Deep learning Approach for Pancreas segmentation using U-Net Architecture

Saraswathi H S
Dept of CSE JIT,
Davangere

Ganesh B S
Dept of CSE JIT,
Davangere

Ganesh G S
Dept of CSE JIT,
Davangere

Manoj k v
Dept of CSE JIT,
Davangere

Madhusudhan H M
Dept of CSE JIT,
Davangere

ABSTRACT

The pancreas is one of the most difficult organs to segment due to its unique shape, position, and size in each individual. With the advancement of machine learning, multiple deep learning algorithms are being used to partition the pancreas among abdominal organs. The U-Net model, one of the convolutional neural networks (CNN) models, is used in this study to segment the pancreas. The findings of pancreatic segmentation done on the pancreatic CT data set collected from The Cancer Imaging Archive (TCIA) database, which contains computed tomography images of 82 patients, are detailed. The results show that the Dice similarity coefficient and the Jaccard similarity coefficient are 0.78 and 0.66, respectively.

INTRODUCTION

The pancreas, one of the abdominal organs, is extremely difficult to segment compared to other fixed organs such as the liver and spleen, because they have variable shapes, sizes, placements, and sizes in humans. Magnetic Resonance Imaging (MRI) or Computerised Tomography (CT) can be used to do advanced imaging of pancreatic disorders.

It is performed by medical imaging methods such as tomography (CT).

When compared to traditional neural network methods, deep learning methods produce better classification results in large data sets. When the studies in the literature are examined, recent studies on deep learning show that progress has been made in pancreas segmentation. In 2015, Ronneberger et al. [1] established the U-Net model for the segmentation of biological pictures, which was one of the most significant steps in this field.

Farag et al.'s study is one of the most current studies on pancreas segmentation in the literature. Farag et al. conducted the study [2]. They obtained the CT images they used as maps in this work [2], which consisted of two parts: a superpixel map and a random forest probability map. In their study, they reached the most astonishing result of 70% dice rate with the cascade structure in the KSA structure utilising four distinct ways. Gibson and co. [3] used a Dense V-Network model that does not require registration to separate several organs in the abdomen area on CT scans acquired from 90 patients. Roth et al. [4] performed three-dimensional segmentation of abdominal organs using a 3D U-Net model with two stages. In the first portion of their model, they roughly segmented the pancreas and used the segmentation maps they produced as the input picture for the second part. In this manner, they attained a Dice Similarity Coefficient of 82%. Those named Man et al. [5] used the Deep Q Network method to find the position of the pancreas in three dimensions, axial, coronal, and sagittal, and the U-Net model to segment it. They achieved an 86% Dice rate by turning three separate maps into a single result using the Majority Voting approach. Yang and coThey used four-fold cross-validation. They attained an 87% Dice similarity rate in this study, which is one of the finest achievements in this discipline.

In this investigation, the area containing the pancreas was determined using Pancreas CT images [7] acquired from The Cancer Imaging Archive (TCIA) database, and the area of interest was defined and the images were cropped. The data augmentation method was then used, which included the use of hyperparameters such as zoom, colour change, and rotation. The pictures were trained using the U-Net model developed as a result of this procedure, and segmentation maps of the pancreas organ were acquired as a result. As calculation criteria, the dice similarity coefficient and the Jaccard similarity coefficient were utilised, with values of 78% and 66%, respectively

MATERIAL AND METHOD

A. Data Set

The Pancreas CT dataset from the TCIA database was used in this investigation [7]. This dataset contains 82 contrast enhanced CT images from 53 men and 27 women aged 18 to 76 years received from the National Institutes of Health (NIH) Clinical Centre in the United States. There are a total of 19328 slices in these CT pictures. As a result, the U-Net network was trained with 14738 slices and tested with 4590 slices. In our study, 62 of the 82 data points in the dataset were set aside for training and the remaining 20 for testing the network.

B. Preprocessing

The region of interest was selected by manually recognising the area containing the pancreas and cropping these areas in all CT images with dimensions of 512x512 allotted for training and testing. All photos were scaled to 256x256 using this method. As a result, the network model that is being trained is precluded from constructing a map by extracting features from areas where there is no pancreatic. Because this process involves cropping the same regions in all images, it can be automated for a different test image. Figure 1 depicts the extraction of the area of interest from the original image.

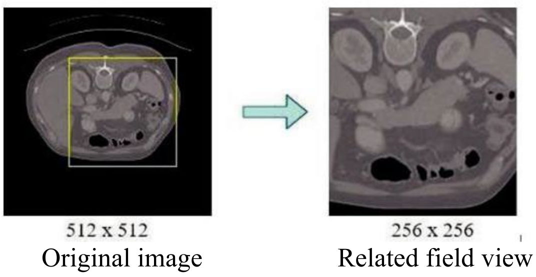


Fig. 1. ROI image extraction from original image

Overfitting occurs in artificial neural networks when the system successfully learns the data used in training but produces poor results after testing. Using the data augmentation method, the number of images was enhanced by mirroring, zooming in, zooming out, panning, and rotating parameters. In this method, the problem of overfitting, which is critical in deep learning, is avoided. Figure 2 shows some photos obtained through data augmentation. Furthermore, failure to prevent overfitting is a problem method was also used. This method is based on the notion of forgetting the deep learning network layers' learned values at a set rate. This strategy was implemented in this work by inserting random forgetting layers between the layers of the U-Net model.

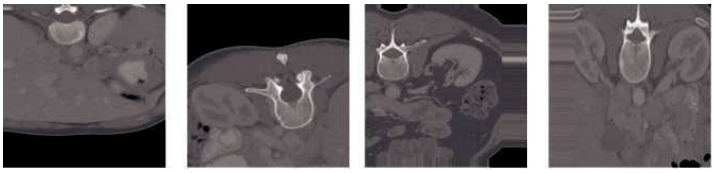


Fig. 2. 20% shift images in the data augmentation method,

(a) upward shifted image,
(b) downward shifted image,
(c) image shifted to the left,
(d) image shifted to the right

C. Model

The photos used as input to the model in KSA Machines learn the properties via convolution and orthogonal layers. The learned features are then sent to an artificial neural network model with fully linked layers for classification. The machine-learned features in the U-Net model are obtained as a segmentation map in the model's last layer without being fed into any fully connected artificial neural network. The U-Net architecture described in Figure 3 is used in this study, together with the Tensorflow and Keras libraries [1].

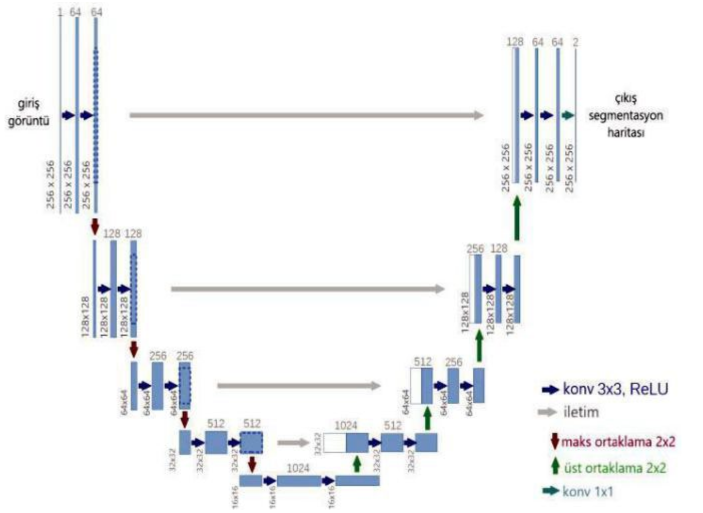


Fig. 3. U-Net Architecture Used

This architecture is divided into two parts: contraction and expansion. The orthogonalisation process is performed after using the convolution layer twice in the contraction section, while the convolution layer is used twice after the upsampling process in the expansion part. In this manner, the information about the object searched in the image is learned more effectively. While performing these operations, the object information containing its location is reduced while the findings of what the object is increased. To avoid this, the feature maps learned during the contraction phase are passed to the expansion phase unchanged. Location information is also saved in this manner. This is where U-Net differs from traditional KSA.

D. Computational Criteria

The Dice similarity coefficient and the Jaccard similarity coefficient were used to calculate the efficiency of the acquired results. A complexity matrix was generated as a result of the trained model, and the formulas in (1) and (2) were utilised for the Dice and Jaccard similarity ratios, respectively.

In Equations (1) and (2), DP represents true positive values, FC fake positive values, and FN false negative values. The complexity matrix is used to calculate these values. In these two equations, the genuine negative (DN) value from the complexity matrix is not employed.

The Structural Similarity Index (SSI) was utilised to assess the picture quality of the pancreas segmentation results acquired in the study, in addition to the Dice and Jaccard similarity coefficients. The SLI, which is given in Equation (3) and assesses the similarity of two pictures, is made up of three parameters: brightness term, contrast term, and structure term.

It corresponds to the terms for brightness, contrast, and structure in equation (3). These terms relate to (4), (5), and It is calculated using the methods in (6).

III. FINDINGS AND DISCUSSION

In the study, the areas of interest in the abdomen CT images were initially identified, and then cropping was done to each of them. By using the data augmentation method, the clipped photos were resized to 256x256 and a new data set for the U-Net model was formed. The study's flowchart is depicted in Figure 4. In this study, a 6 GB NVIDIA GEFORCE GTX 1660 Ti graphics card was used.

The study's U-Net model hyperparameters and data augmentation are optimised.The model was run with different parameter values to determine the state.

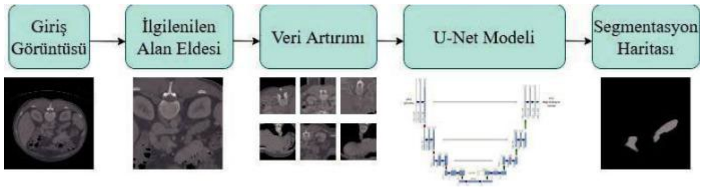


Fig. 4. Flow chart

Hyperparameters such as chunk size, number of rounds, number of iterations, forgetting rate, and data augmentation rate were used to train the U-Net model. The chunk size is the parameter that indicates how many training photos will be taken in a single iteration of the model's training. Because of the video card utilised in this investigation, the maximum number of chunks possible was 7. The number of iterations determines how many times the model's algorithm parameters will be updated in a single round. The number of rounds indicates how many times all of these processes will be carried out. The forgetting rate is the pace at which the layers in the network topology are reset. The shifting and rotation of data are referred to as data augmentation rates.

Hyperparameters such as chunk size, number of rounds, number of iterations, forgetting rate, and data augmentation rate were used to train the U-Net model. The chunk size is the parameter that indicates how many training photos will be taken in a single iteration of the model's training. Because of the video card utilised in this investigation, the maximum number of chunks possible was 7. The number of iterations determines how many times the model's algorithm parameters will be updated in a single round. The number of rounds indicates how many times all of these processes will be carried out. The forgetting rate is the pace at which the layers in the network topology are reset. The shifting and rotation of data are referred to as data augmentation rates.

In our work, we initially attempted to investigate the effect of the number of pieces on pancreas segmentation findings. As a result, the number of rounds and iterations, which are the hyperparameters utilised in model training, are kept constant at 100. Furthermore, a random rate of 20% was used for shifting, rotation, and zooming for data augmentation. The number of chunks was gradually raised using these settings, and the influence on training time and pancreatic segmentation outcome was noted; the findings are shown in Table I.

TABLE I. NUMBER OF STACKS ON THE RESULT OF PANCREAS SEGMENTATION IMPACT

Bulk Tour Training Dice Jaccard YBİ

Number Number Duration (%) (%) (%)

Table:

Stack Number	Type Number	Education Time	Dice(%)	Jaccard(%)	YBI(%)
1	100	23 min	13	7	82
2	100	38 min	21	12	78
4	100	1 s 1 min	61	46	93
6	100	1 s 23 min	67	53	94
7	100	1 s 31 min	67	53	94

When Table I is examined, it can be seen that as the number of chunks is raised while maintaining the data augmentation and model hyperparameters same, the Dice, Jaccard, and YBI ratios increase as a result of segmentation.

The number of chunks was reduced to 7 and the number of iterations to 1000 in the second stage of the study, as were the forgetting rates and hyperparameters of the data augmentation method, and the effects of these parameter changes on the pancreas segmentation maps were investigated; some of the segmentation results obtained are shown in Table II.

TABLE II. OBTAINED BY CHANGING THE HYPERPARAMETERS

PANCREAS SEGMENTATION RESULTS

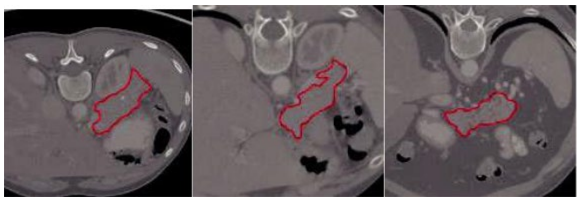
Table

Unutturma Oranı / Veri Artırımı Oranları / Uygulanan Katman Sayısı	Type Number	Education Time	Dice (%)	Jaccard (%)	YBİ (%)
%20/%10/2	100	15 s 9 min	67	54	97
%20/%10/5	10	1 s 20 min	65	51	95
%50/%20/2	22	3 s 34 min	73	60	96
%50/%30/2	36	6 s 1 min	72	59	95
%50/%20/3	45	7 s 16 min	78	66	97

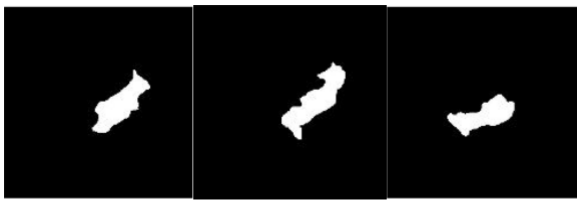
*: The number of rounds varies since training stops when the network parameters in the U-Net model are not updated three times in a row.

As seen in Table II, increasing the forgetting rate raised the Dice rate, but increasing the number of layers applied decreased the Dice rate. Furthermore, when the data augmentation rate was set to 20%, better results were obtained.

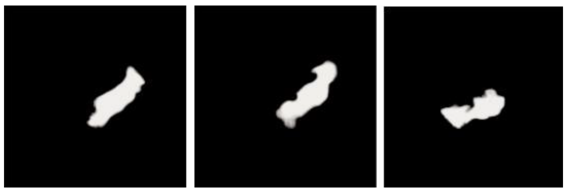
Figure 5 shows a handful of the photographs with the greatest Dice ratios, as well as the true and original images. These photos' dice ratios are shown below the images.



(a)



(b)



(c)

Zar = %93 Zar = %94 Zar = %95 Jaccard = %87 Jaccard = %89 Jaccard = %90

Fig. 5. Some of the segmentation maps obtained, (a) images of the areas of interest, (b) actual images, (c) obtained images.

Table III compares the maximum Dice coefficient obtained in this investigation to the results obtained in other pancreatic segmentation studies in the literature.

As shown in Table III, when compared to other research that used different deep learning algorithms, this study achieved a result that was near to the literature in terms of Jaccard ratio and a significant result in terms of Dice ratio.

TABLE III. COMPARISON OF THE STUDY WITH THE LITERATURE.

Table

Study	Network Model,Algorithm	Data Set Number	Dice (%)	Jaccard (%)
[9]	Superpixel, KSA	82	68	-
[2]	Superpixel, Probability Map KSA	80	68	57
[4]	3 D U-Net	150	82	-
[3]	Dense V Network	90	75	-
[8]	Holisitic-Nested Neural Network	82	81	68
[5]	Deep Q Network, U-Net	82	86	-
[6]	Cascade Neural Network	82	87	-
This Work	U-Net	82	78	66

References

1] O. Ronneberger, P. Fischer, and T. Brox (2015). Convolutional Networks for Biomedical Image Segmentation (U-Net). 243-241 in International Conference on Medical Image Computing and Computer-Assisted Intervention.

[2] Farag, A., Lu, L., Roth, H. R., Liu, J., Turkbey, E. & Summers, R. M.(2017). Using Cascaded Superpixels and (Deep) Image Patch Labelling, a Bottom-Up Approach to Pancreas Segmentation was developed. IEEE Transactions on Image Processing, vol. 26, no. 1, pp. 386-399.

[3] Gibson, E., Giganti, F., Hu, Y., Bonmati, E.,Bandula, S., Gurusamy, K., Davidson, B., Pereira, S. P., Clarkson, M. J. & Barratt, D. C

IV. CONCLUSIONS

As a result of the research, the pancreas organ was segmented using the U-Net model, one of the deep learning methods. When compared to pancreas segmentation studies using deep learning methods in the literature, the study found that a good result could be obtained without using a complex model structure or any post-processing. In future investigations, it is hoped that upgrading the U-Net model structure may yield better outcomes in pancreatic segmentation.

(2018).

[4] Roth, H. R., Oda, H., Zhou, X., Shimizu, N., Yang, Y., Hayashi, Y., Oda, M., Fujiwara, M., Misawa, K. & Mori, K. (2018). An application of cascaded 3D fully convolutional networks for medical image segmentation. Computerized Medical Imaging and Graphics, 66, 90-99.

[5] Man, Y., Huang, Y., Feng, J., Li, X. and Wu F., (2019), Deep Q Learning Driven CT Pancreas Segmentation With Geometry-Aware U-Net, IEEE Transactions on Medical Imaging, 36 (8), 1971-1980.

[6] Yang, Z., Zhang, L., Zhang, M., Feng, J., Wu, Z., Ren, F. and Lv, Y.,(2019), Pancreas Segmentation in Abdominal CT scans using Inter-/Intra-Slice Contextual Information with a Cascade Neural Network, 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany.

[7] Roth, H. R., Farag, A., Turkbey, E. B., Lu, L., Liu, J. & Summers, R. M. (2016). Data From Pancreas-CT. The Cancer Imaging Archive. http://doi.org/10.7937/K9/TCIA.2016.tNB1kqBU.