

Supplementary Material- MedLesSynth-LD: Lesion Synthesis using Physics-Based Noise Models for Robust Lesion Segmentation in Low-Data Medical Imaging Regimes

Methods

Rician distribution probability density function (p.d.f.) can be expressed by

$$f(x|\nu, \sigma) = \frac{x}{\sigma^2} \exp\left(-\frac{(x^2 + v^2)}{2\sigma^2}\right) I_0\left(\frac{xy}{\sigma^2}\right) \quad (1)$$

where $\nu \geq 0$, the distance between the reference point and center of the bivariate distribution, $\sigma \geq 0$, scale and $I_0(z)$ is the modified Bessel function of first order.

Similarly **Gaussian distribution** p.d.f. can be written as

$$f(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (2)$$

where μ stands for mean and σ^2 for variance. Speckle noise has been modelled before using Perlin Noise [2]. Here we make use of the multiple octaves present in Perlin to simulate the noise. Perlin noise construction has been described in [1,3]. Perlin noise is constructed by initially forming an n-dimensional grid, where each point on the grid is associated with a random vector r_g . For the grid vectors, we also determine the displacement vectors r_d based on the offset between grid corners and candidate points of interest. Later, dot products are performed between vectors r_g and the displacement vectors r_d for candidate points. Finally, we interpolate the dot product values using a function which has its first derivative zero at the grid nodes.

We have shown sample noise profiles for the above noise distributions/models in Fig. S1.

Comparison of semi-supervised lesion segmentation using synthetic lesions with baseline

Fig. S2 contains the Receiver Operating Characteristics (ROC) curve for all 3 settings (`F_sup`, `Semi_DA` and `Semi_FT`) which was further used to determine suitable threshold for inference stage. Table S1 reports the result of applying Wilcoxon signed ranked test on comparison of M_{seg} architectures (results of Table 2 from the main manuscript). Models are compared from left to right (arranged in the order of increasing DSC scores).

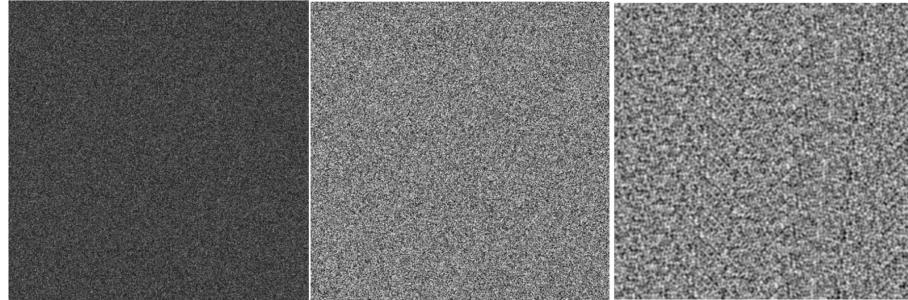


Fig. S1. Examples of the 3 noise patterns: (from left to right) Rician, Gaussian and Perlin.

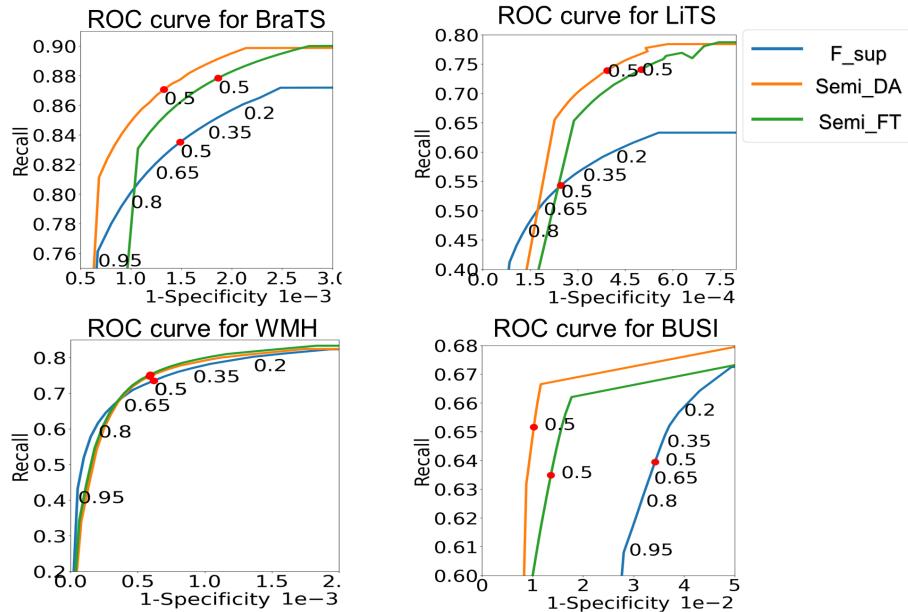


Fig. S2. ROC curve for F_{sup} (blue), Semi_DA (orange) and Semi_FT (green) performance for BraTS, WMH, LiTS and BUSI

Effect of real-world and simulated data proportions on lesion segmentation:

In this section, we report the Wilcoxon signed-rank test results for Fig.4 from the main manuscript depicting the significance between Semi_DA and F_{sup} (Semi_DA – F_{sup}) on different real data proportions in Table S2. Similar test results have been reported for the simulation factor in Table S3. UNet++ architecture was used for M_{seg} for all cases.

Table S1. Wilcoxon signed rank test values for comparison between the performance of candidate M_{seg} models.

Dataset	Slim UNetr	Half UNet	UNet	UNetr	UNet++
DSC					
BraTS	-	1.43x10-133	2.92x10-17	7.35x10-09	6.27x10-50
WMH	-	0.04x10-2	0.02	0.09	0.30
LiTS	-	2.28x10-7	3.27x10-7	0.05x10-2	0.01
BUSI	-	1.186x10-10	0.34	6.58x10-5	0.42
Recall					
BraTS	-	1.74x10-153	5.34x10-84	3.42x10-117	5.20x10-30
WMH	-	0.79	0.14x10-02	0.73	0.57
LiTS	-	0.40	0.18x10-3	3.16x10-6	0.09x10-2
BUSI	-	2.32x10-36	5.14x10-24	8.18x10-37	3.09x10-15
HD95 (norm)					
BraTS	-	5.32x10-80	0.84	0.13	7.64x10-25
WMH	-	0.47	0.76x10-2	0.68	0.26
LiTS	-	0.99	8.19x10-9	0.29x10-3	0.03
BUSI	-	0.35	0.03	3.2x10-4	0.53
Specificity					
BraTS	-	1.97x10-176	3.34x10-7	2.09x10-30	1.29x10-14
WMH	-	0.05x10-2	0.57	0.23	0.42
LiTS	-	2.70x10-9	0.01	0.01x10-2	0.43x10-2
BUSI	-	4.96x10-10	8.27x10-18	9.19x10-51	1.51x10-24

Table S2. Wilcoxon signed rank test values between F_sup and Semi_DA for different (20%, 40%, 60%, 75% and 100%) real data proportions.

Dataset	20%	40%	60%	75%	100%
DSC					
BraTS	9.64x10-48	0.01x10-2	0.01x10-2	3.29x10-7	1.41x10-53
WMH	0.79	0.72	0.05	0.09x10-2	0.04x10-2
LiTS	0.83	0.10	0.68	9.40x10-5	4.28x10-6
BUSI	0.04	0.41	4.15x10-3	4.62x10-11	6.78x10-25
Recall					
BraTS	3.17x10-34	7.08x10-62	2.01x10-65	8.02x10-73	4.71x10-26
WMH	0.68x10-2	0.23	0.92x10-2	0.85	0.05
LiTS	0.70x10-2	0.11	2.93x10-8	1.22x10-5	1.45x10-11
BUSI	2.54x10-21	2.52x10-29	1.92x10-8	1.24x10-5	0.09
HD95 (norm)					
BraTS	4.63x10-26	0.18	3.51x10-179	1.64x10-8	2.76x10-37
WMH	0.01	0.01	0.05	0.12	0.04x10-2
LiTS	0.37	0.90	0.06	0.01	3.56x10-6
BUSI	0.14	0.14	7.82x10-6	4.39x10-6	8.23x10-16
Specificity					
BraTS	5.82x10-130	1.78x10-79	1.46x10-75	7.106x10-42	0.04x10-2
WMH	0.03	0.79	0.34	0.38	0.34
LiTS	0.02x10-2	0.61	2.08x10-8	0.39	0.2x10-2
BUSI	0.25x10-2	6.57x10-10	4.40x10-12	1.25x10-25	1.56x10-42

Ablation study of the proposed lesion simulation:

Table S4 contains the p values from signed rank test between *create_pert* and *create_pert + localise_pert*, followed with comparison between *create_pert + localise_pert* and *create_pert + localise_pert + intensity_blend*. Actual comparison values are shown in the main manuscript Table 4.

Table S3. Wilcoxon signed rank test values between F_sup and Semi_DA for different factors of simulations (1x, 2x and 5x, where x is 100% real-world data).

Dataset	1x	2x	5x
DSC			
BraTS	4.37x10-5	0.22	1.41x10-53
WMH	0.03	0.03	0.04x10-2
LiTS	0.62x10-2	0.15x10-2	4.28x10-6
BUSI	1.25x10-5	1.11x10-9	6.78x10-25
Recall			
BraTS	1.26x10-99	2.76x10-84	4.71x10-26
WMH	0.09x10-2	0.03	0.05
LiTS	0.01x10-2	2.29x10-10	1.45x10-11
BUSI	3.25x10-4	3.76x10-4	0.09
HD95 (norm)			
BraTS	6.22x10-6	0.09	2.76x10-37
WMH	0.12	0.01	0.04x10-2
LiTS	0.24	0.34x10-2	3.56x10-13
BUSI	1.13x10-8	1.46x10-6	8.23x10-16
Specificity			
BraTS	4.25x10-97	2.76x10-58	0.04x10-2
WMH	0.24x10-2	0.97	0.34
LiTS	0.06	3.94x10-5	0.16x10-2
BUSI	2.38x10-13	1.24x10-22	1.56x10-42

Table S4. Wilcoxon signed rank test values between different components of the proposed simulation method.

Method	BraTS	WMH	LiTS	BUSI
DSC				
<i>create_pert</i>	-	-	-	-
<i>+localise_pert</i>	5.84x10-9	0.05	4.30x10-6	0.13x10-2
<i>+intensity_blend</i>	4.58x10-5	0.38	0.51	0.11
Recall				
<i>create_pert</i>	-	-	-	-
<i>+localise_pert</i>	4.12x10-48	0.57	2.39x10-6	4.14x10-5
<i>+intensity_blend</i>	1.04x10-11	0.09	1.86x10-9	5.47x10-3
HD95 (norm)				
<i>create_pert</i>	-	-	-	-
<i>+localise_pert</i>	2.63x10-198	4.80x10-4	3.56x10-13	1.71x10-65
<i>+intensity_blend</i>	0.02x10-2	4.80x10-4	3.56x10-13	0.04
Specificity				
<i>create_pert</i>	-	-	-	-
<i>+localise_pert</i>	3.61x10-90	0.38	0.05	4.82x10-32
<i>+intensity_blend</i>	2.86x10-6	0.34	7.06x10-7	0.14x10-2

Some more visual results comparing F_sup, Semi_DA and Semi_FT

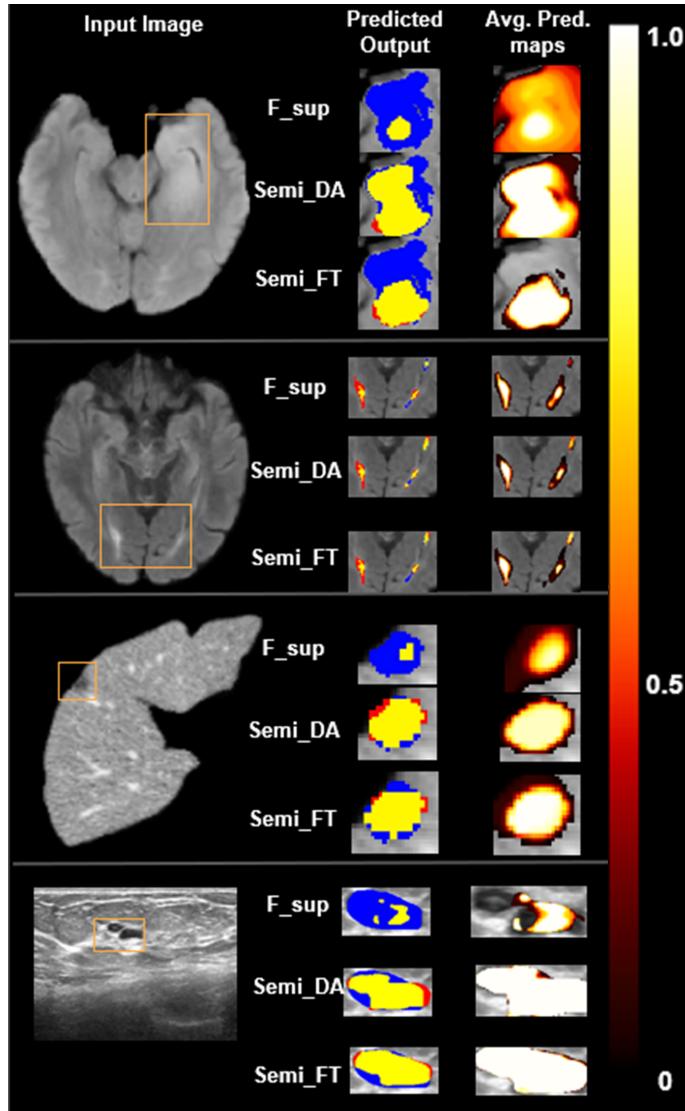


Fig. S3. Comparison of F_{sup}, Semi_DA and Semi_FT lesion segmentation with UNet++ architecture. Input image is shown (left) with F_{sup}, Semi_DA and Semi_FT segmentation results, for BraTS, WMH, LiTS and BUSI (top to bottom respectively). Yellow, blue and red show true positive, false negative and false positive voxels on predicted maps (middle), and average of prediction outputs (softmax values over different blur levels) showing confidence maps (right).

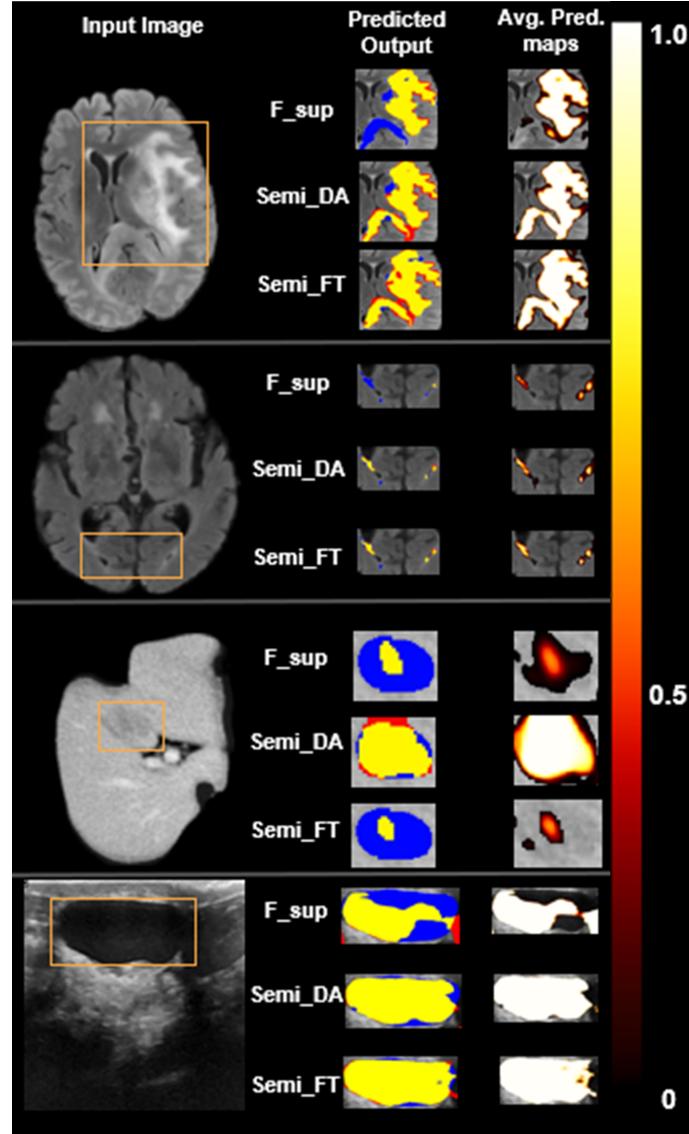


Fig. S4. Comparison of F_sup, Semi_DA and Semi_FT lesion segmentation with UNet++ architecture. Input image is shown (left) with F_sup, Semi_DA and Semi_FT segmentation results, for BraTS, WMH, LiTS and BUSI (top to bottom respectively). Yellow, blue and red show true positive, false negative and false positive voxels on predicted maps (middle), and average of prediction outputs (softmax values over different blur levels) showing confidence maps (right).

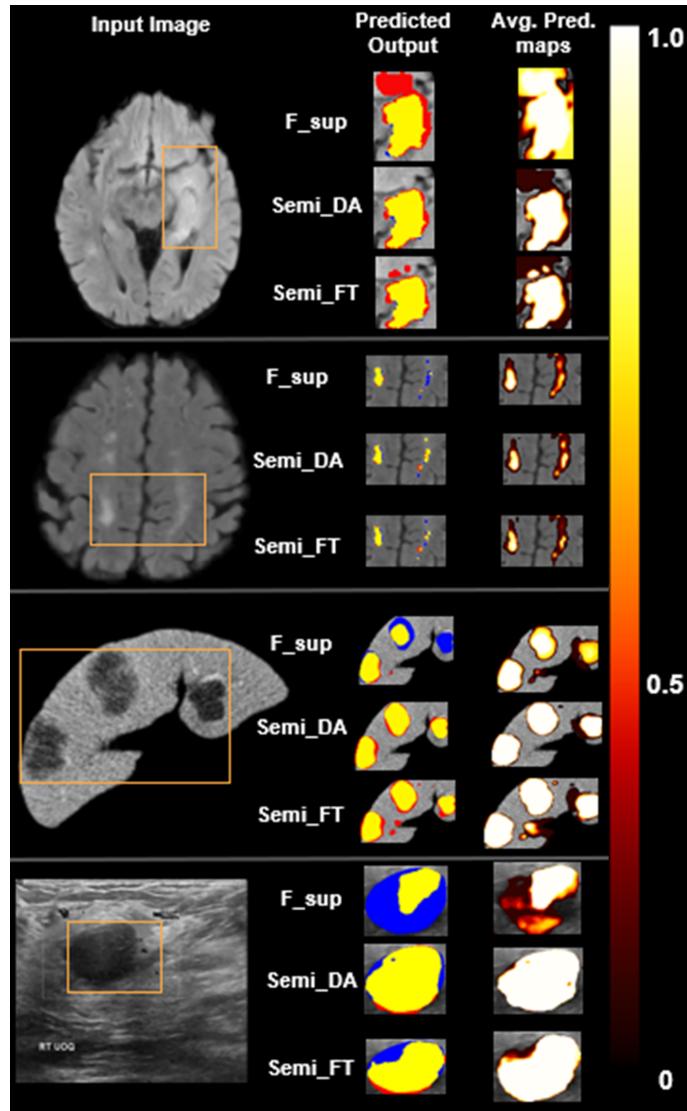


Fig. S5. Comparison of F_{sup}, Semi_DA and Semi_FT lesion segmentation with UNet++ architecture. Input image is shown (left) with F_{sup}, Semi_DA and Semi_FT segmentation results, for BraTS, WMH, LiTS and BUSI (top to bottom respectively). Yellow, blue and red show true positive, false negative and false positive voxels on predicted maps (middle), and average of prediction outputs (softmax values over different blur levels) showing confidence maps (right).

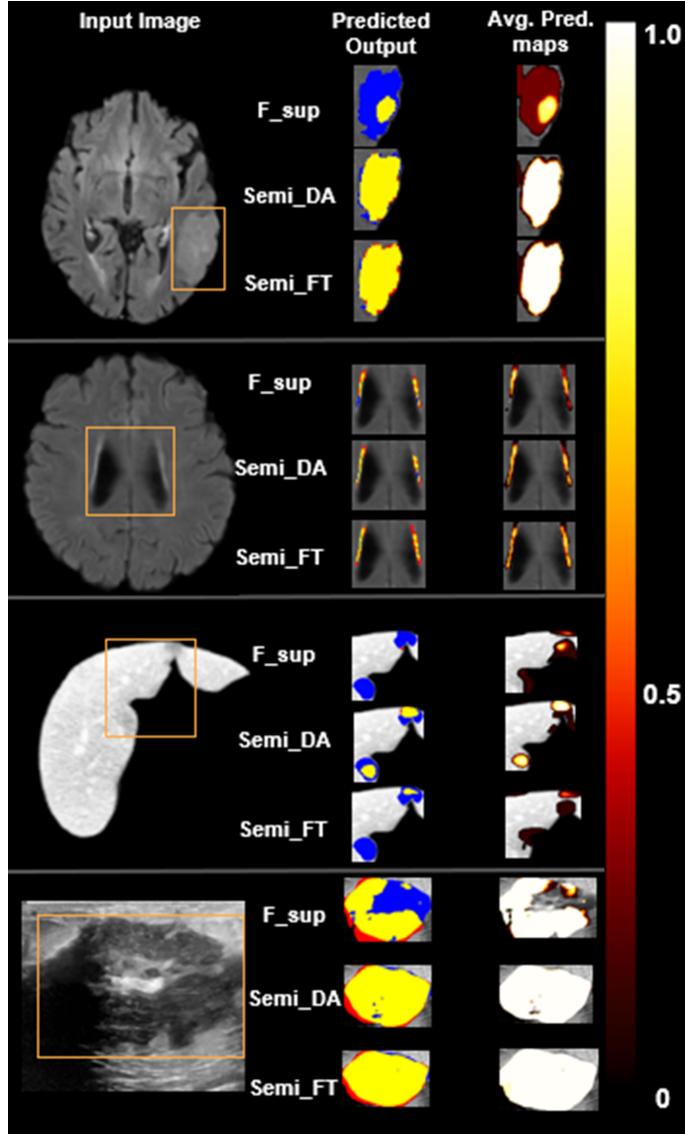


Fig. S6. Comparison of F_{sup}, Semi_DA and Semi_FT lesion segmentation with UNet++ architecture. Input image is shown (left) with F_{sup}, Semi_DA and Semi_FT segmentation results, for BraTS, WMH, LiTS and BUSI (top to bottom respectively). Yellow, blue and red show true positive, false negative and false positive voxels on predicted maps (middle), and average of prediction outputs (softmax values over different blur levels) showing confidence maps (right).

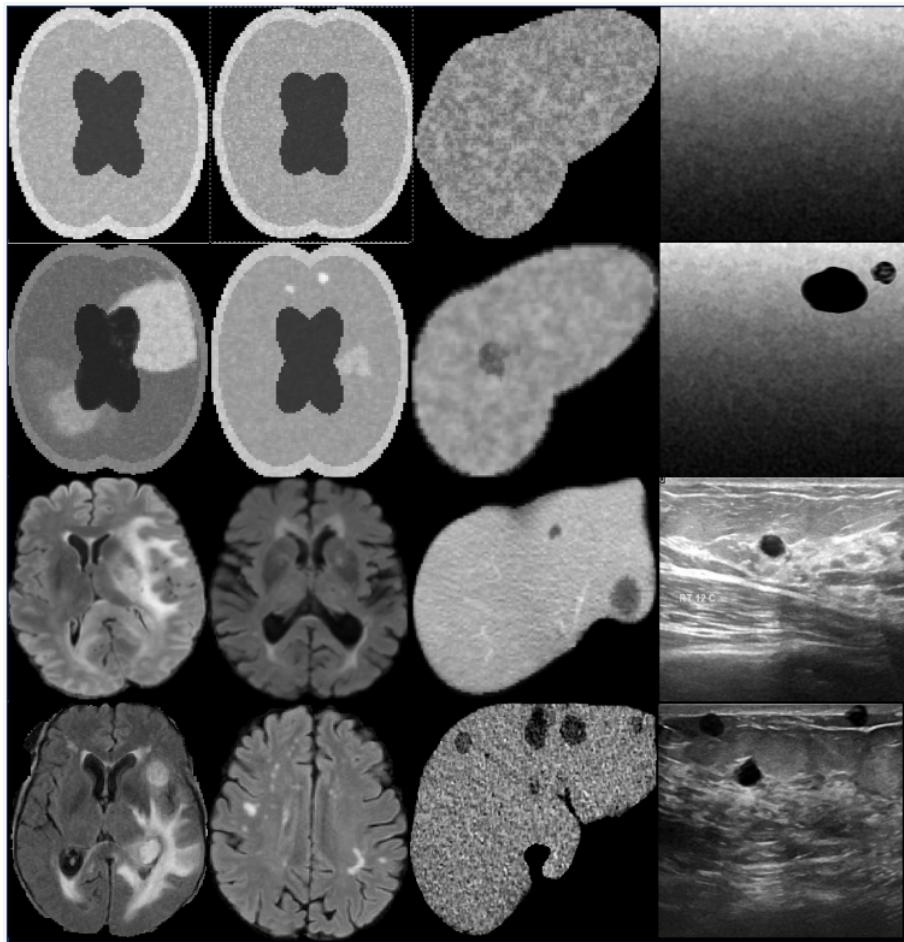


Fig. S7. Examples of simulated and real lesions from different modalities. Top two rows: Generic Simulation maps (GS_{map}) and simulated lesions on the GS_{map} using underlying noise models (Rician, Gaussian and Speckle for MRI, CT and ultrasound scans respectively). Bottom two rows: real-world images and simulated augmented examples.

References

1. Perlin, K.: Improving noise. In: Proc. of the 29th annual SIGGRAPH. pp. 681–682 (2002)
2. Roos-Hoefgeest, S., Roos-Hoefgeest, M., Álvarez, I., González, R.C.: Simulation of laser profilometer measurements in the presence of speckle using perlin noise. Sensors **23**(17) (2023). <https://doi.org/10.3390/s23177624>, <https://www.mdpi.com/1424-8220/23/17/7624>
3. Wikipedia contributors: Perlin noise — Wikipedia, the free encyclopedia (2024), https://en.wikipedia.org/w/index.php?title=Perlin_noise&oldid=1230993513, [Online; accessed 22-July-2024]