

ADVANCED DIAGNOSTIC TECHNIQUES: LEVERAGING GENERATIVE ADVERSARIAL NETWORKS FOR IMPROVED DETECTION AND ANALYSIS OF KIDNEY DISEASE

K. Sumathi and P. Karthika

Department of Computer Science and Engineering, Erode Sengunthar Engineering College, India

Abstract

It is necessary to have a precise evaluation of the morphometry of a kidney tumor in make decisions regarding treatment and diagnosis accurately. It is difficult to undertake quantitative research between kidney tumor morphology and clinical outcomes because there is a lack of data and the need to manually evaluate imaging parameters, which is a lengthy process. A conventional generative adversarial network model serves as the foundation for the STGAN approach, which is an autonomous kidney segmentation method that is proposed in this research as a solution to this problem. The primary foundation of this system is comprised of a completely convolutional generating network that is made up of densely linked blocks and a discriminator network that is equipped with multi-scale feature extraction as well. For conducting quantitative and qualitative comparisons with the STGAN methodology, the medical image segmentation networks U-Net, FCN, and SegAN are deployed. STGAN achieves greater performance in comparison to the other neural networks, the model that we have proposed demonstrates potential for improving the accuracy of CT-based kidney segmentation.

Keywords:

GAN, Diagnosis, Kidney Disease, Segmentation

1. INTRODUCTION

One of the most common forms of cancer that can occur in the urinary tract is a tumor of the kidney. Malignant tumors make up the vast bulk of these tumors, which are becoming more prevalent with each passing year [1]. When compared to the nearly 400,000 persons who passed away from kidney cancer in 2002, the number of people who passed away in 2018 was over 175,000 [2]. In the year 2002, there were a total of 208,000 newly diagnosed cases of kidney cancer and 102,000 deaths. There is a significant increase in the rate between the ages of 60 and 70, and it is higher in industrialized countries than it is in developing countries. A spiral CT scan is frequently used for clinical diagnosis of lesions that occupy renal space because of the short amount of time it takes to complete the scan. CT scans are the most common method for diagnosing renal cell carcinoma. This is accomplished by monitoring the patient's kidneys, which are greatly enlarged. A discernible increase in the thickness of the patient's soft tissue occurs concurrently with the growth and outward expansion of the tumor. What function does adipose tissue serve in the body is not yet fully understood. It is likely that kidney segmentation requires extra refinement in be as precise as feasible. This is since radiation therapy for kidney cancer can have both beneficial and harmful consequences, depending on the quality of the segmentation results.

Renal CT data segmentation was a laborious and time-consuming operation that required the presence of an observer who had received training. Numerous techniques for the

segmentation of renal CT images have been developed as a result. Some research have offered techniques to raise the amount of segmentation automation by integrating existing human anatomy knowledge (for example, the relative location or morphological aspects of the kidney) with the strategy of locating the kidneys first and then segmenting them [3]. This has been done in achieve the goal of increasing the level of segmentation automation. Using an improved area growth algorithm that takes into consideration multi-scale morphology, a labeling method that utilized the spine as a marker, and the connected region labeling methodology that depends on image intensity, In [4] were able to successfully segment the kidney. This was accomplished by using these three methods. The spine was also utilized by Abirami et al. as a marker in determine the location of the kidney through the utilization of the linked region labeling technique. Based on this information, the renal region was removed by the application of the adaptive region expansion approach [5]. Following the removal of the spine through the application of a threshold and many other image modifications, the target region and mark are subsequently collected. Kidney segmentation is accomplished through the utilization of the marking control watershed technique and the image gradient [6]. A random forest technique was developed by the authors of [7] in perform kidney segmentation and tissue categorization. This strategy was formed on the basis of previous information regarding the structure of the kidney as well as high-order feature information. It was hypothesized by Song et al. that a two-stage technique that progresses from coarse to fine would be used to segment the kidney from a CT scan. For estimating the approximate range of the kidney, the fuzzy c-means clustering method, which is founded on spatial information, is utilized. Then, the GrowCut algorithm is utilized in a more refined manner in accomplish the task of fine segmentation [8]. Since a single segmentation approach is not capable of meeting the requirements of CT image kidney segmentation, it is required to implement a complementary strategy that involves the integration of many segmentation methods. In extract renal lesions from CT images in a manner that is both efficient and accurate, Kaur et al. introduced a hybrid segmentation method that combines distance regularized level set evolution (DRLSE) with fuzzy c-means clustering that is based on spatial information. The first shape of the target is created by the utilization of fuzzy c-means clustering, which takes advantage of the spatial information and hesitation displayed in the image. After that, a technique known as distance regularization level set methodology is utilized in segment the focus [9]. Also, it were able to separate the kidney and renal cortex in an improved CT image [10]. This was accomplished with the use of the Intelligent Scissors algorithm, the image cutting algorithm, and various other active appearance model methods. Kim demonstrated a multimap approach that considers limits on the image's intensity and form. Possible applications include the segmentation of the kidneys based on CT scans, the prediction of

renal function based on changes in renal volume, and the provision of support with treatment plans for patients who have undergone partial nephrectomy [11].

The STGAN approach is presented as the architecture for kidney area segmentation [12]. This method takes its cues from the SegAN model. It is possible for the GAN to capture data distribution in part, and it possesses adequate generation capabilities. Through our research, we were able to increase the accuracy of kidney segmentation by refining the objective function and adjusting the structure of the GAN network. The following are some of the distinctive contributions that this study makes:

In the generator network, the generation of the segmentation result images is accomplished through the utilization of an end-to-end complete convolutional network that is modeled after the U-Net design.

By utilizing densely connected blocks between all the front layers and the posterior layers to construct dense connections, we were able to reduce the number of parameters, increase feature propagation, and alleviate the gradient vanishing problem. As a result of the fact that they link the features in the channel dimension, it is feasible to reuse features.

Discriminator networks are trained with multi-scale features to decrease the chance of overfitting and to strengthen the dependability of sparsity guarantees. Currently, the L1 parametric form of the mean absolute error is incorporated into the objective function on a consistent basis.

2. PROPOSED METHOD

A generator network (G) and a discriminator network (D) are the two components that make up our segmentation approach, which is known as STGAN. Specifically, the generator is meant to gather up the data to provide images of the kidney region that are comparable to the actual data distribution. The discriminator generates outputs that are discriminating because it differentiates between the images generated by the generator and the images that represent the ground truth. As a result of returning the discriminating results to the generator, the two networks come into conflict with one another.

2.1 GENERATOR

Generator G refers to a network that divides data into segments from the beginning to the finish. The encoder-decoder structure of the fully convolutional network U-Net serves as the basis for the core design of G. This network employs both up-sampling and down-sampling strategies in its operations. In perform the down-sampling technique, there are three densely linked blocks, three maximum pooling layers, and a convolutional layer that is comprised of three \times three convolution kernels. Utilizing a 1×1 convolution kernel, three blocks that are firmly related to one another, and three deconvolution layers are all components of the up-sampling technique. The network takes on the characteristics of an autoencoder when skip connections are added between the two nodes. This arrangement enables the downsampling of image features at different sizes and the upsampling of the same size as the input image, which ultimately results in the re-creation of the original image at the same size as

the output. Additionally, it can acquire familiarity with representative examples.

The incorporation of a dense block structure into the generator network not only increases network performance but also overcomes the problem of gradient disappearance. This is accomplished by the utilization of channel dimension feature connections, which enables the reuse of features. This structure ensures that each of the front layers is closely connected to the layer that is located behind it. This demonstrates a four-layer structure, which is comparable to the architecture that provides the foundation for DenseNet. The batch normalization, a rectified linear unit, and a 3×3 convolution kernel are all components that are incorporated into each layer.

2.2 DIFFERENTIATOR

Each of the six layers that make up the multi-dimensional feature extraction network is referred to as Discriminator D. Convolutional, Binary Neural Network, and leaky ReLU layers are some of the activation techniques that are incorporated into each layer. The convolutional kernels possess dimensions of 7×7 , 5×5 , 4×4 , and 3×3 , respectively.

3. STGAN

Through the encoding and decoding levels, the generator is the one who is responsible for providing segmentation masks. It is the responsibility of the discriminator to determine and evaluate the legitimacy of a segmentation mask that has specific characteristics. As a result, the adversarial network is instructed to differentiate between genuine and fabricated signals in motivate the generator to produce segmentation masks that are as comparable as they can be.

When calculating the mean absolute error (MAE) loss, which is also referred to as the L1 parametric loss, the average of all absolute discrepancies that exist between the values that were anticipated and those that were observed is computed. In situations in which the distribution of the target variable contains outliers, the MAE loss becomes more resilient. For making the model more educated about the target data, the L1 regularization technique is utilized; the weights are distributed according to a zero-mean Laplace distribution. The complexity of the neural network is effectively reduced and overfitting is prevented by the use of L1 regularization. This is accomplished by making the weights of the neural network as tiny and convergent to zero as possible.

4. DATASET

Kits19, a publicly available dataset, was provided by the Kidney Tumor Segmentation Challenge in 2019, and the model makes use of this dataset. All patients who underwent a kidney tumor removal procedure at the University of Minnesota Medical Center between the years of 2010 and 2018 and underwent either a partial or radical nephrectomy that was performed were eligible to be included in this database. A total of three hundred instances were chosen at random from the list of instances. A group of medical students, under the direction of Dr. Christopher Weight, who served as the clinical chair, were responsible for providing the manual segmentation labels. For training the neural network,

data was collected from 150 participants who were chosen at random. The data that was used to validate the network was collected from sixty subjects, and the data that was used to test the network was collected from the remaining ninety individuals. Once the preprocessing phase has been completed, a binning analysis is performed using $16 \times 128 \times 128$ slices, without any pixel value normalization being performed. Nevertheless, in enhance the data in its whole, overlapping slices measuring $12 \times 32 \times 32$ are utilized. The original scanned image had a resolution of 512 dimensions by 512 dimensions. During training, it is necessary to skip over blank facets and create overlapping facets in accomplish the goal of making an accurate prediction.

5. EVALUATION

During the process of programming the segmentation models that were given in this research, Python, TensorFlow, and Keras were utilized. Every single trial is carried out by a single workstation that is outfitted with a graphics processing unit (GPU) from Nvidia called GeForce RTX 3080. This GPU offers a learning rate of 0.0001, a weight decay of 0.0001, and a momentum of 0.9.

To get started, the weights of STGAN are modified by utilizing the training dataset, and the validation dataset is used to determine the appropriate hyper-parameters. Once it is complete, the performance of STGAN is assessed by making use of the Kits19 and Kits21 test datasets. The STGAN algorithm is one of the better generative adversarial segmentation models that are currently available. The generator makes use of a dense block in guarantee that the segmentation results are created in a manner that is very similar to the information that is present. In achieve more accurate results, it is sent into the discriminator in conjunction with the labels. U-Net, FCN, SegAN, and STGAN are the four network topologies that are analyzed in determine whether or not the improved generative adversarial network architecture has the potential to increase the segmentation performance. It is possible to obtain the final metrics by first calculating each two-dimensional slice and then averaging the results of those calculations.

5.1 DISCUSSION

During a computed tomography (CT) scan of the abdomen, the procedure of renal segmentation involves completely segmenting the margins of the kidneys. After completing the segmentation technique, the results indicate that more sophisticated methods are required for the subsequent analysis of certain images. We constructed a STGAN segmentation model and put it through its paces in comparison to other approaches during this research. The assumption that the entire network is hostile is the foundation around which the paradigm is built. A CT scan is fed into the generator, which then generates segmentation results that are so convincing that they almost look like actual labels. Concurrently with the labels, these results are presented to the discriminator for decision-making purposes. It is possible to train the discriminator network to work against the generator until it reaches a state of equilibrium. The discriminator network receives data, processes it, and then ultimately returns a binary output.

Table.1. Accuracy

Test data	CNN	CRNN	Proposed
30	0.8098	0.8301	0.8908
45	0.7896	0.8098	0.8706
60	0.7592	0.7896	0.8503
75	0.7468	0.7775	0.8389
90	0.7263	0.7570	0.8184
105	0.7059	0.7366	0.7979
120	0.6854	0.7161	0.7775
135	0.6650	0.6956	0.7570
150	0.6445	0.6752	0.7366
165	0.6240	0.6547	0.7161
180	0.6036	0.6343	0.6956

Table.2. Precision

Test data	CNN	CRNN	Proposed
30	0.8605	0.8200	0.9111
45	0.8402	0.8098	0.8908
60	0.8200	0.7997	0.8706
75	0.8082	0.7877	0.8593
90	0.7877	0.7673	0.8389
105	0.7673	0.7468	0.8184
120	0.7468	0.7263	0.7979
135	0.7263	0.7059	0.7775
150	0.7059	0.6854	0.7570
165	0.6854	0.6650	0.7366
180	0.6650	0.6445	0.7161

Table.3. Recall

Test data	CNN	CRNN	Proposed
30	0.7896	0.8605	0.8706
45	0.7592	0.8301	0.8503
60	0.7289	0.8098	0.8301
75	0.7161	0.7979	0.8184
90	0.6956	0.7775	0.7979
105	0.6752	0.7570	0.7775
120	0.6547	0.7366	0.7570
135	0.6343	0.7161	0.7366
150	0.6138	0.6956	0.7161
165	0.5933	0.6752	0.6956
180	0.5729	0.6547	0.6752

Table.4. FPR

Test data	CNN	CRNN	Proposed
30	0.1215	0.1012	0.0810
45	0.1417	0.1215	0.1012
60	0.1620	0.1417	0.1215
75	0.1841	0.1637	0.1432
90	0.2046	0.1841	0.1637
105	0.2251	0.2046	0.1841

120	0.2455	0.2251	0.2046
135	0.2660	0.2455	0.2251
150	0.2864	0.2660	0.2455
165	0.3069	0.2864	0.2660
180	0.3274	0.3069	0.2864

This research has several shortcomings that need to be addressed. Due to the small amount of data and the fact that the study utilized a dataset that was accessible to the public, the method remains restricted in real-world scenarios, even though it is quite accurate. During our future work, we intend to test the network on a variety of datasets, including clinical data, and then adjust it in accordance with the results of those tests. In addition to that, we will try to use the model for MRI image segmentation.

6. CONCLUSION

An innovative form of deep neural network, known as STGAN, is presented in this research. The gist of this research is summarized in the following paragraphs. Within the context of this model, the discriminator network is developed through the utilization of a multi-scale convolutional network, whilst the generator network is developed through the utilization of densely connected blocks and an encoder-decoder structure. In the second piece of guidance, it is recommended to optimize the goal function and enhance the segmentation performance by utilizing the loss functions that correspond to the two networks. The STGAN algorithm is superior to its competitors when it comes to the segmentation of medical images. The STGAN model emerges as the best CT segmentation method for kidneys when its quantitative and qualitative aspects are compared. This makes it a suitable choice for inclusion in applications that are used in the real world.

REFERENCES

[1] J. Xie and C. Wang, "Using Support Vector Machines with a Novel Hybrid Feature Selection Method for Diagnosis of Erythematous-Squamous Diseases", *Expert Systems with Applications*, Vol. 38, No. 5, pp. 5809-5815, 2011.

[2] X. Yanwei, J. Wang, Z. Zhao and Y. Gao, "Combination Data Mining Models with New Medical Data to Predict Outcome of Coronary Heart Disease", *Proceedings of International Conference on Convergence Information Technology*, pp. 868-872, 2007.

[3] Deepika Kancherla, Jyostna Devi Bodapati and N. Veeranjaneyulu, "Effect of Different Kernels on the Performance of an SVM Based Classification", *International Journal of Recent Technology and Engineering*, Vol. 7, No. 4, pp. 1-6, 2019.

[4] Jane Lee Jia Jing and Thiagarajan Madheswaran, "Hand Sanitizers: A Review on Formulation Aspects, Adverse Effects, and Regulations", *International Journal of Environmental Research and Public Health*, Vol. 17, pp. 3326-3335, 2020.

[5] R.D. Aruna and B. Debtera, "An Enhancement on Convolutional Artificial Intelligent Based Diagnosis for Skin Disease Using Nanotechnology Sensors", *Computational Intelligence and Neuroscience*, Vol. 2022, pp. 1-8, 2022.

[6] P. Ghosh, F.J.M. Shamrat and A.A. Khan, "Optimization of Prediction Method of Chronic Kidney Disease using Machine Learning Algorithm", *Proceedings of International Joint Symposium on Artificial Intelligence and Natural Language Processing*, pp. 1-6, 2020.

[7] S. Buyrukoglu, "Early Detection of Alzheimer's Disease using Data Mining: Comparison of Ensemble Feature Selection Approaches", *Konya Journal of Engineering Sciences*, Vol. 9, No. 1, pp. 50-61, 2021.

[8] G. Battineni and F. Amenta, "Improved Alzheimer's Disease Detection by MRI using Multimodal Machine Learning Algorithms", *Diagnostics*, Vol. 11, No. 11, pp. 2103-2109, 2021.

[9] S. Yamini and K.P. Rama Prabha, "A Data Mining with Big Data Disease Prediction", *International Research Journal of Engineering and Technology*, Vol. 5, No. 4, pp. 829-832, 2018.

[10] C. Szegedy and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision", *Proceedings of International Conference on Recent Trends in Machine and Deep Learning*, pp. 156-163, 2016.

[11] Shraddha Subhash Shirsath and Pro. Shubhangi Patil, "Disease Prediction using Machine Learning over Big Data", *International Journal of Innovative Research in Science, Engineering and Technology*, Vol. 7, No. 6, pp. 6752-6757, 2018.

[12] B. Subramanian, V. Saravanan and S. Hariprasath, "Diabetic Retinopathy-Feature Extraction and Classification using Adaptive Super Pixel Algorithm", *International Journal of Engineering and Advanced Technology*, Vol. 9, pp. 618-627, 2019.