

Evaluating the Combined Impact of Denoising and Resampling Methods on Brain Tissue Segmentation

Smith, Chanae

*Artificial Intelligence for Medicine
Medical Image Analysis Lab
chanae.smith@students.unibe.ch*

Torres, Isabella

*Artificial Intelligence for Medicine
Medical Image Analysis Lab
isabella.torresrevelo@students.unibe.ch*

Zaharia, Tudorita

*Artificial Intelligence for Medicine
Medical Image Analysis Lab
tudorita.zaharia@students.unibe.ch*

Abstract—Automated segmentation of brain structures in MRI facilitates the detection and monitoring of neurodegenerative diseases such as Alzheimer’s disease. Despite advances in acquisition protocols, segmentation accuracy remains dependent on effective preprocessing. In this study, we evaluated whether combining denoising via a Wiener filter (WF) and resampling through linear interpolation (LI) yields additive improvements compared to applying each step individually. Three experiments were conducted: (1) WF-based denoising alone, (2) LI-based resampling alone, and (3) the combination of WF and LI. Using the Human Connectome Project dataset, we assessed segmentation performance with Dice Similarity Coefficient (DSC), Hausdorff Distance (HD), and Structural Similarity Index Measure (SSIM). The results demonstrated that Wiener denoising alone achieved superior DSC scores and lower HDs across most structures. These findings underscore the sequence sensitivity of preprocessing and emphasize that WF, rather than a combined strategy, offers the most consistent improvements for segmentation in this context.

Index Terms—medical image analysis, denoising, wiener filter, linear interpolation, resampling, brain tissue classification

I. INTRODUCTION

Neurodegenerative diseases, such as Alzheimer’s disease (AD), pose a significant challenge in clinical practice. AD is an increasingly prevalent form of dementia, affecting approximately 15% of individuals over the age of 65 and nearly half of those beyond 85 [1]. Characterized by neuronal loss in regions like the hippocampus, AD often goes undiagnosed until advanced stages. This clinical context highlights the need for reliable and sensitive biomarkers, especially in early stages of the disease. Magnetic resonance imaging (MRI)-based automated segmentation techniques provide a promising approach to identify reliable biomarkers by delineating and quantifying brain structures. By accurately segmenting these tissues, clinicians and researchers can track subtle changes that may indicate disease progression or therapeutic outcomes [2].

MRI-based segmentation for isolating brain structures has been playing an increasingly important role in neurology and radiology. Automated methods have gained widespread use due to their ability to reduce time, cost, and inter-observer variability compared to manual delineation [3]. However, these approaches depend heavily on input image quality, making preprocessing a critical step. In this context, preprocessing typically includes (i) background removal to isolate relevant anatomy, (ii) noise reduction to improve the signal-to-noise

ratio (SNR), (iii) intensity normalization across varying imaging protocols, and (iv) resampling to ensure uniform resolution and facilitate registration or comparison across datasets.

From a technical standpoint, the Human Connectome Project (HCP) dataset, comprising T1-weighted (T1w) and T2-weighted (T2w) scans, already benefits from advanced 3T MRI protocols [4]. Nonetheless, residual noise and variations in voxel dimensions remain challenges for accurate segmentation. While normalization is often useful for datasets with heterogeneous acquisition protocols, the HCP’s relatively uniform imaging parameters mean that additional intensity normalization would likely yield diminishing returns. In contrast, denoising and resampling can still confer substantial benefits, particularly in downstream tasks such as feature extraction and classification.

In MRI-based segmentation tasks, denoising serves to reduce random variations introduced by hardware constraints or acquisition settings while preserving critical structural details. This process improves the SNR, clarifies boundaries, and ultimately enhances the reliability of segmentation algorithms [5]. Resampling, in turn, ensures consistent voxel dimensions and spatial resolution across multiple scans, thereby facilitating accurate inter-subject comparisons and image registration [6].

In this study, we adopt the Wiener filter (WF) for denoising because it adaptively suppresses Gaussian noise based on local SNR while preserving anatomical structures. This targeted smoothing helps retain edges and reduces high-frequency distortions, leading to clearer MRI data. For resampling, we employ linear interpolation (LI), a commonly-used technique that allows precise adjustment of voxel sizes (VS) and orientations, improving uniformity among scans and simplifying downstream processing [7].

We hypothesize that combining denoising and resampling provides an additive improvement in segmentation performance over using each method independently. The rationale is twofold: reducing noise clarifies tissue boundaries for the classifier, and resampling images to a more uniform grid can enhance both inter-subject comparability and registration accuracy. To test this hypothesis, we conduct a series of experiments that isolate the effects of WF-based denoising and LI-based resampling, followed by a combined approach to evaluate the synergy of these preprocessing steps.

II. MATERIALS & METHODS

Detailed below is the overall pipeline used for the brain tissue segmentation task, the baseline, and the three main experiments conducted. Fig. 1 provides an overview of the steps involved in each experiment.

A. Pipeline Overview

The pipeline provided by the course organizers follows a sequence of registration, pre-processing, feature extraction, classification, and post-processing. First, T1w and T2w images, along with their corresponding ground truth (GT) labels and masks, are aligned to a common coordinate space. Next, images undergo pre-processing to improve quality. The pre-processed images are then converted into meaningful features for classification. Finally, post-processing refines the segmentation by removing artifacts and smoothing boundaries. All 3 experiments share these core steps but differ in the specific pre-processing strategy (denoising, resampling, or both).

B. Baseline

The Baseline pipeline applies no additional preprocessing beyond standard registration, skull stripping, and intensity normalization. First, the GT labels are registered, then the binary brain mask is aligned to the same space as the T1w/T2w images. Next, registration is performed between the T1w and T2w volumes to ensure spatial alignment. Following this is skull stripping to focus on the relevant intracranial regions, and finally intensity normalization to account for potential scanner or subject-level variations. No denoising or resampling is applied at this stage, allowing the baseline to serve as a reference for segmentation performance.

C. Experiment 1: Denoising Only (Wiener Filter)

In Experiment 1, a WF is introduced to reduce noise while retaining important anatomical boundaries. As in the baseline, the GT labels and mask are first registered and aligned to the subject's T1w/T2w images. Subsequently, the two MRI modalities undergo skull stripping. The critical addition here is that the T1w/T2w images are processed with the WF with varying kernel sizes ($KS = 3, 5, 7, 9$). Following this, intensity normalization is applied. By isolating the effect of noise reduction, Experiment 1 evaluates how suppressing high-frequency noise influences segmentation outcomes.

D. Experiment 2: Resampling Only

Experiment 2 focuses on the impact of voxel-size resampling. The GT labels are first registered and then resampled using nearest-neighbor interpolation (NNI) to preserve their integer values. A similar approach is applied to the brain mask, ensuring it remains strictly binary. After T1w/T2w registration and skull stripping, the images themselves are resampled with LI using various scaling factors (upsampling by 1.1, 1.2, and downsampling by 0.8, 0.9). Finally, intensity normalization is carried out on the resampled images.

Several other resampling methods were initially tested, however, these methods were excluded from the final analysis

due to unsatisfactory performance. Additionally, resampling was tested both before and after registration. While performing resampling before registration theoretically reduces computational load and speeds up preprocessing, it resulted in significantly longer training times in practice, taking approximately an hour. Consequently, this approach was discarded, and the final pipeline was configured with resampling after registration. The advantages of using LI after registration is that it preserves data integrity and allows for more accurate alignment, although it offers less flexibility for post-registration adjustments. By investigating voxel-size adjustments alone, Experiment 2 provides insight into the relationship between spatial resolution and segmentation performance.

E. Experiment 3: Combination of Denoising & Resampling

In Experiment 3, both denoising and resampling are combined to examine potential synergistic benefits. First, the GT is registered and resampled via NNI, while the mask undergoes the same process for consistency. Once T1w/T2w registration and skull stripping have aligned and isolated the brain region, the best performing WF is employed. Following denoising, the images are resampled via the best performing LI scaling factor. A final intensity normalization step completes the preprocessing chain. By jointly applying noise reduction and resampling, Experiment 3 explores whether this dual approach yields additive improvements in segmentation performance.

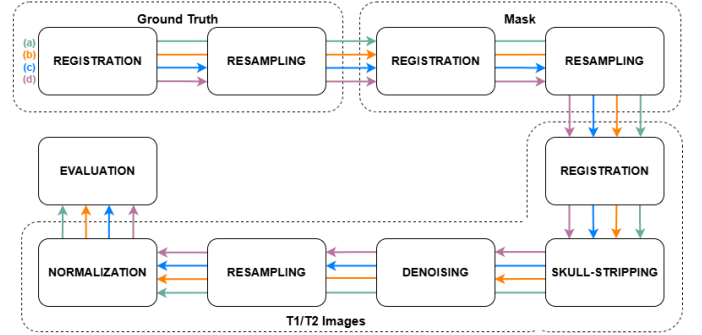


Fig. 1. Steps involved in each experiment. (a) Baseline, (b) Experiment 1, (c) Experiment 2, (d) Experiment 3.

F. Evaluation Metrics

Three metrics were employed to assess segmentation quality. Dice Similarity Coefficient (DSC) measures volumetric overlap by comparing the intersection of the predicted and GT volumes against their average size, offering a direct gauge of segmentation accuracy. Hausdorff Distance (HD) quantifies the greatest point-to-point discrepancy between two boundaries, thus capturing fine-grained shape errors and outliers. Finally, the Structural Similarity Index Measure (SSIM) evaluates perceptual image fidelity by incorporating luminance, contrast, and structural information, which is useful for detecting subtle distortions arising from preprocessing. In combination, these metrics provide a comprehensive view of both volumetric and structural consistency.

III. RESULTS

The results obtained after the three experiments were evaluated in terms of HD, DSC, and SSIM. The best performing filter size from Experiment 1, and the best performing combination of resampling method and voxel size from Experiment 2, were used for Experiment 3.

A. Experiment 1: Denoising Only (Wiener Filter)

To select the optimal kernel size (KS) for WF, the result of the median value for each metric was evaluated for different kernel sizes, as seen in Table I. The lowest median HD was observed for $KS = 9$ ($HD = 16.05$), with HD generally decreasing as kernel size increased from 3 to 9. The DSC remained relatively stable across kernel sizes ($DSC = 0.49, 0.50$), though smaller structures like the Amygdala and Hippocampus show lower values ($DSC \leq 0.31$), highlighting challenges in segmenting small or low-contrast regions (see Table IV in Section VI). The median SSIM is high ($SSIM = 0.99$) for all kernel sizes, indicating minimal structural distortion.

Metric	Median			
	KS = 3	KS = 5	KS = 7	KS = 9
Hausdorff Distance	20.31	22.16	19.13	16.05
DICE Score	0.50	0.50	0.49	0.50
SSIM	0.99	0.99	0.99	0.99

TABLE I
MEDIAN VALUES FOR DIFFERENT KERNEL SIZES.

Box plot analysis (see Fig. 5 in Section VI) confirms that with a kernel size of 9, HD is below the median for 3 of 5 structures (except the Amygdala and Grey Matter), while DSCs are above the median for the same proportion (except the Amygdala and Hippocampus). Based on these results, $KS = 9$ was selected as optimal, achieving the lowest HD, competitive DSCs, and good SSIM, balancing accuracy and structural fidelity.

B. Experiment 2: Resampling Only

Four different voxel sizes ($VS = 0.8, 0.9, 1.1, 1.2$) were tested using three resampling methods: LI, B-spline Interpolation (BSI), and NNI, as summarized in Table II. The experiment included both upsampling (reducing the voxel size and interpolating intermediate voxels) and downsampling (reducing the image resolution). NNI generally produced HD values slightly higher than LI. It demonstrates a more conservative change compared to BSI, which tended to result in slightly higher HD values. This suggests that BSI may lead to greater misalignment between predicted and true segmentation boundaries at these sizes. All 3 methods yielded similar DICE scores and high SSIM values. Among the scaling factors tested, $VS = 0.9$ for LI produced the best balance between HD and DSC, while maintaining a near-ideal SSIM. For this reason, LI with $VS = 0.9$ was chosen as the best-performing as it achieved the most balanced trade-off between boundary accuracy and structural fidelity.

Metric	VS = 0.8	VS = 0.9	VS = 1.1	VS = 1.2
Median (Linear Resampling)				
Hausdorff Distance	31.79	30.69	30.42	31.56
DICE Score	0.25	0.25	0.24	0.24
SSIM	0.97	0.97	0.97	0.97
Median (B-spline Resampling)				
Hausdorff Distance	32.21	32.51	31.90	32.51
DICE Score	0.25	0.25	0.23	0.24
SSIM	0.97	0.97	0.97	0.96
Median (Nearest Neighbour Resampling)				
Hausdorff Distance	31.72	33.30	31.26	31.33
DICE Score	0.25	0.25	0.24	0.24
SSIM	0.98	0.98	0.97	0.97

TABLE II
MEDIAN VALUES FOR DIFFERENT VOXEL SIZES (VS) WITH LINEAR, B-SPLINE, AND NEAREST NEIGHBOUR INTERPOLATION.

C. Experiment 3: Combination of Denoising & Resampling

In this experiment, WF ($KS = 9$) from Experiment 1 was combined with LI ($VS = 0.9$) from Experiment 2. As seen in Fig. 2, denoising with WF alone (Experiment 1) achieved the highest DSCs across all brain structures, consistently surpassing the baseline, resampling with LI (Experiment 2), and the combined approach (Experiment 3). Experiment 3 produced DSCs comparable to those of the baseline and Experiment 2, with no evidence of additive improvements. This finding highlights that the denoising step was the primary contributor to segmentation accuracy, particularly for small structures such as the Amygdala and Hippocampus, which exhibited lower scores due to their inherent anatomical complexity.

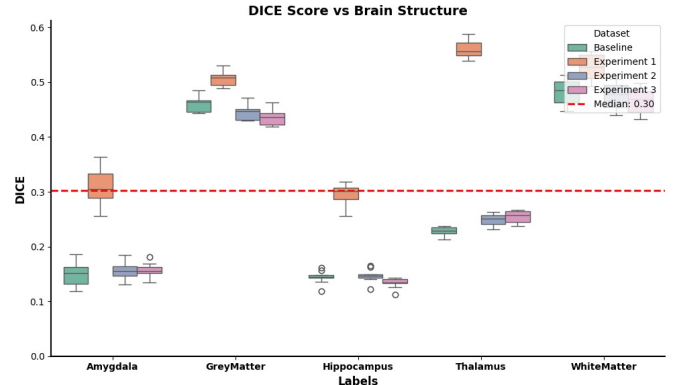


Fig. 2. Dice Scores vs. Anatomical Area for all Experiments

Fig. 3 shows that Experiment 1 consistently demonstrated superior results, yielding the lowest HD values across all anatomical structures. Experiment 3 produced slightly improved HD values for Grey and White Matter compared to the baseline and Experiment 2. However, for other brain regions, Experiment 3 failed to outperform the denoising-only approach, reinforcing the dominant role of WF in reducing boundary irregularities and enhancing structural alignment. Overall, Experiment 3 demonstrated that, despite marginal HD gains in specific brain structures, combining WF with LI did not confer broad segmentation advantages beyond WF alone.

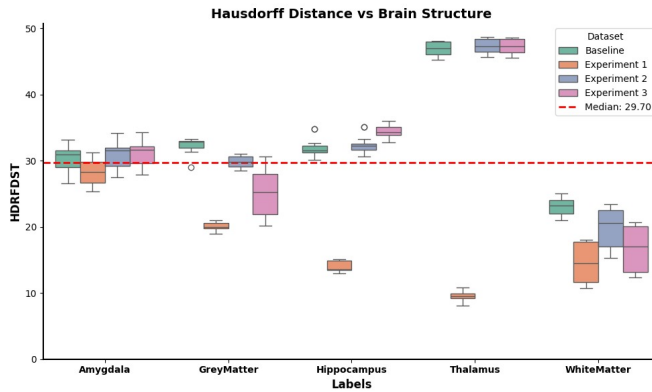


Fig. 3. Hausdorff Distance Scores vs. Anatomical Area for all Experiments

IV. DISCUSSION

The results of the experiment underscore the importance of carefully selecting preprocessing strategies for segmentation in MRI-based analyses. Experiment 1 showed that increasing the KS generally reduced HD, improving alignment between predicted and reference boundaries. This is due to the fact that larger KSs suppress high-frequency noise, resulting in smoother boundaries and reduced irregularities [3]. However, small structures such as the amygdala retained lower DSCs and higher HDs, indicating that denoising alone cannot fully address the challenges of segmenting subtle anatomical regions. Given that the best-performing KS for the WF was the largest examined, a promising future direction is to explore even larger filter sizes. This could further reduce noise and improve segmentation, particularly for smaller structures, but the risk of oversmoothing emphasizes the need for careful parameter tuning and robust cross-validation.

Experiment 2 revealed that voxel-size modifications exert non-uniform effects across different brain structures. While LI offers a straightforward mechanism to match image resolutions, it can blur fine details and potentially alter noise distributions in ways detrimental to certain tissues. This structure-dependent response underscores the limitations of a single interpolation technique in a heterogeneous anatomical context. Similar observations have been noted in previous work, where interpolation methods can negatively affect small-region segmentation [8].

Experiment 3 did not achieve the hypothesized additive improvement. Instead, resampling after denoising reduced segmentation accuracy relative to using denoising alone. The results suggest that while LI may provide marginal benefits for specific structures, it does not enhance overall segmentation outcomes when paired with WF. One plausible explanation is the sequence sensitivity of preprocessing steps: WF effectively removed high-frequency noise at the native resolution, yet subsequent LI may have reintroduced artifacts or altered residual noise, diminishing the net benefit. Additionally, both WF and LI are known to attenuate high-frequency components of the image, which can undesirably merge small structural

boundaries that are crucial for accurate segmentation. These findings collectively highlight that applying multiple enhancements does not inherently guarantee cumulative performance gains; rather, each step must be validated in the broader pipeline context.

The Wiener filter's pronounced performance in the thalamus (Fig. 2 and Fig. 3), highlight the structure-dependent responses. Anatomically, the thalamus is relatively large and has more uniform intensity on T1w/T2w MRI compared to smaller, complex regions. Because the filter adaptively smooths homogeneous areas while preserving edges, clear boundaries and consistent signal levels benefit most. In the thalamus, this effectively reduces noise without obscuring well-defined edges, yielding higher DSC and lower HD. In contrast, smaller or irregularly shaped regions remain challenging to segment even after denoising, given the influence of partial-volume effects and heterogeneous intensities.

SSIM remained high across all experiments, indicating that the applied preprocessing methods preserved the overall structure of the images effectively, though larger structures like Grey Matter and White Matter show slightly lower SSIMs, likely due to greater smoothing effects (Fig. 6 and Fig. 9).

Future research could explore adaptive preprocessing techniques that adjust parameters based on the characteristics of individual images or structures, potentially improving generalizability across datasets. Additionally, investigating advanced denoising methods, such as those based on machine learning, and their integration with resampling techniques could provide further insights.

V. CONCLUSION

Overall, our work emphasizes the nuanced interplay between denoising and resampling for MRI segmentation. Our initial hypothesis posited that combining denoising and resampling would yield additive gains in segmentation performance compared with applying each step individually. Contrary to expectations, the results did not substantiate this claim. Instead, WF alone produced the highest DSCs across all examined brain structures and yielded the lowest HDs. These results suggest that WF alone confers the most reliable gains, while combining LI can inadvertently degrade performance unless meticulously optimized for specific anatomical regions. These findings effectively disprove the hypothesis of additive improvements when combining noise reduction and resampling. Consequently, the main takeaway is that careful consideration must be given to the selection and sequencing of preprocessing steps in MIA pipelines.

REFERENCES

- [1] K. Gustaw-Rothenberg et al., "Biomarkers in Alzheimer's disease: past, present and future," *Biomarkers in Medicine*, vol. 4, no. 1, pp. 15–26, Feb. 2010, doi: <https://doi.org/10.2217/bmm.09.86>
- [2] S. Pereira, A. Pinto, J. Oliveira, A. M. Mendrik, J. H. Correia, and C. A. Silva, "Automatic brain tissue segmentation in MR images using Random Forests and Conditional Random Fields," *Journal of Neuroscience Methods*, vol. 270, pp. 111–123, Sep. 2016, doi: <https://doi.org/10.1016/j.jneumeth.2016.06.017>

- [3] L. Lenchik et al., "Automated Segmentation of Tissues Using CT and MRI: A Systematic Review," *Academic Radiology*, vol. 26, no. 12, pp. 1695–1706, Aug. 2019, doi: <https://doi.org/10.1016/j.acra.2019.07.006>
- [4] D. C. Van Essen et al., "The Human Connectome Project: A data acquisition perspective," *NeuroImage*, vol. 62, no. 4, pp. 2222–2231, Oct. 2012, doi: <https://doi.org/10.1016/j.neuroimage.2012.02.018>. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3606888/>
- [5] S. V. Mohd Sagheer and S. N. George, "A review on medical image denoising algorithms," *Biomedical Signal Processing and Control*, vol. 61, p. 102036, Aug. 2020, doi: <https://doi.org/10.1016/j.bspc.2020.102036>
- [6] J. A. Parker, R. V. Kenyon, and D. E. Troxel, "Comparison of Interpolating Methods for Image Resampling," *IEEE Transactions on Medical Imaging*, vol. 2, no. 1, pp. 31–39, Mar. 1983, doi: <https://doi.org/10.1109/tmi.1983.4307610>
- [7] T. M. Lehmann, C. Gonner, and K. Spitzer, "Survey: interpolation methods in medical image processing," *IEEE Transactions on Medical Imaging*, vol. 18, no. 11, pp. 1049–1075, 1999, doi: <https://doi.org/10.1109/42.816070>
- [8] A. Paiement, Majid Mirmehdi, X. Xie, and Mark, "Integrated Segmentation and Interpolation of Sparse Data," *IEEE transactions on image processing*, vol. 23, no. 1, pp. 110–125, Jan. 2014, doi: <https://doi.org/10.1109/tip.2013.2286903>

VI. APPENDIX

A. Experiment 1: Tables and Boxplots

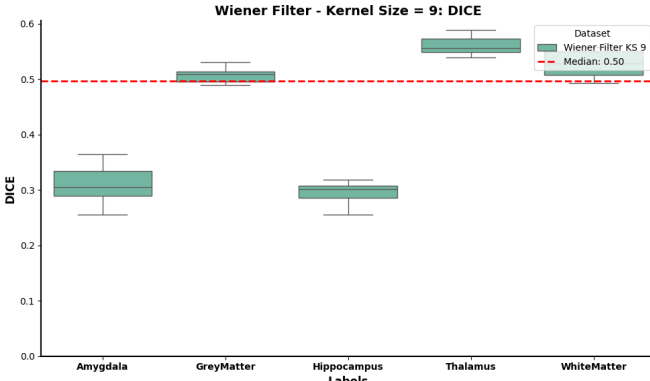


Fig. 4. DSC Scores vs. Anatomical Area for Wiener Filter (KS=9)

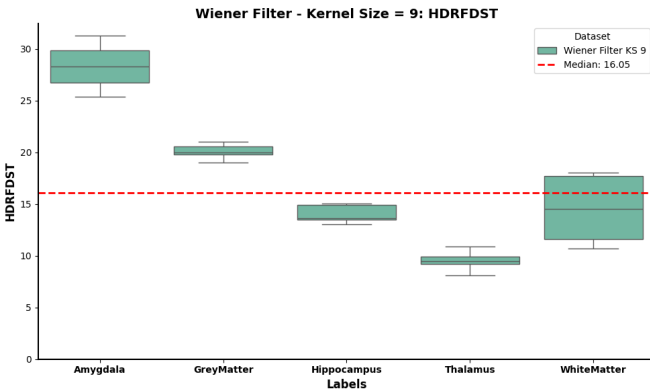


Fig. 5. HD Scores vs. Anatomical Area for Wiener Filter (KS=9)

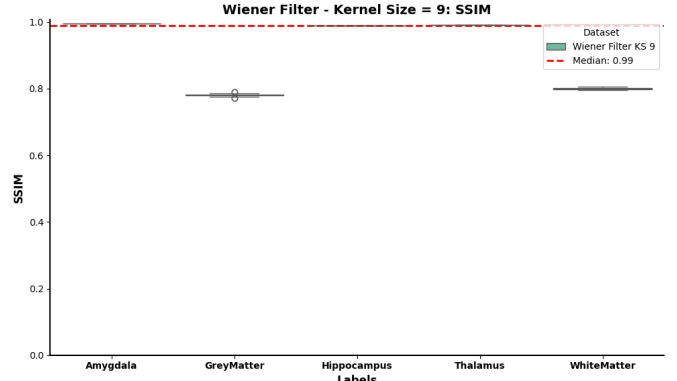


Fig. 6. SSIM Scores vs. Anatomical Area for Wiener Filter (KS=9)

Structure	Hausdorff Distance			
	KS=3	KS=5	KS=7	KS=9
Amygdala	0.27±0.18	0.28±0.18	0.29±0.01	0.28±0.19
GreyMatter	0.20±0.60	0.24±0.46	0.19±0.73	0.20±0.62
Hippocampus	0.24±0.87	0.22±0.67	0.21±0.74	0.13±0.72
Thalamus	0.11±0.11	0.09±0.73	0.09±0.45	0.09±0.66
WhiteMatter	0.13±0.04	0.15±0.38	0.14±0.35	0.14±0.30

TABLE III

HAUSDORFF DISTANCE FOR DIFFERENT KERNEL SIZES.

Structure	DICE Score			
	KS=3	KS=5	KS=7	KS=9
Amygdala	0.26±0.02	0.28±0.03	0.28±0.03	0.31±0.03
GreyMatter	0.50±0.01	0.50±0.01	0.50±0.01	0.50±0.01
Hippocampus	0.27±0.01	0.27±0.01	0.26±0.01	0.29±0.01
Thalamus	0.58±0.01	0.56±0.01	0.55±0.01	0.55±0.01
WhiteMatter	0.53±0.02	0.52±0.02	0.52±0.02	0.52±0.02

TABLE IV

DICE SCORE FOR DIFFERENT KERNEL SIZES.

Structure	Structural Similarity Index			
	KS=3	KS=5	KS=7	KS=9
Amygdala	0.99±0.00	0.99±0.00	0.99±0.00	0.99±0.00
GreyMatter	0.78±0.00	0.78±0.00	0.77±0.00	0.78±0.00
Hippocampus	0.98±0.00	0.98±0.00	0.98±0.00	0.98±0.00
Thalamus	0.98±0.00	0.99±0.00	0.99±0.00	0.99±0.00
WhiteMatter	0.79±0.00	0.79±0.00	0.79±0.00	0.79±0.00

TABLE V

SSIM FOR DIFFERENT KERNEL SIZES.

B. Experiment 2: Tables and Boxplots

Structure	DICE Score			
	VS = 0.8	VS = 0.9	VS = 1.1	VS = 1.2
Amygdala	0.116274	0.144142	0.125741	0.131193
GreyMatter	0.437371	0.437409	0.438306	0.430815
Hippocampus	0.142299	0.140867	0.140448	0.144838
Thalamus	0.254698	0.253924	0.239465	0.243507
WhiteMatter	0.477808	0.476408	0.476079	0.478838

TABLE VI

DICE SCORES FOR DIFFERENT VOXEL SIZES WITH NEAREST NEIGHBOUR RESAMPLING.

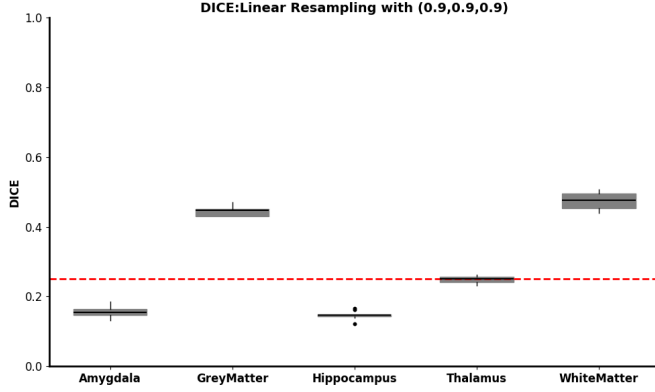


Fig. 7. DSC Scores vs. Anatomical Area for Linear Interpolation (VS=0.9)

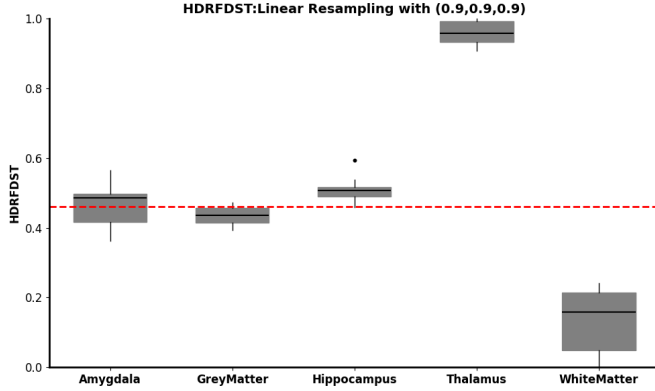


Fig. 8. HD Scores vs. Anatomical Area for Linear Interpolation (VS=0.9)

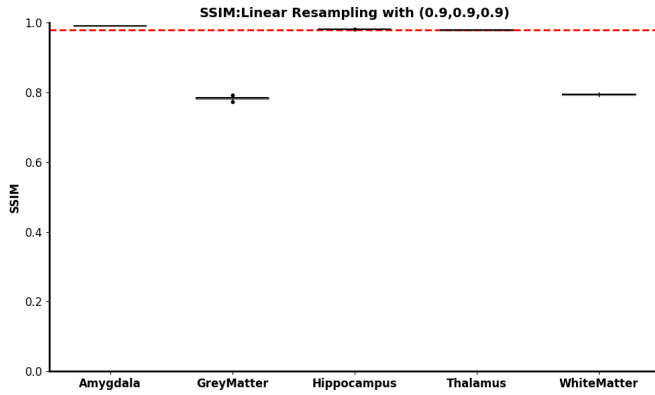


Fig. 9. SSIM Scores vs. Anatomical Area for Linear Interpolation (VS=0.9)

Structure	Hausdorff Distance			
	VS = 0.8	VS = 0.9	VS = 1.1	VS = 1.2
Amygdala	31.404458	31.564220	31.994687	31.200000
GreyMatter	32.800000	33.300000	22.000000	30.758413
Hippocampus	32.139695	33.300000	33.109817	32.555184
Thalamus	45.830994	45.979343	46.173802	45.993043
WhiteMatter	18.880678	18.900000	23.360437	23.082461

TABLE VII

HAUSDORFF DISTANCE FOR DIFFERENT VOXEL SIZES WITH NEAREST NEIGHBOUR RESAMPLING.

Structure	Structural Similarity Index			
	VS = 0.8	VS = 0.9	VS = 1.1	VS = 1.2
Amygdala	0.991243	0.990740	0.988065	0.988963
GreyMatter	0.797059	0.789114	0.771972	0.765475
Hippocampus	0.982440	0.977141	0.974482	0.977782
Thalamus	0.980127	0.978507	0.976150	0.974791
WhiteMatter	0.802925	0.794316	0.778066	0.767848

TABLE VIII

SSIM FOR DIFFERENT VOXEL SIZES WITH NEAREST NEIGHBOUR RESAMPLING.

Structure	DICE Score			
	VS = 0.8	VS = 0.9	VS = 1.1	VS = 1.2
Amygdala	0.105737	0.130257	0.118975	0.119385
GreyMatter	0.416659	0.415247	0.408897	0.411800
Hippocampus	0.126130	0.129706	0.136436	0.134536
Thalamus	0.257338	0.262885	0.234878	0.251641
WhiteMatter	0.460725	0.462051	0.459043	0.459940

TABLE IX

DICE SCORES FOR DIFFERENT VOXEL SIZES WITH B-SPLINE RESAMPLING.

Structure	Hausdorff Distance			
	VS = 0.8	VS = 0.9	VS = 1.1	VS = 1.2
Amygdala	32.984845	33.409280	33.164590	33.557115
GreyMatter	18.434750	26.100000	20.900000	31.107555
Hippocampus	35.063371	34.972847	33.182827	35.211362
Thalamus	47.328638	47.110508	47.529675	47.517576
WhiteMatter	14.377761	15.948041	24.423964	24.000000

TABLE X

HAUSDORFF DISTANCE FOR DIFFERENT VOXEL SIZES WITH B-SPLINE RESAMPLING.

Structure	Structural Similarity Index			
	VS = 0.8	VS = 0.9	VS = 1.1	VS = 1.2
Amygdala	0.990439	0.989875	0.987165	0.987064
GreyMatter	0.784014	0.773914	0.752328	0.745540
Hippocampus	0.975444	0.975106	0.979283	0.967714
Thalamus	0.980465	0.979364	0.979983	0.979689
WhiteMatter	0.805763	0.799278	0.786249	0.781303

TABLE XI

SSIM FOR DIFFERENT VOXEL SIZES WITH B-SPLINE RESAMPLING.

Structure	DICE Score			
	VS = 0.8	VS = 0.9	VS = 1.1	VS = 1.2
Amygdala	0.126622	0.131094	0.122835	0.126199
GreyMatter	0.429193	0.429828	0.434044	0.431709
Hippocampus	0.131413	0.146582	0.141376	0.132762
Thalamus	0.252455	0.254070	0.242404	0.246613
WhiteMatter	0.469683	0.475584	0.478338	0.471729

TABLE XII

DICE SCORES FOR DIFFERENT VOXEL SIZES WITH LINEAR RESAMPLING.

Structure	Hausdorff Distance			
	VS = 0.8	VS = 0.9	VS = 1.1	VS = 1.2
Amygdala	23.224566	25.170312	25.000000	23.446450
GreyMatter	15.874338	20.000000	17.173612	22.000000
Hippocampus	31.113334	35.000000	30.285384	32.196331
Thalamus	45.726687	46.249073	45.800421	46.348264
WhiteMatter	17.418455	19.302557	20.000000	20.000000

TABLE XIII

HAUSDORFF DISTANCE FOR DIFFERENT VOXEL SIZES WITH LINEAR RESAMPLING.

Structure	Structural Similarity Score			
	VS = 0.8	VS = 0.9	VS = 1.1	VS = 1.2
Amygdala	0.987566	0.985704	0.986401	0.986207
GreyMatter	0.771993	0.759937	0.742345	0.736139
Hippocampus	0.973775	0.974244	0.973670	0.972678
Thalamus	0.979445	0.978485	0.976368	0.975325
WhiteMatter	0.795340	0.786301	0.772778	0.766755

TABLE XIV

SSIM FOR DIFFERENT VOXEL SIZES WITH LINEAR RESAMPLING.