

Original Article

A Hybrid Semantic Model for MRI Kidney Object Segmentation with Stochastic Features and Edge Detection Techniques

Sitanaboina S L Parvathi¹, Harikiran Jonnadula²

^{1,2}School of Computer Science and Engineering, Vellore Institute of Technology, VIT-AP University, Vijayawada.

¹Corresponding Author : s.srilakshniparvathi@gmail.com

Received: 19 July 2022

Revised: 10 August 2022

Accepted: 03 September 2022

Published: 30 September 2022

Abstract - Because of its importance in disease diagnosis, medical image analysis has been a prominent research topic for over a decade. Advancements in deep learning methodologies made computer-aided disease diagnosis feasible from medical images. Object semantic segmentation is the primary activity in medical image analysis. Various deterministic deep learning models for semantic segmentation of objects (organs) from medical images are introduced. Generally, the medical image objects (i.e., kidneys) contain the routine shape, size, and brightness which help the deterministic models for efficient segmentation. In the case of chronic diseases such as cancer, the objects in medical images contain tumors, which appear with a high degree of uncertainty in object properties. Due to the uncertainty in object shape, size, and brightness, former deep learning models performed less accurately in diseased object segmentation. In this paper, we proposed a hybrid semantic segmentation model with stochastic feature mapping techniques for the accurate segmentation of medical image objects under uncertainty. The location-dependent split method is used for seeded region marking and approximating object location. Stochastic neural networks with feature mapping techniques are introduced to localize the target objects without deterministic modeling. The recursive block-wide segmentation process is used to lineate the target objects from boundary elements. We tested the proposed stochastic segmentation model and its deep learning counterparts on a human kidney tumor MRI dataset. The experimental results show that the proposed stochastic segmentation model outperformed the segmentation of diseased kidney tumor objects with high accuracy and reliability.

Keywords - Stochastic Feature Mapping, Block Segmentation, Kidney Object Segmentation, Deep Neural Networks, Medical Image Analysis.

1. Introduction

Medical imaging equipment (such as X-rays, CT scans, and MRIs) scans the body's internal organs and generates medical images for clinical analysis and disease diagnosis. Deep learning-based Computer-Aided Diagnosis (CAD) has been an efficient methodology for performing Medical Image Analysis (MIA) for over a decade. Object (organ) segmentation is the fundamental task in computer vision applications, which designs the computation models to lineate the objects from the images or video frames. Before the medical image analysis, the image should get segmented into various objects based on their properties and features.

According to the world kidney day theme "2020 WKD Theme [1]," 850 million people worldwide suffer from non-communicable kidney diseases. Chronic kidney disease affects one out of every ten senior citizens, and treatment costs approximately 2-3 percent of the country's annual health budget. The GBD Chronic Kidney Disease Collaboration published an article [2] in 2020, estimating that global kidney transplantation will reach 5.4 million by 2030, indicating the need to prevent kidney diseases at

early stages. MRI [3] has recently proven to be the best imaging method for kidney scanning, with high 3D resolution and a clear separation of normal and diseased tissues. Medical MR images contain various objects (organs) that must be segmented accurately with boundaries for further analysis. For example, abdominal and pelvic MRI frames contain organs such as the liver, kidneys, pancreas, spleen, and adrenal glands. To scrutinize the organs and extract the features from them, the objects of MRI should be segmented. This segmentation will group a set of interrelated pixels (as an object) to draw separation boundaries and label it for further image analysis tasks. Many computer vision applications extensively use object segmentation extensively, such as face recognition, medical image analysis, text recognition, image search, and self-driving technology.

In recent times, deep learning emerged as a promising technology for computer vision applications (i.e., medical image analysis), which works at the pixel level to produce accurate and reliable results. Popular object segmentation models in computer vision include edge-based



segmentation, region-based segmentation, threshold-based segmentation, and deep neural networks-based segmentation.

Choi [4] et al. proposed CNN based deep learning model for automated striatum segmentation from brain MR images. At first, using the global CNN model, they identified the approximate striatum region (location) information and extracted the bounding boxes from the input MR volumes. Later the global CNN-generated bounding boxes and the cropped MR regions were evaluated using the local CNN model for accurate segmentation of the objects and the labeling process. Guo [5] et al. proposed the combination of deep feature learning and sparse patch matching method for the MRI prostate localization process. Initially, they used the Stacked Sparse Auto-Encoder (SSAE) model to extract hidden knowledge, which is used to localize deformable prostate objects in MR images. Later, the sparse patch matching method is used to label the new images based on the inter-patch matching results from SSAE. Ibragimov [6] et al. used the deep CNN models to segment the organs from head and neck Computed Tomography (CT) images. They trained the CNN models with positive and negative intensive patches around the voxels to find consistent intensity patterns, which aids in drawing boundaries around various organs in CT images. They passed the training patches through network layers to find the edges, corners, border points, and location information for accurate localization of the object of interest for efficient segmentation. Kline [7] et al. used the widely known CNN architecture U-Net [8] to segment polycystic disease kidney images from MR images. To handle the deformable shapes and volumes of the diseased kidneys, they captured the local and global contexts of images and applied the contracting path and symmetric expanding paths for voxel-level accurate classification of the kidney images.

In general, healthy kidney images that appear in the form of a bean with a regular size are located in a standard location of the MR images. The diseased (especially cancer) kidney images, on the other hand, have an irregular shape, size, and location due to the disease's severe effect on the kidneys. Most of the former segmentation models [4, 5, and 6] were designed to lineate the objects from various MR images using deep learning classification models. Due to some intrinsic limitations (discussed in section II), segmenting diseased kidney objects from MR images is more complex than other segmentation operations. These limitations include uncertainty, irregularity, and chaotic behavior of the object properties. Although many researchers [7, 9, 10, and 11] concentrated on diseased kidney object segmentation from medical images using various deep learning modalities, they are still suffering from several limitations that are i) uncertain Shape and volume of the diseased kidneys ii) indivisible boundary points (i.e., similar intensities) iii) overlapped objects and iv) unclear boundary points of the object.

To address the above limitations in diseased kidney object segmentation from MR images, we proposed a hybrid semantic model for MRI kidney object segmentation using Stochastic Feature mapping Neural Networks (SFMNN) in this paper. Initially, the seeded regions will get identified using the location-dependent split method to mark the approximate location of the target kidney object. Later the seeded regions are fed into the stochastic feature mapping neural networks for target object mapping. SFMNN uses the image features variability and probability distribution functions for diseased kidney object detection to overcome uncertainty and irregularity issues in segmentation. Finally, the recursive block-wide segmentation process will get applied to SFMNN localized target objects for an accurate segmentation process. Experimental results on tumor kidney MR images proven that the SFMNN-based segmentation model recorded high accuracy in segmentation compared to the traditional deep neural network architectures.

2. Related Work

2.1. Background

Biomedical imaging with ultrasound, CT, and MRI technologies has become reliable and is now playing a vital role in disease diagnosis. Because of the high contrast imaging with greater clarity and the lack of harmful dye for contrast, MRI has become a better and safer option, particularly for kidneys. Kidney stones, glomerulonephritis, polycystic kidney disease, and cancer tumors are certain chronic kidney diseases that can lead to death in humans. According to health organization recommendations [1 and 2], it is critical to diagnose these kidney diseases at an early stage to prevent people from ailments and mortality within a low budget. Traditional methods for diagnosing kidney disease and monitoring growth from medical images require more human resources and a time-consuming process.

For the past few years, various deep learning modalities [4, 5, 6, 8, and 9] have been demonstrated as reliable mechanisms to deal with automated kidney disease diagnosis from medical images. Figure 1 depicts abdominal medical images such as CT and MRI that include several other objects besides kidneys. They are segmented accurately from medical images to extract disease information and parameters from kidneys. Kidney segmentation is a process of locating, mapping, and labeling kidney objects from medical images. Segmentation returns the prominent features (size, volume, and intensity) of the object of interest, which aids in the object localization process for locating diseased lesions.

2.2. Limitations

As stated in section 1, kidney object segmentation becomes too complex due to several inherent limitations in medical images. Figure 1 defines four abdominal cancerous kidney MR images that depict the issues involved in kidney object segmentation. From these images, we observed that the shape and volume of cancer-

diseased kidney images are different (non-uniform) from each other. Unlike the healthy kidney images, the cancerous images contain tumors and appear in an irregular shape. Template matching [12 and 13] is a widely used model to recognize the objects from the image based on the trained templates. In case of shape and volume inconsistency, the kidney object detection [14], recognition, and labeling process become too complex and return inaccurate results. Tumors are a type of kidney object that exists in and around the kidneys in various sizes. Apart from the shape and volume, the intensity contrast (brightness) is also different at various locations of the kidney image. In Figures 1a and 1b, the brightness of the tumor pixels is different from the actual kidney tissues. This variation in pixel brightness values [5] causes

ambiguity in determining the boundaries of the kidney object with tumors. Because of variations in brightness, deep learning models [4, 6, and 9] frequently fail to extract the kidney object as well as its tumors. In general, boundary points help determine the semantic boundaries between the objects in medical images. The irregular shape and volume of kidney objects in Figures 1a and 1d made the bounding points unclear and harder to detect. Similarly, the connected and overlapped objects surrounding the kidney image in Figures 1a and 1c will result in inaccurate lineation. Inconsistent location of the kidney object and the intensity contrast in local and global contexts of the MR images are the other hurdles we found in the segmentation of kidney cancer images.

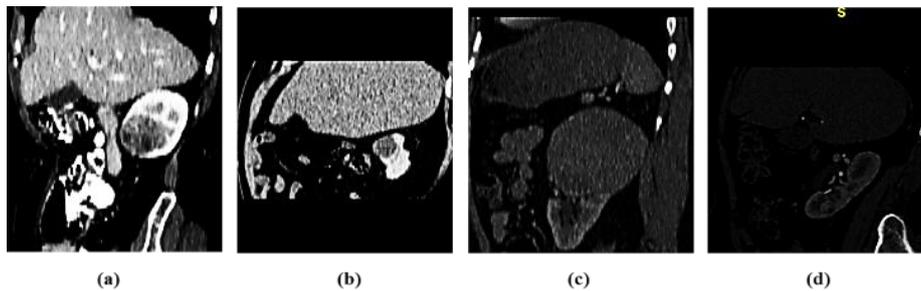


Fig. 1 Representing kidney objects from abdominal MR images

2.3. Image Segmentation Techniques

Different types of segmentation models are available in image processing to segment the objects from the images. Pixel similarity-based segmentation models [15], object structure-based segmentation models, boundary-based segmentation models, and stochastic features-based segmentation models are reliable segmentation models. Pixel similarity-based segmentation model marks the regions [16] of an image based on the threshold values. With the help of the feature similarity techniques, the objects are identified and lineated from the marked regions. The pixel similarity model is best suitable to separate the backgrounds and identify various regions from the images. For object detection and segmentation, an object structure-based segmentation model [17] has given training using target object features such as size, shape, pixel density, and location values. This model is optimal for the images' standard or consistently shaped object detection. Boundary-based segmentation model [18] is similar to the edge detection model and evaluates the pixels' dissimilarity for the object segmentation process. This model is suitable for the segmentation of noisy and inconsistent objects from images as it does not require any prior information about the target object. The stochastic feature-based segmentation model [19] uses the distinct patches of the target object for the segmentation process. This model depends only on discrete features of an image for segmentation. So, this model is the fastest and most lightweight compared to its counterparts in segmentation. Supervised and unsupervised deep image learning algorithms illustrate the stochastic segmentation models. These models are efficient and adaptive to process the high-volume images due to cost and time-saving.

3. Hybrid Semantic Segmentation Model With Stochastic Features and Edge Detection Techniques

Irregular-shaped or diseased kidney object segmentation from the MR images differs from the regular-shaped kidney object segmentation. Section-1 and 2 thoroughly discussed the hurdles involved in diseased kidney segmentation. As the diseased kidneys are inconsistent, varied, and irregular in shape, size, contrast, and patches, none of the individual segmentation techniques (discussed in section 2) are adequate for accurate segmentation. Pixel-based segmentation models [15, 16] are unsuitable for kidney segmentation because they can only concentrate on selecting regions of interest rather than specific objects. Pixel-based segmentation models [15 and 16] are unsuitable for kidney segmentation because they focus only on region selection rather than specific objects. Likewise, object structure-based segmentation models [17] are unsuitable because diseased kidneys with tumors have irregular structures. Boundary-based segmentation models [18] are adaptable for diseased kidney segmentation, but they are not enough for the whole segmentation process. Similarly, stochastic feature-based models [19] are incapable of performing the entire segmentation process accurately.

The object segmentation process is generally categorized into three sequential stages: image classification, object localization, and boundary lineation. Initially, the medical image is classified into various regions in this process (i.e., background separation or target area mapping). Later the target objects are localized

based on the feature analysis. Finally, the edge detection methodologies are applied to accurately line the target objects. None of the individual technology alone can perform this whole object segmentation process from the medical images. We proposed a "hybrid semantic segmentation model with stochastic features and edge detection techniques" to address the limitations of diseased kidney image segmentation from medical MRI images. First, the location-dependent split method is applied to identify the seeded region containing the target diseased kidney object. Later the stochastic feature mapping techniques are applied using the deep learning algorithms to localize the target object. Finally, the boundary detection methodologies would be used to accurately lineate diseased kidney object with tumors.

3.1. Location-Dependent Split Method

Object detection from the bio-medical image is possible with the image analysis using the traditional image segmentation methods like seeded region growing [22 and 23] and region-splitting and merging [21]. These methods will use the supervised seeded points from the training knowledge to adjoin the neighbors (pixels) with similar features to outline the regions from the images. But in our diseased kidney segmentation model, the seeded points are hard to detect due to the inconsistent brightness values of kidney images with tumors. Hence it is required to analyze the whole image with random seeded points, which requires a high amount of memory and processor resources [21].

To address processing level infrastructure limitations and simplify the processing load, we used the location-dependent split method, which evaluates and extracts the possible region for the presence of the kidney object. Instead of detecting seeded points, we encircle seeded regions on medical images using supervised training knowledge from ground truth labels, as discussed in the seeded region detection algorithm.

Algorithm-1: Seeded Region Detection Algorithm

```
// local seeded region detection
for each  $G_i$  from  $G$ 
 $X_b = X_{begin}$  and  $X_e = X_{end}$ 
 $Y_b = Y_{begin}$  and  $Y_e = Y_{end}$ 
 $X_c = \frac{X_b + X_e}{2}$  and  $Y_c = \frac{Y_b + Y_e}{2}$ 
 $C_i = (X_c, Y_c)$ 
 $V_1 = (X_b, Y_b)$  and  $V_2 = (X_e, Y_e)$ 
 $V_3 = (X_b, Y_e)$  and  $V_4 = (X_e, Y_b)$ 
 $d_1 = dist(C_i, V_1)$ , and  $d_2 = dist(C_i, V_2)$ 
 $d_3 = dist(C_i, V_3)$  and  $d_4 = dist(C_i, V_4)$ 
 $R_i = \max(d_1, d_2, d_3, d_4)$ 
```

```
 $G.add\{G_i(C_i, R_i)\}$ 
end
// global seeded region detection
for each  $G_i(C_i, R_i)$  from  $G$ 
 $A_b = C_i - R_i$ 
 $A_e = C_i + R_i$ 
end
 $G_c = \frac{\min(A_b) + \max(A_e)}{2}$ 
 $G_r = \max[(G_c - \min(A_b)), (\max(A_e) - G_c)]$ 
return  $(G_c, G_r)$ 
```

Inspired by the seeded points-based region exploration process [22] and the split and merge-based image segmentation process [21], we proposed the seeded region-based object detection and segmentation method. The main aim of the seeded region detection algorithm is to encircle the approximate region of the medical image in which the target kidney object has existed. Instead of analyzing the whole image for kidney object detection, our proposed method analyzes only the seeded region. As a result, the processing load is reduced by more than half, and object detection will get completed at high speed. As discussed in algorithm -1, our method learns about the object location from the random ground truth labels in a supervised manner to outline the seeded region for object detection. Encircling the seeded region from medical images is accomplished in two phases: local seeded region detection and global seeded region detection. Local seeded region detection will encircle the object region at each ground truth label level before integrating the complete local information to encircle the global seeded region, where the test image kidney objects will have appeared.

The group ' G ' is a collection of ground truth labels containing the manually segmented kidney images for the training dataset. Each label G_i will get presented on a graph as part of the local seeded region detection process, and the kidney object-related most significant quad points are collected as $X_b, X_e, Y_b,$ and Y_e , which are the kidney object-related beginning and endpoints of X and Y-axis (shown in figure 2. A and 2. B). The median of the X and Y axis related quad points X_c and Y_c are used to find the object centroid as $C_i = (X_c, Y_c)$. We configured the four vertices V_1 to V_4 and evaluated the distance from C_i to V_i using the centroid C_i of the kidney object ground truth label, and the quad points $X_b, X_e, Y_b,$ and Y_e , and the maximum distance is selected as the object encircling radius R_i for the label G_i . The same process is iteratively executed with all ground truth labels from G_1 to G_n (as shown in figure-2), and finally, the local seeded region information $G_i(C_i, R_i)$ is added to ' G '.

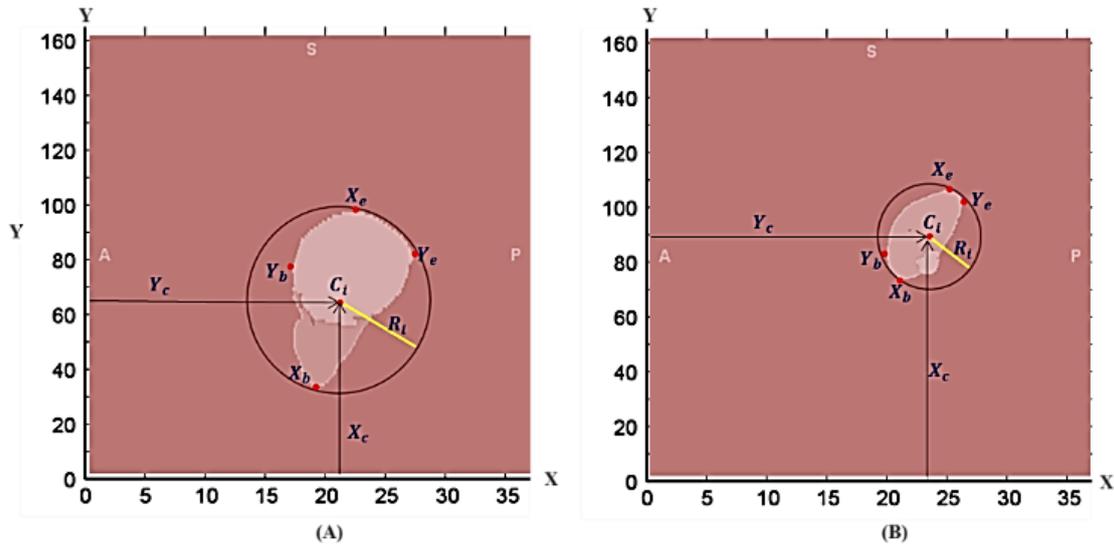


Fig. 2 Seeded Region Detection from the Ground Truth Labels

Each label-related local seeded region value $G_i(C_i, R_i)$ is extracted iteratively from G , adding their beginning and end points distance from the local centroid C_i to the arrays A_b and A_e to encircle the global seeded region. Now the global centroid value G_c is evaluated from the $\min(A_b)$ and $\max(A_e)$ values from the arrays. Finally, the global radius G_r will get calculated using the global centroid G_c , $\min(A_b)$, and $\max(A_e)$ values as shown in algorithm-1. The seeded region of the medical image is now in hand, and the next step is to apply stochastic feature mapping techniques to the seeded region for object detection and further lineation.

3.2. Stochastic Feature Mapping Neural Networks (SFMNN)

In recent years, stochastic feature modeling [24, 25] has gained popularity in the medical image object segmentation process to deal with the target object's inherent uncertainty, inconsistent measures, and ambiguous boundaries. Stochastic models are widely used in image feature extraction, segmentation, and restoration activities due to their low classification error value, uncertainty handling, and low computational complexity. In image processing, standard deterministic classification models are limited to extracting the results with exact matches, whereas stochastic models are able to extract the results from the uncertain (inconsistent) data using the features variability and probability distribution.

In this research, we used the Stochastic Feature Mapping Neural Networks (SFMNN) to localize and segment the target kidney object from the seeded regions of the MR images. Unlike the pixel-based traditional supervised segmentation models (i.e., U-Net, DCN, and FCNN), our model uses the discrete pixel group (features) mapping for localization and segmentation of the target kidney object from the seeded regions. In contrast to traditional neural networks that predict only one admissible hypothesis, our SFMNN predicts multiple allowable hypothesis predictions, making it well suited to

simultaneously predicting both kidney and tumor tissue. To accomplish this, the SFMNN gets trained using a training dataset with mask information, and then the trained model knowledge maps the input image seeded region features against spatial domains of the trained objects.

3.2.1. Training

MRI image set ' T ' with its associated manual ground truth (mask) images set ' M ' is given as training data to train the SFMNN. Each binary image $t_i \in T$ contains K pixels with values ranging from 0 to 255. Based on the metadata from M , label the regions of t_i as either 0 or 1, where $t_i = 0$ means non-target object and $t_i = 1$ is the target object. After the labeling process, the training image gets divided into a $S \times S$ grid, where each cell is a set of discrete pixel groups termed a feature, which is randomly extracted [26] from the masked regions and their surroundings of t_i to train the model with both positive and negative classification capability. It allows us to determine the object regions in both directions, which enhances the prediction capabilities of the low-contrast medical image under uncertainty. Each feature extracted from t_i is labeled into different classes $t_i(C) = \{\bar{V}_i\}^C$ and stored in vector as where class label set $C = \{c_1 \dots c_n\}$. After the feature extraction and vector class labeling processes, the redundant classes in the vectors get eliminated, and the remaining classes are grouped as positive (\bar{Q}) or negative (\bar{H}) vectors. This grouping is done based on feature logistic probabilities and correlations, which aids in overcoming overfit issues when using neural networks for classification. Now, these vector classes have to be ordered with weights $\omega = [0, 1]$ to maximize the prediction capabilities of the likelihood information [27] from testing data. After assigning the initial base weights to 0, they will be adjusted around the iterations based on the distance between the label target values (0, 1) and the feature logistic values (ϵ), using the log function with gradient ascent [28] is as shown

$$L = \log \prod_{\bar{V}_i \in T} P(\bar{V}_i = \varepsilon_i * \bar{v}_i) \text{ or } \sum_{\bar{V}_i \in T} \log P(\bar{V}_i = \varepsilon_i * \bar{v}_i)$$

Partial derivatives [37] play a vital role in assessing the border directions and their extents based on color intensity variations along the curves. To evaluate the partial derivatives in uncertain images, their underlying pixels behavior should be modeled before. The SFMNN performs stochastic simulations for each $t_i \in T$ to evaluate the partial derivatives from the feature logistic values (ε) of the image class label vector \bar{V} . The partial derivatives of the likelihood are modeled in our SFMNN using the log function L , and the weights are defined as follows:

$$\frac{\partial L}{\partial \omega_i} = \sum_{\bar{V}_i \in T} \sum_Q P(\bar{Q}_i = \bar{q}_i * \varepsilon_i | \bar{V}_i = \varepsilon_i * \bar{v}_i) * \sigma(-\bar{q}_i * \sum_{x < i} \bar{q}_i * \bar{\omega}_{i*x})$$

Across the iterations of the learning process, the labeled vector class weights are risen by the proportional average of the positive class vectors $\sigma(-\bar{q}_i * \sum_{x < i} \bar{q}_i * \bar{\omega}_{i*x})$ and gradient ascents [28]. Labeled vector classes are ordered as deterministic feature vectors \bar{D} , based on tuned weights to evaluate the local maxima of the likelihood information. This fine-grained training knowledge now applies to the seeded region containing target object localization and lineation.

3.2.2. Processing

After completing the training process, the SFMNN neural networks will map the test (seeded region) image random features $\{x_0 \dots x_n\}$ against the trained label class vectors to determine the output $\{y_0 \dots y_n\}$. The test images are divided into $S \times S$ grids for this purpose, and each grid is a set of pixels containing the spatial information known as feature space. SFMNN randomly selects the feature grids from the seeded input vector set \bar{D} with input values $\{x_0 \dots x_n\}$, where each feature space has spatial grid properties u^{th} row, v^{th} cell, weight ω , global centroid G_c and global radius G_r as $x_i = \{x_u, x_v, x_\omega, G_c, G_r\}$. SFMNN chooses a test image-related seeded vector \bar{d}_i with feature space inputs $\{x_0 \dots x_n\}$ as the main input, and each input x_i is mapped against the labeled class vector \bar{V} with the probability $P(x_i|\bar{V})$ calculated as follows.

$$P(x_i|\bar{V}) = \frac{P(x_i) * P(\bar{V})}{P(\bar{V})} \quad \text{where } 0 \leq P(x_i|\bar{V}) \leq 1$$

Similarly, the probabilities $P(x_i|\bar{V})$ are identified with all feature space inputs of \bar{d}_i , and the probabilities that are equal or greater than the threshold value (δ) are added to the intermediate positive results vector \bar{R}_p , while the probabilities that are less than the threshold (δ) are added to the intermediate negative results vector \bar{R}_N simultaneously, as follows:

$$\forall x_0 \dots x_n \in \bar{d}_i \text{ if } P(x_i|\bar{V}) \geq \delta \text{ then } \bar{R}_p.add(P(x_i|\bar{V}))$$

$$\forall x_0 \dots x_n \in \bar{d}_i, \text{ if } P(x_i|\bar{V}) < \delta \text{ then } \bar{R}_N.add(P(x_i|\bar{V}))$$

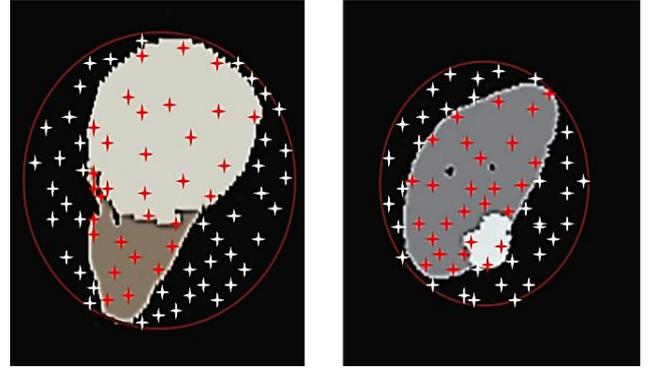


Fig. 3 Stochastic Features Mapping with seeded regions of MR images

The positive results vector probabilities \bar{R}_p are mapped against the ordered positive labels vector \bar{Q} containing feature probabilities, and the negative results vector \bar{R}_N probabilities are mapped against the ordered negative labels vector \bar{H} containing feature probabilities. Figure 3 presents the stochastic probabilities obtained from the positive (red) and negative (white) result vectors and their mapping against the seeded regions of the MR images. The intermediate results vector elements are ranked into $P_{max}, P_{mid},$ and P_{min} based on the probabilities of \bar{Q} and \bar{H} with respect to \bar{R}_p and \bar{R}_N . An element \bar{R}_p of rank P_{max} indicates that this element recorded the maximum similarity with the positive likelihood information. Likewise, an element of \bar{R}_p of rank P_{min} specifies that the element recorded the minimum similarity with the positive likelihood information. The similarity chances of the likelihood information are uncertain in the case of rank P_{min} elements of the intermediate results vector, and the variability of the features is high. Our SFMNN interprets the location-based conditions for accurate classification of the seeded region features to overcome the ambiguity in feature similarity decision-making when the rank is P_{min} . For this, the feature element $x_i \in \bar{R}_p$ or \bar{R}_N position from the centroid G_c is identified, and the distance is compared to G_r . If the feature element x_i with rank P_{min} is outside the circle area of (G_c, G_r) , it simply indicates that the x_i does not belong to the likelihood area; otherwise, the x_i may belong to the likelihood area, as confirmed by the surrounding neighbor elements. In this way, the features of the seeded region are mapped with the likelihood area of localizing the target object.

3.3. Boundary Segmentation

Soon after the target object localization process completes, the boundary segmentation [30] begins to lineate the target object from uncertain seeded image regions. For this, we used the recursive block-wise segmentation process [38], which is independent of object properties and capable of detecting boundaries with a low-intensity variation of border pixels. Instead of scanning the

whole image pixels for boundary segmentation, our block-wise segmentation model identifies the border blocks or elements from \bar{R}_p and \bar{R}_N vectors based on their

probability ranks. As shown in figure-4, it begins labeling some border blocks as boundary blocks based on their feature probability variation with neighbors.

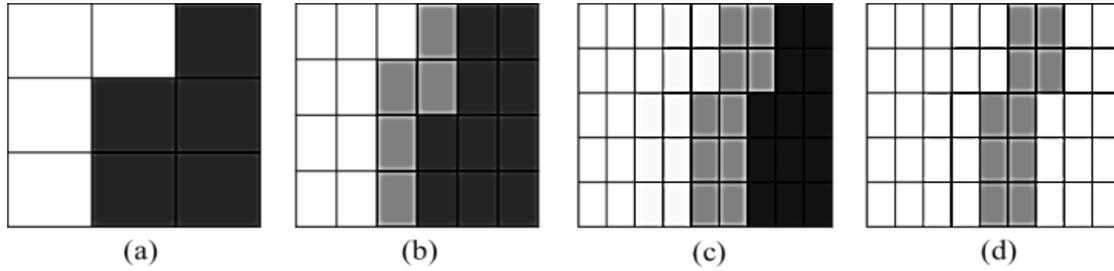


Fig. 4 Recursive block wise boundary segmentation (a) boarder blocks labeling (b) boundary blocks detection (c) adjacent boarder (gray) pixels detection (d) boarder pixels presentation

After labeling the boundary blocks from the seeded region, the pixels level scanning will execute to identify the actual pixels that are part of the boundary. The process will be done recursively with all blocks inter-connected to the borderline. Finally, adjacent blocks with inter-connected border pixels are highlighted to show the boundaries and finish the segmentation process.

4. Experiments

In this section, we discuss the results of the experiments conducted on our proposed stochastic segmentation model SFMNN in detail. A set of 30 human kidney tumor MR images from the TCGA-KIRC dataset [32] is selected with their respective ground truth labels for experiments. The images and ground labels of the dataset with size 350 x 350 pixels are further classified into the training set, and test set for model learning and results from evaluation. To execute the segmentation process, a system with hardware resources Intel i5-4590K 3.70 GHz processor, 8 GB Ram, 1 TB hard disk, and NVIDIA GTX Titan Xp GPU is utilized in experiments. The proposed stochastic segmentation model SFMNN prototype was implemented and executed using python Keras (2.4.3) and TensorFlow (2.3.0) libraries.

We chose the prominent deep convolutional architecture U-Net [8] to compare the results of our SFMNN-based segmentation. U-Net is selected as a counterpart because this is a baseline (core) project for all advanced semantic segmentation models. To uncover the differences between the two models, we used the same dataset, software, and hardware setups to train them. We used three evaluation indicators [39] linked to segmentation to determine the performance of these models: Pixels Segmentation Accuracy (PSA), Intersection over Union (IoU) with Jaccard Index [34], and Dice Coefficient (DCE) with F1 score [35]. Pixel segmentation accuracy evaluates the percentage of the pixels classified correctly from the test images over the ground truth labels. Although this statistic indicates pixel classification accuracy, it may be influenced by class imbalance difficulties when segmenting small-sized target objects. Intersection over Union (IoU) [34] is the most reliable and commonly used Jaccard index-based model designed for

evaluating the object segmentation accuracy. IoU evaluates the overlapped (intersection) area of the prediction image and ground truth images, which is divided against the union area the both, for IoU calculation as follows:

$$IoU = \frac{I_p^i \cap I_G^i}{I_p^i \cup I_G^i} = \frac{TP}{TP + FP + FN}$$

Here the I_p^i is i^{th} the prediction image, and I_G^i is the i^{th} the ground truth image, which are evaluated using the union and intersection mnemonic values. As illustrated above, the confusion matrix connected to TP, TN, FP, and FN values is used to implement the same in reality. Similarly, the Dice Coefficient (DCE) with an F1 score is used to assess segmentation accuracy in a different dimension. The twofold overlapped area is divided by the entire pixel count of the prediction and ground truth images in DCE [35]. The following is the DCE formulation with mnemonic and confusion matrix-related truth labels:

$$DCE \text{ or } DICE = \frac{2|I_p^i \cap I_G^i|}{|I_p^i| + |I_G^i|} = \frac{2TP}{2TP + FP + FN}$$

The images are reshaped into 320x320 pixel size with 3x3 kernels to make U-Net convolutions possible, and the sigmoid activation is applied with 1x1 convolutions for hidden feature extraction. Similarly, the SFMNN images are changed with the same configuration, and the stochastic networks are designed with gradient ascent and partial derivatives for object semantic segmentation that is not dependent on properties. Figure-5 presents the kidney object with tumors segmentation from MR images using the U-Net and SFMNN. All images that appeared in the first column are the original input images given for testing, the second column with their ground truth mask images, the third column with U-Net segmented kidney objects, and the fourth column presenting the proposed SFMNN segmented kidney objects.

We can observe from figure-5 that the U-Net shows efficient segmentation accuracy under the routine shape and size of kidney objects (1 and 2) with low variations. While processing the MR images (3 and 4) with irregular shapes and sizes, the U-Net is unable to segment the

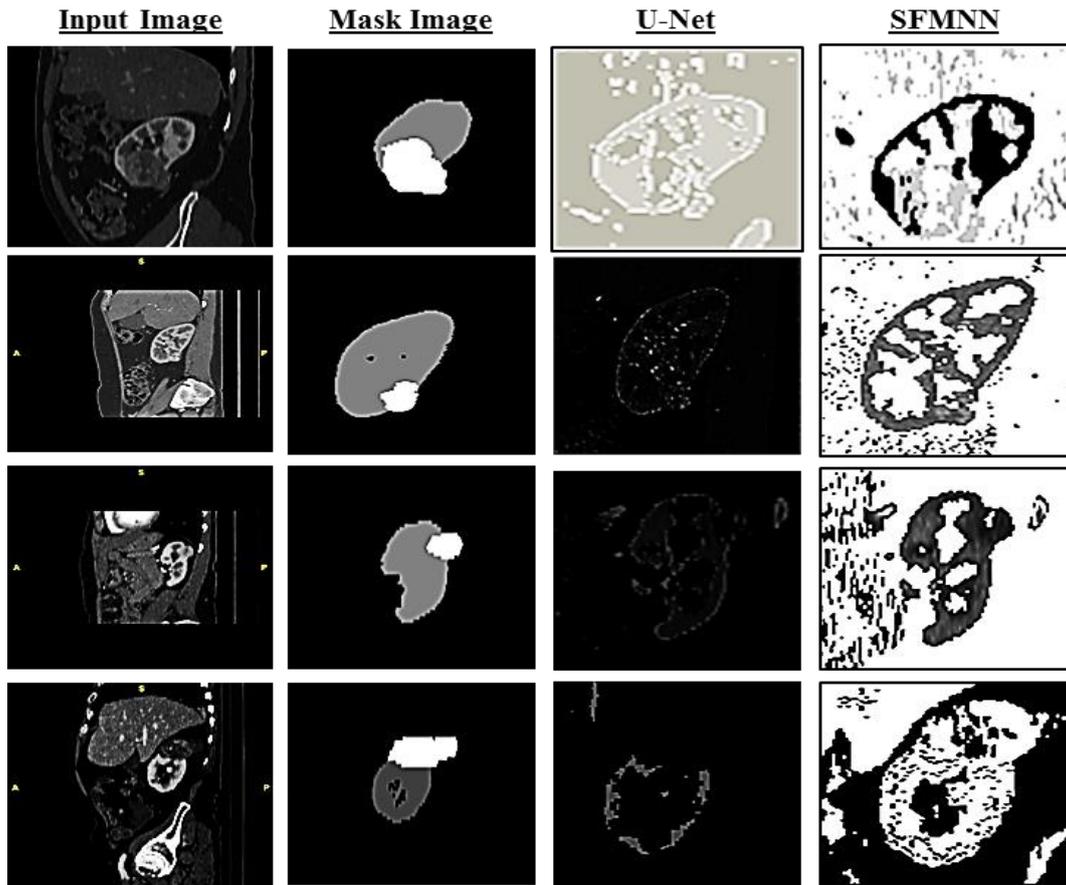


Fig. 5 MRI Kidney object segmentation results using U-Net and SFMNN

kidney objects with tumors properly. Contrary to U-Net, our SFMNN model efficiently segmented the kidney with tumors, regardless of variability in object properties like shape, size, and location. Table 1 presents the evaluated segmentation metrics on U-Net and SFMNN with respective statistics values.

Table. 1 MRI Kidney Object Segmentation Metrics Comparison

		PSA	IoU	DCE
MRI-1	U-Net	91.12%	85.11%	89.07%
	SFMNN	92.42%	86.05%	90.80%
MRI-2	U-Net	96.44%	93.08%	90.54%
	SFMNN	95.02%	92.76%	90.03%
MRI-3	U-Net	81.47%	78.07%	80.74%
	SFMNN	86.95%	85.14%	83.43%
MRI-4	U-Net	75.60%	72.74%	70.95%
	SFMNN	88.94%	87.18%	86.03%

5. Conclusion

Due to the uncertainty, irregularity, and chaotic nature of the object properties, Diseased medical image object

segmentation is a rather complex operation compared to routine objects. Traditional deep learning models provide less accuracy when segmenting diseased objects (i.e., kidneys with tumors). Inconsistent shape, volume, brightness, overlapping, and location issues made the segmentation of diseased objects more complex than routine objects. This paper proposed a stochastic feature mapping-based semantic segmentation model with hybrid technologies. Location-dependent split method, stochastic feature mapping neural networks, and recursive block-wise segmentation methods are integrated to accomplish the object segmentation from the uncertain medical images. A real-life medical dataset with kidney MR images is selected for conducting the experiments using the proposed stochastic model. Popular deep learning architecture U-Net was selected for results comparison with SFMNN. Union over Intersection (IoU) and Dice Coefficient (DCE) standard metrics are selected to evaluate the segmentation results. The analysis results show that the proposed SFMNN outperforms the U-Net when segmenting the diseased MR kidney image with tumors.

References

- [1] World Kidney Day- 12th March 2020, “2020 WKD Theme”, <https://www.Worldkidneyday.org/2020-Campaign/2020-Wkd-Theme/>
- [2] GBD Chronic Kidney Disease Collaboration, “Global, Regional, and National Burden of Chronic Kidney Disease, 1990– 2017: A Systematic Analysis for the Global Burden of Disease Study 2017,” 2020, DOI: [https://doi.org/10.1016/S0140-6736\(20\)30045-3](https://doi.org/10.1016/S0140-6736(20)30045-3)
- [3] N. M. Selby Et Al., “Magnetic Resonance Imaging Biomarkers for Chronic Kidney Disease: A Position Paper from the European Cooperation in Science and Technology Action PARENCHIMA,” *Nephrology Dialysis Transplantation*, vol. 33, no. 2, pp. 4–14, 2018.

- [4] Choi H, Jin KH, "Fast and Robust Segmentation of the Striatum Using Deep Convolutional Neural Networks," *Journal of Neuroscience Methods*, vol. 274, pp. 146-153, 2016. DOI: 10.1016/J.jneumeth.2016.10.007. PMID: 27777000.
- [5] Y. Guo, Y. Gao and D. Shen, "Deformable MR Prostate Segmentation Via Deep Feature Learning and Sparse Patch Matching," *In IEEE Transactions on Medical Imaging*, vol. 35, no. 4, pp. 1077-1089, 2016. Doi: 10.1109/TMI.2015.2508280.
- [6] Ibragimov B, Xing L, "Segmentation of Organs-At-Risks in Head and Neck CT Images Using Convolutional Neural Networks," *Medical Physics*, vol. 44, no. 2, pp. 547-557, 2017. Doi: 10.1002/Mp.12045. PMID: 28205307; PMCID: PMC5383420.
- [7] Kline TL, Korfiatis P, Edwards ME, Blais JD, Czerwiec FS, Harris PC, King BF, Torres VE, Erickson BJ, "Performance of an Artificial Multi-Observer Deep Neural Network for Fully Automated Segmentation of Polycystic Kidneys," *Journal of Digital Imaging*, vol. 30, no.4, pp. 442-448, 2017. Doi: 10.1007/S10278-017-9978-1. PMID: 28550374; PMCID: PMC5537093.
- [8] Ronneberger, Olaf & Fischer, Philipp & Brox, Thomas, "U-Net: Convolutional Networks for Biomedical Image Segmentation," 2015.
- [9] Daza, Laura & Gomez, Catalina & Arbelaez, Pablo, "Semantic Segmentation of Kidney Tumor Using Convolutional Neural Networks," 2019. 10.24926/548719.077.
- [10] N. Goceri and E. Goceri, "A Neural Network Based Kidney Segmentation From MR Images," *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, pp. 1195-1198, 2015. Doi: 10.1109/ICMLA.2015.229.
- [11] W. Thong, S. Kadoury, N. Piché, C.J. Pal, W. Thong, S. Kadoury, N. Piché, C.J. Pal, "Convolutional Networks for Kidney Segmentation in Contrast-Enhanced CT Scans," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 1163, pp. 1-6, 2016.
- [12] R. Prevost, B. Mory, R. Cuingnet, J.M. Correas, L.D. Cohen, R. Ardon, "Kidney Detection and Segmentation in Contrast-Enhanced Ultrasound 3D Images," *Abdomen and Thoracic Imaging: An Engineering & Clinical Perspective*, pp. 37-67, 2014.
- [13] B. Mory, O. Somphone, R. Prevost, R. Ardon, "Real-Time 3D Image Segmentation By User-Constrained Template Deformation," *In: Medical Image Computing and Computer Assisted Intervention*, pp. 561-568, 2012.
- [14] R. Ardon, R. Cuingnet, K. Bacchuwar, V. Auvray, "Fast Kidney Detection and Segmentation with Learned Kernel Convolution and Model Deformation In 3D Ultrasound Images," *In: International Symposium on Biomedical Imaging*, pp. 268-271, 2015.
- [15] Abubakar, Fari, "A Study of Region-Based and Contour Based Image Segmentation," *Signal & Image Processing :An International Journal*, vol. 3, pp. 15-22, 2012. 10.5121/Sipij.2012.3602.
- [16] Wahba Marian, "An Automated Modified Region Growing Technique for Prostate Segmentation" *In Trans Rectal Ultrasound Images, Master's Thesis, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada, 2008.*
- [17] Zhang L, Gao Y, Xia Y, Et Al, "Representative Discovery of Structure Cues for Weakly-Supervised Image Segmentation," *IEEE Transactions on Multimedia*, vol. 16, no. 2, pp. 470-479, 2014.
- [18] Zhu, Guopu & Zhang, Shuqun & Zeng, Qingshuang & Wang, Changhong, "Boundary-Based Image Segmentation Using Binary Level Set Method," *Optical Engineering - OPT ENG*. 46. 10.1117/1.2740762. 2007.
- [19] El-Baz, Ayman & Gimel'farb, Georgy & Suri, Jasjit, "Stochastic Modeling for Medical Image Analysis," 10.1201/B19253.
- [20] I.N. Manousakas, P.E. Undrill, G.G. Cameron, T.W. Redpath, "Split-and-Merge Segmentation of Magnetic Resonance Medical Images: Performance Evaluation and Extension to Three Dimensions, Computers and Biomedical Research," vol. 31, no. 6, pp. 393-412, 1998. ISSN 0010-4809, <https://doi.org/10.1006/Cbmr.1998.1489>.
- [21] Szenasi, Sandor, "Medical Image Segmentation with Split-and-Merge Method," *Óbuda University, John Von Neumann Faculty of Informatics*, Budapest, Hungary, 2013.
- [22] Gómez, Octavio & Gonzalez, Jesus & Morales, Eduardo, "Image Segmentation Using Automatic Seeded Region Growing and Instance-Based Learning," vol. 4756, pp. 192-201, 2007. 10.1007/978-3-540-76725-1_21.
- [23] Fan, Minjie & Lee, Thomas, "Variants of Seeded Region Growing," *IET Image Processing*, vol. 9, 2015. 10.1049/Iet-Ipr.2014.0490.
- [24] M. Welling and Y. W. Teh, "Bayesian Learning Via Stochastic Gradient Langevin Dynamics," *In Proceedings of the 28th International Conference on International Conference on Machine Learning*, Ser. ICML'11. Omnipress, pp. 681-688, 2011.
- [25] G. Wang, W. Li, M. Aertsen, J. Deprest, S. Ourselin, and T. Vercauteren, "Aleatoric Uncertainty Estimation with Test-Time Augmentation for Medical Image Segmentation with Convolutional Neural Networks," *Neurocomputing*, vol. 338, pp. 34-45, 2019.
- [26] Yuan, X.; Huang, T, "Spatial Domain-Based Nonlinear Residual Feature Extraction for Identification of Image Operations," *Applied Science*, vol. 10, pp. 5582, 2020. <https://doi.org/10.3390/App10165582>
- [27] Tang, Yichuan and Ruslan Salakhutdinov, "Learning Stochastic Feed Forward Neural Networks," *Advances in Neural Information Processing Systems-26, NIPS*, 2013.
- [28] Hinton, G. E, "Training Products of Experts By Minimizing Contrastive Divergence," *Neural Computation*, vol. 14, pp. 1771-1800, 2002.
- [29] A.Pushpa Athisaya Sakila Rani, "Reversible Data Hiding for Medical Images Using Segmentation and Prediction," *SSRG International Journal of Computer Science and Engineering*, vol. 6, no. 11, pp. 70-77, 2019. Crossref, <https://doi.org/10.14445/23488387/IJCSE-V6I11P115>.
- [30] Bhalerao, A. and R. Wilson, "Unsupervised Image Segmentation Combining Region and Boundary Estimation," *Image and Vision Computing*, vol. L9, pp. 353-368, 2001.

- [31] Nageswari P, Rajan S, Manivel K, "Medical Image Segmentation Approaches: A Survey," *SSRG International Journal of Electronics and Communication Engineering*, vol. 7, no. 7, pp. 1-3, 2020.
Crossref, <https://doi.org/10.14445/23488549/IJECE-V7I7P101>
- [32] Akin, O., Elnajjar, P., Heller, M., Jarosz, R., Erickson, B. J., Kirk, Filippini, J, "Radiology Data From the Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma [TCGA-KIRC] Collection," *The Cancer Imaging Archive*.
<http://doi.org/10.7937/K9/TCIA.2016.V6PBVTDR>
- [33] Shivani, Er. Harjeet Singh, "The Performance Analysis of Edge Detection Algorithms for Image Processing Based on Improved Canny Operator," *International Journal of Computer Trends and Technology*, vol. 68, no. 10, pp. 29-34, 2020.
Crossref, 10.14445/22312803/IJCTT-V68I10P106.
- [34] Cardenes R, De Luis-Garcia R, Bach-Cuadra M, "A Multidimensional Segmentation Evaluation for Medical Image Data," *Computer Methods and Programs in Biomedicine*, vol. 96, no. 2, pp. 108–24, 2009.
- [35] Zou KH, Warfield SK, Baharatha A, Tempny C, Kaus MR, Haker SJ, et al, "Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index," *Academic Radiology*, vol. 11, pp. 178–89, 2004.
- [36] Kumar, K. K., Ramaraj, E., & Geetha, P, "Multi-Sensor Data Fusion for an Efficient Object Tracking in Internet of Things (Iot)," *Applied Nanoscience*, pp. 1-11, 2021.
- [37] Tomasi, Carlo, "Convolution, Smoothing, and Image," 2003. Derivatives
<https://Courses.Cs.Duke.Edu/Compsci527/Fall14/Notes/Filtering.Pdf>.
- [38] Won, C.S., "A Block-Based MAP Segmentation for Image Compressions," *IEEE Transaction Circuit and Systems for Video Technology*, vol. 8, no. 5, pp. 592-601, 1998.
- [39] Taha AA, Hanbury A, "Metrics for Evaluating 3D Medical Image Segmentation: Analysis, Selection, and Tool," *BMC Med Imaging*, vol. 15, no. 29, 2015. Doi: 10.1186/S12880-015-0068-X. PMID: 26263899; PMCID: PMC4533825.