Medical Image Segmentation Using Machine Learning

Dr.Amit Gupta[1], Bonam Raghavaiah[2], Shashi Kant Dargar[3], Abha Dargar[4]

Professor, NEC, Narasaraopet, AP
Associate Professor, SVIET, Nandamuru,, AP
Associate Professor, KARE, Tamilnadu
Assistant Professor, KARE, Tamilnadu

Abstract—- Since the last decade, most medical image segmentation applications have relied on fully convolutional neural networks (CNNs) with contracting and expanding pathways. As an element of CNNs, encoders must learn both global and native features and contextual representations in semantic output prediction to use by decoders. Because the convolutional layers are so provincial in CNNs, they're unable to be told long-range spatial relationships. For volumetric (3D) medical picture segmentation, we take inspiration from NLP's recent success in long-range sequence learning by rethinking the matter as a sequence-to-sequence prediction challenge.

To search out sequence representations of the input volume and successfully capture global multiscale information, we develop a unique architecture, called CNET Transformers (CNETS), which uses a transformer because the encoder and follows the successful "U-shaped" network design for the encoder and decoder.

The last word semantic segmentation output is computed via skip connections between the encoder and decoder at different resolutions. Multi Atlas Labeling Beyond The Cranial Vault dataset (BTCV) and Medical Segmentation Decathlon (MSD) datasets are accustomed test our method's performance on various multi-organ segmentation tasks. On the BTCV leaderboard, our benchmarks show new state-of-the-art performance.

Keywords:Key Words: Fully Convolutional Neural Networks, Natural Language Processing, Beyond The Cranial Vault, Medical Segmentation Decathlon

I. Introduction:

Medical image analysis relies heavily on picture segmentation because the initial stage within the study of anatomical features. "U-shape" encoder-decoder architectures just like the FCNN have produced the only ends up in several medical semantic segmentation tasks since deep learning came into play in 2015.

The encoder gradually down samples the extracted features to seek out out global context representations, while the decoder up-samples these representations to the input resolution for pixel/voxel-wise semantic prediction. When down sampling, spatial information is lost. Therefore skip connections blend the encoder and decoder output at separate keys, with the recovery of spatial data. A failure to capture information at several scales ends up in inaccurate segmentation of structures of varying sizes and forms (e.g., brain lesions with different sizes). Atrous convolutional layers are utilized in several attempts to increase the receptive fields. Localization of receptive fields in convolutional layers indeed means they'll only study relatively limited areas.

Non-local modeling may possibly be enhanced by combining self-attention modules with convolutional layers, per one theory. Some of the best language Processing (NLP) results could even be achieved using transformer-based models. The self-attention mechanism of transformers makes it possible to stress the foremost significant aspects of word sequences dynamically.

Adopting transformers as a backbone encoder in computer vision is advantageous because of its capacity to elucidate long-range relationships and capture global context. Transformers encode pictures as a succession of 1D patch embedding's and use self-attention modules to be told a weighted sum of values calculated from hidden layers, unlike the local formulation of convolutions. Consequently, this versatile formulation makes it possible to be told long-term data efficiently.

Additionally, Vision Transformer (Vit) and its variations have demonstrated exceptional skills in learning pre-text tasks which will be translated to downstream applications.

UNEt transformers may well be a brand new architecture we propose using transformers for volumetric medical picture segmentation during this paper (UNETR). as an example, when it involves 3D segmentation, we employ a transformer encoder to search out out context from the embedded input patches used as input. Although they have a significant capacity to accumulate global knowledge, transformers cannot adequately achieve localized information.

T. HASSANZADE et. al. reported that Manually designing a CNN could be a time–consuming process with reference to thevarious layers that it can have, and therefore the sort of parameters that has got to be founded. Increasing the complexityof the network structure by employing various forms of connections makes designing a network even morechallenging.

X. Liu at et. al. reported that Despite the great achievements of medical image segmentation inrecent years, medical image segmentation based on deep learning has still encountered difficulties inresearch. For example, the segmentation accuracy is not high, the number of medical images in thedata set is small and the resolution is low. The inaccurate segmentation results are unable to meetthe actual clinical requirement..

Image segmentation based on medical imaging is the use of computer image pro-cessing technology to analyze and process 2D or 3D images to achieve segmentation,extraction, three-dimensional reconstruction investigated by Hu, P.[11

Traditional image segmentation methods can no longer be compared with the segmentation methods based on deep learning in effect, but the ideas are still worthlearning [12–13]

Han et al. [85] developeda deep convolutional neural network method, which belongs to the category of "fullyconvolutional neural networks". The DCNN model takes a bunch of adjacent slices asinput and generates a segmentation map corresponding to the central slice.[14].

In this paper, a new automated computer aided model has been developed to recognize lung cancer on the applied CT images. The presented model comprises a set of four main steps namely pre-processing, segmentation, feature extraction and classification. here, a set of two methods are applied namely watershed segmentation and random forest classification model. For simulation processes, a benchmark

.

Methodology:

The entire working process of the presented method is shown in Fig. 1. As shown in figure, the presented model consists a series of processes which are discussed in the s . The input CT scan lung images are provided as input to the presented model. In the initial stage, the pre-processing takes place by the use of median filter and Gaussian filter.

Once the input image is pre-processed, the segmentation of images will take place by watershed segmentation algorithm which produces the output as segmented image in the binary form. Atthe next stage, a collection of important features gets extracted from the segmented image namely centroid, area, perimeter,eccentricity, pixel mean intensity and diameter.

Then, the classification of images will be carried out using randomforest (RF) classification model which finally provides the output as classified image into 'normal' or 'abnormal'.

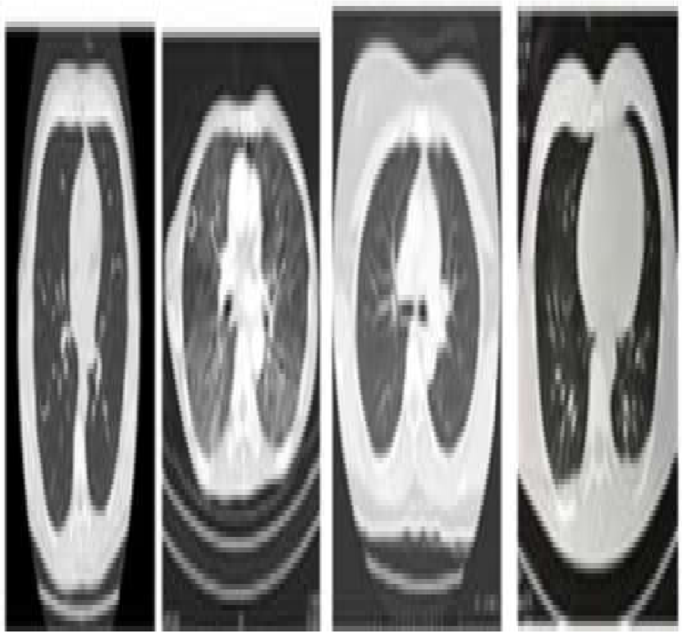


Figure 1. Sample Test Images

We didn't utilize pre-trained weights for our transformer backbone (like ViT on ImageNet) because they didn't increase performance. We tested our model and other baselines in the Standard and Free Competitions of the leader board for the BTCV dataset.

Additional data from the same cohort was used for the Free Competition, bringing the total number of training instances to 80.

We used five-fold cross-validation with a 95:5 ratio in all of our trials. Additional data augmentation techniques included random rotations of 90, 180, and 270 degrees, random flips in all three axes, and random scale and shift intensity shifts in all three planes of the MRI scan. To combine the results of four distinct five-fold cross-validations, we employed assembling.

In the MSD dataset, we divided the data into three sets: training, validation, and test, using an 80:15:5 ratio.

When it comes to BTCV Standard and Free Competitions, UNETR surpasses the current state-of-the-art approaches.

A Dice score of 0.899 ranks UNETR at the top of the leaderboard. Table 1 shows that methods increased by 1.238 percent, 1.696 percent, and 5.269 percent in the Free Competition.

The second, third, and fourth most highly regarded We evaluated the performance of UNETR against CNN and transformer-based baselines in the Standard Competition. The average Dice score for all organs for UNETR is 85.3%, which is a new record high for the organization.

Our technique beats the second-best baseline Dice score for significant organs, such as the spleen, liver, and by 1.043 percent, 0.830 percent, and 2.125 percent, respectively, in terms of the stomach. Furthermore, our technique greatly surpasses the second-best in the segmentation of tiny organs.
For both Standard and Free Competitions on the BTCV, UNETR surpasses the state-of-the-art approaches.

A Dice score of 0.899 ranks UNETR at the top of the leaderboard. 1.238 percent, 1.696 percent, and 5.269 percent are the percentages of methods in the Free Competition displayed in Table1. The second, third, and fourth most highly regarded

We evaluated the performance of UNETR against CNN and transformer-based baselines in the Standard Competition. The average Dice score for all organs for UNETR is 85.3%, which is a new record high for the organization.

Our technique surpasses the second-best baseline Dice score for significant organs such as the spleen, liver, and stomach by 1.043 percent, 0.830 percent, and 2.125 percent, respectively. The Dice score of our technique greatly exceeds the second-best baselines in the segmentation of tiny organs,such as the gallbladder and adrenal glands.

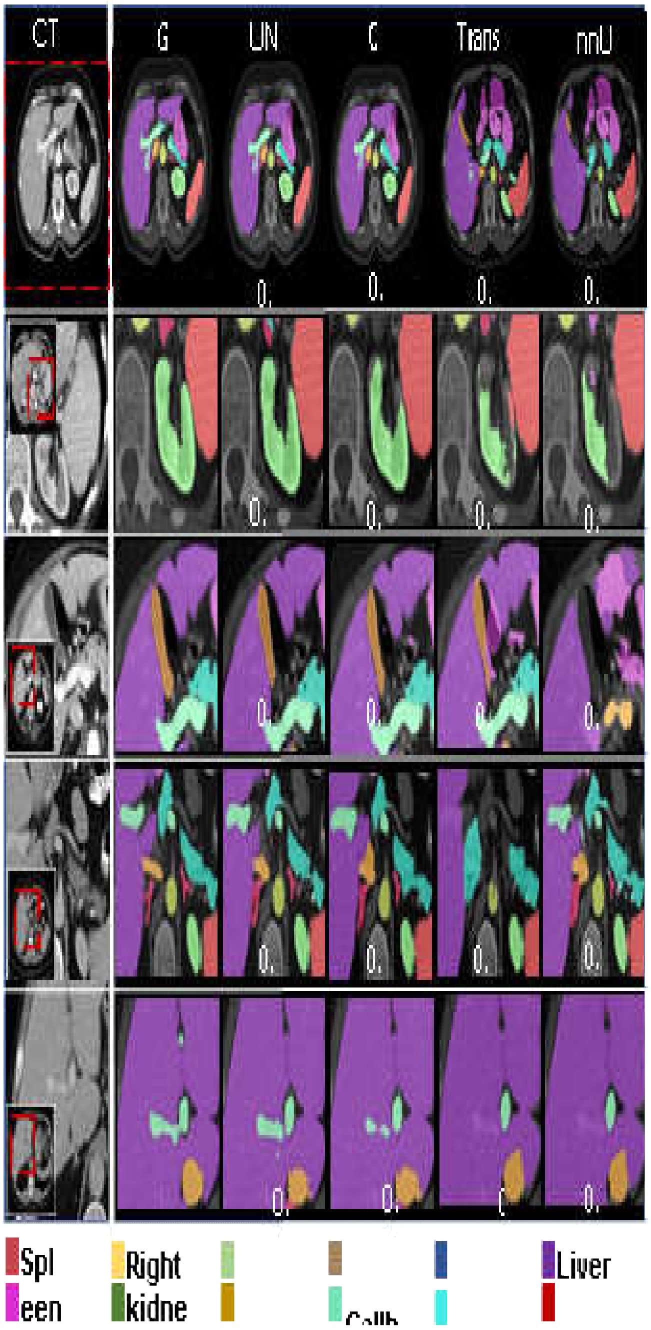
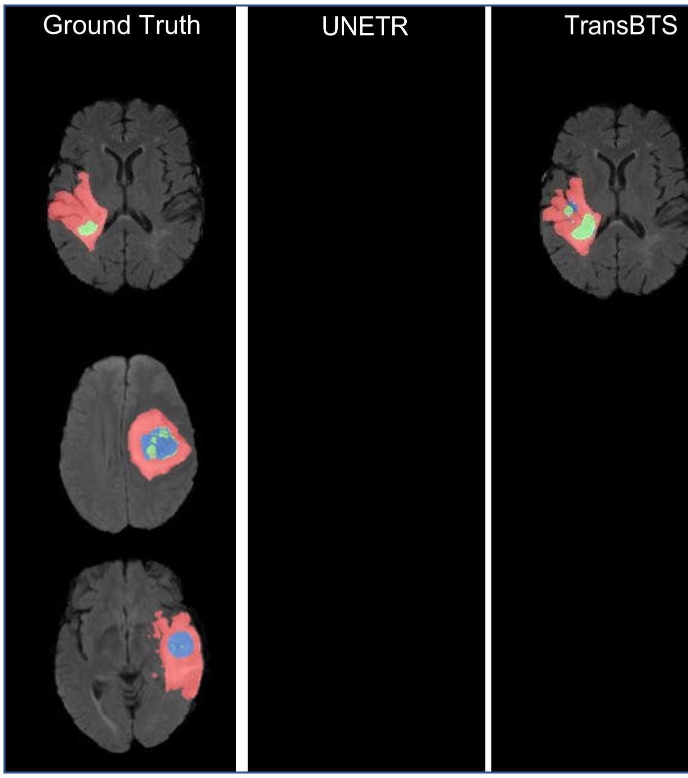


Figure 2. Cross-validation using multiple BTCV starting

points allows for qualitative comparisons (Figure 3). The first row depicts a CT slice that is indicative of the rest of the images in the collection. Here, we present four zoomed-in subjects, where our technique provides a visual improvement in the segmentation of the kidney and spleen (row 2); pancreas and adrenal gland (row 3); the gallbladder (row 4); and the portal vein (row 5). (row 5). Each sample shows the average Dice score for each subject.



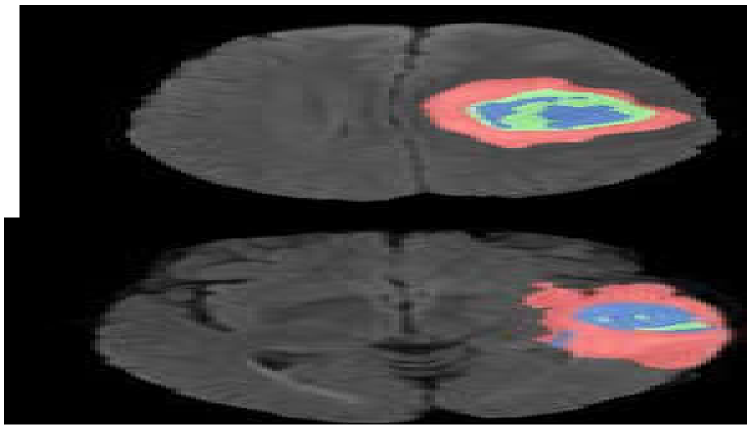


Figure3: UNETR efficiently captures segmentation outputs. Colors red, blue, and green form the Whole Tumor (WT). In the Tumor Core (TC), red and blue zones are combined. These are the green sections of the Enhancing Tumor core (ET).

Organ		Spleen	Brain			
Resolution		Spleen	WT	ET	TC	All
31	0.93	0.75	0.50	0.72	0.65	

Models	#Params(M)	FLOPs(G)	Inference Time(s)
nnUNet	20.07	415.65	17.28
CoTr	41.51	397.21	17.21
TransUNet	94.07	44.34	23.97
ASPP	42.92	42.87	24.47
SETR	84203	41.49	22.86
UNETR	92.58	42.19	14.08

Conclusion:

Our work reframes the challenge of volumetric medical picture semantic segmentation as a 1D sequence to sequence prediction problem and presents the "UNETR" architecture for transformer-based segmentation. Modeling long-range relationships and capturing global context at many scales can be improved by using a transformer encoder, which we suggested.

We tested UNETR's performance on a variety of volumetric segmentation tasks using CT and MRI. UNETR surpasses competing algorithms for a brain tumor and spleen segmentation on the MSD dataset in the Standard and Free Competitions on the BTCV leaderboards for multi-organ segmentation.

UNETR has demonstrated the ability to successfully understand the essential anatomical relationships displayed in medical pictures, to sum it up. A new class of transformer-based segmentation models in medical image analysis might be built on the presented technique.

References:

[1] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. Layer normalization.2016.

[2] SpyridonBakas, Mauricio Reyes, et Int, and BjoernMenze. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. In*arXiv:1811.02629*,2018.

[3] Nayak, R. Sambasiva, and D. R. Singh. "Performance and mprovement of Antenna Designs in Modern Wireless Communication System." *Journal of Telecommunications System & Management* 7.1 (2018): 1000156.

[4] Nayak RS, Bajidvali SK, Gupta A, Venkatarao V (2021) High Isolation Eight-Element Mimo Antenna Array Tightly Arranged Based on Embedded Meta-Material Cells. Indian Journal of Science and Technology 14(15): 1168-1176. https ://doi.org/10.17485/IJST/v14i15.2319.

[5] R. SambasivaNayak, Dr.R.P. Singh, Dr.M. SatyaSai Ram, Dr. G.R. Selokar, Dr. Pushpendra Sharma, Dr.SonalBharti, Alka Thakur "Beak-Shaped Monopole-Like Slot UWB Antenna for Modern Wireless Communication Systems IJR: International Journal of Research ,Volume 6, IssueXII,December/2017,ISSN:2236-6124,pp:43-60. doi:16.10089.IJR.2017.V6112.236456.2596.

[6] HangboBao,LiDong,andFuruWei.Beit:Bertpre-trainingof imagetransformers.*arXivpreprintarXiv:2106.08254*,2021.

[7] Irwan Bello. Lambdanetworks: Modeling long-range interactions without attention. In *International Conferenceon Learning Representations*,2020.

[8] NicolasCarion,FranciscoMassa,GabrielSynnaeve,Nicolas Usunier,AlexanderKirillov,andSergeyZagoruyko.End-to-end object detection with transformers. In *European Confer-enceonComputerVision*,pages213–229.Springer,2020.

[9] MathildeCaron,HugoTouvron,IshanMisra,HervéJégou, JulienMairal,PiotrBojanowski,andArmandJoulin.Emerg-ing properties in self-supervised vision transformers. *arXiv preprint arXiv:2104.14294*,2021.

[10] Jieneng Chen, YongyiLu, QihangYu, XiangdeLuo, EhsanAdeli, Yan Wang, Le Lu, Alan L Yuille, and YuyinZhou. Transunet: Transformers make strong encoders for medical imagesegmentation.*arXivpreprintarXiv:2102.04306*,2021.

[11] Jianxu Chen, Lin Yang, Yizhe Zhang, Mark Alber, and DannyZChen.Combiningfullyconvolutionalandrecurrent neural networks for 3d biomedical image segmentation. In *Proceedings of the 30th International Conference on NeuralInformationProcessingSystems*,pages3044–3052,2016.

[12] Liang-Chieh Chen, George Papandreou, IasonasKokkinos, KevinMurphy,andAlanLYuille.Deeplab: Semanticimage segmentationwithdeepconvolutionalnets,atrousconvolution, and fully connected crfs. *IEEE transactions on pattern analysisandmachineintelligence*,40(4):834–848,2017.

[13] Liang-ChiehChen,YukunZhu,GeorgePapandreou,Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *ProceedingsoftheEuropeanconferenceoncomputervision (ECCV)*, pages 801–818,2018.

[14] R.SambaSivaNayak, Dr, R .P. Singh, Dr. M. SatyaSai Ram, "performance and improvement of ARS ultra wideband antenna" international journal of advanced trends in engineering and technology, ISSN: 2456-1126, Volume 03 Issue 04 July 2018, pp: 01-05.

[15] R.SambaSivaNayak, Dr, R .P. Singh, Dr. M. SatyaSai Ram, Dr. G.R. Selokar, Dr. Pushpendra Sharma,, Dr. SonalBharti& Prof. Alka Thakur "Eye Shaped MIMO Antenna for UltraWideband Applications" Global Journal of Computer Science and Technology: E Network, Web & Security Type: Double Blind Peer Reviewed International Research Journal Publisher: Global Journals Online ISSN: 0975-4172 & Print ISSN: 0975-435, Volume 18 Issue 3 Version 1.0 Year

2018, pp: 14-20.

[16] Hu, P.; Cao, Y.; Wang, W.; Wei, B. Computer Assisted Three-Dimensional Reconstruction for Laparoscopic Resection in AdultTeratoma. J. Med. Imaging Health Inform. 2019,9, 956**–96.**

[17] Ma, Z.; Tavares, J.M.R.S.; Jorge, R.M.N. A review on the current segmentation algorithms for medical images. In Proceedings ofthe 1st International Conference on Imaging Theory and Applications, Lisbon, Portugal, 5–8 February 2009.

[18] Ferreira, A.; Gentil, F.; Tavares, J.M.R.S. Segmentation algorithms for ear image data towards biomechanical studies. Comput.MethodsBiomech. Biomed. Eng. 2014,17, 888–904.

[19] Ma, Z.; Tavares, J.M.R.S.; Jorge, R.N.; Mascarenhas, T. A review of algorithms for medical image segmentation and theirapplications to the female pelvic cavity.Comput.Methods Biomech.Biomed. Eng. 2010,13, 235–246.