MSc Project: Integrating Location Context in Patch-based 3D Medical Image Segmentation

Abstract:

Deep learning-based semantic segmentation has demonstrated exceptional capabilities in automatic detection and segmentation of tumors, anomalies, or organs-at-risk. While new approaches are largely derived from the Computer Vision (CV) domain, the medical imaging field presents unique challenges compared to general CV. Medical imaging datasets, particularly those from CT and MRI scans, often consist of large 3D volumes. Processing these volumes in their entirety can be computationally prohibitive due to memory constraints. To address this, patch-based methods are employed in state-of-the-art segmentation models like nnUNet [1] and SwinUNETR [2]. These models divide the large 3D volumes into smaller, more manageable subvolumes or patches. During training, patches are extracted from the original 3D volume, typically chosen based on the input size the network can handle without encountering memory issues. The size of these patches is crucial; they must be large enough to provide sufficient context for the network to learn meaningful features, yet small enough to fit within computational limits. The network must also learn to understand the relative position of the patches within the body, inferring their location in the broader anatomical structure. To optimize this process, the network often utilizes the largest patch size that hardware can support, maximizing the global context available in each patch. Structures of interest, such as tumors, are often located in specific body regions. To further enhance efficiency, a region-of-interest (ROI) cropping step can be applied before patch extraction, using location information of the target structure to focus the patch extraction process on relevant areas.

Another approach to integrating location information is by providing the network with additional location input. In patch-based segmentation, the 3D location from which the patch is created can be used as input, either as image coordinates [3] or physical scanner coordinates [4]. An even more sophisticated alternative is to use a body-part regression tool [5], as seen in BLE-U-Net [6]. This tool regresses a score for each slice, as a form of a "relative human body coordinate system", offering high-level anatomical context beyond simple image coordinates.

The goal of the thesis is to test different forms of location information input and methods to integrate them into the CNNs and Transformers in a variety of 3D patch-based segmentation tasks (Medical Segmentation Decathlon [7]). This methods may include: coordinate images as additional channels, CoordConv layers, attention mechanism on information context or positional embeddings.

Requirements:

- Prior experience and good understanding in machine learning and computer vision
- Good programming skills in Python and PyTorch
- Interest in medical imaging

Affiliation:

Prof. Dr. Julia Schnabel
Informatik 32 - Lehrstuhl für Computational Imaging and AI in Medicine

Supervision: Stefan Fischer

Application:

Please send a mail, involving a CV, a current transcript of records and a brief statement on why you are interested in the project, to stefan.mi.fischer@tum.de.

References

[1] Isensee, Fabian, et al. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation." Nature methods 18.2 (2021): 203-211.

[2] Tang, Yucheng, et al. "Self-supervised pre-training of swin transformers for 3d medical image analysis." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022.

[3] El Jurdi, R., et al. "CoordConv-unet: investigating CoordConv for organ segmentation." IRBM 42.6 (2021): 415-423.

[4] Das, Badhan Kumar, et al. "Co-ordinate-based positional embedding that captures resolution to enhance transformer's performance in medical image analysis." Scientific Reports 14.1 (2024): 9380.

[5] Yan, Ke, Le Lu, and Ronald M. Summers. "Unsupervised body part regression via spatially self-ordering convolutional neural networks." 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE, 2018.

[6] Nag, Manas K., et al. "Body location embedded 3D U-Net (BLE-U-Net) for ovarian cancer ascites segmentation on CT scans." 18th International Symposium on Medical Information Processing and Analysis. Vol. 12567. SPIE, 2023.

[7] Antonelli, Michela, et al. "The medical segmentation decathlon." Nature communications 13.1 (2022): 4128.