

2025 Autumn Semester

Investigation of cortical thinning in patients with
neuropathic pain based on gray matter segmentation
by embedding vision transformers into U-net
frameworks

Name: Yu-Pei Lai

Student Number: 111261003

Supervisor: Ming-Chang Chiang, MD, Phd

I. Abstract

Cortical thinning is a common feature of aging and neurodegenerative diseases, , and detection of changes in cortical thickness may provide an imaging biomarker for disease progression and prognosis prediction. Nevertheless, accurate measurement of cortical thickness relies on accurate segmentation of the gray matter of the cerebral cortex. In the current project, I will apply a deep-learning method, known as the Swin-UNETR model, which combines the Vision Transformer (ViT) to encode self-attention of imaging features, and U-net frameworks to allow for high-resolution segmentation of 3D MRI scans. Previous studies show that the Swin-UNETR model achieved high accuracy in the segmentation of brain tumors. In the current project, I will apply this method to segment cortical gray matter from the structural MRI scans of the patients with neuropathic pain and healthy adults as the controls. I expect that this approach will facilitate the detection of cortical thinning in patients with neuropathic pain, as a manifestation of maladaptive cortical plasticity following chronic pain.

II. Background

The cerebral cortex is a thin ribbon of gray matter approximately 2.5 mm in thickness, forming the outermost layer of the brain parenchyma [1]. The thickness of the cerebral cortex decreases with normal aging, and this reduction becomes more pronounced in neurodegenerative diseases [2]. For example, patients with Alzheimer's disease show cortical thinning in the temporal and prefrontal regions [3]. Other neurodegenerative diseases, such as frontotemporal dementia, multiple system atrophy, corticobasal degeneration (CBD), and progressive supranuclear palsy, are also associated with cortical thinning [4-6]. Measurement of cortical thickness from magnetic resonance (MR) images provides a sensitive biomarker for these neurodegenerative diseases, facilitating early diagnosis and monitoring of disease progress [7].

An important prerequisite for cortical thickness measurement is an accurate segmentation of cortical gray matter. Detecting disease early and preventing severe cases are crucial. Software like FreeSurfer is frequently used to calculate cortical thickness [8]. While deep learning (DL) shows promise for medical image analysis [9], challenges with interpretability and clinical trustworthiness have limited the adoption of deep neural networks in medical applications [10].

The Vision Transformer (ViT), initially introduced by Google for machine translation in 2017, is a model that embeds a self-attention module into a Transformer-like architecture for image classification and segmentation [11]. Here, an image is divided into fixed-size patches that are combined with position-embedding and classification tokens before entered into a standard Transformer encoder [11]. Transformer architecture uses an entirely attention-based approach, bypassing traditional CNN and RNN architectures. In 2020, Google extended Transformers to image classification with ViT, marking a breakthrough in computer vision due to its simplicity, scalability, and effectiveness.

In the current project, I will apply an algorithm that implements the vision transformer under the U-net scheme, named Swin-UNETR [12], to segment the gray matter of the cerebral cortex and detect cortical thinning in neuropathic pain. Cortical thinning has been documented in patients with neuropathic pain caused by diabetes mellitus [13-15], and I expect that this vision transformer method will provide a higher accuracy for cortical segmentation than conventional atlas-based

methods [16] or recently developed convolutional neural network (CNN) approaches [17], facilitating the identification of structural alterations of the cerebral cortex following chronic neuropathic pain.

III. Literature Review

Although cortical thickness has been considered as a sensitive biomarker for neurodegenerative diseases, estimation of cortical thickness with sub-voxel accuracy from MRI scans is difficult because an accurate estimation of cortical thickness relies on an accurate segmentation of cerebral gray matter [18].

Conventional brain segmentation methods may be divided into the categories: (1) direct segmentation methods, and (2) registration-based methods.

Among direct segmentation methods, FreeSurfer is the most popular and serves as the “gold standard” method for cortical segmentation and thickness estimation [1, 8]. Freesurfer segments cortical gray matter out of the brain by computing gradients of image intensity to delineate the gray-white matter junction. However, brain segmentation using FreeSurfer is very time-consuming – processing a single MRI image can take several hours to several days even with high-performance computers. This poses significant challenges for studies requiring large-scale data processing, such as group analyses of a large sample size. Furthermore, FreeSurfer has high memory requirements and demands substantial hardware resources, particularly when handling high-resolution images.

Another direct segmentation method is SPM (Statistical Parametric Mapping), which is also widely used for image segmentation. SPM employs a Gaussian Mixture Model (GMM) to segment T1 MRI images into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF), generating probability maps for each tissue type. SPM excels in group-level studies with standardized workflows. However, it is less accurate for pathological images and is sensitive to noise or motion artifacts. Processing is computationally intensive and requires substantial time and resources.

For registration-based methods, ANTS (Advanced Normalization Tools) is widely applied. It employs template-based segmentation, mapping target images to standard spaces and generating labels. For template construction, ANTS generates group-specific templates used for population analysis and reference normalization. The key algorithm in ANTS is the SyN Algorithm, which is renowned for fast and precise registration [19]. Brain segmentation with ANTS is probability-based, generating probability maps for tissues such as gray matter, white matter, and cerebrospinal fluid (CSF). However, there are some limitations. Nonlinear registration, particularly with the SyN algorithm, requires significant computational resources and is slow for large datasets. The method is also sensitive to noise or artifacts, which can affect registration and segmentation accuracy. Most importantly, template-based segmentation may be less accurate when handling pathological or atypical structures.

To address the drawbacks of conventional brain segmentation methods, in the current project, I will apply a deep learning model, Swin-UNETR [12], to segment cortical gray matter from MRI scans. In **Figure 1**, Swin-UNETR consists of two important features, the Swin transformer and the U-net structure. The Swin transformer comprises a hierarchical vision transformer computing self-attention in an efficient shifted window partitioning scheme [20]. dividing the input image into

smaller non-overlapping windows and computes self-attention locally within each window. This approach reduces computational complexity from quadratic to linear time with respect to the image size, while still capturing long-range dependencies across spatial neighborhoods using shifted windows [20, 21]. Image features encoded by the Swin transformer are then entered into a U-net structure, which is a convolutional neural network (CNN)-based decoder at different resolutions via skip connections. Unlike traditional CNNs, which often require large kernel sizes or deep convolution stacks to learn global features, Swin Transformer reduces feature resolution progressively, and captures multi-scale information efficiently with fewer layers. The model also benefits from skip connections between the encoder (Swin Transformer) [22] and decoder (CNN-based) as inherent in a U-net scheme [23], allowing for effective fusion of low-level and high-level features without the heavy computational burden of deep convolutional operations [21]. The linear computational complexity of Swin Transformer combined with its ability to extract rich features using fewer resources, makes Swin-UNETR especially suitable for handling high-resolution 3D medical images like MRI scans. In contrast, traditional CNN models struggle with the increased computational load required to process such data.

Swin-UNETR has been shown able to segment brain tumors with high accuracy [12]. In the current project I will apply Swin-UNETR to segment the cerebral cortex. I expect that this deep learning method will be able to capture the complex and convoluted geometry of the central cortex by learning directly from MRI data, to reduce pre-processing needs for cortical thickness estimation.

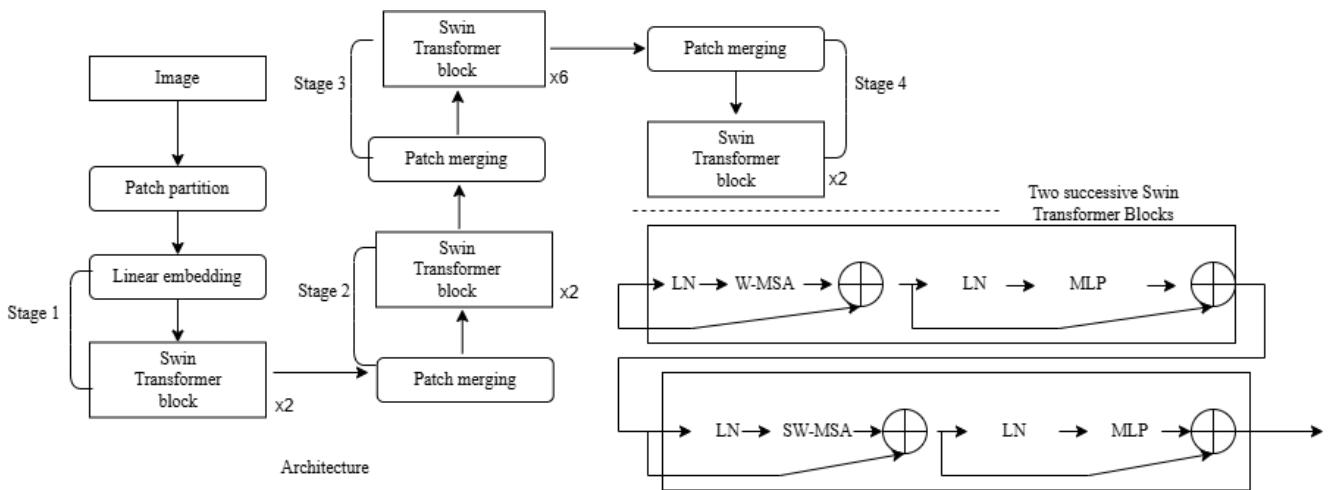


Figure 1. The structure of Swin Transformer blocks. Each Swin Transformer contains the window-based (W-MSA) and the shifted window-based (SW-MSA) multi-head self-attention modules. [20]

IV. Method

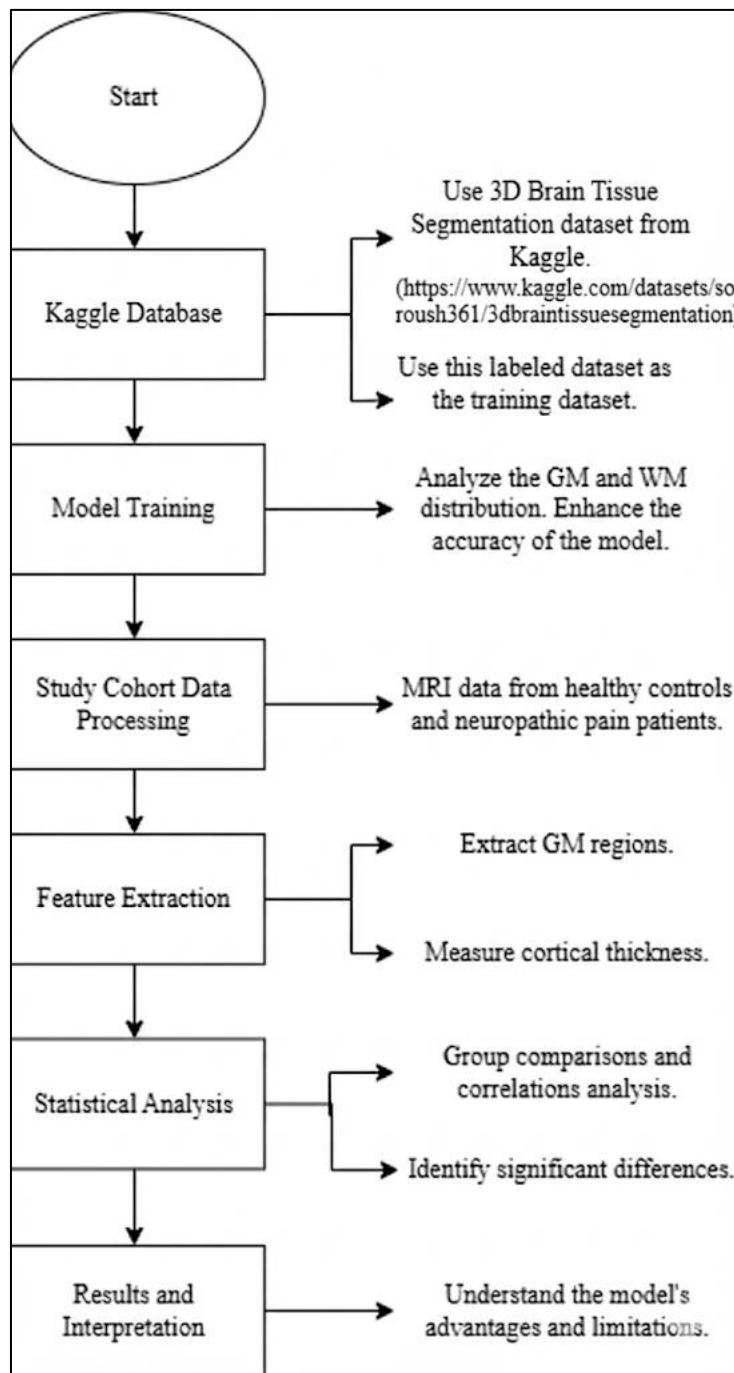


Figure 2. The study protocol of the current project.

The study protocol of the current project is summarized in **Figure 2**. Briefly, I trained the Swin-UNETR model using an open dataset, which can be found on Kaggle(<https://www.kaggle.com/datasets/soroush361/3dbraintissuesegmentation>) , which consists of a cohort of preprocessed T1-weight MRI images to provide the gold standard of gray matter segmentation of the brain. The trained Swin-UNETR model is then planned to be applied to our cohort of patients with neuropathic pain and healthy controls, to facilitate the detection of cortical thinning in the patients as a manifestation of maladaptive brain plasticity following chronic neuropathic pain. However, due to the small cohort of this dataset and the test dice coefficient too low, I did not further apply this model to the clinical data. The detail of the study protocol is described as follows.

Training of the SWIN-UNETR model:

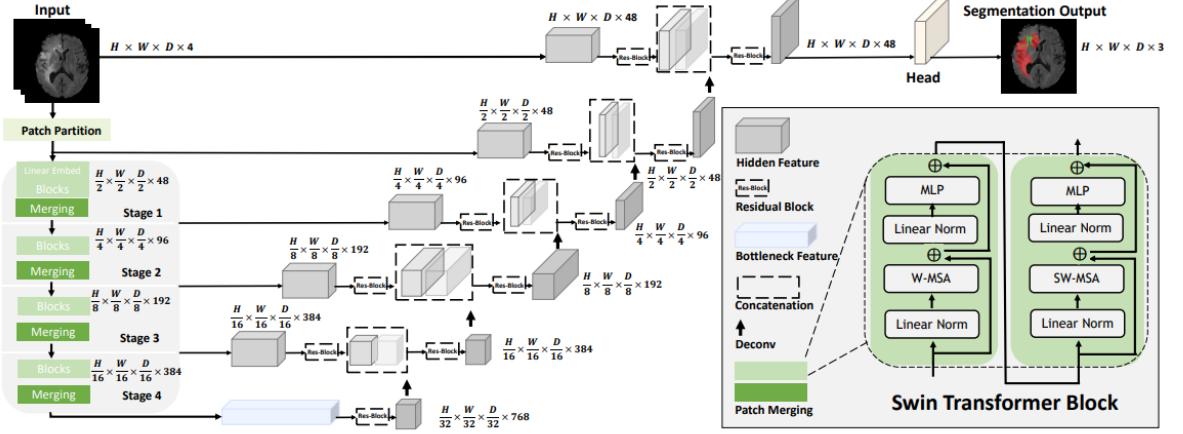


Figure 3. The architecture of the Swin UNETR model. The Swin UNETR model partitions the input 3D multi-modal MRI images into non-overlapping patches and utilizes a patch partition layer to generate windows of a specified size for self-attention computation. Encoded features from the Swin Transformer are connected to a CNN-based U-net structure with skip connections at multiple resolutions. The feature maps of different resolutions are then up-sampled and concatenated to generate the final segmentation output [21].

The cerebral cortex will be segmented out from the T1- weighted MR image using the SWIN-UNETR model [21] as demonstrated in **Figure 3**. During the encoding session, the input for the model, $X \in R^{H \times W \times D \times S}$, will be a token with patch resolution (H_0, W_0, D_0) and dimension $H_0 \times W_0 \times D_0 \times S$. There will be a patch layer that creates a sequence of 3D tokens with dimension $\frac{H}{H_0} \times \frac{W}{W_0} \times \frac{D}{D_0}$, embedded into a space of dimension C, set to 48 here. This patch layer will be followed by a self-attention layer shifting across the entire image with non-overlapping windows, or the Swin Transformer. This encoding process consists of 4 stages, each containing 2 transformer blocks, for a total of L=8 layers, and the 3D T1- weighted MR image will be entered into the encoder with 4 channels using a patch size of $2 \times 2 \times 2$ and a feature dimension of $2 \times 2 \times 2 \times 4 = 32$.

After the encoding process, there will be 4 stages of linear embedding layers. In stage 1, the embedding layer generates $\frac{H}{2} \times \frac{W}{2} \times \frac{D}{2}$ 3D tokens. To maintain the hierarchical structure, a patch merging layer will be employed to reduce the feature resolution by a factor of 2 after each stage. This layer integrates patches of size $2 \times 2 \times 2$ into a $4C$ -dimensional feature embedding, which is reduced to $2C$ using a linear layer. Stages 2, 3, and 4 follow the same design, with the resolution set to $\frac{H}{4} \times \frac{W}{4} \times \frac{D}{4}$, $\frac{H}{8} \times \frac{W}{8} \times \frac{D}{8}$, and $\frac{H}{16} \times \frac{W}{16} \times \frac{D}{16}$, respectively.

At each resolution, the encoders and decoders will be linked using a U-net structure with skip connections. At each stage i ($i \in [2, 3, 4]$) and the bottleneck ($i = 5$), the output features will be reshaped to $\frac{H}{2^i} \times \frac{W}{2^i} \times \frac{D}{2^i}$ and processed through a residual block with two $3 \times 3 \times 3$ convolutional layers normalized by instance normalization. During the decoding step, the resolution will be doubled stage by stage using a deconvolutional layer, and the output of each decoding stage will be concatenated with the outputs of the corresponding encoder stage. The concatenated features will then be processed by a residual block. Final

segmentation will be obtained using a $1 \times 1 \times 1$ convolutional layer followed by a sigmoid activation function. The loss function will be the soft-dice function, computed in a voxel-wise manner.

The SWIN-UNETR model is trained using a dataset on Kaggle, (<https://www.kaggle.com/datasets/soroush361/3dbraintissuessegmentation>), an open dataset that consists of preprocessed T1-weight MRI images to provide the gold standard of gray matter (GM) and white matter (WM) segmentation of the brain.

V. Result

The SWIN-UNETR model [12] was trained to perform automated segmentation of cerebral cortical gray matter to enable subsequent cortical thickness estimation. Following model training, segmentation performance was evaluated using both the internal training dataset and an independent test dataset. **Figure 4.** visualizes the training performance of the Swin UNETR model over 100 epochs, specifically tracking the Epoch Average Loss (training performance) and the Validation Mean Dice (generalization performance) at validation intervals.

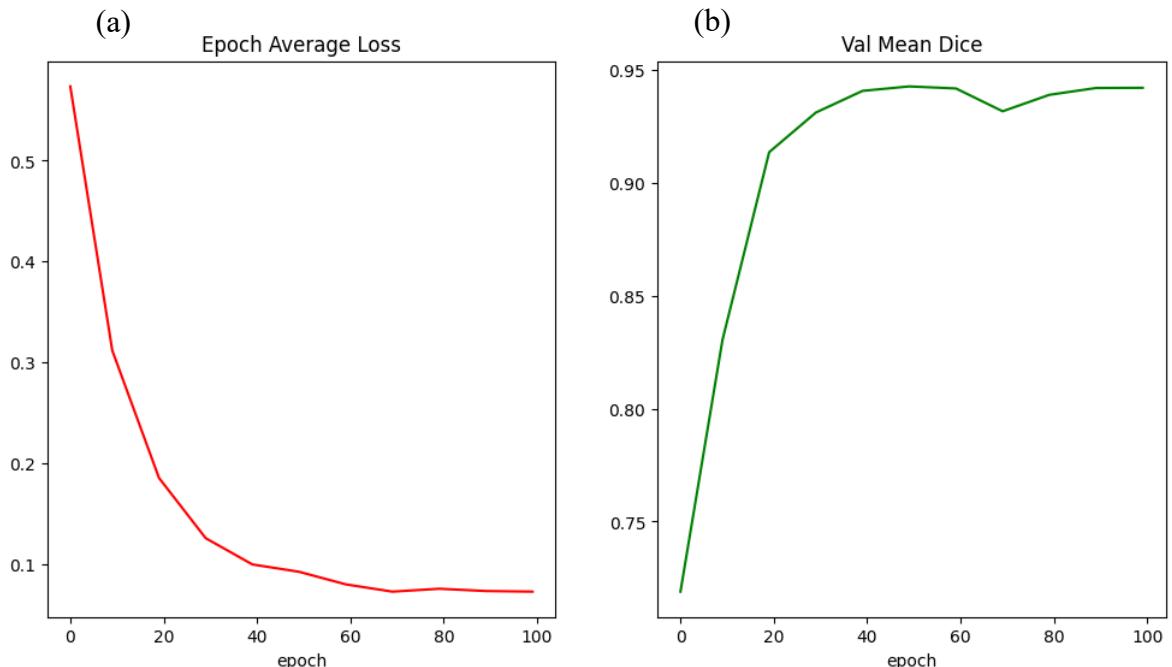


Figure 4. Epoch average loss and Validation mean dice. **(a)** representing the Epoch Average Loss (calculated using the Dice Loss), shows a steep initial decrease during the first approximately 20 epochs (from ≈ 0.55 to ≈ 0.20). **(b)** representing the Validation Mean Dice score, is the primary metric for evaluating the model's segmentation accuracy on unseen data.

The rapid drop in **Figure 4.(a)** indicates that the model is learning effectively and quickly adjusting its weights to minimize the loss on the training data. After the initial drop, the loss continues to decrease but at a much slower rate, eventually stabilizing and remaining low (below 0.1) from around epoch 70 onwards. This suggests the model has largely converged on the training dataset.

In **Figure 4. (b)**, the Dice score exhibits a strong and rapid increase from the beginning, rising from ≈ 0.72 to ≈ 0.90 within the first 20 epochs. The score

continues to climb and reaches a peak performance of 0.945 around epoch 50 to 60. A higher Dice score indicates better overlap between the predicted segmentation and the true label. In the later epochs (after epoch 60), the Validation Mean Dice remains consistently high, hovering between 0.93 and 0.945. The generally stable and high value of the validation Dice score, even as the training loss flattens, indicates that the model is generalizing well and is not significantly overfitting to the training data. The slight drop around epoch 70 followed by a recovery suggests minor fluctuations in the validation performance, but the overall trend confirms robust training.

After training the model, segmentation validation was performed. Each predicted probability map was first binarized using a threshold of 0.5. To refine the segmentation, small, isolated regions were removed using connected-component analysis, retaining only components larger than a predefined minimum size. This procedure reduced spurious predictions and ensured more anatomically consistent gray-matter masks. Segmentation performance was evaluated using the Dice coefficient and the 95th percentile Hausdorff Distance (HD95).

Given the relatively small size of the dataset and the modest performance observed on the test set, the model was not applied to the clinical data in the current stage of the project. This decision was made to avoid potential biases in cortical thickness estimation and to prevent mis-segmentation-related errors from propagating into subsequent clinical analyses.

VI. Discussion

In this study, I applied the SWIN-UNETR model to segment cortical gray matter from structural MRI scans. Although the model achieved a high Dice coefficient on the training dataset (0.94), performance on the independent test set was substantially lower (Dice = 0.67, HD95 = 3.78 mm), indicating limited generalization.

The primary factor contributing to this discrepancy is likely the complexity of the model relative to the small dataset. SWIN-UNETR is a transformer-based architecture with many parameters, which enables it to capture intricate spatial features but also increases the risk of overfitting when training data are limited. The high training Dice coupled with a lower test Dice suggests that the model memorized the training examples rather than learning features that generalize well. Future work could consider regularization strategies (e.g., weight decay, dropout) or adopting lighter-weight transformer variants to mitigate overfitting.

Data characteristics and boundary complexity further challenge cortical segmentation. MRI scans in this study may come from different scanners, protocols, and resolutions, introducing variability that can affect model generalization. Moreover, the cortical gray matter is a thin, elongated structure, making segmentation highly sensitive to boundary accuracy. In our pipeline, predicted probability maps were thresholded to generate binary masks, followed by removal of small, isolated regions. While these steps help reduce noise, they may inadvertently remove small true cortical regions, leading to reduced Dice scores and increased HD95. Our per-sample quality control analysis revealed that some samples exhibited notably higher HD95 values, reflecting the sensitivity of surface-based metrics to even minor boundary deviations.

The evaluation metrics themselves also introduce limitations. Dice coefficient, though widely used, can be disproportionately affected by small errors in thin cortical structures. HD95, while providing information on the largest deviations, is particularly sensitive to outlier points at the surface. These observations suggest that additional metrics, such as mean surface distance, volumetric error, or boundary-aware measures, may provide a more comprehensive assessment of cortical segmentation quality.

Overall, the combination of small dataset size, model complexity, cortical boundary sensitivity, and metric limitations likely contributed to the observed performance gap between training and test sets. Future work should focus on enlarging the dataset, standardizing MRI acquisition, refining preprocessing and postprocessing (e.g., adaptive thresholding, conditional random fields), and exploring alternative evaluation metrics to better capture segmentation quality. Furthermore, improving model performance could enable its use in clinically relevant applications, such as cortical thickness estimation and the study of neuropathic pain, ultimately supporting insights into structural alterations associated with chronic conditions.

VII. Future work

While the current model demonstrates potential, further optimization is required to reach the segmentation accuracy and robustness necessary for clinical application. Future work will focus on refining the model architecture to handle the heterogeneity of clinical data more effectively. Once the model achieves a satisfactory performance metric (e.g., a Dice coefficient above 0.85), we will validate its efficacy using an independent, in-house dataset.

Recruitment of neuropathic pain patients and healthy controls:

The MRI data of 30 patients with neuropathic pain due to peripheral neuropathy and 30 healthy adults as controls will be acquired. The diagnosis will be established by neurologists at National Taiwan University Hospital.

The inclusion criteria for the patients are:

1. Presence of neuropathic pain symptoms, including burning, electric shocks, tingling, sharp pain, aching, cold or freezing, and numbness (paresthesia/dysesthesia), distributed in a symmetric graded stocking-glove pattern.
2. Evidence of large-fiber neuropathy confirmed by nerve conduction studies, or small-fiber neuropathy confirmed by abnormal thermal threshold at the foot on quantitative sensory testing (QST) or a reduced intraepidermal nerve fiber density (IENFd) on skin biopsy at the distal leg.
3. Age 20 or older and 90 or younger.
4. Ability to give informed consent.
5. Agree not to take caffeine, alcohol, tea, or drugs with significant nervous system effects for 48 hours before each study session because these agents can modify brain activity and may confound outcome measures.

The exclusion criteria for the patients are:

1. Presence of systemic diseases, including poorly controlled diabetes mellitus (AC sugar > 200 mg/dl or HbA1c > 10%), severe ischemic heart disease or congestive heart failure (left ventricular ejection fraction < 30%), severe hypertension (systolic blood pressure > 180 mmHg or diastolic blood pressure > 110 mmHg under medication control), severe lung diseases with dyspnea, severe generalized edema, systemic infection, intracranial hemorrhage, brain tumor, and head trauma
2. History of seizures
3. Presence of major depression or any major mental disorders
4. Presence of a cardiac pacemaker, brain stimulator, aneurysm clips, cochlear implant, perfusion pump, external cerebrospinal fluid drainage, or metal implants, unless these devices are certified to be MR-compatible
5. Pregnancy
6. Addiction to smoking or drinking
7. Claustrophobia

Image acquisition:

MRI scanning will be performed on a Siemens-Prisma 3T scanner (Erlangen, Germany) using a 32-channel head coil. The 3D T1-weighted image will be acquired for cortical segmentation and thickness estimation, using a magnetization-prepared rapid gradient-echo (MP-RAGE) sequence (TR/TI/TE = 2530/1100/2.27 ms; flip angle = 9°; FoV = 251 × 220 mm; acquisition matrix = 256 × 224 × 176; voxel dimensions = 0.98 × 0.98 × 1 mm).

Estimation of cortical thickness:

After the cortical gray matter is extracted using the above SWIN-UNETR model, the thickness of the cerebral cortex will be estimated using the projection-based thickness method [24]. Briefly, the white matter distance (WMD) of each gray matter voxel, defined as the distance from the inner boundary of gray matter (gray-white junction) to that voxel, will be estimated. Cortical thickness is then defined as the WMD value of voxels at the outer boundary of the gray matter (pial surface), which represents the distance between the outer and inner boundary of the gray matter.

Statistical analysis:

To assess cortical thickness differences between neuropathic pain patients and healthy controls, several statistical analyses will be performed. The primary focus will be on group comparisons of cortical thickness measurements across different brain regions. The following methods will be employed:

1. Descriptive statistics: Basic descriptive statistics (mean, standard deviation, median, range) will be calculated for all cortical thickness measurements to provide an overview of data distribution.
2. Group comparisons: A two-sample t-test will be used to compare the cortical thickness between neuropathic pain patients and healthy controls.
3. Correlation Analysis: Pearson correlation analyses will be performed to explore the relationships between cortical thickness measurements and clinical variables, such as pain intensity scores, duration of neuropathic pain, and the degree of degeneration of large and small-fiber peripheral nerves.

VIII. Conclusion

In this study, the combination of MRI data heterogeneity, boundary thinness, and sensitivity of Dice and HD95 contributed to variability in segmentation results across samples. This study provides a systematic framework for cortical segmentation and quality assessment, including voxel-level Dice, surface-based HD95, and small region filtering. This framework lays the groundwork for future improvements, such as enlarging the dataset, adopting adaptive thresholding or boundary-aware postprocessing, and exploring alternative metrics for thin-structure evaluation.

Overall, the findings underscore the importance of balancing model complexity with dataset size, careful preprocessing, and metric selection when segmenting fine cortical structures. With further refinement, automated cortical segmentation has the potential to support quantitative analyses of cortical thickness and provide structural insights into neurological conditions such as neuropathic pain.

IX. Reference

1. Fischl, B. and A.M. Dale, Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A*, 2000. **97**(20): p. 11050-5.
2. Salat, D.H., et al., Thinning of the cerebral cortex in aging. *Cereb Cortex*, 2004. **14**(7): p. 721-30.
3. Fjell, A.M., et al., Accelerating cortical thinning: unique to dementia or universal in aging? *Cereb Cortex*, 2014. **24**(4): p. 919-34.
4. Kertesz, A., et al., The corticobasal degeneration syndrome overlaps progressive aphasia and frontotemporal dementia. *Neurology*, 2000. **55**(9): p. 1368-75.
5. Mochizuki, Y., et al., The somatosensory cortex in multiple system atrophy. *J Neurol Sci*, 2008. **271**(1-2): p. 174-9.
6. Constantinides, V.C., et al., Corticobasal degeneration and corticobasal syndrome: A review. *Clin Park Relat Disord*, 2019. **1**: p. 66-71.
7. Worker, A., et al., Cortical thickness, surface area and volume measures in Parkinson's disease, multiple system atrophy and progressive supranuclear palsy. *PLoS One*, 2014. **9**(12): p. e114167.
8. Fischl, B., FreeSurfer. *Neuroimage*, 2012. **62**(2): p. 774-81.
9. Yao, A.D., et al., Deep Learning in Neuroradiology: A Systematic Review of Current Algorithms and Approaches for the New Wave of Imaging Technology. *Radiol Artif Intell*, 2020. **2**(2): p. e190026.
10. Ching, T., et al., Opportunities and obstacles for deep learning in biology and medicine. *J R Soc Interface*, 2018. **15**(141).
11. Dosovitskiy, A., et al., An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *ArXiv*, 2020. **abs/2010.11929**.
12. Hatamizadeh, A., et al. UNETR: Transformers for 3D Medical Image Segmentation. in 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). 2022.
13. Zhang, Y., et al., Sensorimotor and pain-related alterations of the gray matter and white matter in Type 2 diabetic patients with peripheral neuropathy. *Hum Brain Mapp*, 2020. **41**(3): p. 710-725.
14. Galiero, R., et al., Peripheral Neuropathy in Diabetes Mellitus: Pathogenetic Mechanisms and Diagnostic Options. *Int J Mol Sci*, 2023. **24**(4).
15. Hostrup, S.N., et al., Altered surface-based brain morphometry in type

- 1 diabetes and neuropathic pain. *Neuroscience*, 2024.
16. Takahashi, S., et al., Comparison of Vision Transformers and Convolutional Neural Networks in Medical Image Analysis: A Systematic Review. *J Med Syst*, 2024. **48**(1): p. 84.
17. Huo, Y., et al. Vision Transformer (ViT)-based Applications in Image Classification. in 2023 IEEE 9th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS). 2023.
18. Rebsamen, M., et al., Direct cortical thickness estimation using deep learning-based anatomy segmentation and cortex parcellation. *Hum Brain Mapp*, 2020. **41**(17): p. 4804-4814.
19. Avants, B.B., et al., Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal*, 2008. **12**(1): p. 26-41.
20. Liu, Z., et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. in 2021 IEEE/CVF International Conference on Computer Vision (ICCV). 2021.
21. Hatamizadeh, A., et al., Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images, in Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 7th International Workshop, BrainLes 2021, Held in Conjunction with MICCAI 2021, Virtual Event, September 27, 2021, Revised Selected Papers, Part I. 2021, Springer-Verlag. p. 272–284.
22. Tang, Y.C., et al., Self-Supervised Pre-Training of Swin Transformers for 3D Medical Image Analysis. 2022 Ieee/Cvf Conference on Computer Vision and Pattern Recognition (Cvpr 2022), 2022: p. 20698-20708.
23. Hatamizadeh, A., et al., Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Brainles 2021, Pt I, 2022. **12962**: p. 272-284.
24. Dahnke, R., R.A. Yotter, and C. Gaser, Cortical thickness and central surface estimation. *Neuroimage*, 2013. **65**: p. 336-48.