Institute of Advanced Research
The University for Innovation

Size Penalty Loss Function

Exploiting exponential function behavior for segmenting stars & lesions

Mohit Joshi[1,2], Dr Kshitij Jadhav[1]

mpjoshi2425@gmail.com, kshitij.jadhav@iitb.ac.in

[1] KCDH IIT Bombay

[2] Institute of Advanced Research

Project Page: Link

May 15, 2026

Contents

Size Penalty Loss Function

Size Penalty Loss Function: Notations

$$\mathcal{L}_{\text{size}} = \exp\left(-\frac{V_g}{\tau_s}\right) \cdot \frac{|V_p - V_g|}{V_g + \varepsilon}$$

τ_s = size-decay constant controlling "how strongly" small lesions are weighted.

ε = small numerical constant for stability.

$$V_p = \sum_{i \in \Omega} p_i$$

V_p = predicted lesion volume (soft sum of sigmoid probabilities)

$$V_g = \sum_{i \in \Omega} y_i$$

V_g = ground-truth lesion volume (voxel count)

Size Penalty Loss Function: Individual Terms

$$W(V_g) = \exp\left(-\frac{V_g}{\tau_s}\right)$$

Continuous Size Weight Term

(exponential term effect)

Interpretation

Strong penalty for tiny lesions.

Moderate penalty.

Negligible penalty for large lesions.

$$E_{rel} = \frac{|V_p - V_g|}{V_g + \varepsilon}$$

Relative Error Term

This measures the "relative mismatch" between predicted and true lesion size.

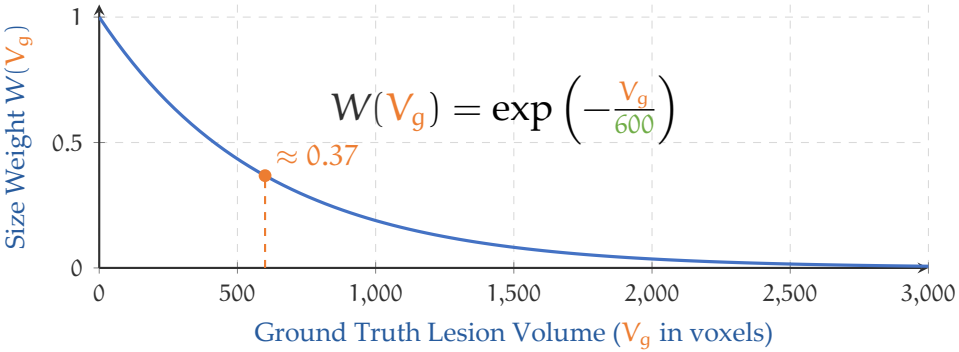
$$\mathcal{L} = \min\left(W(V_g) \cdot E_{rel},\ 10\right)$$

Size Clamping

Clamps the "Error" at 10 prevents excessively large gradients while preserving gradient direction.

Purpose: Prevents Gradients blow off for Small Sub 500 Voxels Segmentation Task due to large Prediction error.

Size Penalty Loss Function: Graph



Size Penalty Loss Function: Key Properties

$$\mathcal{L}_{\text{size}} = \exp\left(-\frac{V_g}{\tau_s}\right) \cdot \frac{|V_p - V_g|}{V_g + \varepsilon}$$

Fully differentiable: No hard thresholding on lesion size.

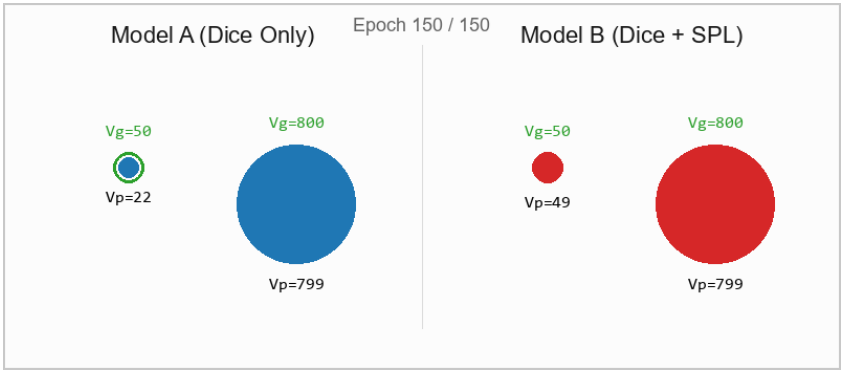
Small-lesion emphasis: Tiny lesions receive stronger supervision.

Scale-adaptive: Penalty smoothly decays with lesion size.

Stable optimization

- sigmoid soft volume
- denominator smoothing
- gradient clamp

Size Penalty Loss Function: Visualisation



Dice Blindspot Simulation: Model A (Dice) vs Model B (Dice + SPL)

Project Page: Link

Training

Training: Methods

Model: Attention UNET 3D (Base Filters = 32)

Input Channels: 3 (ADC, DWI & Mismatch)

Augmentation: Spatial, Cutout, Intensity Augmentation

Loss Functions:

- 1×Dice Loss
- 1×Tversky Loss
- 0.3× "Size Penalty Loss"
- 1×Focal Loss

Optimizer: AdamW optimizer

Early Stopping: Patience 40 epochs, Metric: EMA smoothed composite Val Score

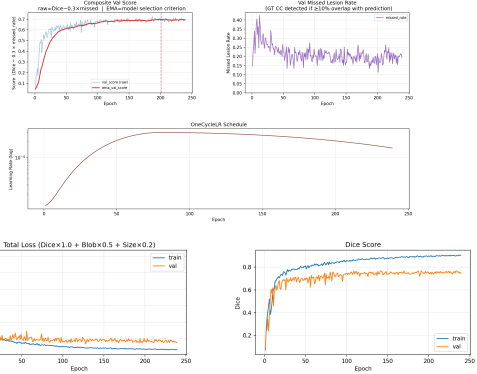
Training Metrics: Batch Size (18), Max Epochs (400), Patch Sampling (Large 128, 128, 32 and Small 64, 64, 32)

Inference Metrics: Sliding window, TTA (8 flip), Min detection Volume (5 Voxel)

Data Sampler: Lesion weighted Loading

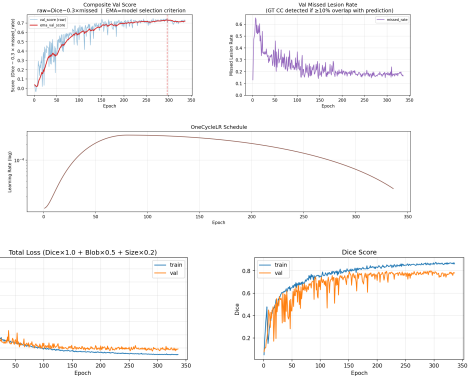
- Tiny (<500 Vox): 5
- Small (500–1000 Vox): 3
- Large (>5000): 2

Training: Proprietary Dataset



Tversky + Dice + Size Penalty Loss Model Training

Training: ISLES Dataset



Focal + Dice + Size Penalty Loss Model Training

Dataset Description

Dataset Description: ISLES 2022

187 Total Number of Dataset

Dataset Split: 80 (Train), 10 (Validation), 10 (Validation)

ADC-DWI: Depth:Width||112-112, Slices||72, Thickness: 2mm

Mean infarct volume per scan: 22.53 ± 44.76 ml (Range: 0 – 482.152 ml).

Mean lesion-wise volume: 2.42 ± 15.80 ml.

Mean number of lesions per scan: 9.30 ± 13.50 (Range: 0 – 126 disconnected lesions).

Temporal Stages: Covers hyper-acute, acute (0–7 days), and sub-acute (1–3 weeks) stages.

Morphology & Distribution: Includes large territorial infarcts, small punctate (micro-embolic) lesions, posterior circulation infarcts, and cases with no visible infarct (n=5). Strong class imbalance between lesion and background voxels.

Dataset Description: Proprietary Dataset

210 Total Number of Dataset

Dataset Split: 80 (Train), 10 (Validation), 10 (Validation)

ADC-DWI: Depth:Width||256-256, Slices||25 (padded 7 0s to 32), Thickness: 6mm

Metric	Mean (Std)	[Min, Max]
Total scan infarct volume (ml)	32.799 ($\pm$ 55.172)	[0.131, 370.492]
Total scan infarct volume (voxels)	6,772 ($\pm$ 11,392)	[27, 76,499]
Number of unconnected ischemia per scan	5.900 ($\pm$ 6.091)	[1, 36]
Lesion-wise infarct volume (ml)	5.559 ($\pm$ 25.048)	[0.005, 369.955]
Lesion-wise infarct volume (voxels)	1,148 ($\pm$ 5,172)	[1, 76,388]
Largest single lesion per scan (ml)	29.536 ($\pm$ 54.568)	[0.131, 369.955]
Lesion bounding box spatial extent (voxels)	59,609 ($\pm$ 65,514)	[30, 523,089]

Observations

Observations: Proprietary Dataset

D+T+S is strongest for proprietary sub-mL lesions. For <1 mL lesions, **D+T+S** improves PQ, SQ, RQ, Dice, and volume error against **D+T+F**.

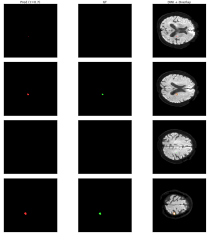
D+T+S reduces missed lesions versus **D+T+F** on proprietary Dataset. Mean missed lesions improve from 6.15 to 6.03, and **D+T+S** wins 4/5 folds.

D+T+S improves recognition-side instance behavior versus **D+T+F**. RQ improves from 0.3873 to 0.4022; PQ improves from 0.2903 to 0.2940; ASSD improves from 3.9814 to 3.5665.

Size Bin	Loss Function Combination	Number of Samples	Panoptic Quality (PQ)	Segmentation Quality (SQ)	Recognition Quality (RQ)	DICE
<1 ml	D+T+S	17	0.2961	0.4355	0.4286	0.5119
<1 ml	D+T+F	17	0.2885	0.4140	0.4033	0.4948
<1 ml	T+F+S	17	0.2501	0.3669	0.3549	0.4919
<1 ml	F+S+D	17	0.2225	0.3474	0.3034	0.4428

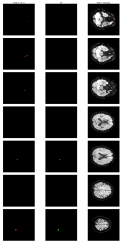
Observations: Proprietary Dataset

D+T+S Loss Combination



Ground Truth: 0.63ml lesion
Predicted Volume: 0.73ml
DICE: 0.733

D+T+F Loss Combination



Ground Truth: 0.63ml lesion
Predicted Volume: 1.49ml
DICE: 0.5434

Observations: ISLES Dataset

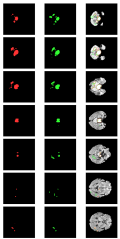
F+S+D is the best ISLES size-penalty configuration. It has the best Dice, volume error, PQ, SQ, RQ, and ASSD on the single split.

All ISLES size-penalty combinations reduce false positives versus **D+T+F**. **D+T+S** gives the largest reduction, while **F+S+D** gives the best full metric balance.

Loss Function Combination	Size Penalty Loss Present	Panoptic Quality (PQ)	Segmentation Quality (SQ)	Recognition Quality (RQ)	DICE
F+S+D	Yes	0.3882	0.6801	0.5074	0.7774
D+T+S	Yes	0.3389	0.6596	0.4440	0.7514
D+T+F		0.3486	0.6596	0.4712	0.7503
T+F+S	Yes	0.3600	0.6791	0.4999	0.7472

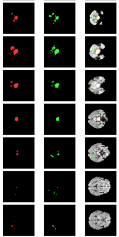
Observations: ISLES Dataset

F+S+D Loss Combination



Ground Truth: 24.33ml lesion
Predicted Volume: 21ml
DICE: 0.8412

D+T+F Loss Combination



Ground Truth: 24.33ml lesion
Predicted Volume: 22.98ml
DICE: 0.8287

Observations: Generalized

S+T and **S+D** are strong double combinations. They are among the highest Dice models and remain stable, but they are not superior on all panoptic metrics.

The best size-penalty partner is dataset-dependent. **D+T+S** is strongest for the sub <1ml Stroke lesions; **F+S+D** is strongest on ISLES.

Size penalty cannot stand alone. S alone collapses spatial segmentation and produces extreme false positives.

Size penalty is best treated as an auxiliary volume regularize. It needs Dice/Tversky/Focal partners to provide spatial gradients.

D+T+F remains better than **D+T+S** on lesions above 10ml, False Positives minimization, and Segmentation Quality.

References

References

Floran Kofler. **Panoptica Metrics**.
https://arxiv.org/pdf/2312.02608

Muhammad Febrian Rachmadi. **Instance Level Loss function**.
https://doi.org/10.1016/j.compbiomed.2024.108414

Ezequiel de la Rosa. **DeepISLES**.
https://www.nature.com/articles/s41467-025-62373-x

Hoel Kervadec. **Constrained CNN loss Function**.
https://openreview.net/pdf?id=BkIBHb2sG

ISLES'22 SEALS.
https://github.com/Tabrisrei/ISLES22_SEALS

Tianyi Ren. **ISLES'24 Preprocessing challenge**.
https://arxiv.org/pdf/2505.18424#page=3.78