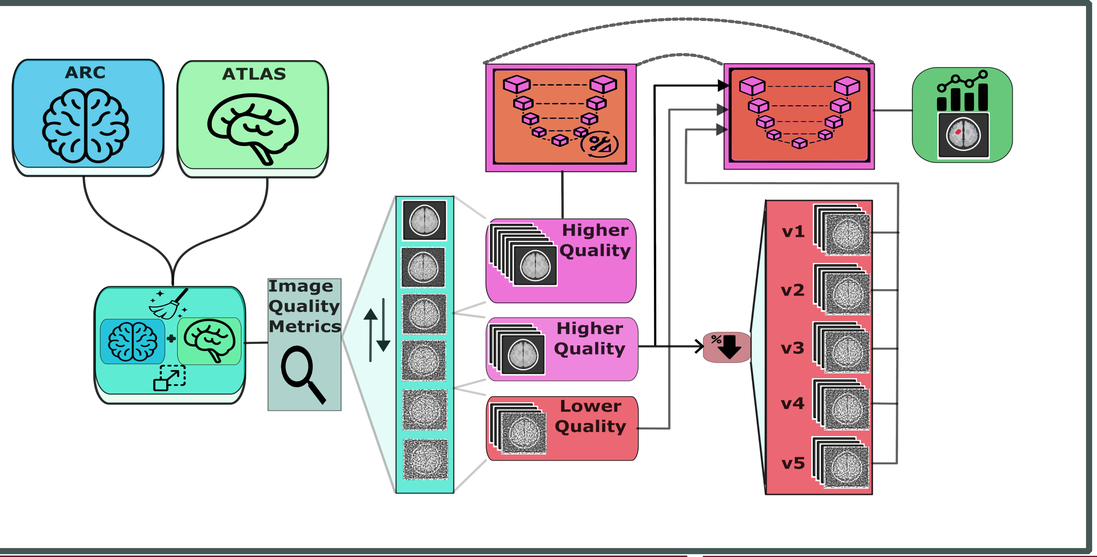
Assessing the Impact of MRI Resolution on the Clinical Applicability of Lesion Mapping Models

Rustin Bielski and Lucy L. W. Owen
University of Montana

Overview

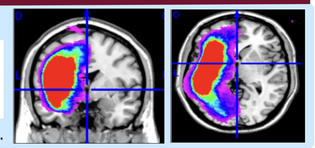


Data

ARC: 230 chronic stroke T1w MRI's w/ lesion masks

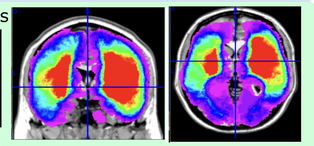
neurolabusc/
AphasiaRecoveryCohort...
Aphasia Recovery Cohort (ARC) Demonstration of Processing and Inference

Right images show distribution of lesions.

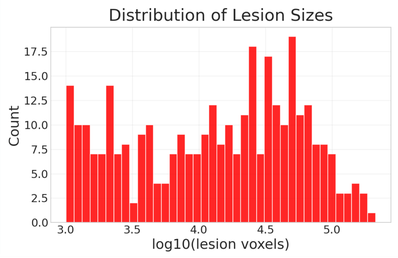


ATLAS: 655 chronic stroke T1w MRI's w/ lesion masks



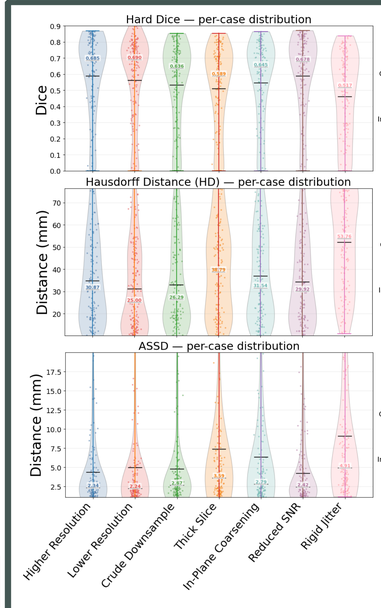


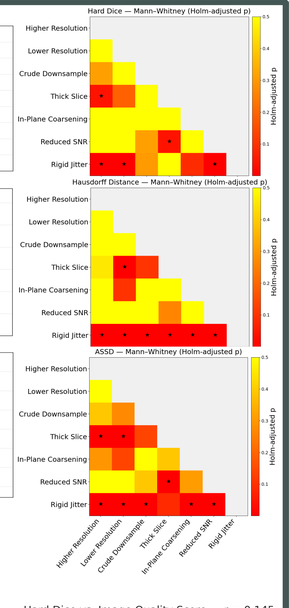
Right images show distribution of lesions.

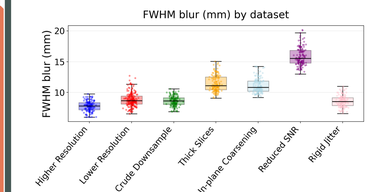




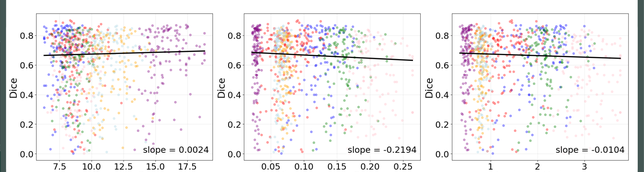
Results











Background

Question:
How does MRI image quality influence a model's ability to detect and quantify chronic stroke lesions in T1-weighted images?

Why:

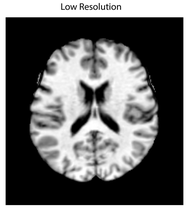
➢Clinical vs. research scans:
- MRI image quality varies widely across scanners, protocols, and sites.Models trained on pristine research-grade data often fail on lower-quality clinical scans.

➢Relevance to stroke imaging:
- Acute stroke: Radiologists use diffusion and perfusion imaging (DWI, FLAIR) for treatment decisions within hours of onset.
- Chronic stroke: T1-weighted MRI is preferred weeks to months later for identifying tissue loss and scarring.

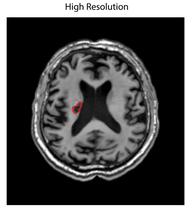
Why segmentation matters:

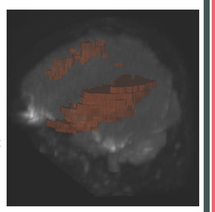
➢Automated lesion segmentation enables:
- Outcome prediction, rehabilitation and neuromodulation planning, lesion–symptom mapping, and quantitative research endpoints

Low Resolution



High Resolution



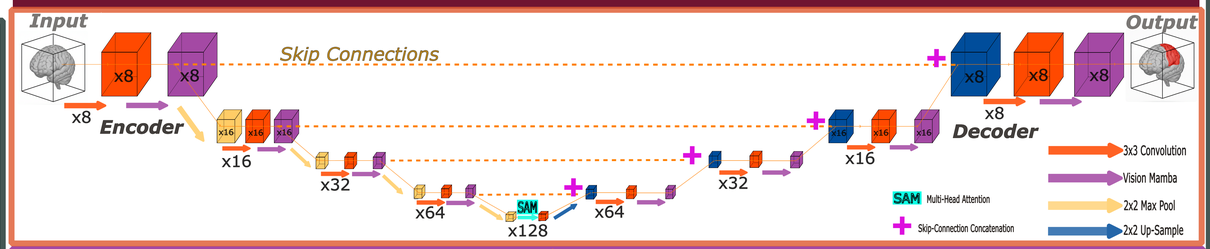


Left images show lowest resolution image compared to highest resolution image in the dataset. Right shows thickness of slices in the z axis.

Gap:

➢Few studies systematically test how spatial or slice resolution affects model accuracy in the chronic phase.
➢Clinical imaging trades off resolution for shorter scan time and patient comfort, yet this trade-off's impact on model robustness and generalization remains poorly quantified.

Methods



Model:
➢A hybrid 3D U-Net backbone with skip-connections, Vision-Mamba–style blocks after every convolution for linear-time long-range mixing in 3D; a single bottleneck multi-head attention layer (SAM-2–inspired memory) restores global coherence before decoding. This keeps U-Net's localization while adding efficient context and low GPU footprint.

Downsampling:

1. Crude
➢Reduces the image resolution equally in x, y, and z by an integer stride, then brings it back to the original matrix size using nearest-neighbor repetition.
Simulates very coarse voxel grids from fast protocols without anti-alias filtering.

2. Thick Slices
➢Smoothes only along the slice (z) direction with a Gaussian blur, keeping every k-th slice, then repeat slices to return to the original matrix size.
Mimics anisotropic stacks with thick slices that shorten scan time and introduce through-plane partial volume.

3. In-Plane Coarsening
➢Smoothes within the axial plane (x and y) with a Gaussian blur, keeps every m-th pixel in x and y, then repeats in-plane pixels to return to the original matrix size.
Emulates reduced in-plane matrix to cut scan time while maintaining thin slices.

4. Reduced SNR
➢Adds independent zero-mean Gaussian noise to synthetic real and imaginary channels and then takes the magnitude.
Approximates magnitude-MR noise from fewer averages, higher acceleration, smaller voxels, or lower field strength.

5. Motion Jitter
➢For each axial slice, applies a small random in-plane rotation and pixel shift, then re-stacks the slices.
Recreates common 2D multi-slice misalignment from patient motion that degrades boundaries.

Image Quality Metrics:

1. Guassian Blur FWHM
➢Fits how "blurry" the image is by assuming a Gaussian blur and estimating its full width at half maximum (FWHM).

2. High Frequency Energy Fraction
➢Transforms the image to frequency space and compute what fraction of total energy lies above a chosen high-frequency cutoff (i.e., fine detail/edges).

3. Laplacian Variance
➢Applies a 3D Laplacian (edge/second-derivative) to the brain and measure the variance of that response.

Evaluation Metrics:

1. Dice Similarity Coefficient
➢Dice measures volumetric overlap between the predicted segmentation and the expert ("ground truth") segmentation mask.
Higher Dice indicates better agreement (1 = perfect overlap, 0 = no overlap)

2. Hausdorff's Distance
➢HD measures the maximum surface discrepancy between the predicted and expert lesion masks (lower values mean closer boundary agreement).

3. Average Symmetric Surface Distance
➢ASSD represents the average boundary distance between the two masks, giving a smoother estimate of spatial accuracy.

Summary of Findings

➢Across the 5 derived (downsampled) variants of the higher resolution images, the paired test showed performance differences across few of the of the groups. Notably, only the synthetic motion jitter showed consistently lower performance.
➢While the reduced SNR dataset had the highest noise of all, it did not seem to affect the models performance, and across all datasets there didn't seem to be any correlation between image quality and the models ability to accurately predict lesions.

References/Acknowledgments

➢I would like to thank Lucy and the rest of the brAIn Lab for believing in me, Rebecca Bryan for helping during the early stages of the project and Danielle Fahey for her support and motivation.

Gibson, M., Newman-Norlund, R., Bonilha, L., Fridriksson, J., Hickok, G., Hillis, A. E., den Ouden, D.-B., … Rorden, C. (2024). The Aphasia Recovery Cohort, an open-source chronic stroke repository. Scientific Data, 11, 981. https://doi.org/10.1038/s41597-024-03819-7
Liew, S.-L., Anglin, J. M., Banks, N. W., Sotiropoulos, S. N., … (2022). A large, curated, open-source stroke neuroimaging dataset to improve lesion segmentation algorithms (ATLAS v2.0). Scientific Data, 9, 320. https://doi.org/10.1038/s41597-022-01401-7