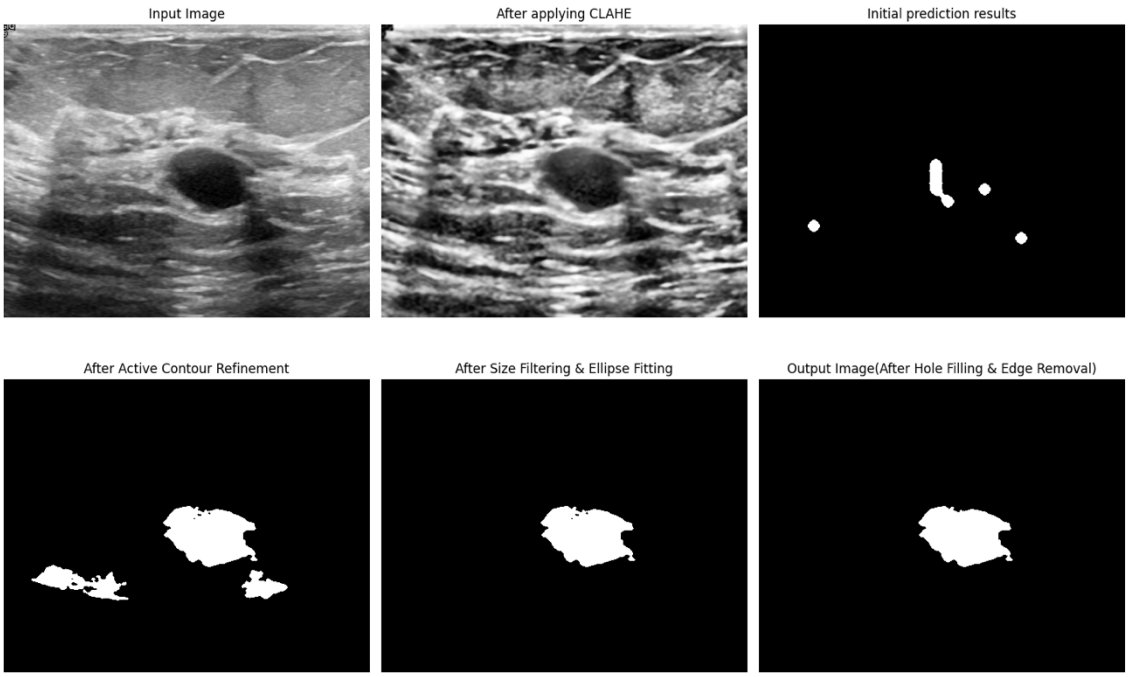



Fig 9 Results of each step in successful case of Experiment1

Tumor detection in this study heavily relies on the initial predicted regions generated by the CNN. Consequently, as shown in Fig 10, when the initial predictions fail to capture the actual tumor regions, the overall performance of the algorithm is significantly affected, leading to inaccurate segmentation. Therefore, relying solely on Model 1 results in insufficient accuracy of the initial predictions, ultimately compromising the final segmentation outcomes.



doi:10.6342/NTU202500232

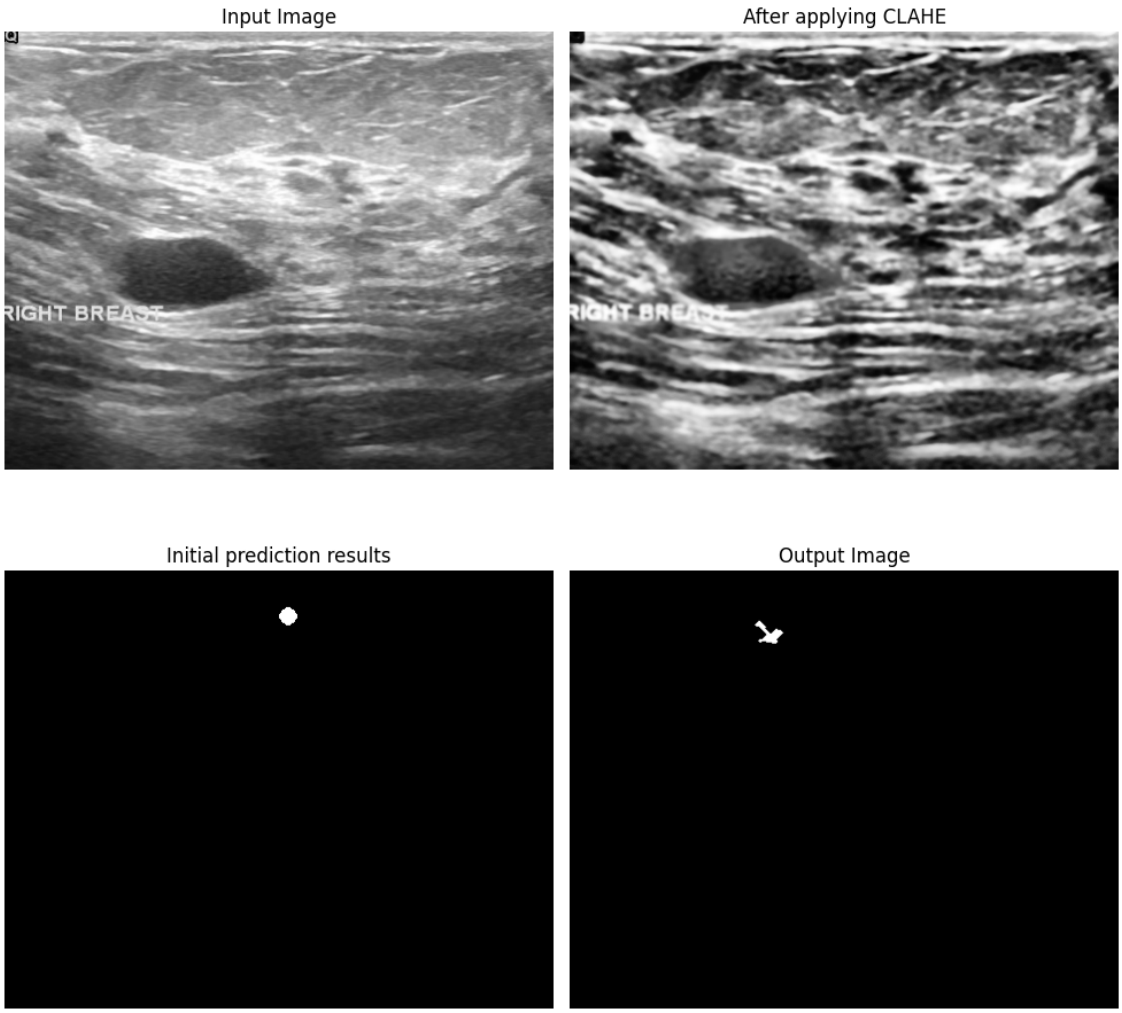



Fig 10 Results of each step in failure case of Experiment1

## 5.2 Tumor Segmentation Using Model 2

While Model 1 employs CLAHE preprocessing and a CNN-based initial prediction, Model 2 introduces noise addition techniques and iterative refinement to complement Model 1's predictions



doi:10.6342/NTU202500232

### 5.2.1 Experimental Methodology 2

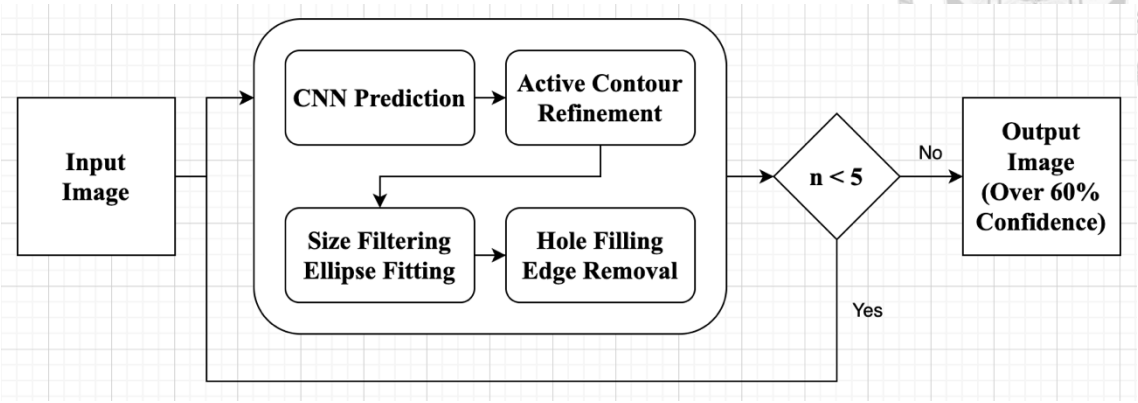



Fig 11 Flow Chart of Tumor Segmentation Using Model2

- Tumor Prediction Using CNN Model 2

  Model 2 applies a different preprocessing strategy compared to Model 1. Gaussian noise with a mean (μ) of 0 and a standard deviation (σ) of 30 is added to the input grayscale medical images to simulate real-world imaging variations. After noise augmentation, Gaussian filtering is performed to reduce noise, like Model 1. However, unlike Model 1, Model 2 does not enhance contrast through CLAHE. Instead, it retains the original contrast of the images for tumor prediction, allowing the model to focus on inherent image features.

  The threshold for tumor prediction remains the same as in Model 1, where grids with a confidence score greater than 0.3 are classified as tumor regions.

- Iterative Detection



doi:10.6342/NTU202500232

To improve the accuracy of tumor detection, Model 2 uses the Chan-Vese Active Contour Model, similar to Model 1. However, due to the lack of contrast enhancement, the results after applying the Active Contour Model exhibit some variability. To address this, the process is repeated iteratively over five iterations.

Each iteration mirrors the process of Model 1 and performs the following: Size Filtering, Ellipse Fitting, Hole Filling, Edge Removal.

The final tumor masks from all iterations are combined into a single cumulative mask. The pixel intensity of this mask represents the cumulative confidence of tumor detection across the five iterations. A threshold of 60% confidence is applied to obtain the final tumor prediction results for Model 2.

### 5.2.2 Results and Analysis 2

The performance of Model 2, evaluated independently, highlights the limitations of relying solely on this approach for tumor segmentation. The average F1 Score was 0.2343, and the average IoU Score was 0.1870, both significantly lower than the scores achieved by Model 1.

| | |
|---|---|
| Average F1 Score | 0.2343 |
| Average IoU Score | 0.1870 |

Table 4 Average F1 and IoU Score using Model2



doi:10.6342/NTU202500232

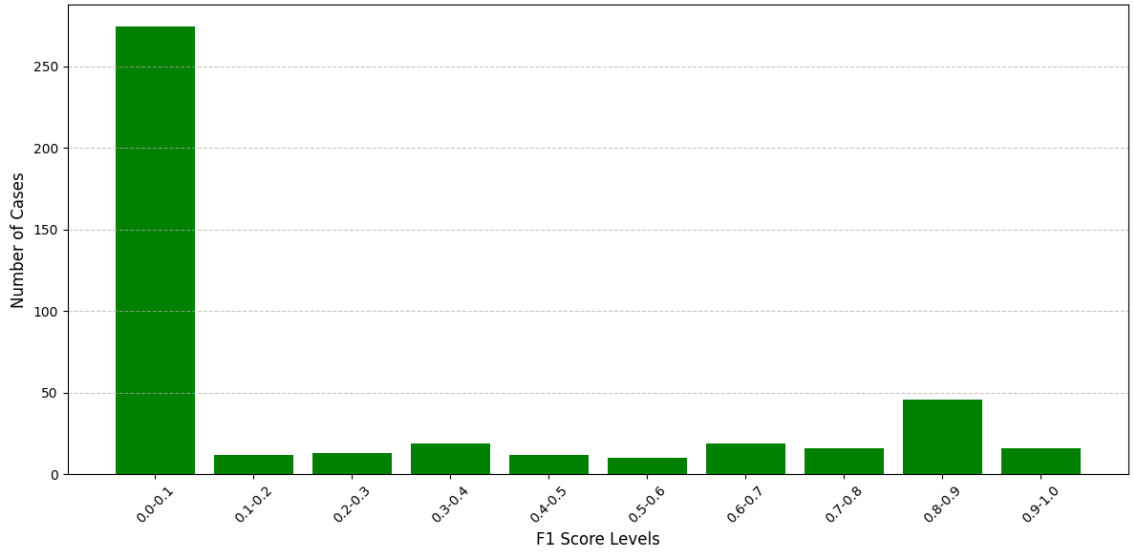



Fig 12 F1 score distribution using Model 2

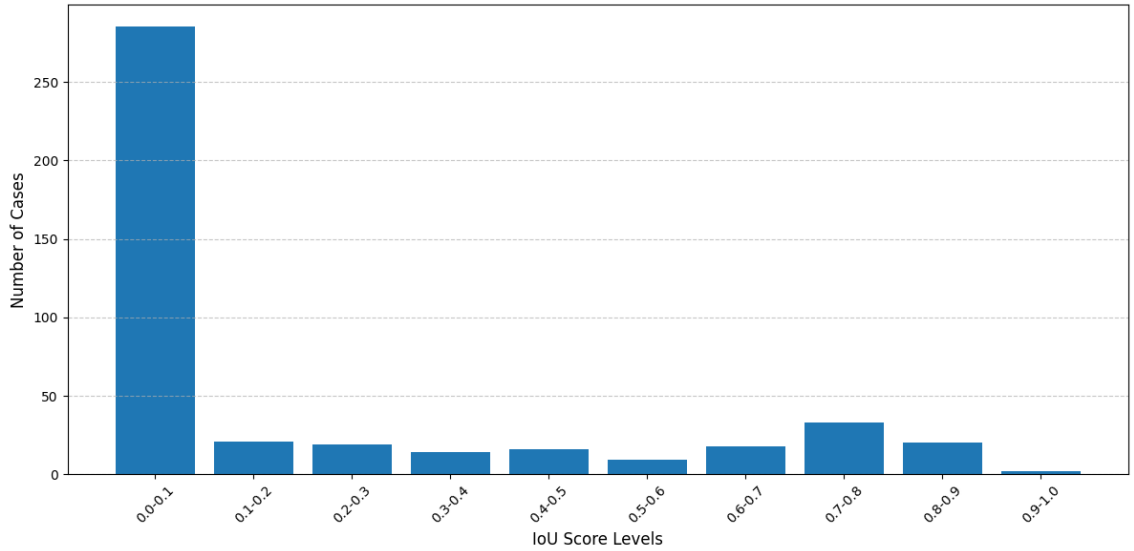



Fig 13 IoU score distribution using Model 2

The distributions of the F1 and IoU scores reveal that the majority of cases fall into the 0.0–0.1 range, indicating limited segmentation accuracy across most test cases. However, in certain cases, superior segmentation results compared to those achieved by Model 1 were observed. Fig 14 presents an example of such a case, illustrating the transitions through the results of five iterative refinements, the overlay of these results, and the final output retaining only regions with a confidence level of 60% or higher.



doi:10.6342/NTU202500232

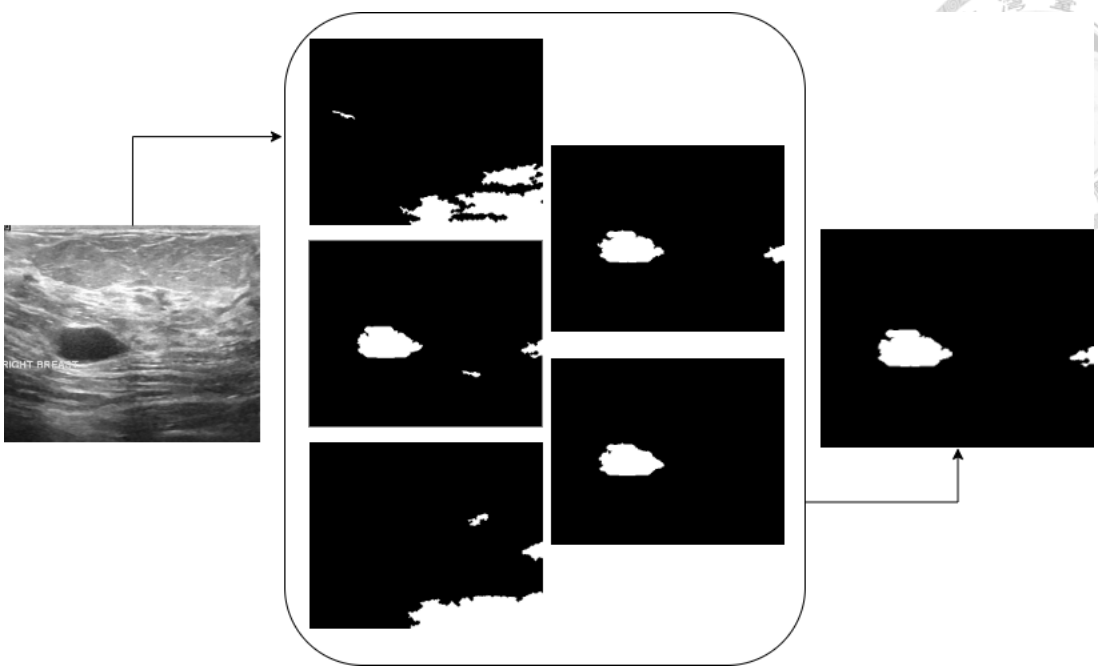


Fig 14 Successful iterative refinement and final segmentation results using Model 2

The primary reason for Model 2's lower performance is the absence of contrast enhancement during preprocessing. Unlike Model 1, which uses CLAHE to improve local contrast and enhance tumor visibility, Model 2 relies on the original image contrast, making it less effective in distinguishing subtle tumor boundaries. This limitation is particularly evident in cases with low-contrast or irregular tumor regions, where Model 2 struggles to produce accurate predictions.

Despite its lower standalone performance, Model 2's iterative detection and noise-augmented preprocessing offer complementary strengths that can be leveraged when combined with Model 1. While Model 2 alone is insufficient for reliable tumor segmentation, its ability to capture additional features and refine predictions can serve as a valuable supplementary component within a hybrid framework.



doi:10.6342/NTU202500232

These results underscore the importance of using Model 2 as a supporting method rather than as a standalone solution.

## 5.3 First Merged Models

To enhance tumor detection accuracy, Model 1 and Model 2 were integrated into a combined framework.

### 5.3.1 Experimental Methodology 3

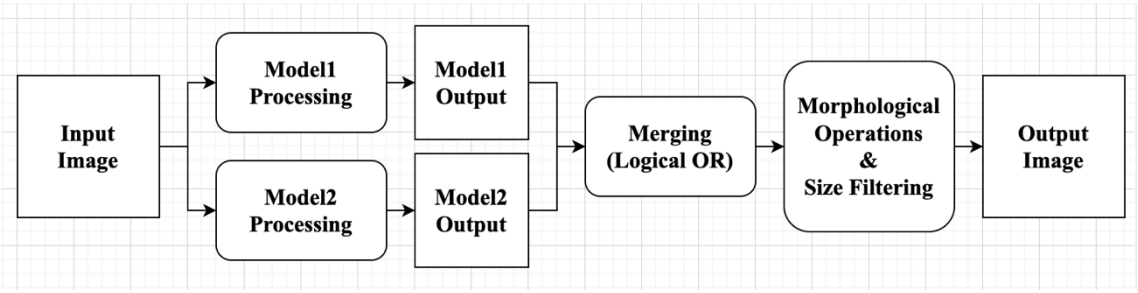



Fig 15 Flow Chart of Tumor Segmentation Using First Merged Model

The binary tumor masks generated by Model 1 and Model 2 are merged using a logical OR operation. This ensures that any tumor region detected by either model is included in the final segmentation.

To refine the merged result, tumor regions near the edges are re-evaluated using morphological operations such as erosion and dilation to minimize false positives. Additional size filtering is applied to ensure that the final segmentation contains accurate and reliable tumor regions.



doi:10.6342/NTU202500232

By integrating the outputs from both models, the combined framework leverages the complementary strengths of Model 1 and Model 2, resulting in a more robust and comprehensive tumor segmentation.

### 5.3.2 Results and Analysis 3

The experimental results of tumor segmentation merging Model 1 and Model 2 are evaluated based on two metrics: F1 score and Intersection over Union (IoU). The detailed results and their distributions are as follows:

| | |
|---|---|
| Average F1 Score | 0.4714 |
| Average IoU Score | 0.3874 |

Table 5 Average F1 and IoU Score merging Model 1 and Model 2

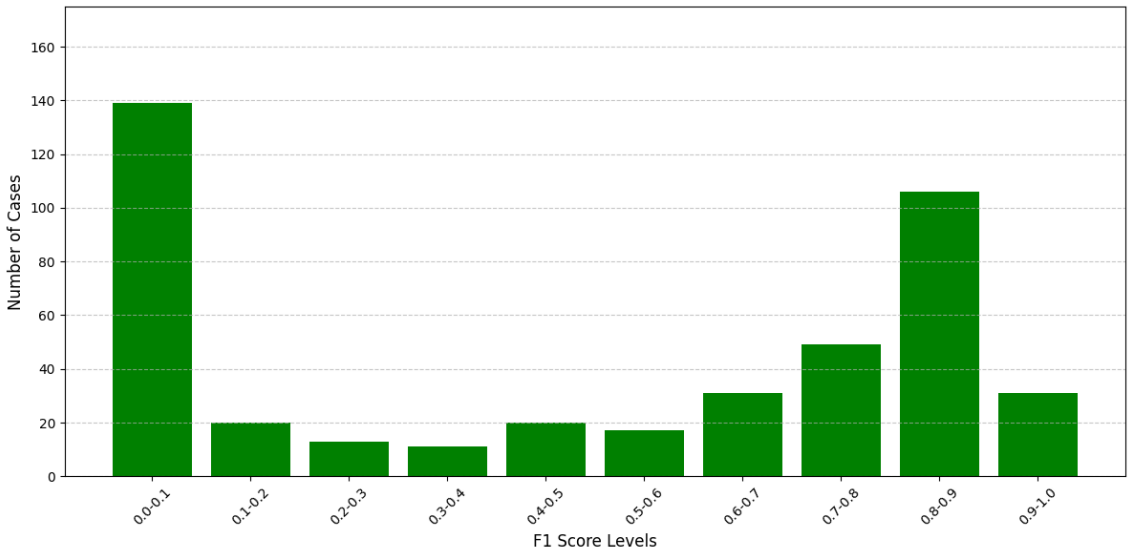



Fig 16 F1 score distribution merging Model 1 and Model 2



doi:10.6342/NTU202500232

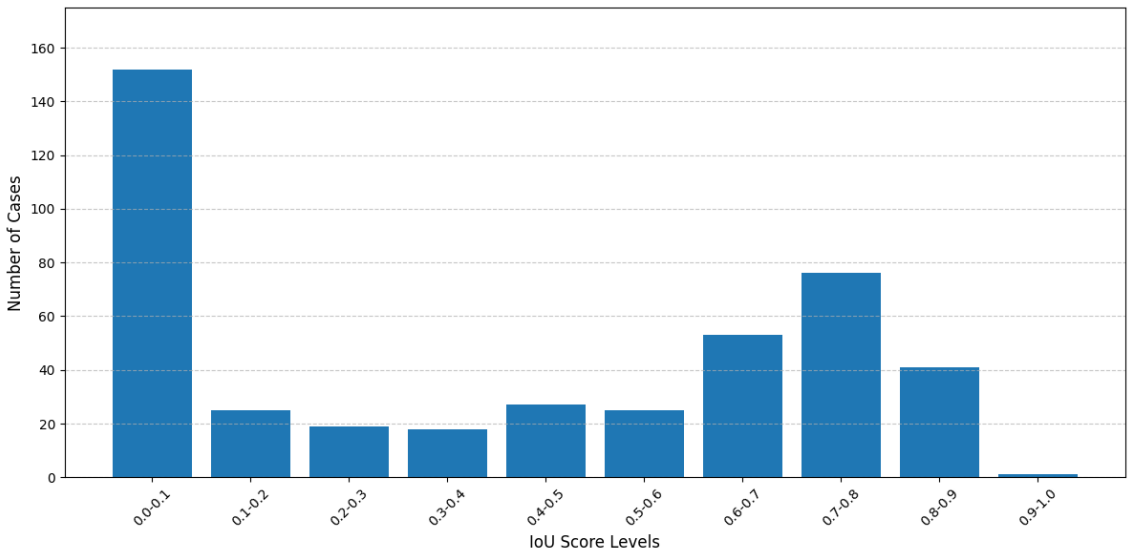



Fig 17 IoU score distribution merging Model 1 and Model 2

The combined model achieved an average F1 Score of 0.4714, demonstrating an improvement in the balance between precision and recall compared to using Model 1 alone.

The average IoU Score was 0.3874, reflecting better overlap and alignment between the predicted tumor regions and the ground truth.

While the majority of cases are still concentrated in the 0.0–0.1 range, there is a noticeable increase in cases achieving higher IoU scores (e.g., 0.6-0.7, 0.7-0.8 and 0.8-0.9), reflecting more accurate tumor boundary segmentation in these instances.

By merging Model 1 and Model 2, the prediction of tumor presence and location became more accurate across a greater number of cases compared to using either model individually. However, as shown in Fig 18, there were cases where the active contour model failed to accurately capture the tumor's contours, indicating room for improvement in the region expansion methodology.



doi:10.6342/NTU202500232

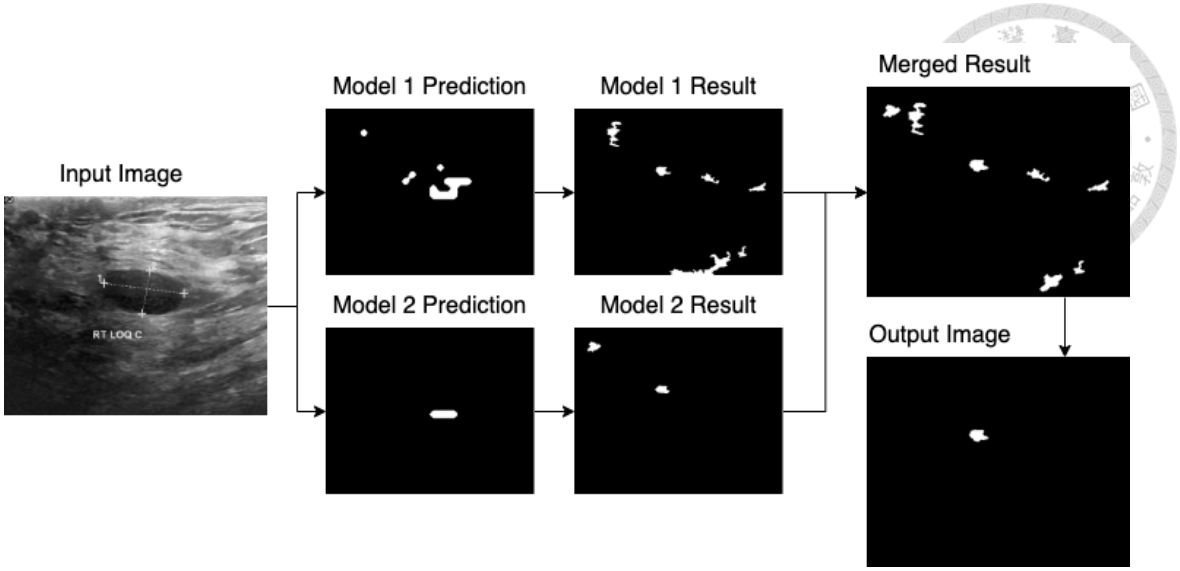



Fig 18 Example of a case where the active contour model does not work well

## 5.4 Second Merged Models

### 5.4.1 Experimental Methodology 4

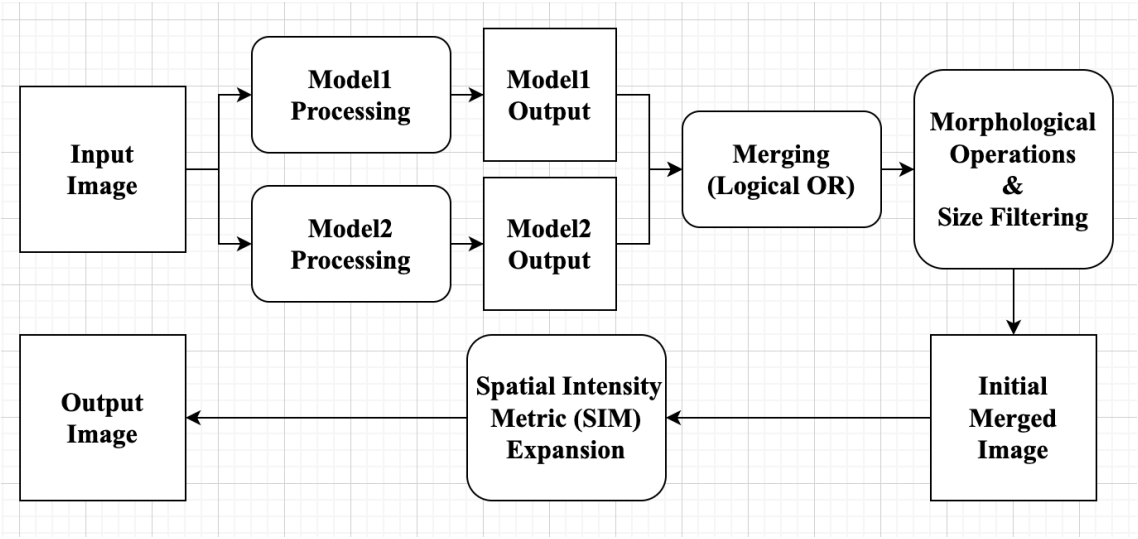



Fig 19 Flow Chart of Tumor Segmentation Using Second Merged Model

The initial merged segmentation mask is generated by combining the outputs of Model 1 and Model 2:

Model 1: Detects tumor regions using CLAHE-enhanced preprocessing, CNN-based predictions, and active contour refinement.



doi:10.6342/NTU202500232

Model 2: Complements Model 1 by introducing noise-augmented preprocessing, iterative detection, and cumulative confidence scoring.

Merging: The binary masks from both models are merged using a logical OR operation, followed by post-processing steps such as edge refinement and size filtering.

While the merged mask provides accurate segmentation, some tumor regions may remain under-segmented, particularly near region boundaries. To address this, the SIM expansion is applied to further refine the mask.

The SIM expansion algorithm operates on the merged segmentation mask to enhance tumor detection by leveraging spatial and intensity relationships.

spatial distance + intensity difference < threshold is added to the expanded mask.

The expanded mask is combined with the original binary mask using a logical OR operation.

Connected components are identified, and the largest connected region is retained.

The final region is refined by filling holes to ensure continuity and morphological completeness.

### 5.4.2 Results and Analysis 4

The experimental results of the enhanced merging model, incorporating the Spatial Intensity Metric (SIM) expansion, demonstrate significant improvements in segmentation accuracy. Below, the evaluation metrics and their distributions are summarized.



doi:10.6342/NTU202500232

| | |
|---|---|
| Average F1 Score | 0.5450 |
| Average IoU Score | 0.4460 |

Table 6 Average F1 and IoU Score incorporating the Spatial Intensity Metric (SIM)

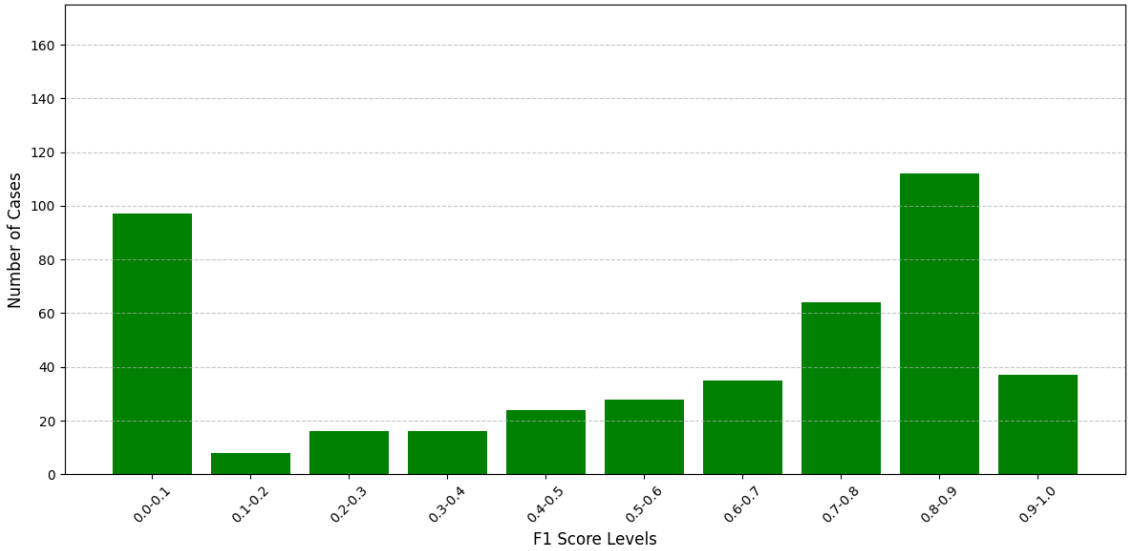



Fig 20 F1 score distribution incorporating the Spatial Intensity Metric (SIM)

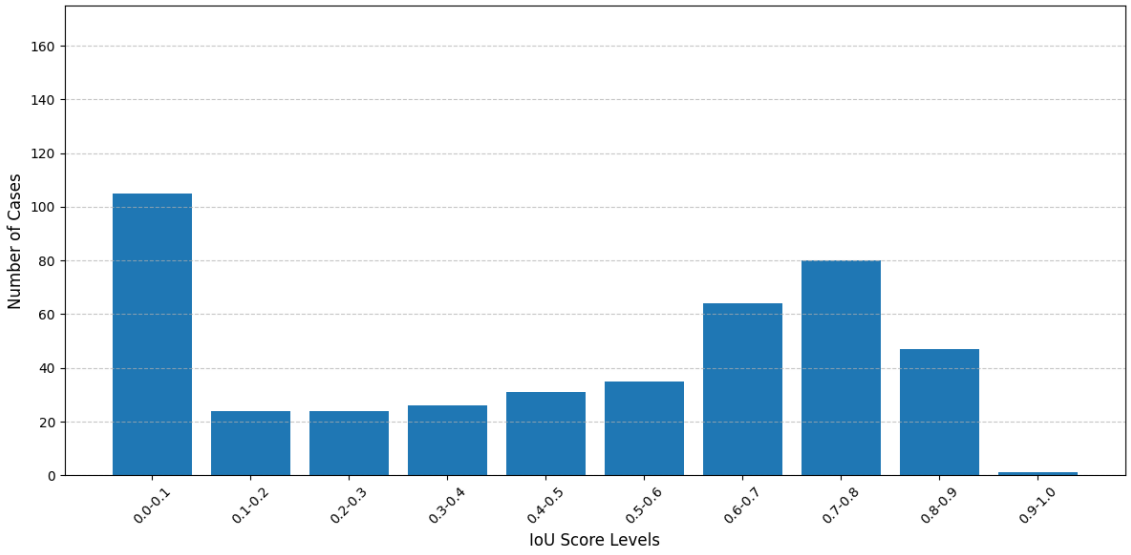



Fig 21 IoU score distribution incorporating the Spatial Intensity Metric (SIM)

The enhanced model achieved an average F1 Score of 0.5450, indicating improved balance between precision and recall compared to the previous merging model.

The average IoU score increased to 0.4460, reflecting a more precise overlap between the predicted tumor regions and the ground truth.



doi:10.6342/NTU202500232

The distribution shows a marked improvement in high-performing cases.

A substantial increase in the number of cases achieving both scores in the 0.8–0.9 and 0.9–1.0 ranges is observed.

Although the proportion of cases in the lowest range (0.0–0.1) has decreased, some low-scoring cases remain, indicating challenges in handling certain tumor patterns.

Fig 22 illustrates the processing transitions in cases where the introduction of the SIM effectively expanded regions that the active contour model failed to capture, successfully delineating the tumor boundaries.

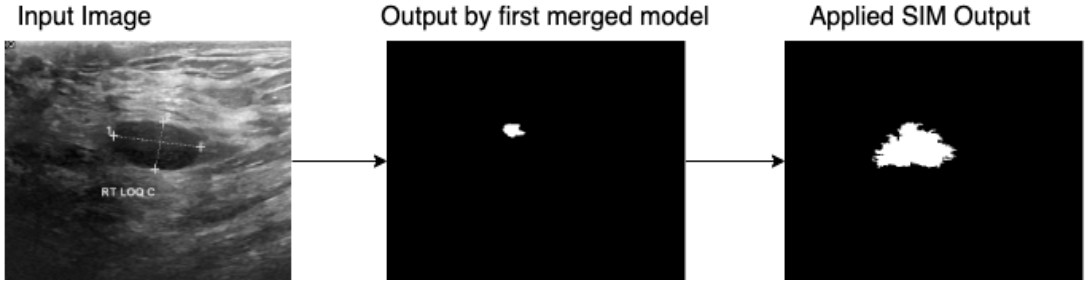



Fig 22 Example of a case where the SIM is working effectively



doi:10.6342/NTU202500232

# Chapter 6 Conclusion and future work

## 6.1. Conclusion

This study presents a progressive approach to tumor segmentation in medical imaging, integrating multiple models and refining methodologies to achieve enhanced accuracy and reliability. Initially, the performance of Model 1 highlighted the importance of robust preprocessing and active contour refinement but also revealed the limitations of relying solely on a single model. Although Model 1 effectively provided initial tumor predictions, its inability to capture complex tumor boundaries and subtle intensity variations led to insufficient segmentation performance, as reflected in lower F1 and IoU scores.

To address these challenges, Model 2 was introduced, leveraging noise-augmented preprocessing and iterative refinement to complement Model 1's predictions. The iterative process of Model 2, combined with cumulative confidence scoring, demonstrated the capacity to detect tumor regions more consistently. However, while the combined framework of Model 1 and Model 2 significantly improved segmentation accuracy, it still struggled with under-segmented regions and boundary inconsistencies, particularly in cases with low contrast or irregular tumor shapes.



doi:10.6342/NTU202500232

To overcome these limitations, an enhanced merging model was developed, incorporating a Spatial Intensity Metric (SIM) expansion. This approach refined and expanded the major tumor regions by accounting for spatial and intensity relationships in the surrounding areas. The introduction of SIM expansion led to significant improvements in segmentation accuracy, as evidenced by higher average F1 and IoU scores. Furthermore, the refined boundaries and expanded regions effectively addressed the issues of under-segmentation and incomplete delineation, making the enhanced model more robust and reliable.

Overall, the results demonstrate that a hybrid approach combining complementary model predictions, iterative refinement, and advanced post-processing techniques can achieve superior tumor segmentation. While the proposed methodologies show promising results, the remaining challenges in handling low-contrast and irregular tumor regions highlight the need for further enhancements. Additionally, as predictions are made on a 16×16 grid basis, further improvements are required to effectively address cases involving large tumor regions that span multiple grids.



doi:10.6342/NTU202500232

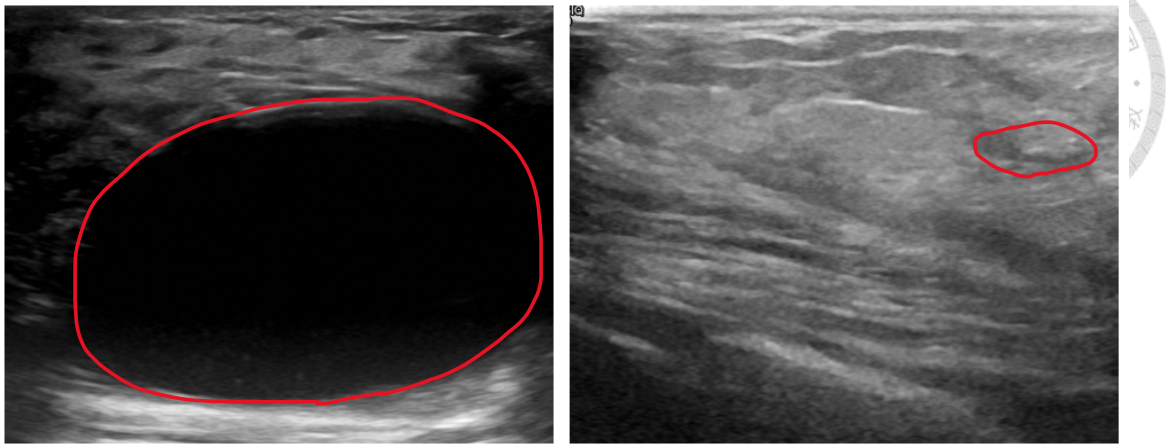


Fig 23 Examples of cases where improvements are needed

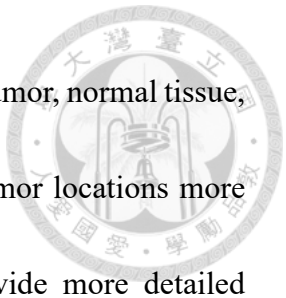
## 6.2. Future Work

The current methodology relies on a grid-based approach, dividing images into 16×16 pixel grids to classify each grid as either tumor or normal tissue. While this approach provides a structured way to segment tumors, its effectiveness diminishes when dealing with large tumors that span broad regions. In such cases, the grid resolution may fail to adequately capture the tumor's full extent, resulting in under-segmentation or boundary inaccuracies. Future efforts could explore increasing grid resolution or implementing adaptive grid sizes based on tumor characteristics to better accommodate varying tumor sizes.

Additionally, the current labeling strategy assigns each grid a binary label based on the proportion of tumor pixels it contains. While effective for general classification, this method does not account for the nuanced transitions between tumor and normal tissues.



doi:10.6342/NTU202500232



Expanding the labeling system to include additional classes, such as tumor, normal tissue, and boundary regions, could enhance the model's ability to infer tumor locations more precisely from the initial predictions. This refinement would provide more detailed guidance for subsequent refinement processes, potentially improving overall segmentation accuracy.

Furthermore, as the grid-based approach becomes more precise, there is potential to extend the methodology to incorporate local texture analysis for distinguishing between malignant and benign tumors. High-resolution grids combined with advanced feature extraction techniques could facilitate simultaneous tumor segmentation and malignancy assessment, providing more comprehensive diagnostic insights. This expansion would significantly enhance the clinical utility of the model, moving beyond segmentation to offer valuable prognostic information.



doi:10.6342/NTU202500232

# References

[1] Kaus, M.R.; Warfield, S.K.; Nabavi, A.; Black, P.M.; Jolesz, F.A.; Kikinis, R. Automated Segmentation of MR Images of Brain Tumors. Radiology 2001, 218, 586–591.

[2] Prastawa, M.; Bullitt, E.; Ho, S.; Gerig, G. A brain tumor segmentation framework based on outlier detection. Med. Image Anal. 2004, 8, 275–283.

[3] Hui, J.; Xixi, J.; Lei, Z. Clustering based content and color adaptive tone mapping. Comput. Vis. Image Underst. CVIU 2018, 168, 37–49.

[4] Nanda, A.; Barik, R.C.; Bakshi, S. SSO-RBNN driven brain tumor classification with Saliency-K-means segmentation technique. Biomed. Signal Process. Control 2023, 81, 104356.

[5] Usman Akbar, M., Larsson, M., Blystad, I. et al. Brain tumor segmentation using synthetic MR images - A comparison of GANs and diffusion models. Sci Data 2024, 11, 259.

[6] Wang, T.; Cheng, I.; Basu, A. Fluid vector flow and applications in brain tumor segmentation. IEEE Trans. Biomed. Eng. 2009, 56, 781-789.



doi:10.6342/NTU202500232

[7] Khotanlou, H.; Colliot, O.; Atif, J.; Bloch, I. 3D brain tumor segmentation in MRI using fuzzy classification, symmetry analysis and spatially constrained deformable models. Fuzzy Sets Syst. 2009, 160, 1457-1473.

[8] Kermi, A.; Andjouh, K.; Zidane, F. Fully automated brain tumour segmentation system in 3D-MRI using symmetry analysis of brain and level sets. IET Image Process. 2018, 12, 1964-1971.

[9] Khosravanian, A.; Rahmanimanesh, M.; Keshavarzi, P.; Mozaffari, S.; Kazemi, K. Level set method for automated 3D brain tumor segmentation using symmetry analysis and kernel induced fuzzy clustering. Multimed. Tools Appl. 2022, 22, 89.

[10] El-Dahshan, E.S.A.; Hosny, T.; Salem, A.B.M. Hybrid intelligent techniques for MRI brain images classification. Digit. Signal Process. 2010, 20, 433-441.

[11] Abd-Ellah, M.K.; Awad, A.I.; Khalaf, A.A.M.; Hamed, H.F.A. Design and implementation of a computer-aided diagnosis system for brain tumor classification. J. Comput. Sci. 2016, 16, 1-13.

[12] Zhang, Y.; Dong, Z.; Wu, L.; Wang, S. A hybrid method for MRI brain image classification. Expert Syst. Appl. 2011, 38, 10049-10053.

[13] Devasena, T.; Hemalatha, M. A hybrid approach for brain tumor detection and classification using support vector machine and artificial neural networks. Int. J. Comput. Appl. 2014, 93, 21-25.



doi:10.6342/NTU202500232

[14] Sarith, M.; Sreeja, S. A novel approach for brain tumor detection using hybrid techniques. Int. J. Comput. Appl. 2015, 116, 1-5.

[15] Khairandish, M.; Khosravi, A.; Nahavandi, S.; Nguyen, T. A hybrid deep learning model for brain tumor classification. Comput. Methods Programs Biomed. 2020, 190, 105-123.

[16] Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Int. Conf. Med. Image Comput. Comput.-Assist. Interv. 2015, 9351, 234-241.

[17] Milletari, F.; Navab, N.; Ahmadi, S.-A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. Int. Conf. 3D Vis. 2016, 10, 565-571.

[18] Isensee, F.; Jaeger, P. F.; Kohl, S. A.; Petersen, J.; Maier-Hein, K. H. nnU-Net: A Self-Configuring Method for Deep Learning-Based Biomedical Image Segmentation. Nat. Methods 2021, 18, 203-211.

[19] Xu, Y.; Xu, C.; Zou, L.; Li, X.; Zhao, Y.; Cai, J. Hybrid Transformer and Convolutional Neural Network for Medical Image Segmentation. Med. Image Anal. 2022, 79, 102472.

[20] Selvaraju, R. R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. IEEE Int. Conf. Comput. Vis. 2017, 618-626.



doi:10.6342/NTU202500232