Universal Lesion Segmentation Challenge 2023: A Comparative Research of Different Algorithms

Kaiwen Shi, Yifei Li, Binh Ho, Jovian Wang, Kobe Guo

Vanderbilt University

Project Debrief

INTRODUCTION

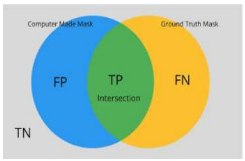
Our project addresses the challenge of medical image segmentation, significantly advanced by technologies such as CNN[1], U-Net variants, transformer-based models like Medical Transformers and SwinUnet, and innovative methods like the knowledge embedding network (KEN)[2]. Despite these advancements, the universal applicability of these networks across different tissue types remains limited. We participated in the Universal Lesion Challenge '23 to develop a solution that balances precision, robustness, and efficiency. We evaluated algorithms including SwinUnet, Medical Transformer. DeepLabV3+, and TransUnet on the Bone Lesion dataset, with TransUnet emerging as the most effective. Our efforts include fine-tuning TransUnet with sophisticated data augmentations, contributing to the field by addressing the critical need for adaptable and efficient segmentation models.

DATASET

The dataset utilized in our study is referred to as ULS23 Data, which consists of two novel datasets annotated by radiologists from the Radboud Medical Center. The primary focus of testing was on bone lesions derived from the ULS23 data provided by Radboudmc. The imaging data is stored in .nii.gz format, with accompanying annotations preserved in .nii.gz.zip files. This dataset comprises volumes of interest (VOIs) meticulously cropped around each individually annotated lesion. These VOIs were selected based on a criterion that the largest axial diameter of the lesion should be equal to or greater than 5 millimeters. The dimensions of each VOI are standardized at 256x256x128 volumetric pixels, ensuring a consistent field of view across all samples. Furthermore, the centering of each VOI is determined by a randomly sampled lesion volumetric pixel.

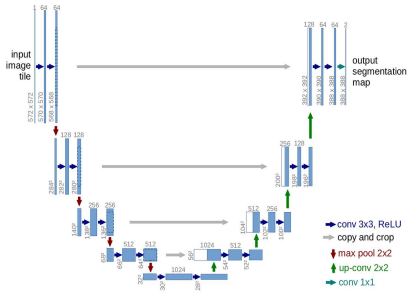
DICE SCORE

Our project uses the dice score metric to evaluate our models. Dice score measures the similarity between two sets of samples, with a 1 indicating a perfect overlap and 0 indicating no overlap.

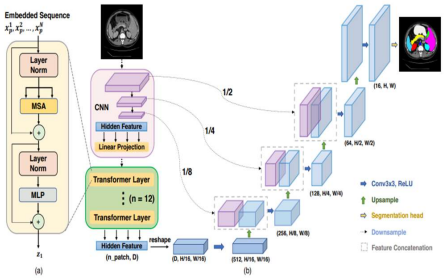


Experiment

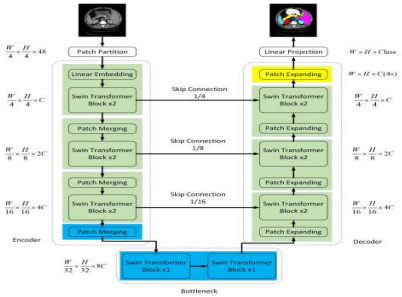
UNet Model (Baseline)[3]



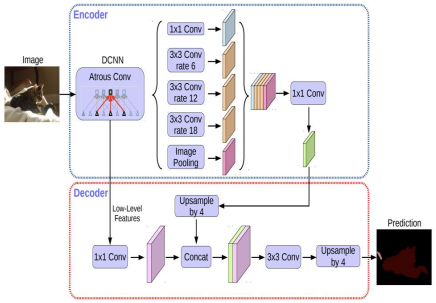
TransUNet Model[2]



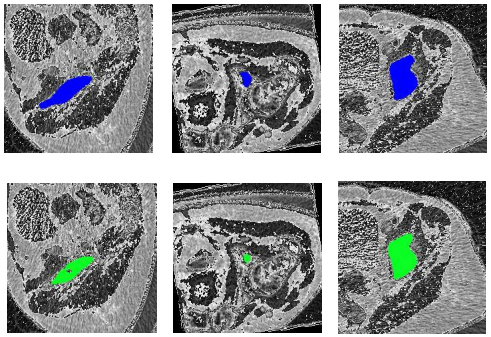
SwinUNet Model[4]



DeepLabV3+ Model[5]



Lesion Segmentation Images



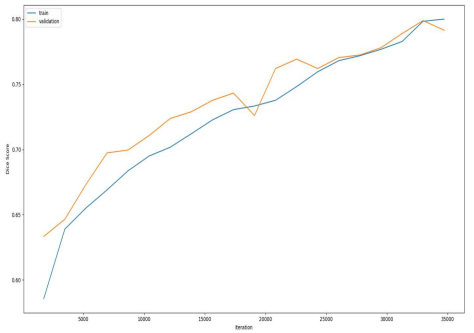
Blue = ground truth
Green = TransUNet

Citations

1. Chen et al., 2017, "Rethinking atrous convolution for semantic image segmentation," CoRR.
2. Chen et al., 2021, "Transunet: Transformers make strong encoders for medical image segmentation," CoRR.
3.Isensee et al., 2020, "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," Nature Methods.
4.Cao et al., 2021, "Swin-unet: Unet-like pure transformer for medical image segmentation."
5. Sandler et al., 2018, "ModileNetV2: Inverted Residuals and Linear Bottlenecks," CoRR.

Results

TransUnet Training Dice Score



Model Comparison on Novel Bone Lesion Dataset

	Baseline	Candidate Algorithms		
Model	Residual 3D U-net	Modified DeepLabV3+	Swin-Unet	TransUNet
Dice Score	0.680 ± 0.24	0.823	0.579	0.864

CONCLUSION

From our experimentation, our models suffer from out-of-memory errors. Many of our models fail to run because the initialization exceeds our current 50-GB capacity. The massive memory consumption is caused by the scale-up to accommodate ULS challenge data. Remedies to this include data parallelism, pipeline partition, and model parallelism so that the model can be split across many devices and allow the training of a more complex model that can provide better results. Additionally, model limitations in capturing tissue-type variations hint at a need for architectures capable of contextual adaptation. Our Knowledge Embedding Network (KEN) has been an attempt to address this, but initial results show inadequate training and validation losses. Further investigation is needed to determine if the current architectures have limited capabilities for encoding contextual information.

ACKNOWLEDGEMENTS

Special thanks to Prof. Ipek Oguz and TA Chaoquan Cai for their support throughout this project. Vanderbilt's ACCRE provided essential computational resources for this project. This work contributes to the Universal Lesion Challenge '23.