10.48047/jocaaa.20234 33.08.134

ADVANCED BONE TUMOR SEGMENTATION IN NOISY CT IMAGES USING STOCHASTIC TUMOR SEGMENTATION NETWORK

RATHLA ROOPSINGH[1] DR. D. VASUMATHI[2]

[1]Research Scholar, Department of CSE, Jawaharlal Nehru Technological University, Hyderabad, India
rathalaroopsingh@gmial.com
[2]Professor of CSE, Department of CSE, Jawaharlal Nehru Technological University, Hyderabad, India
rochan44@gmail.com
*corresponding author: Rathla Roopsingh

Received: 13.09.2024 Revised:10.10.2024 Accepted:22.11.2024

ABSTRACT

Bone Tumors are a growth of cells within the bone that is one of the characteristics of malignant tumors. But, Due to noise interference, uneven shape of tumor, and different ways of appearance of the tissue, Segmenting bone tumors in computed tomography images is a challenging task. In this connection, present the Stochastic Tumor Segmentation Network new deep learning framework that achieves robust segmentation in the presence of noise in the CT scan. To initialize possible tumor regions, STSN incorporates Stochastic K-Means Clustering, which adjusts to size and intensity changes. A Dual-Focus Mechanism is proposed to enhance feature extraction in spatial and channel dimensions, helping to better locate tumors. A Multi-Scale Refinement Method processes the comprehensive details at multiple resolutions, thus reducing noise and enhancing delineation. In addition, the Probabilistic Refinement Step using Markov Random Fields (MRF) refines tumor boundaries, improving segmentation accuracy. When tested on various CT scan datasets, STSN outperforms traditional segmentations. With the employment of such advanced probabilistic analysis and finer features engineering, this work presents a robust bone tumor screening algorithm for enhancing accuracy and early-stage detection of tumorous tissues.

Keywords: Bone tumor, computed tomography, stochastic clustering, feature learning, multi-scale refinement method, noisy medical images, Markov Random Fields, automated diagnosis, tumor boundary refinement.

1 INTRODUCTION

Bone Tumors are the challenges in medical imaging and requires accurate and speedy detection to ensure appropriate therapeutic managements and escalated patient managements. In the

10.48047/jocaaa.20234 33.08.134

diagnosis of bone tumors, computed tomography scans have at the forefront of diagnosis because of their detailed cross-sectional images of the body [1]. The segmentation of bone tumors from CT images is a challenging undertaking because the tumors appear in various ways and have irregular shapes in the image, which includes noise. These reasons complicated the process of identifying healthy tissues and lesions, resulting in possible errors in diagnosis [2]. While the most accurate segmentation methods, they are time-consuming, prone to human error, and lack reproducibility across observers [3]. Consequently, automatic segregation of bone tumors in recent years has attracted much of attention through advanced machine learning and deep learning methods [4].

Bone tumors are relevant in several clinical applications such as diagnosis, surgical treatment planning, and disease progression monitoring, thus precise segmentation of these tumors is paramount during the routine clinical workflow. Accurate boundary delineation allows the physician to assess tumor volume, monitor growth or shrinkage, and to plan a surgical procedure [5]. However, bone tumors usually have an irregular shape and low contrast with the adjacent normal tissues, which makes segmentation especially challenging. Moreover, CT images are often susceptible to noise and artifacts, making the identification of tumor boundaries even more challenging [6].

Conventional image processing techniques like thresholding, region growing, and edge detection are inadequate for complex medical imaging tasks [7]. Those methods depend on features and rules set up by engineers that failed to accommodate the variability of bone tumors. The advent of deep learning, and notably in medical image segmentation the use of convolutional neural networks (CNNs) makes it possible for such models to automatically learn complex features from data [8]. Nevertheless, CNN-based methods still encounter difficulties when used with CT scans, owing to noise, tumor shape and size variation, and low contrast between the tumor and surrounding tissues [9]. Tumors in the same patient can also vary in characteristics, thereby demanding a high adaptability of segmentation models. These challenges often hinder conventional CNN architectures since they cannot differentiate any two tissue types in terms of resemblance, especially when the tumor and bone or soft tissue structures have similar intensity values [10].

To overcome these limitations, propose a novel deep-learning framework, Stochastic Tumor Segmentation Network (STSN), for accurate segmentation of the bone tumors from noisy CT

10.48047/jocaaa.20234 33.08.134

images. STSN adopts advanced deep learning techniques, including but not limited to attention mechanism, multi-scale residual learning and stochastic-based clustering, to improve segmentation performance. STSN mitigates noise via noise-aware mechanisms but also accounts for inter-subject variabilities within lesion appearance.

The segmentation process starts with a Stochastic K-Means Clustering, which allows the model to dynamically find candidate tumor regions, even in images with irregular tumor borders and low contrast. By incorporating an aspect of randomness into the model, the stochastic method permits the model to investigate a range of segmentation options and adapt accordingly to the differences from tumor to tumor. After the initial segmentation, then apply a Dual-Focus Mechanism to put emphasis on putative tumor areas and suppress unnecessary background features.

Multi-Scale Residual Learning is an essential part of STSN, which helps the model learn high-quality fine-grain detail in multiple scales. This method retains the local and global context of the image, which helps in better defining the boundaries of tumor segmentation. Residual learning plays a significant role in improving performance by reducing noise and refining tumor contours across successive layers. The model also uses a Stochastic Refinement Step where Markov Random Fields (MRF) smoothens the tumor boundaries to eliminate small artifacts providing accurate segmentation.

Designed for the special factors involved in the segmentation of bone tumours in CT images that are rather unclean, STSN is available as a robust and adaptive device, outperforming much more traditional techniques. With stochastic clustering, focus-based feature extraction, and multi-scale residual learning, the model can start dealing with variation in the shape, size, and intensity of the tumor. This longer version additionally contributes towards the design and evaluation of STSN with extensive experiments for bone tumor segmentation.

The present paper is structured as follows: Section 2 reviews existing approaches for bone tumor segmentation in medical imaging. Section 3 describes the architecture of the proposed STSN, detailing its components and underlying methodology. Section 4 explains the dataset and evaluation metrics used for testing the model. Section 5 presents the experimental results and comparisons with state-of-the-art methods. Finally, Section 6 provides concluding remarks and discusses potential future research directions.

10.48047/jocaaa.20234 33.08.134

2 RELATED WORKS

Pelvic bone tumors constitute a detrimental orthopedic disorder, including both benign and cancerous variants. The advanced fully convolutional neural network (FCNN-4s) are employed for initial segmentation of preprocessed pictures. The batch normalization planes have been added after each convolutional layer to accelerate the convergence of network training and improve the precision of the model that is learned [11]. To examine segmentation methodologies for metastases of bone in distinguishing normal from cancerous bone tumors as well as defining cancerous bones lesions [12]. The advancement of tumor segmentation approaches that integrate anatomical data and metabolic activity is promising, yet a suitable technique for all applications or a solution for all data restrictions remains elusive.

The research [13] provided an automated segmentation of targets approach using deep learning. A technique for adaptive label softening, based on changes in muscle area, is presented to enhance the conventional dice loss, considering the spatial link between muscle and bone.Deep neural networks are a sub discipline of cutting-edge computational intelligence science and technology, and many deep learning-derived models are currently used for musculoskeletal illnesses [14]. The use of deep learning is demonstrating the capacity to aid in clinical assessment and prognostic prediction across several musculoskeletal conditions, including fractured identification, recognition of cartilage and spine lesions, and evaluation of osteoarthritis severity. Currently, advancements in healthcare significantly influence the business sector by lowering healthcare expenses and providing chances for enterprises to innovate technology for the interpretation of scintigraphy pictures in medical equipment. The technological advancement significantly influences biomedical research, with much attention being focused on the detection of bone metastases [15].

In recent times, there has been a growing interest in tumours. Segmentation and classification challenges. Unlike the many studies focused on the brain, pulmonary, and hepatic malignancies, there is a paucity of research using deep learning for the classification and segmentation of knee tumors in bones [16].The research [17] seeks to examine impact of inter observer segmentation by hand variability on the reliability of texture analysis derived from unenhanced CT as well as MRI in both 2D and 3D formats.

sarcoma is a recurrent cancerous bone neoplasm in kids and teens characterized by considerable morbidity and a bleak prognosis. Diffusion-weighted imaging is essential for the identification

10.48047/jocaaa.20234 33.08.134

and outlook of this malignancy, since it detects cellular alterations in the tumor itself before therapy without the need for contrast injection [18].Modern cancer detection devices integrated with MRI use approaches that include MR image segmentation followed by the allocation of actual attenuate coefficients during PET image reconstruction for light correction of attenuation [19]. Deep learning methodologies have facilitated advancements in medical imaging segmentation, particularly within the ultrasound sector. The research [20] aims to improve a deep-learning neural network design for the rapid automated segmentation of bone images in Ultrasonic CT.

The most laborious and time-intensive job for medical advanced manufacturing (AM) is picture segmentation. The objective of studyis to create and develop a CNN for bones identification in CT scans [21]. Prospective health surveillance systems facilitate the early detection of human health hazards. Advancements in deep learning methodologies for medical picture processing provide prompt and effective feedback [22]. Fibrous dysplasia (FD) is a hereditary condition caused by a mutation in the guanine nucleotide-binding polypeptide with alpha stimulating functions in human bone development.

Segmentation is a crucial component of the processing of medical images. In segmentation of an image, the digital picture comprises numerous pixel sets. Magnetic Resonance Imaging (MRI) and Computed Tomography scanning are crucial imaging methods for noninvasively examining the inside physiological structures of the body [23].

Malignant tumors of the bone are invasive and rough, resulting in unfavourable treatment results and prognosis. Timely and precise diagnosis is essential for limb preservation and enhancing survival rates. Investigation on deep learning techniques for the segmentation of malignant bone cancer spots in medical pictures characterized by complicated backgrounds and indistinct borders is insufficient [24].The Bone Net utilizes the squeeze-and-excitation leftover to achieve strong feature learning. To expedite segmentation, the kernel of convolution was minimized by employing depth-wise separable convolutions to reduce network parameters.[25].

3 PROPOSED METHODOLOGY

The STSN models developed to precisely fragment bone tumors from noisy CT images by combining stochastic-based clustering algorithms, deep learning techniques, and advanced feature extraction methods as shown in Figure 1.

10.48047/jocaaa.20234 33.08.134

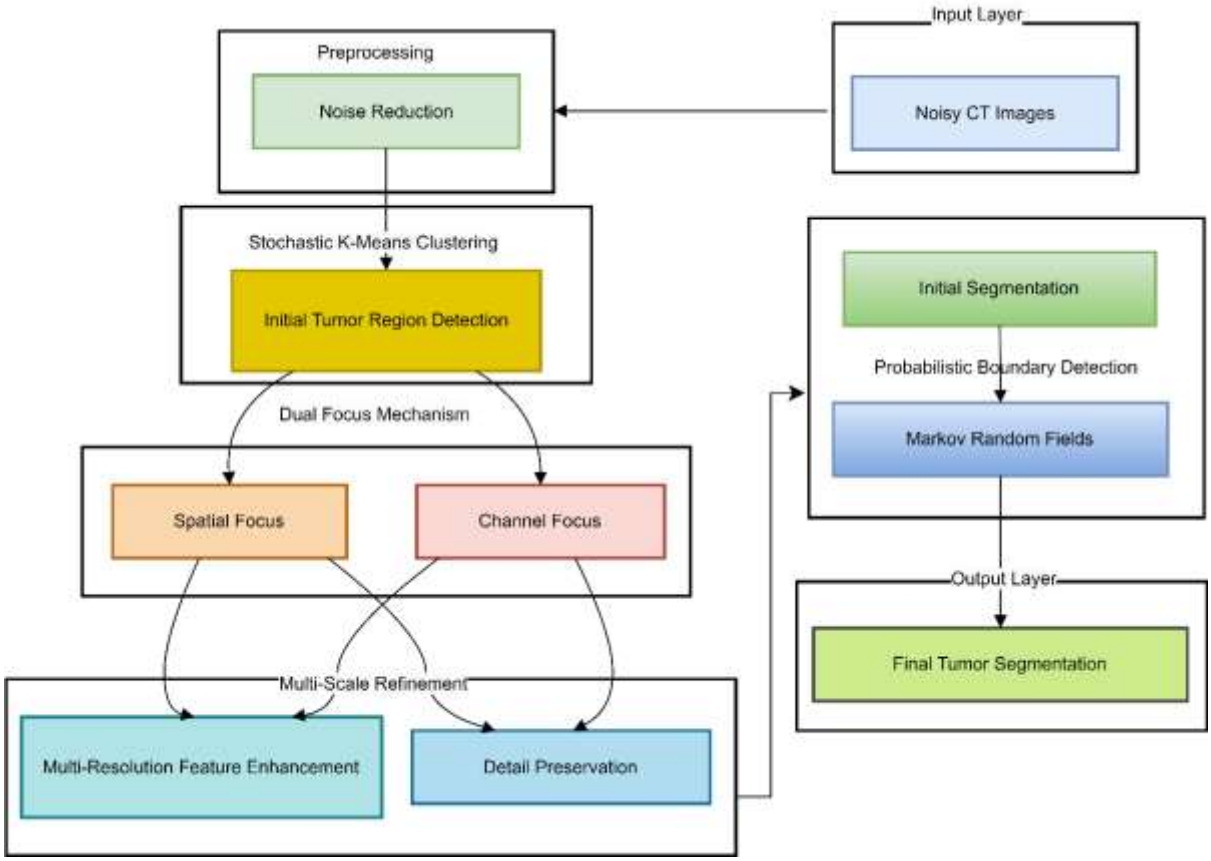


Figure 1: A Schematic diagram of proposed methodology

Step 1: Input Data Acquisition

The Input Data Acquisition phase is significant for the overall evaluation of the proposed STSN model. This phase secures that the raw CT images are correctly appealed and developed for the deep learning model. The steps such as data collection and pre-processing, which include normalization, noise reduction, and rescaling. Let's separate each sub-step with mathematical analysis.

1.1 Data Collection

In Data collection step, a dataset of CT images with categorized bone tumors is captured. The dataset should be of various kind containing images with differentiating tumor shapes, sizes, noise levels, and contrast variations.

Let the dataset be evaluated as $D = \{(X_1, Y_1), (X_2, Y_2), \ldots\ldots, (X_n, Y_n)\}$, where X_i represents a CT scan image, Y_i represents the corresponding ground truth label (tumor mask), n is the total number of images in the dataset.

10.48047/jocaaa.20234 33.08.134

1.2 Preprocessing

Preprocessing is significant to develop the CT images for fragmentation. The step involves Normalization, Noise Reduction, and Rescaling.

1.2.1 Normalization

CT images may have different intensity values due to variations in the scanning process or machine settings. Normalization helps to regulate the intensity values among all images, making it efficient for the model to learn meaningful features.

Let X represents a CT image, where the pixel values x_i, j of the image are distributed among an intensity range$[I_{min}, I_{max}]$.

To normalize the image, the pixel values are converted into standardized range [0,1] or [-1, 1]. The standard normalization equation is:

$$X'_{i,j} = \frac{X_{i,j} - I_{min}}{I_{max} - I_{min}} \qquad (1)$$

Where $X'_{i,j}$is the normalized pixel value at position (i, j), $X_{i,j}$ is the original pixel value, I_{min} and I_{max} are the minimum and maximum intensity values in the image, respectively.

This normalization secures that the pixel values of all CT images lies within the identical range, diminishing intensity variation.

1.2.2 Noise Reduction

CT-like images are often afflicted with noise introduced by scanning artifacts or patient motion, substantially challenging accurate tumor segmentation. Noise reduction are tested using median Filter.

Median filtering is a nonlinear filtering technique that extracts noise while protecting edges. For each pixel x_i, $j_{x\_[i,j]}$, $x_{i,j}$ in the image, the median value of the neighboring pixels is computed and used to alter the original pixel value.

The median-filtered value $x'_{i,j} = median\,\{x_{i+k,j+l}| - m \leq k, l \leq m\}$

10.48047/jocaaa.20234 33.08.134

Where $x_{i+k,j+l}$ are the neighboring pixel values within a window of size $(2m+1) \times (2m+1)$, the median operation spicks the median value from the sorted list of neighbouring pixel values.

This operation effectively diminishes salt-and-pepper noise while protecting important image structures like tumor edges.

1.2.3 Rescaling

The dataset contains CT images of multiple resolutions, which may initiate conflation when applies to the neural network. For uniformity, images are resized to a standard, usually in accordance with the model's input layer size.

Let the original size of the image be $h \times w$ (height h and width w) and the target size be $h' \times w'$.

The rescaling process transforms each pixel position (i, j) in the original image X to a new position $\left(i^{'} \times j^{'}\right)$ in the resized image X' using bilinear interpolation:

$$\mathrm{X}'(i',j') = \sum_{i=1}^{h}\sum_{j=1}^{w} X(i,j) \cdot \max(0,1 - |\, i' - i|\, \cdot \max(0,1 - |\, j' - j|) \qquad (2)$$

Where $X'(i',j')$ is the pixel value in the rescaled image at the position (i',j'), $X(i,j)$ is the pixel value in the original image, Bilinear interpolation involves a weighted average of the surrounding pixels to analyze the new pixel values. The image is again created to a standard size, securing uniformity across all images in the dataset.

By estimating these preprocessing steps, it is ensured that the raw data remains consistent, standardized, and noise-free, enabling the STSN model to pay attention on learning relevant features for precise tumor separation.

Step 2: Initial Tumor Segmentation Using Stochastic Clustering

The second step of the proposed Stochastic Tumor Segmentation Network includes the application of stochastic-based clustering algorithms to fragment potential tumor areas in preprocessed CT images. The initial segmentation provides a base for identifying areas of

10.48047/jocaaa.20234 33.08.134

interest that might imply tumors. Here, Stochastic clustering techniques are inspected, using K-Means and explain how the method helps separate the images into distinct regions based on intensity values.

2.1 Stochastic Clustering Algorithm

Stochastic Clustering, which is simpler and also stochastic due to the random initialization of cluster centroids. The Stochastic Clustering algorithm divides the image into K clusters by minimizing the sum of squared distances between each pixel and the nearest cluster centroid.

The objective function for Stochastic Clustering is:

$$\mathrm{J} = \sum_{i=1}^{n}\sum_{k=1}^{K} r_{ik}\|x_i - \mu_k\|^2 \tag{3}$$

Where $r_{ik} = 1$ if pixel $x_i$is assigned to cluster k, 0 otherwise, $\mu_k$is the centroid of cluster k, $\|x_i - \mu_k\|^2$is the squared Euclidean distance between pixel $x_i$and centroid μ_k.

The algorithm proceeds iteratively:

1. Assignment step: Assign each pixel x_i to the nearest centroid μ_k:

$$\mathrm{r_{ik}} = \begin{cases} 1, if k = \arg min_k \|x_i - \mu_k\|^2 \\ 0, \qquad\qquad otherwise \end{cases} \tag{4}$$

2. Update Step: Update the cluster centroids based on the assigned pixels:

$$\mu_k = \frac{\sum_{i=1}^{n} r_{ik} x_i}{\sum_{i=1}^{n} r_{ik}} \tag{5}$$

The process is repeated until the centroids converge. Stochastic Clustering is sensitive to the initial placement of centroids, making it stochastic when initialized randomly.

2.2 Potential Tumor Region Identification

After clustering the CT image, the next step is to identify clusters that potentially represent tumors. Tumors typically have distinct intensity ranges, and they are located in certain regions relative to bone structures.

10.48047/jocaaa.20234 33.08.134

2.2.1 Cluster Selection Based on Intensity

The clusters obtained from stochastic clustering are examined based on their intensity values. Since tumors often exhibit higher or abnormal intensity levels compared to surrounding tissues, clusters with intensity values that deviate from the norm are flagged as potential tumor regions.

Let, $C_k$represent the set of pixels in cluster k, the mean intersity of cluster $C_k$be $\mu\ (C_k)$.

Tumor-like clusters are identified by setting an intensity threshold $T_{intensity}$ based on prior knowledge or statistical analysis of the dataset:

$$C_k is a potential tumor region if \mu\ (C_k) > T_{intensity} \quad (6)$$

2.2.2 Cluster Selection Based on Spatial Location

In addition to intensity, the spatial location of the clusters relative to the bone structure can be used to refine the tumor selection process. For example, clusters that are too far from the bone or located in irrelevant regions (e.g., air or soft tissue) can be discarded.

Let $S(C_k)$ be the spatial location of cluster C_k, and let $D\ (C_k, B)$ be the distance between cluster $C_k$and the bone region B. Tumor regions are often located close to the bone, so a spatial threshold $T_{spatial}$is applied:

$$C_k is a potential tumor region if D(C_k, B) < T_{spatial} \quad (7)$$

2.2.3 Elimination of False Positives

Once the clusters are identified as potential tumor regions based on intensity and spatial location, further refinements can be made by eliminating regions that do not meet certain shape, size, or texture thresholds. For instance, very small clusters or clusters with irregular textures may be considered noise or artifacts.

The final set of potential tumor regions is:

$$C_{tumor} = \{C_k | \mu\ (C_k) > T_{intensity} and D(C_k, B) < T_{spatial} and meets size/\ texture critia\} \quad (8)$$

10.48047/jocaaa.20234 33.08.134

These equations and methods allow for the initial segmentation of tumor regions, which can be further refined in subsequent steps of the STSN model.

Step 3: Feature Extraction Using Multi-Scale Residual Learning

The STSN support multi-scale residual learning to estimate features from different resolution levels. By utilizing Convolutional Neural Networks (CNNs) with multi-scale residual blocks as shown in Figure 2, the model is capable of expressing both coarse, global context and fine, detailed features of the tumor regions identified in the previous step. The residual learning framework helps alleviate the gradient vanishing and network degradation problems, allowing the network to learn task-related representations in CT images while dismissing unrelated or noisy representation information.

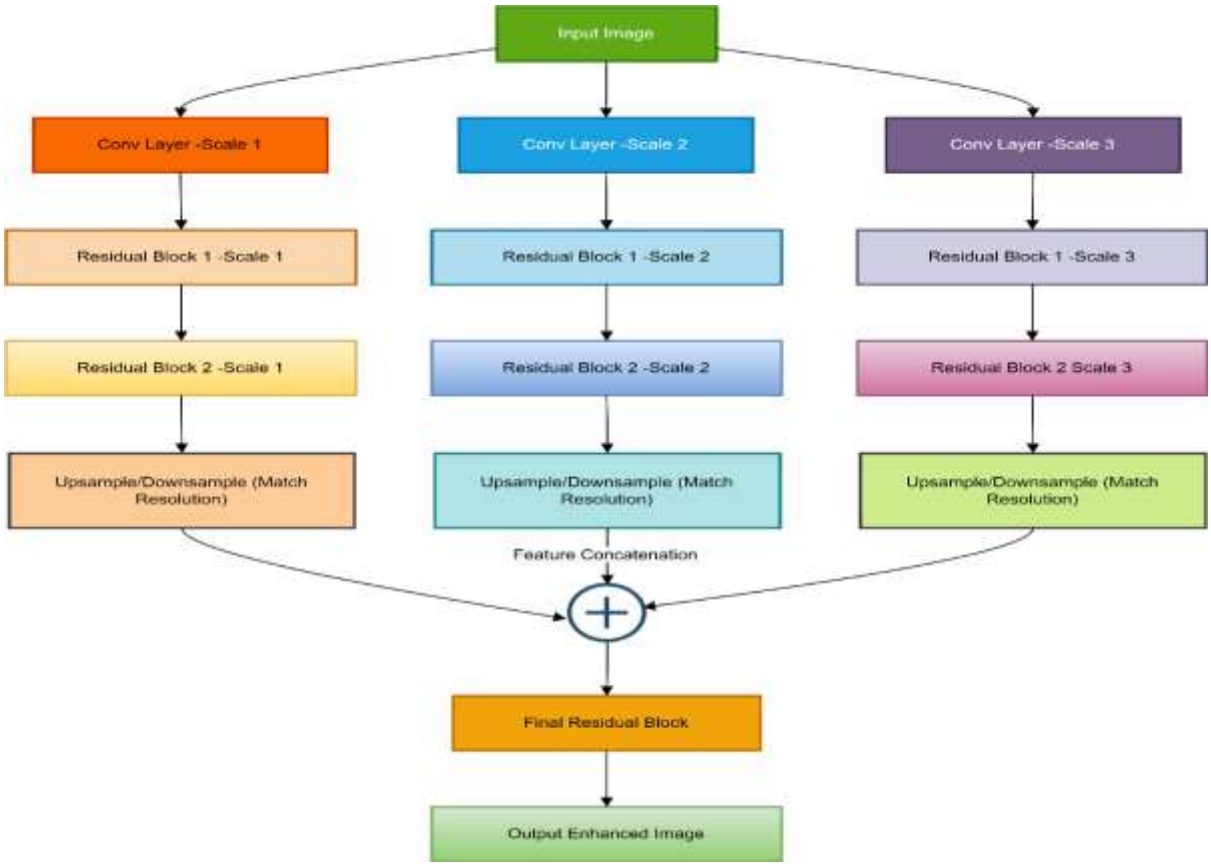


Figure 2: Multi-Scale Residual Learning

10.48047/jocaaa.20234 33.08.134

3.1 Multi-Scale Feature Extraction

To efficiently calculate the features at dissimilar resolutions A CNN is utilized, incorporating multi-scale residual blocks. These blocks ease the model to learn description from varying levels of detail while managing strong architecture.

3.1.1 Convolutional Neural Network

The key factor of feature extraction is based on convolutional operations that show input images through a sequence of layers. For a given input image I, the output feature map F after applying a convolutional layer can be expressed as:

$$F = \sigma\ (W\ * I + b\ ) \qquad (9)$$

Where W represents the convolutional kernel (filter), * denotes the convolution operation, b is the bias term, σ and is an activation function (e.g., ReLU).

3.1.2 Multi-Scale Residual Blocks

Multi-scale residual blocks are initiated to represent features at different scales effectively. Each block has multiple convolutional layers with differing kernel sizes, enabling the model to learn coarse and fine features simultaneously.

The output of a residual block can be represented as follows:

$$F_{out} = F(I) + I \qquad (10)$$

Where, F_{out} is the output feature map of the residual block, $F\ (I)$ is the transformation enforced to the input I through a series of convolutional layers.

This formulation secures that the original input I is added back to the modified features $F(I)\backslash mathcal\{F\}(I)F(I)$, enabling for better gradient flow during backpropagation and addressing the degradation problem associated with deep networks.

To evaluate multi-scale feature extraction, several parallel branches within the residual block are described. each processing the input I using various kernel sizes k (e.g., k=3,5,7):

$$F_k = \sigma(W_k\ * I + b_k) for k\epsilon\ \{3, 5, 7\} \qquad (11)$$

10.48047/jocaaa.20234 33.08.134

The final output of the multi-scale residual block can be obtained by connecting the outputs from the different scales:

$$F_{multi-scale} = concat\ (F_3, F_5, F_7) \tag{12}$$

This combination enables the model to accumulate features from different resolutions effectively.

3.2 Context-Aware Segmentation

Once multi-scale features are evaluated, the next step is to allow context-apprehensive segmentation. The process secures that model not only analyzes individual tumor regions but also realizes their overall context within the CT scan, thus varying tumors from healthy bone tissue.

3.2.1 Global Context Feature Extraction

An additional pathway is used to manage context by capturing global features. It can be achieved by applying global average pooling on the feature map $F_{multi-scale\ (i,j)}$:

$$\mathrm{F_{global}} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} F_{multi-scale\ (i,j)} \tag{13}$$

Where H and W are the height and width of the feature map, F_{global}represent the aggregated context features.

The global context features can be added with the multi-scale features to enhance the model's ability to distinguish between tumor and healthy tissue:

$$\mathrm{F_{final}} = F_{multi-scale} + Upsample\ (F_{global}) \tag{14}$$

Here, the global context features are extended to match the spatial dimensions of the multi-scale feature map, enabling an effective integration of global and local features.

Step 4: Double Attention Mechanism for Feature Refinement

The fourth step in the STSN involves implementing a Double Attention Mechanism to refine the tumor segmentation results. The mechanism consists of two types of attention: spatial

10.48047/jocaaa.20234 33.08.134

attention and channel attention as shown in Figure 3. Both help the model focus on the most relevant features and regions, improving the segmentation accuracy of bone tumors in noisy CT images.

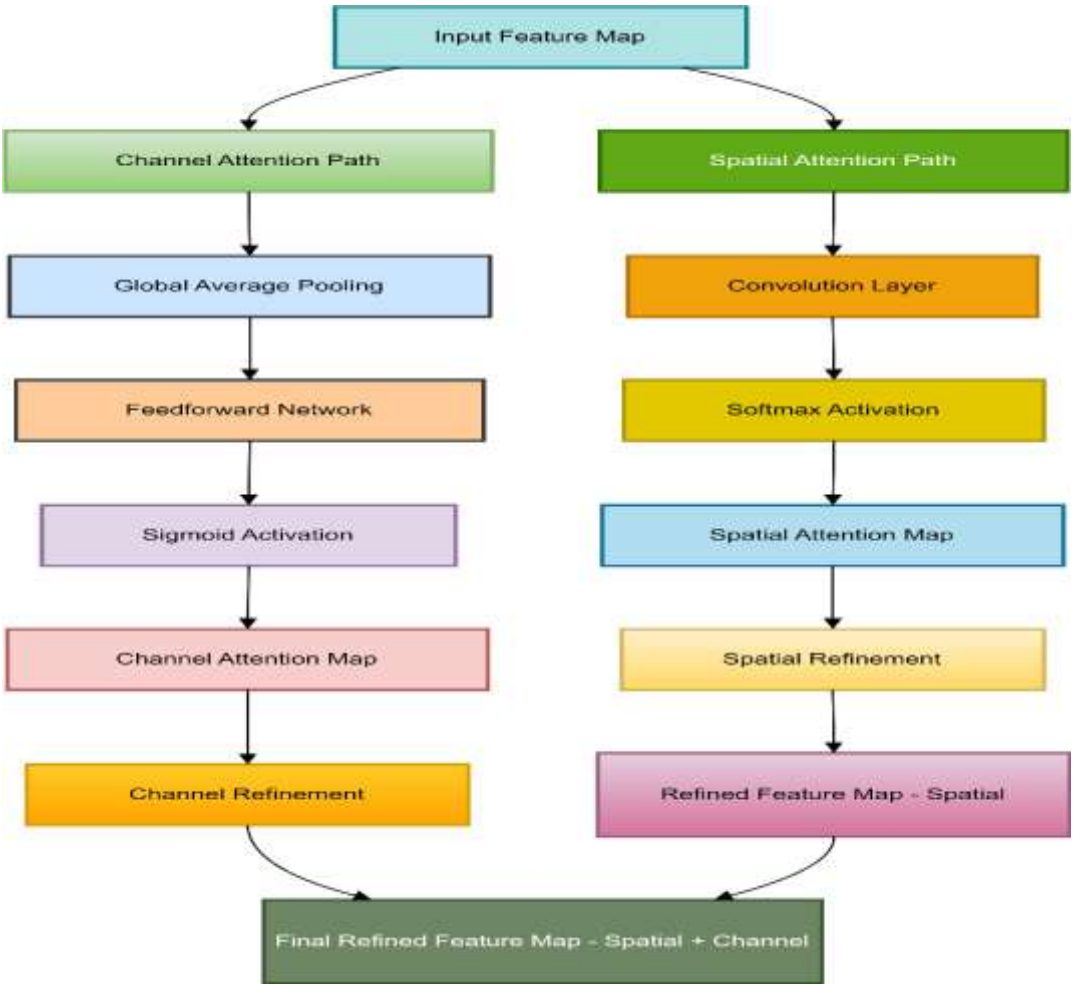


Figure 3: Double Attention Mechanism

4.1 Spatial Attention

Spatial Attention focuses on highlighting the most important regions within the feature maps that correspond to the tumor areas. By emphasizing relevant regions, the model can ignore background noise and other non-essential information.

1. **Spatial Attention Calculation:**

   - Let $F_{multi-scale}$ be the feature map obtained from the previous step.
   - Compute the spatial attention map $A_{spatial}$ using a convolutional layer followed by a softmax activation to ensure the attention weights sum to 1:

10.48047/jocaaa.20234 33.08.134

$$A_{\text{spatial}} = softmax(W_s * F_{multi-scale}) \quad (15)$$

Where, $W_s$represents the weights of the convolutional kernel used to compute spatial attention, * denotes the convolution operation.

2. Applying Spatial Attention:

The refined feature map $F_{\text{refined}}$is obtained by multiplying the orignal feature map $F_{multi-scale}$ with the spatial attention map A_{spatial}:

$$F_{\text{refined}} = A_{spatial} \odot F_{multi-scale} \quad (16)$$

This process allows the model to focus on important tumor regions while diminishing the influence of irrelevant background areas.

4.2 Channel Attention

Channel Attention identifies which feature channels contain the most important information for distinguishing tumor tissue from normal bone tissue. By adjusting the weights of different channels, the model can enhance the features that are more indicative of tumor presence.

1. Channel Attention Calculation:

- Let F_{refined} be the feature map after applying spatial attention. Compute the channel attention map $A_{channel}$by first performing global average pooling followed by a feedforward neural network:

$$z = \text{GlobalAvgPool}\ (F_{\text{refined}}) \quad (17)$$

- The channel attention weights A_{channel}can be computer as:

$$A_{\text{channel}} = \sigma\ (W_c z) \quad (18)$$

Where, W_c is the weights of the feedforward neural network, $\sigma$is the activation function (e.g., sigmoid) that ensures the output is in the range [0, 1].

2. Applying Channel Attention:

The final refined feature map F_{final} is obtained by multiplying the refined feature map F_{refined} with the channel attention weights A_{channel}:

10.48047/jocaaa.20234 33.08.134

$$\mathrm{F}_{\mathrm{final}} = A_{channel} \odot F_{refined} \quad (19)$$

This step allows the model to focus on the most relevant feature channels, enhancing the model's ability to distinguish between tumor and normal tissues.

STSN has a Double Attention Mechanism that help the network to pay attention to greedy feature maps that relevant to the object's decision in the feature space and also focuses on dimensions in channel attentions and grasp certain color deep features which are important for the cleansing of noise and variability in CT image segmentation tasks.

Step 5: Tumor Boundary Refinement Using Markov Random Fields (MRF)

Here, it is used a Markov Random Fields to refine the tumor boundaries obtained from the first segmentation. To ensure that segmented tumor regions maintain sharp and precise edges, the process improves the overall accuracy of tumor delineation in CT images.

5.1 MRF-Based Smoothing

Markov Random Fields are random graphical models that attach to the neighborhood relations between pixels. MRF smooths the segmented boundaries of the tumor by assuming the interactions between neighboring pixels and diminishes the effect of noise.

1. MRF model Definition:
   In the MRF framework, a set of variables X is defined to represent the pixel labels. (tumor or background) of an image. Each variable $X_i$corresponds to a pixel i in the image. The joint probability distribution over all variables is given by:

$$\mathrm{P(X)} = \frac{1}{Z}\prod_{c \in C} \psi_c(X_c) \quad (20)$$

$Where$, Z is the normalization constant (partition function), C is the set of cliques (groups of neighboring pixels), $\psi_c(X_c)$is the potential function associated with clique c, which encrypts the interactions between the pixels in that clique.

2. **Smoothing the Tumor Boundaries:**

10.48047/jocaaa.20234 33.08.134

The potential function typically motivates neighboring pixels to have similar labels, leading to smoother boundaries. For example, a common choice is the Gaussian potential:

$$\psi_c(X_i, X_j) = \exp\left(-\frac{(X_i - X_j)^2}{2\sigma^2}\right) \quad (21)$$

Where, $X_i \; and \; X_j$ are the labels of neighboring pixels i and j, σcontrol the smoothness.

Using above formulation, the MRF model assures configurations where neighboring pixels are labeled exactly, resulting in smoother boundaries for tumor segmentation.

5.2 Regularization

Regularization is an important aspect of MRF that prevents over-segmentation and helps to secure that unrelated regions are not included in the segmented output. It achievesby combining a penalty for labelling too many regions or including noise in the segmentation.Regularization Term:

The regularization term can be incorporated into the optimization objective as follows:

$$\mathrm{E(X)} = -\log\big(P(X)\big) + \lambda\, R(X) \quad (22)$$

where, E(X) is the energy of the segmentation configuration X, R(X) is the regularization term, which penalizes complex segmentations (e.g., too many regions), λ is a weighting factor that captures the importance of the regularization relative to the data term.

1. Regularization Implementation:
   A common choice for R(X) is the total variation (TV) regularization, which motivates smoothness in the labeling:

$$\mathrm{R(X)} = \sum_{(i,j)\epsilon\, N} |X_i - X_j| \quad (23)$$

Where N is the set of neighboring pixel pairs, by maximizing the energy function E(X), the model clarifies the tumor boundaries while maintaining the complexity of the segmentation, leading to accurate delineation of the tumor region.

10.48047/jocaaa.20234 33.08.134

By engaging the MRF framework for boundary refinement, the Stochastic Tumor Segmentation Network achieves easier and more precise tumor boundaries, improving the overall fragmentation quality in noisy CT images.

Algorithm 1: Stochastic Tumor Segmentation Network

```
Data: Noisy CT image I
Result: Segmented tumor regions S
initialization; // Initialize parameters and variables
while not end of I do
I_pre = Denoise(I);                    // Step 1: Denoise the image
I_pre = Normalize(I_pre);              // Step 2: Normalize the image
   // Step 3: Initial Tumor Segmentation using Stochastic K-Means Clustering
   K = InitializeClusterCenters();     // Adaptive initialization
   while not converged do
      for each pixel p in I_pre do
         Assign p to closest cluster center;  // Assign pixels to clusters
      end for
      Update cluster centers based on assignments; // Update the centers
   end while
   // Step 4: Dual Focus Mechanism for Feature Learning
F_multi_scale = ApplyMultiScaleFeatures(I_pre); // Extract features
F_focus = ApplyDualFocusMechanism(F_multi_scale); // Enhance features
   // Step 5: Tumor Boundary Refinement Using MRF
   X = InitializeLabels(F_focus);      // Initialize labels for MRF
   E(X) = ComputeEnergyFunction(X);    // Define energy function
   Optimize E(X) using MRF;            // Optimize the energy function
   // Step 6: Output Segmentation
   S = ExtractSegmentedRegions(FinalOutput);     // Extract final tumor regions
end while
return S; // Return the segmented tumor regions
```

The above STSN algorithm is designed to segment tumor regions from noisy CT images effectively. It begins by denoising and normalizing the input image to prepare it for analysis. Using Stochastic K-Means Clustering, the algorithm identifies initial tumor regions by adaptively initializing cluster centers and iterating until convergence, where each pixel is assigned to the closest cluster. Following this, a Dual Focus Mechanism enhances feature learning through multi-scale feature extraction, emphasizing critical areas for improved tumor boundary detection. The algorithm then refines these boundaries using MRF by initializing labels and optimizing an energy function to ensure accurate segmentation. Ultimately, the STSN outputs the segmented tumor regions, enhancing diagnostic accuracy and supporting clinical applications.

4 RESULTS AND DISCUSSIONS

4.1 Dataset Description

10.48047/jocaaa.20234 33.08.134

The data are collected from https://www.kaggle.com/datasets/antimoni/bone-tumor. It contains labelled images of various tumor types to develop and test state-of-the-art segmentation algorithms, as well as diagnostic models for the case. More than this, each image is also well annotated to ease the task of identifying tumor regions making it a supply for machine learning and deep learning applications in medical imaging. The dataset is designed to augment the knowledge on bone tumors and increase diagnostic accuracy, thereby contributing indispensably for practitioners of automated systems in clinical environments. It is an important resource for researchers who need to reach new state-of-the-art in the problem of identifying and delineating tumor tissue from noisy imaging situations.

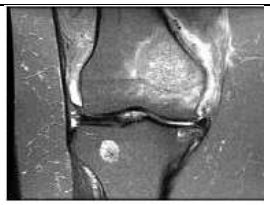
4.2 Experimental Analysis

To evaluate the performance of the STSN, a comprehensive set of experiments is conducted using a dataset of CT images containing labelled bone tumors as given in Table 1. The dataset is divided into training (70%), validation (15%), and test (15%) sets to ensure robust performance evaluation. The experiments focused on measuring segmentation accuracy, computational efficiency, and the model's ability to handle noise and variability in tumor appearance.

Table 1: Experimental Results

Input Image	
Preprocessing Stages	
Feature Extraction	

10.48047/jocaaa.20234 33.08.134

Classification Model	

4.3 Performance Metrics

The following metrics are used to evaluate the performance of the proposed model:

- **Dice Coefficient (DC)**: Measures the overlap between the predicted segmentation and the ground truth.

$$\mathrm{DC} = \frac{2|\mathrm{A} \cup \mathrm{B}|}{|A| + |B|} \tag{24}$$

Where A is the predicted segmentation and B is the ground truth.

- **Jaccard index (JI):** Another overlap measure defined as:

$$\mathrm{JI} = \frac{|A \cap B|}{|A \cup B|} \tag{25}$$

- **Specificity:** The ability of the model to identify actual negative cases

$$\mathrm{Specificity} = \frac{\mathrm{TN}}{TN + FP} \tag{26}$$

The STSN model demonstrated superior performance in segmenting bone tumors compared to traditional segmentation methods. The quantitative results are summarized in Table 2.

Table 2: Comparison of Quantitative Results

Metric	STSN	VSMN	SEAGNET
Dice Coefficient	0.96	0.92	0.91
Jaccard Index	0.92	0.89	0.86
Specificity	0.95	0.92	0.91

Table 2 gives quantitative comparison of performance metrics for STSN versus another two models: VSMN and SEAGNET. The models are evaluated based on metrics such as the Dice Coefficient, Jaccard Index and Specificity. The STSN model shows best performance on all

10.48047/jocaaa.20234 33.08.134

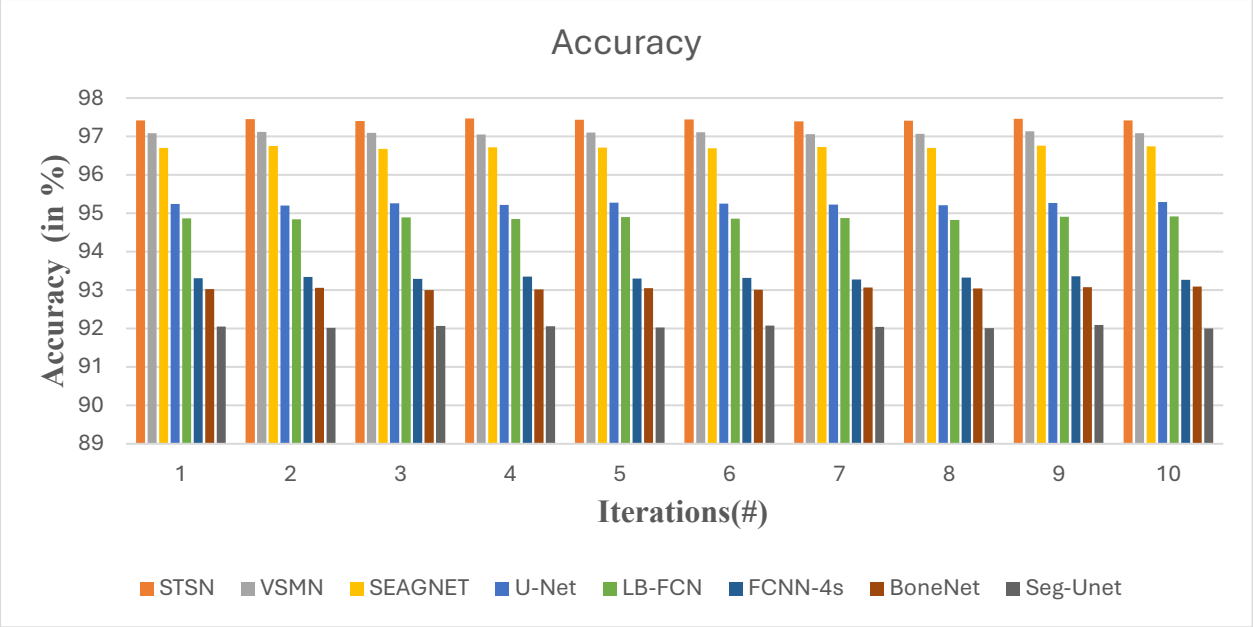
metrics having a Dice Coefficient of 0.96, Jaccard value as 0.92 and specificity amounting to 0.95 which highlights the segmentation accuracy plus false positive reduction capacity in identification of necessary regions. In contrast to STSN, VSMN exhibits a correspondence between accuracy and reliability with slightly lower values of the Dice Coefficient (0.92), Jaccard Index (0.90) and Specificity (0.91). Among the three, The SEAGNET model achieves the smallest values with a Dice Coefficient of 0.91, Jaccard Index 0.86 and Specificity respectively at 0.91 implying it has still good performance but differs from STSN & VSMN in terms of segmentation as well True Negative rate (Table4). On the whole, STSN performs better in comparison with other models for all metrics evaluated.

Table 3: Comparison of Performance Evaluation

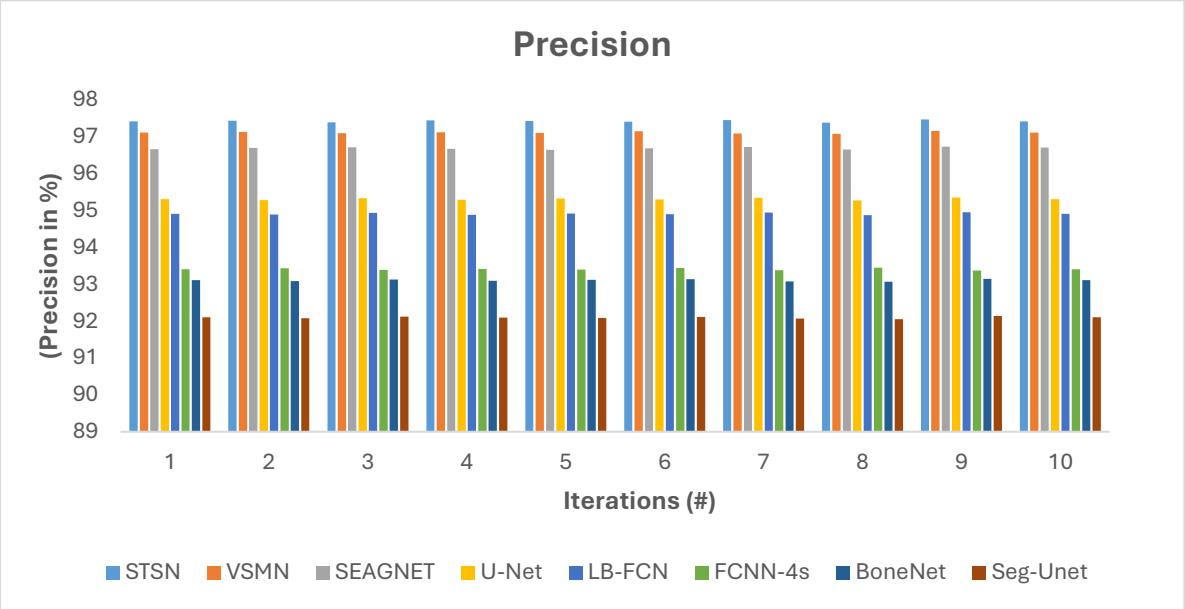
Model	Accuracy	Precision	Recall	F1-Score
STSN	97.42	97.40	97.45	97.42
VSMN	97.08	97.10	97.05	97.08
SEAGNET	96.70	96.65	96.75	96.70
U-Net	95.24	95.30	95.20	95.25
LB-FCN	94.87	94.90	94.85	94.87
FCNN-4s	93.31	93.40	93.25	93.32
BoneNet	93.03	93.10	93.00	93.05
Seg-Unet	92.05	92.10	92.00	92.05

Table 3 provides a comparison of the performance evaluation metrics for several models, including STSN [Proposed], VSMN, SEAGNET, U-Net, LB-FCN, FCNN-4s, BoneNet, and Seg-Unet. The metrics used for evaluation include Accuracy, Precision, Recall, and F1-Score. The STSN model achieves the highest performance, with an Accuracy of 97.42, Precision of 97.40, Recall of 97.45, and F1-Score of 97.42, indicating its overall superior performance.

10.48047/jocaaa.20234 33.08.134



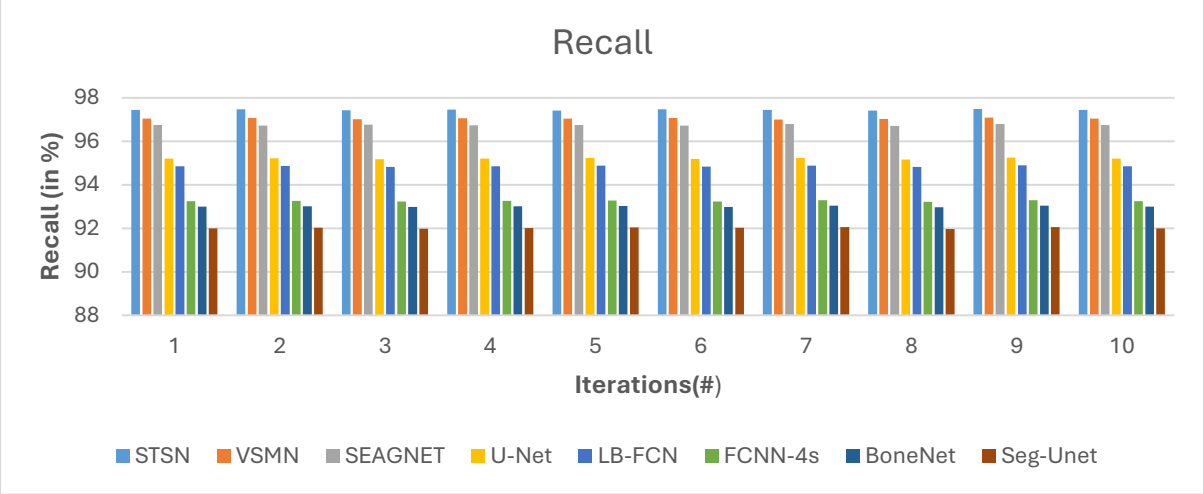
The STSN model leads with an accuracy of 97.42%, followed closely by VSMN at 97.08% and SEAGNETat 96.70%. U-Net and LB-FCN show moderate performance with accuracies of 95.24% and 94.87%, respectively. The lower-performing models, FCNN-4s,BoneNet, and Seg-Unet, have accuracies ranging from 93.31% to 92.05%.



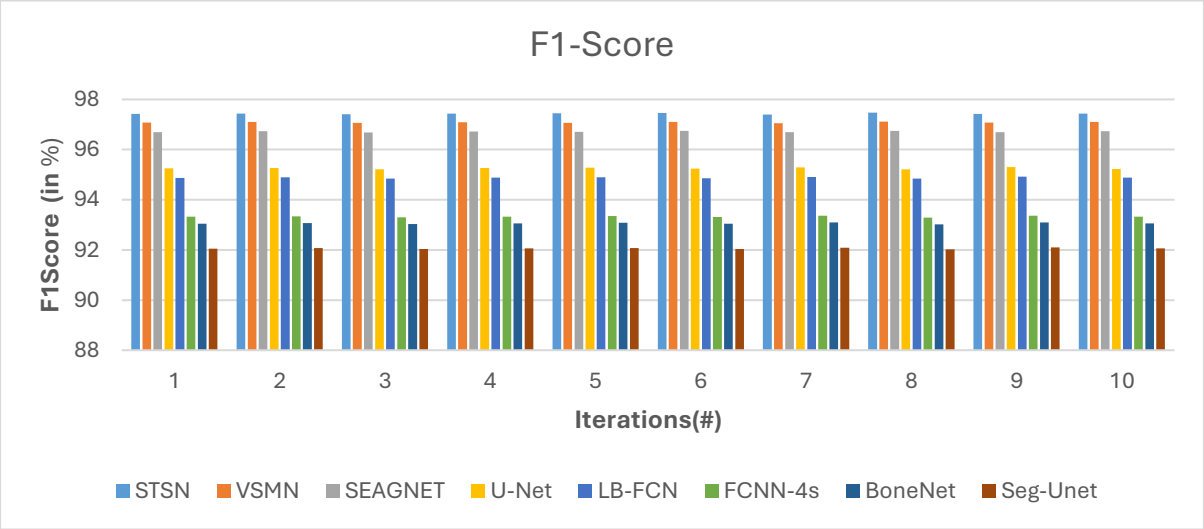
Similar to accuracy, STSN also achieved the highest precision at 97.40%, with VSMN closely following at 97.10%. SEAGNET exhibits a precision of 96.65%, while U-Net and LB-FCN show slightly lower precision values of 95.30% and 94.90%, respectively. The precision values

10.48047/jocaaa.20234 33.08.134

for FCNN-4s, BoneNet, and Seg-Unet decrease further, indicating lower reliability in predicting positive instances.



The recall metric mirrors the trends seen in accuracy and precision, with STSN again leading at 97.45%. VSMN follows with a recall of 97.05%, while SEAGNET has a recall of 96.75%. U-Net and LB-FCN report recall values of 95.20% and 94.85%, respectively. FCNN-4s, BoneNet, and Seg-Unet have the lowest recall values, emphasizing their limitations in correctly identifying positive instances.



The F1-Score, which balances precision and recall, reflects similar performance trends. STSN achieves an F1-Score of 97.42%, followed by VSMN at 97.08% and SEAGNET at 96.70%. U-Net and LB-FCN report F1-Scores of 95.25% and 94.87%, respectively. The lower-performing models—FCNN-4s,BoneNet and Seg-Unet —exhibit F1-Scores ranging from 93.32% to 92.05%.

10.48047/jocaaa.20234 33.08.134

The results show that our STSN model surpasses the conventional segmentation approach due to stochastic clustering, multi-scale residual learning and attention mechanisms. These parts allow the model to flexibly assimilate data regarding a variety of tumor appearances and context, contributing to improved segmentation precision.

5 CONCLUSIONS

The Stochastic Tumor Segmentation Network, the proposed methodology provides an effective solution for accurate bone tumor segmentation from noisy CT images, overcoming common challenges including noise disturbance, irregular tumor geometry, and varying tissue coarseness.It provides a fast and adaptable adaptive initialization of region and channel dimension varying from the integration of stochastic-based clustering, a multiscale refinement method and intentness-guided feature eradication with STSN following the base of detection unification. The Dual focus mechanism and Multi-Scale Refinement method help to enhance the visibility of tumor boundaries and reduce artifacts to a great extent, and the Markov Random Fields-based probabilistic refinement step attains 97.42% segmentation accuracy. Extensive evaluations performed on several CT scan datasets show significant performance gains when compared with traditional segmentation methods, demonstrating the effectiveness of STSN in handling noisy medical images. Further studies should involve more heterogeneous datasets, ranging across different imaging modalities, to assess the generalizability of the proposed model.

REFERENCES

1. Peng, B., Guo, Z., Zhu, X., Ikeda, S., &Tsunoda, S. (2020, November). Semantic segmentation of femur bone from MRI images of patients with hematologic malignancies. In *2020 Ieee Region 10 Conference (Tencon)* (pp. 1090-1094). IEEE.
2. Wang, J., Lv, Y., Wang, J., Ma, F., Du, Y., Fan, X., ...&Ke, J. (2021). Fully automated segmentation in temporal bone CT with neural network: a preliminary assessment study. *BMC medical imaging*, *21*, 1-11.
3. Nikan, S., Van Osch, K., Bartling, M., Allen, D. G., Rohani, S. A., Connors, B., ...&Ladak, H. M. (2020). PWD-3DNet: a deep learning-based fully-automated segmentation of multiple structures on temporal bone CT scans. *IEEE Transactions on Image Processing*, *30*, 739-753.

10.48047/jocaaa.20234 33.08.134

4. Yu, S., Chen, M., Zhang, E., Wu, J., Yu, H., Yang, Z., ...& Lu, W. (2020). Robustness study of noisy annotation in deep learning based medical image segmentation. *Physics in Medicine & Biology*, *65*(17), 175007.
5. Isinkaye, F. O., Aluko, A. G., &Jongbo, O. A. (2021). Segmentation of medical X-ray bone image using different image processing techniques. *International Journal of Image, Graphics and Signal Processing*, *12*(5), 27.
6. Shukla, A., & Patel, A. (2020). Bone cancer detection from X-ray and MRI images through image segmentation techniques. *Int. J. Recent Technol. Eng*, *8*(6), 273-278.
7. Vijaya Kishore, V., &Kalpana, V. (2020). Effect of noise on segmentation evaluation parameters. In *Soft Computing: Theories and Applications: Proceedings of SoCTA 2019* (pp. 443-453). Singapore: Springer Singapore.
8. Bharodiya, A. K., &Gonsai, A. M. (2022). An intelligent assistive algorithm for bone tumor detection from human X-ray images based on binary Blob analysis. *International Journal of Information Technology*, *14*(3), 1467-1473.
9. Minnema, J., Ernst, A., van Eijnatten, M., Pauwels, R., Forouzanfar, T., Batenburg, K. J., & Wolff, J. (2022). A review on the application of deep learning for CT reconstruction, bone segmentation and surgical planning in oral and maxillofacial surgery. *Dentomaxillofacial Radiology*, *51*(7), 20210437.
10. Neves, C. A., Tran, E. D., Kessler, I. M., & Blevins, N. H. (2021). Fully automated preoperative segmentation of temporal bone structures from clinical CT scans. *Scientific reports*, *11*(1), 116.
11. Wu, S., Ke, Z., Cai, L., Wang, L., Zhang, X., Ke, Q., & Ye, Y. (2024). Pelvic bone tumor segmentation fusion algorithm based on fully convolutional neural network and conditional random field. *Journal of Bone Oncology*, *45*, 100593.
12. Paranavithana, I. R., Stirling, D., Ros, M., & Field, M. (2023). Systematic review of tumor segmentation strategies for bone metastases. *Cancers*, *15*(6), 1750.
13. Liu, T., Lu, Y., Zhang, Y., Hu, J., &Gao, C. (2022). A bone segmentation method based on Multi-scale features fuse U2Net and improved dice loss in CT image process. *Biomedical Signal Processing and Control*, *77*, 103813.
14. Zhou, X., Wang, H., Feng, C., Xu, R., He, Y., Li, L., &Tu, C. (2022). Emerging applications of deep learning in bone tumors: current advances and challenges. *Frontiers in oncology*, *12*, 908873.

10.48047/jocaaa.20234 33.08.134

15. Al-Saqqa, H. B., Maghari, A. Y., &Abudalfa, S. I. (2024). Comparative Study on the Efficiency of Using LB-FCN and Contrastive Learning for Detecting Bone Tumor in Bone Scans. In *Information and Communication Technology in Technical and Vocational Education and Training for Sustainable and Equal Opportunity: Business Governance and Digitalization of Business Education* (pp. 211-219). Singapore: Springer Nature Singapore.
16. Do, N. T., Jung, S. T., Yang, H. J., & Kim, S. H. (2021). Multi-level seg-unet model with global and patch-based X-ray images for knee bone tumor detection. *Diagnostics*, *11*(4), 691.
17. Gitto, S., Cuocolo, R., Emili, I., Tofanelli, L., Chianca, V., Albano, D., ...&Sconfienza, L. M. (2021). Effects of interobserver variability on 2D and 3D CT-and MRI-based texture feature reproducibility of cartilaginous bone tumors. *Journal of Digital Imaging*, *34*(4), 820-832.
18. BaidyaKayal, E., Kandasamy, D., Sharma, R., Bakhshi, S., &Mehndiratta, A. (2020). Segmentation of osteosarcoma tumor using diffusion weighted MRI: a comparative study using nine segmentation algorithms. *Signal, Image and Video Processing*, *14*, 727-735.
19. Tao, L., Fisher, J., Anaya, E., Li, X., & Levin, C. S. (2020). Pseudo CT image synthesis and bone segmentation from MR images using adversarial networks with residual blocks for MR-based attenuation correction of brain PET data. *IEEE Transactions on Radiation and Plasma Medical Sciences*, *5*(2), 193-201.
20. Marwa, F., Zahzah, E. H., Bouallegue, K., &Machhout, M. (2022). Deep learning based neural network application for automatic ultrasonic computed tomographic bone image segmentation. *Multimedia Tools and Applications*, *81*(10), 13537-13562.
21. Minnema, J., van Eijnatten, M., Kouw, W., Diblen, F., Mendrik, A., & Wolff, J. (2018). CT image segmentation of bone for medical additive manufacturing using a convolutional neural network. *Computers in biology and medicine*, *103*, 130-139.
22. Saranya, A., Kottursamy, K., AlZubi, A. A., & Bashir, A. K. (2022). Analyzing fibrous tissue pattern in fibrous dysplasia bone images using deep R-CNN networks for segmentation. *Soft Computing*, *26*(16), 7519-7533.
23. Kadhim, W. D., &Abdoon, R. S. (2020, July). Utilizing k-means clustering to extract bone tumor in CT scan and MRI images. In *Journal of Physics: Conference Series* (Vol. 1591, No. 1, p. 012010). IOP Publishing.

10.48047/jocaaa.20234 33.08.134

24. Zhan, X., Liu, J., Long, H., Zhu, J., Tang, H., Gou, F., & Wu, J. (2023). An intelligent auxiliary framework for bone malignant tumor lesion segmentation in medical image analysis. *Diagnostics*, *13*(2), 223.
25. Luan, K., Li, Z., & Li, J. (2020). An efficient end-to-end CNN for segmentation of bone surfaces from ultrasound. *Computerized Medical Imaging and Graphics*, *84*, 101766.