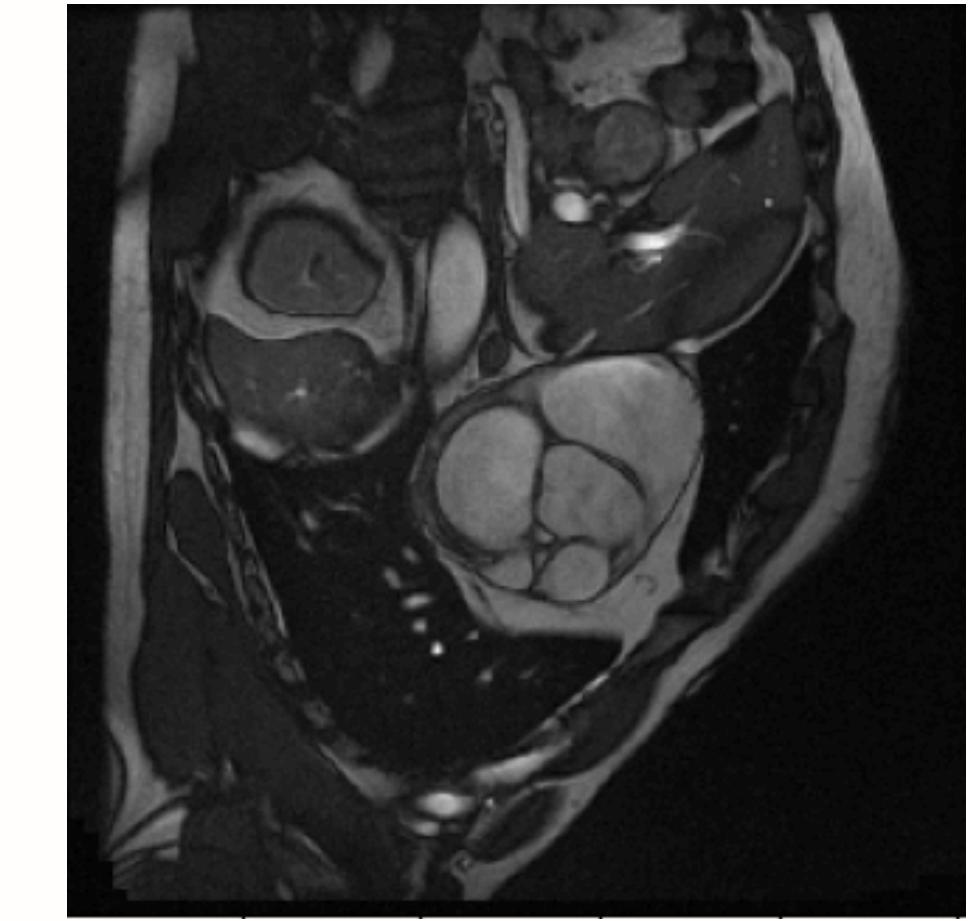
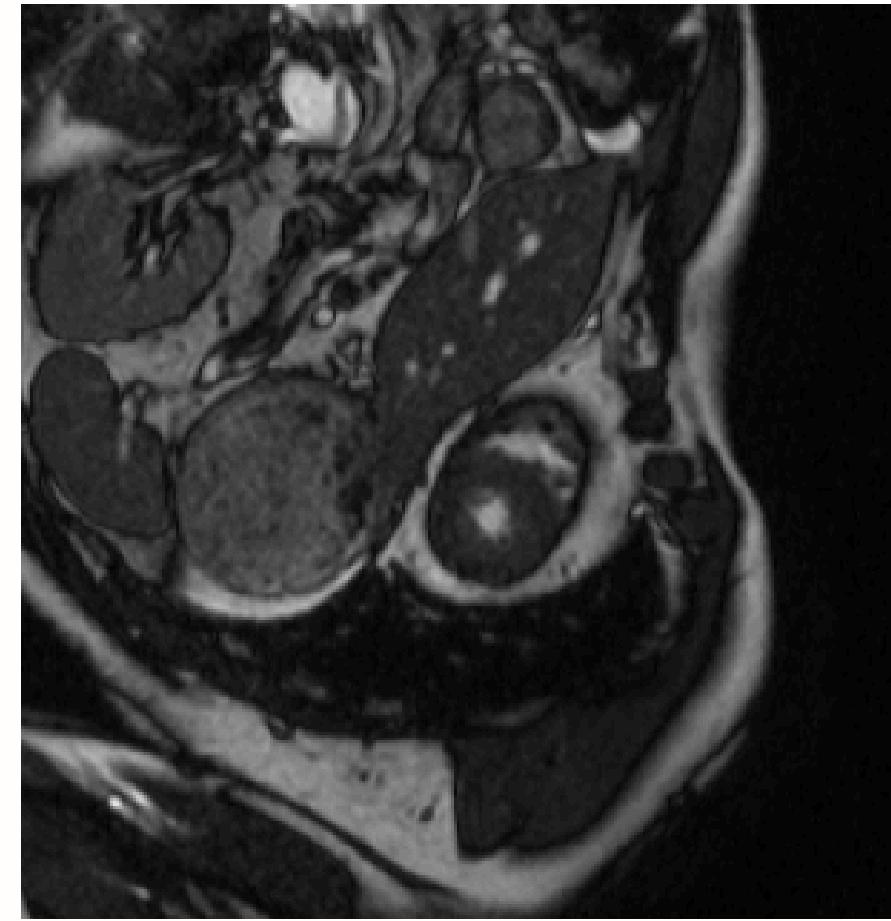




UCD School of Medicine
Scoil an Leighis UCD

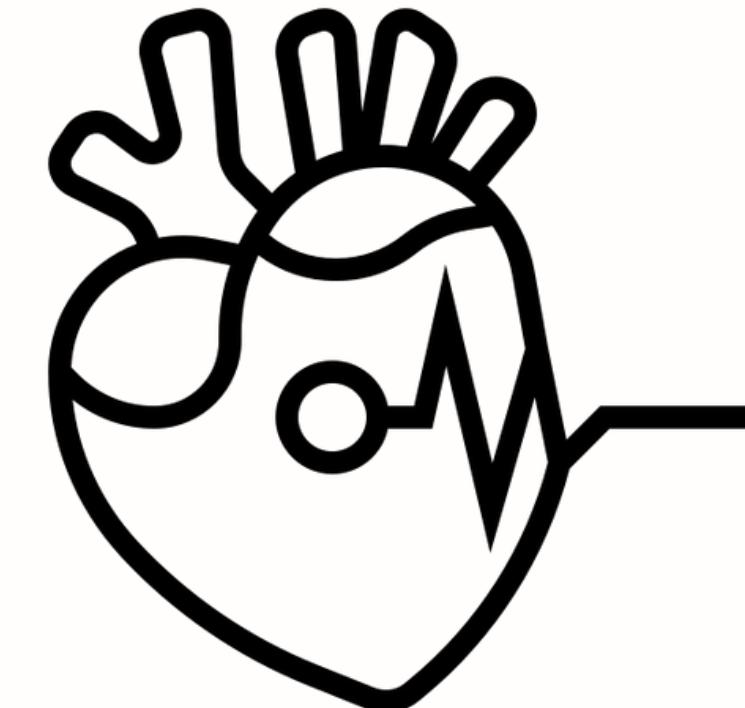
Right Ventricle Segmentation in Cardiac MRI

Anastasiia Deviataieva
Ehab Patrick Issa
Dylan Forde
Dónal Heelan
Julie Sanchis



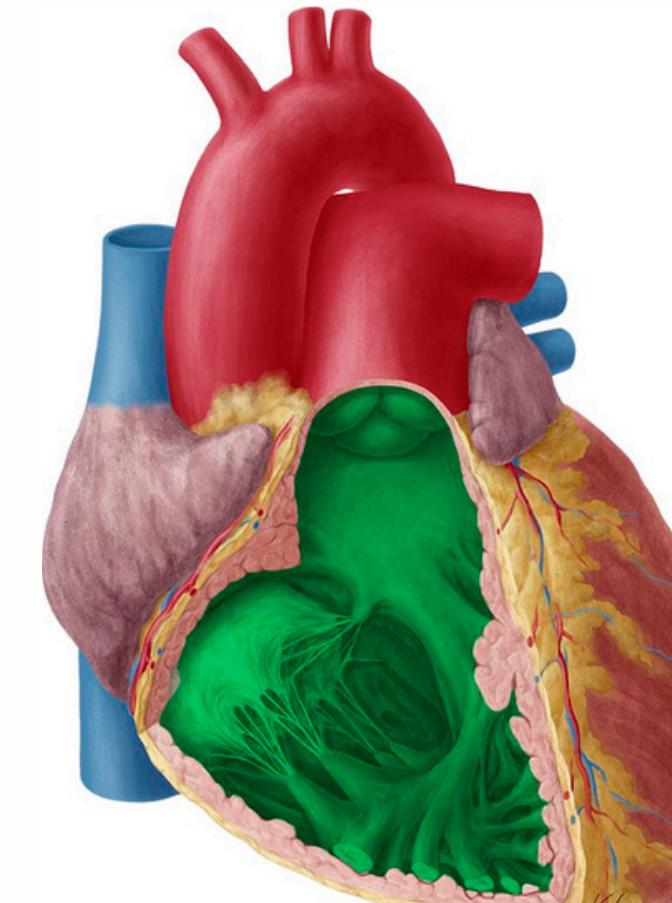
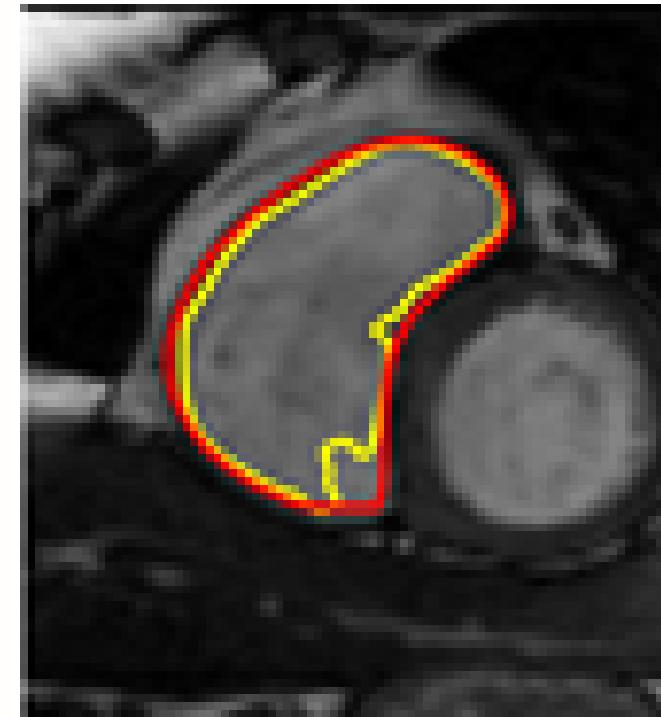
Introduction

- Computer vision has revolutionised medical image analysis – early disease detection and diagnosis
- Cardiac MRI – gold standard for capturing the structure of the heart (Tseng et al, 2016)
- The field has shifted from traditional segmentation methods to machine learning (ML) techniques



Right Ventricle Segmentation

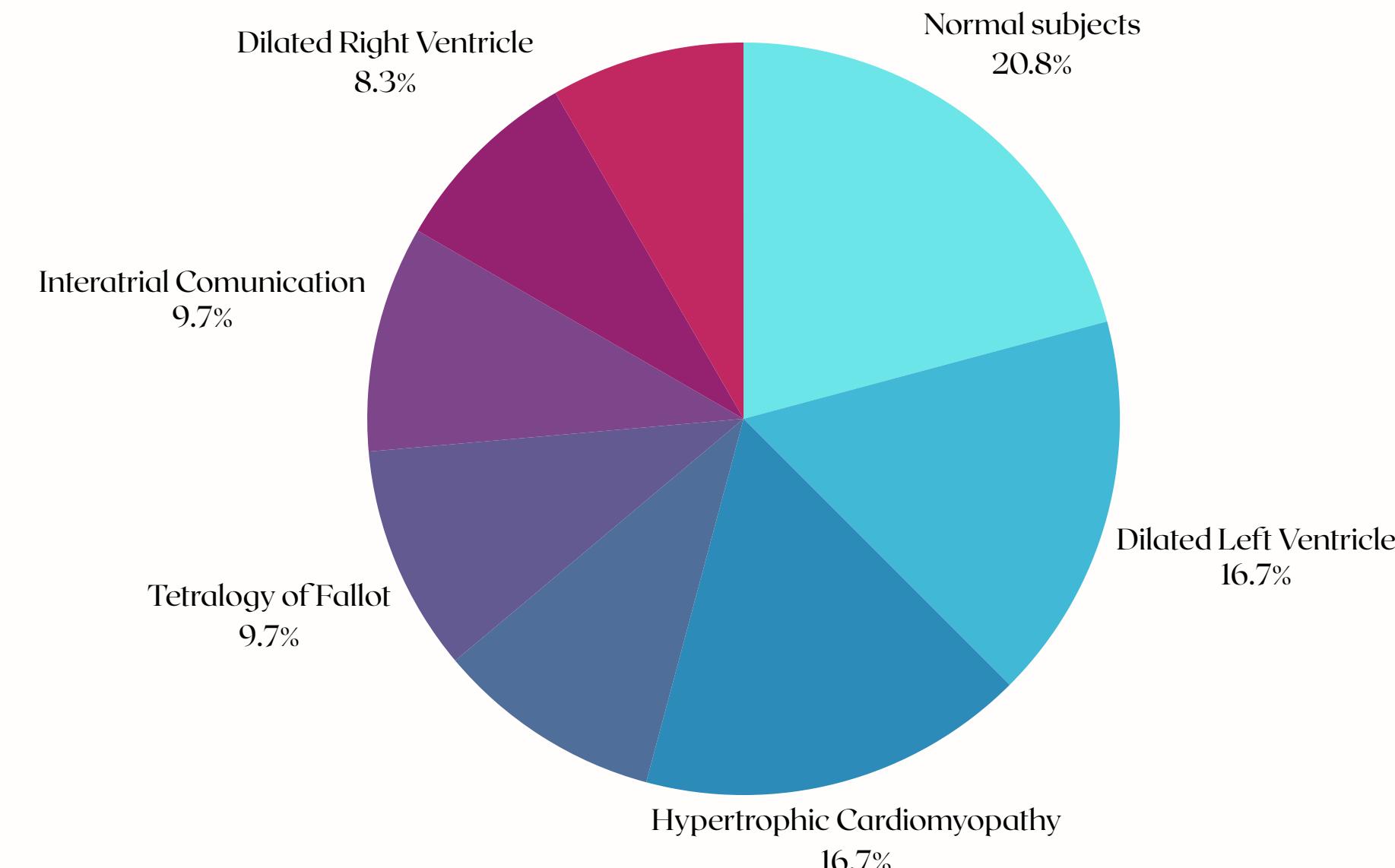
- Historically segmentation tasks have focused on the left ventricle – considered more diagnostically relevant and geometrically easier to segment (Hamosh & Cohn, 1971)
- However segmenting the right “forgotten” ventricle enables early detection of diseases including congenital arrhythmia and right ventricular cardiomyopathy (Tretter & Redington, 2018)



The M&Ms - 2 Dataset

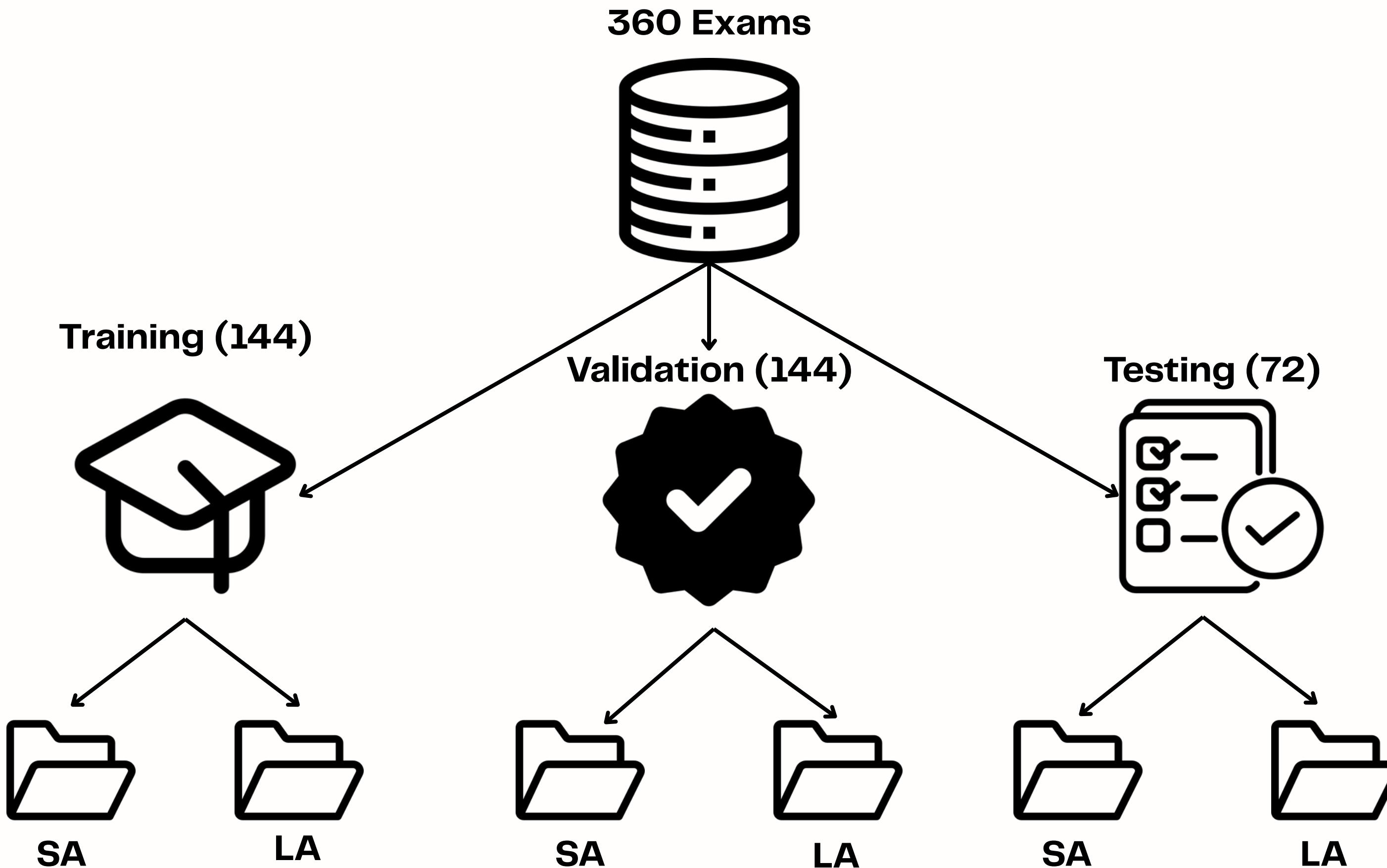
University of Barcelona

- Curated a dataset of 360 cardiac MRIs from scans performed across 3 healthcare centers
- Proposed a challenge asking participants to develop ML models for right ventricle segmentation
- Inspired numerous papers which apply different models to the segmentation task





DataLoaders



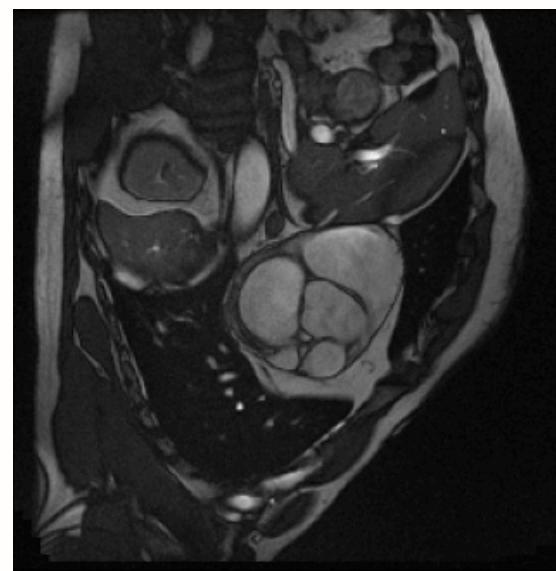
Data Preparation

Augmentation

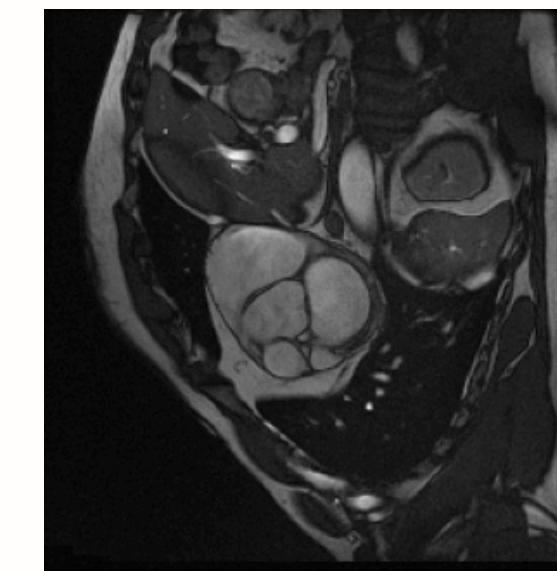
- Training data augmented to prevent overfitting
- Random chance that each image could be modified using TorchIO augmentations

Preprocessing

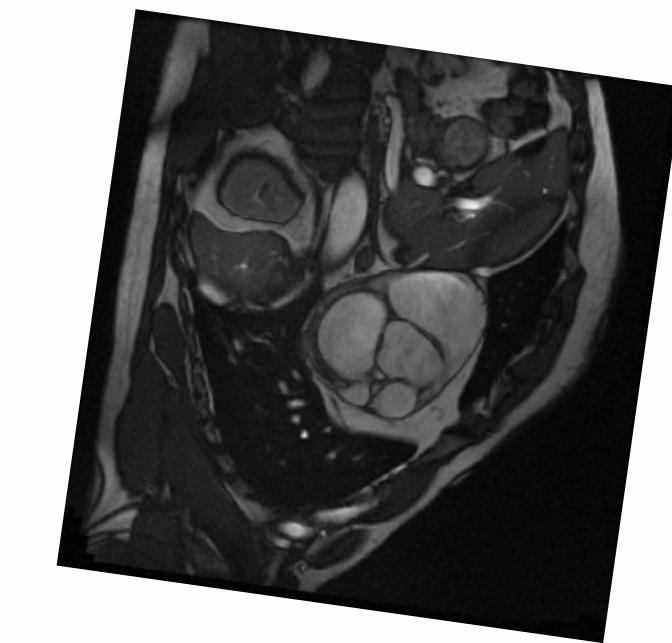
- Scans cropped/padded to uniform size of 256 x 256 x 1
- Histogram of scans standardised using Z-score normalisation
- Crucial due to the heterogeneity of the input data



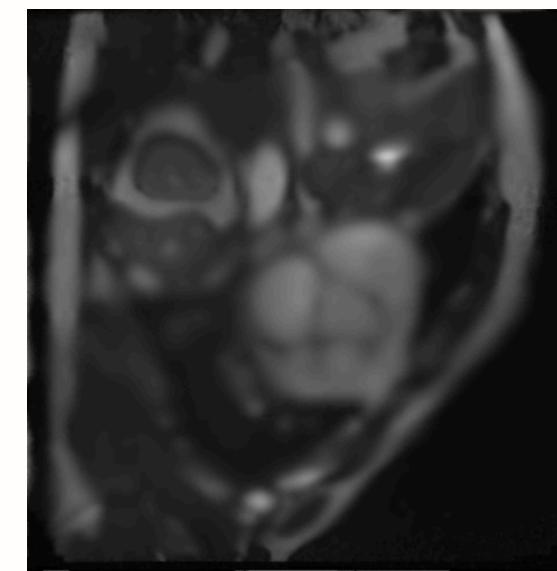
1. Original Image



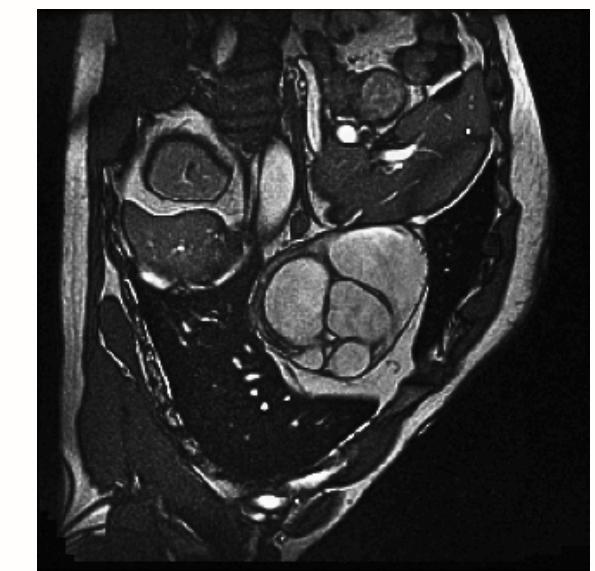
2. RandomFlip



3. RandomAffine



4. RandomNoise



5. RandomGamma



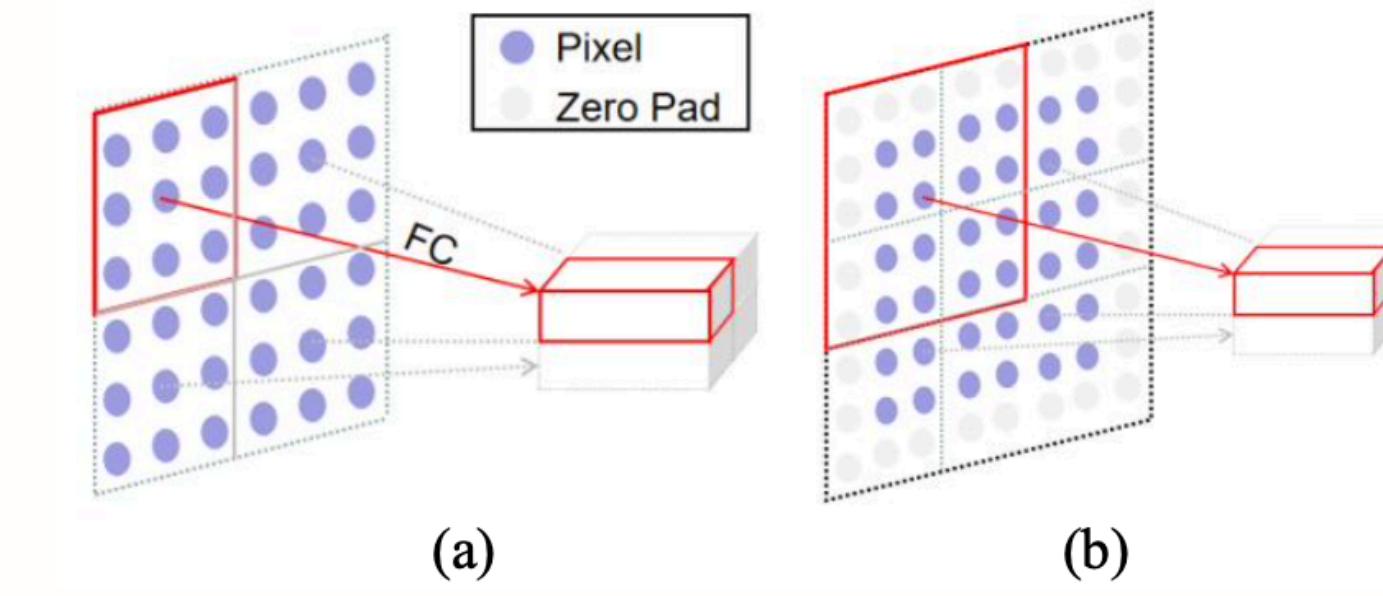
The Segmentation Model

A compact 3D segmentation model inspired by SegFormer3D by Perera, Navard and Yilmaz (2024).

Pipeline: 3D MRI → 4-stage Transformer encoder → all-MLP decoder → RV segmentation mask (upsampled to original input size).

Encoder

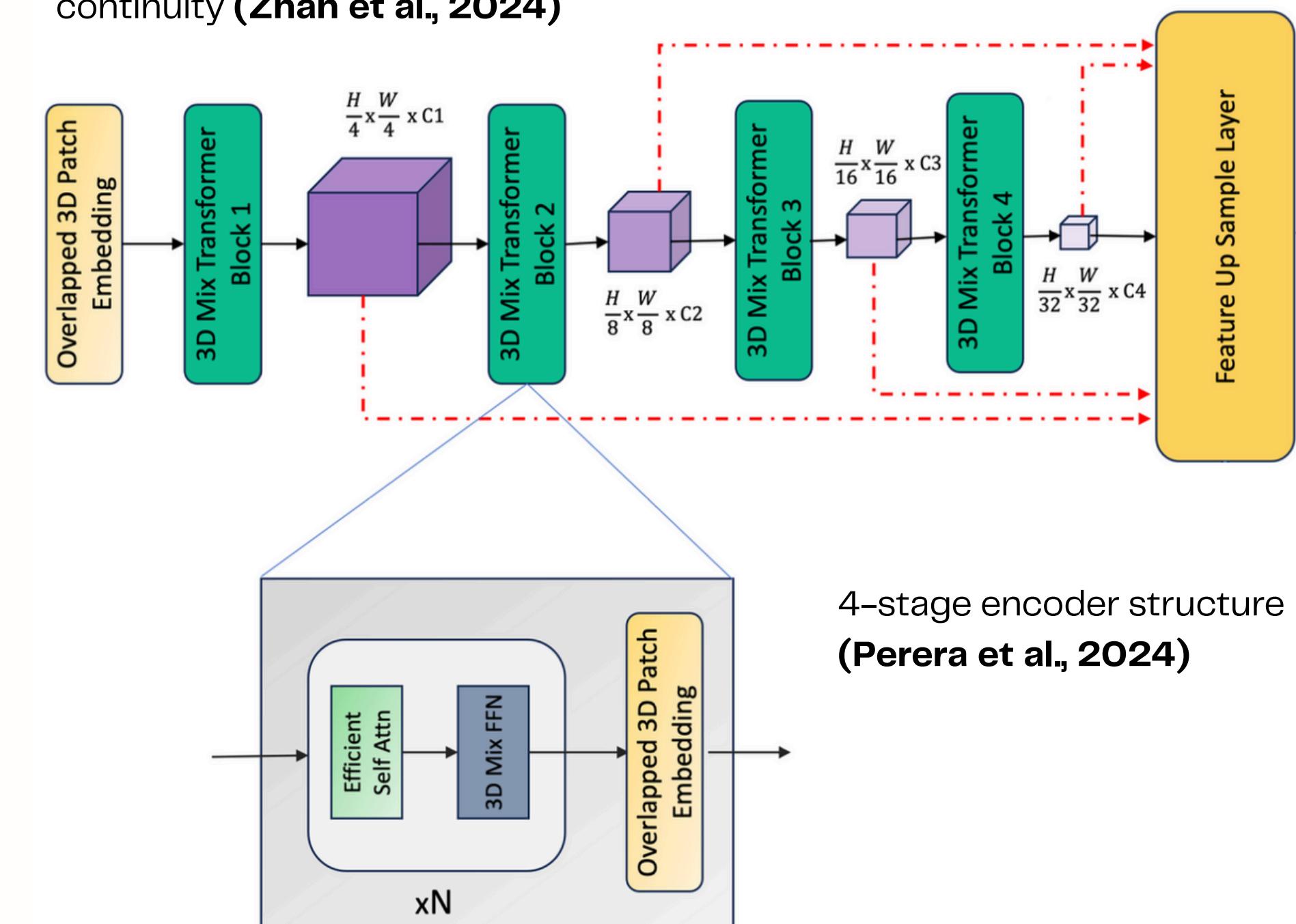
- 4-stage MiT pyramid: channels $32 \rightarrow 64 \rightarrow 160 \rightarrow 256$.
- OverlapPatchEmbed3D: $3 \times 3 \times 3$, stride $(2,2,1)$ + LN.
- Blocks ($\times 2/\text{stage}$):
 - LN \rightarrow MHSA (full) \rightarrow +res;
 - LN \rightarrow MLP ($r=4$) \rightarrow +res.
- Regularization: DropPath; no positional encodings.
- Shapes (D=1): $\sim 1 \times 128 \times 256 \rightarrow \dots \rightarrow 1 \times 16 \times 256$.
- Diff vs Perera et al., (2024): no efficient attention / no Mix-FFN.



(a)

(b)

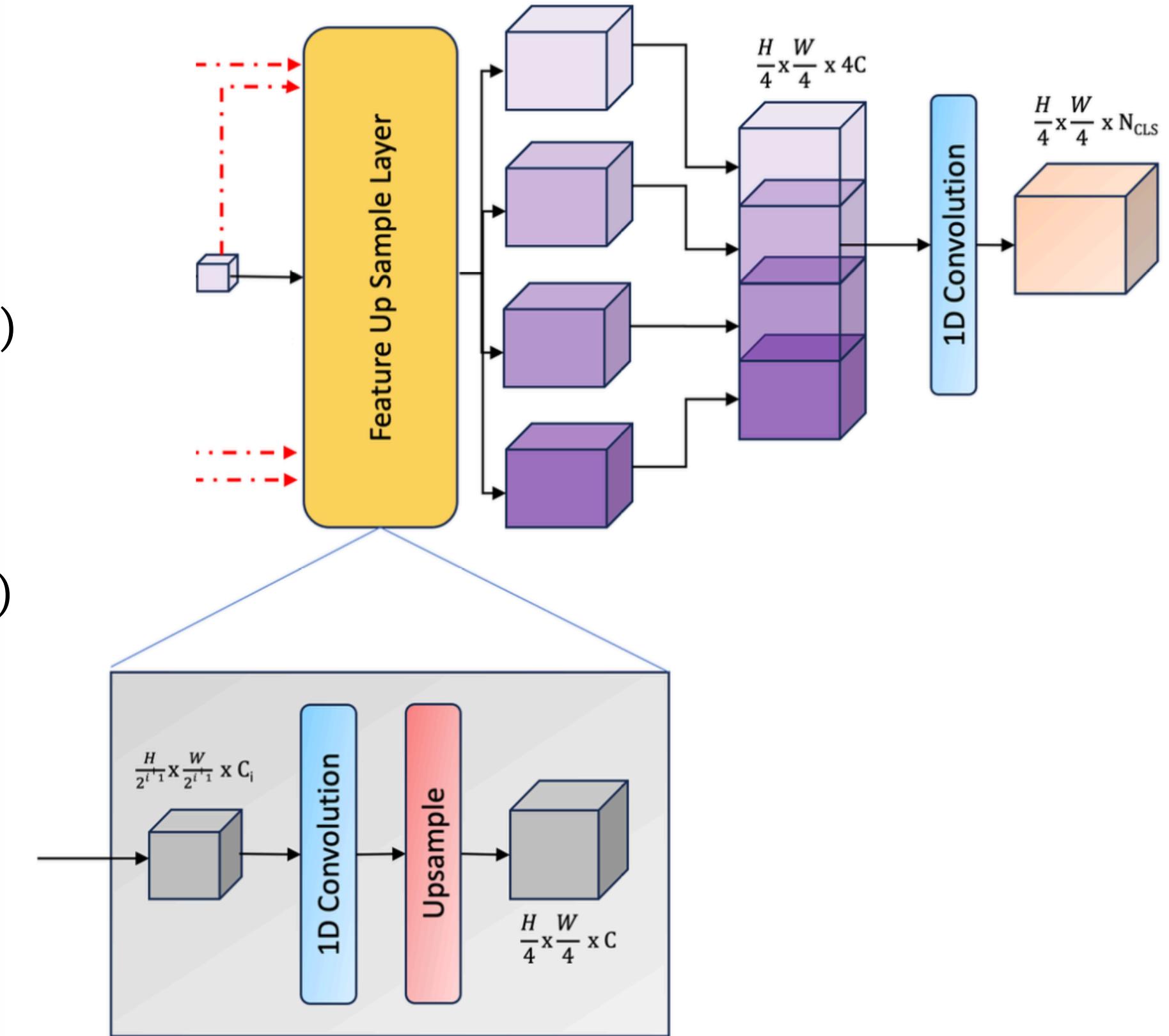
Overlapping (b) vs. non-overlapping patches (a); overlap preserves border continuity (**Zhan et al., 2024**)



4-stage encoder structure
(**Perera et al., 2024**)

Decoder

- Inputs: 4 encoder maps (32/64/160/256 ch)
- Project: each via $1 \times 1 \times 1 \rightarrow 128$ ch
- Upsample: trilinear to Stage-1 size
- Fuse: concat $\rightarrow 512$ ch
- Head: $1 \times 1 \times 1$ $512 \rightarrow 128 \rightarrow 1$ (BN/ReLU/Dropout)
- Output: logits \rightarrow upsample to input size
- Design: no UNet blocks; MLP-style fusion



Decoder All-MLP fusion (**Perera et al., 2024**)



Training: Losses, Metrics, Logging

- CombinedLoss: $2 * \text{DiceLoss} + \text{Focal}(\text{BCEWithLogits})$ ($\alpha=0.5, \gamma=2.0$).
- Dice: sigmoid to logits \rightarrow threshold 0.5 \rightarrow overlap over (D,H,W), averaged across batch.
- Focal: built on BCEWithLogitsLoss, down-weights easy examples.
- TensorBoard: logs Loss/train and Loss/val per epoch.

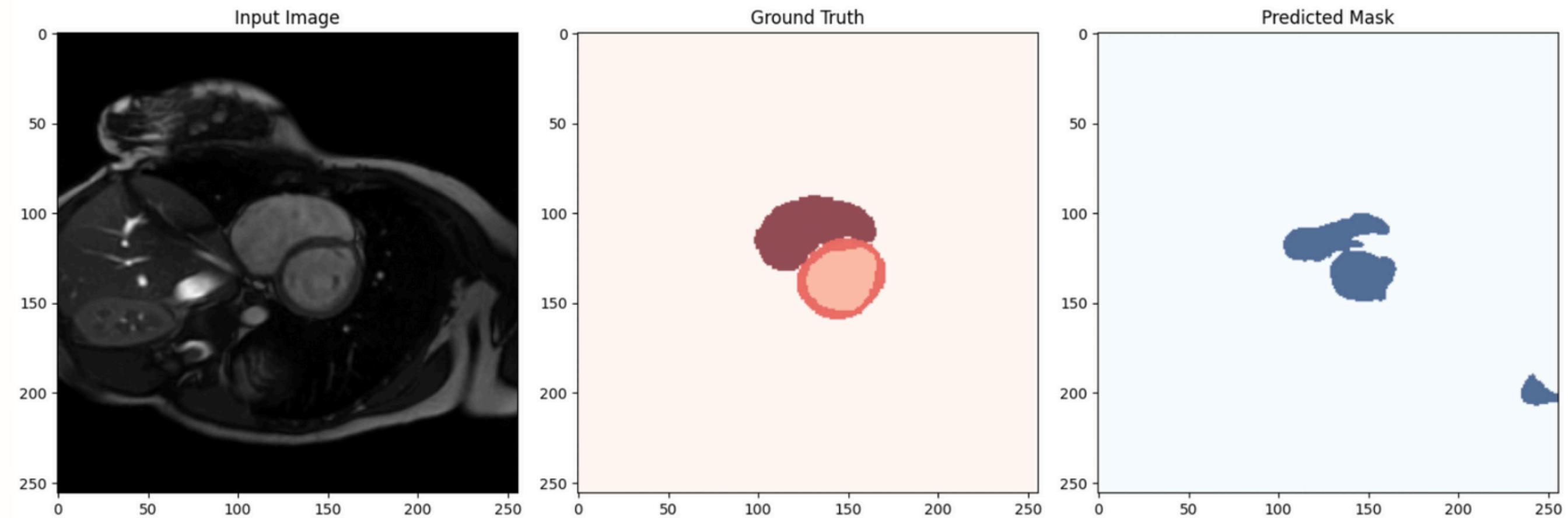


Training: Training Loop & Optimization

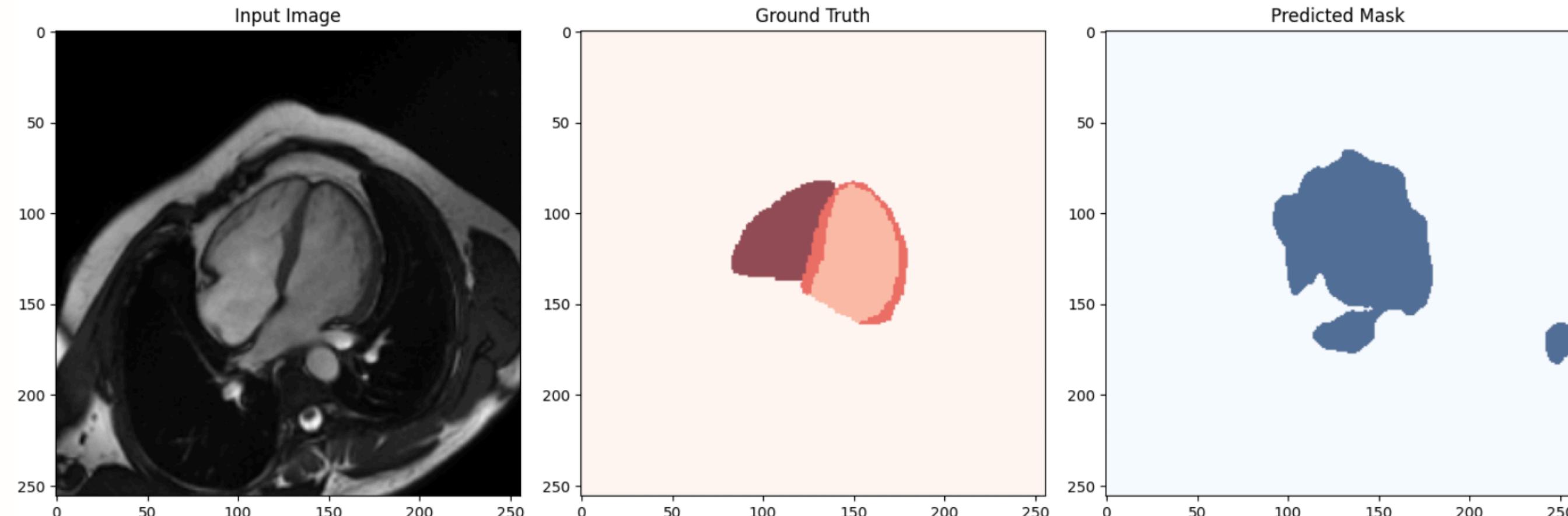
- SA and LA are trained separately on SegFormer3D (quasi-2D setup: `img_size=(256,256,1)`), with `batch_size=1`.
- Optimizer: AdamW($\text{lr}=1\text{e}-4$).
- LR scheduler: ReduceLROnPlateau – halves LR after 3 stagnant epochs (on val loss)
- Gradient accumulation (4 steps) – simulates a larger batch under memory limits.
- Early stopping: stop after 3 epochs with no val loss improvement.
- Checkpointing: saves the best model (lowest val loss).
- Epoch flow (50 epochs): `model.train()` → train with backprop & accumulation; `model.eval()` → validation under `no_grad`.

Training: Predictions visualisation

- SA model



- LA model

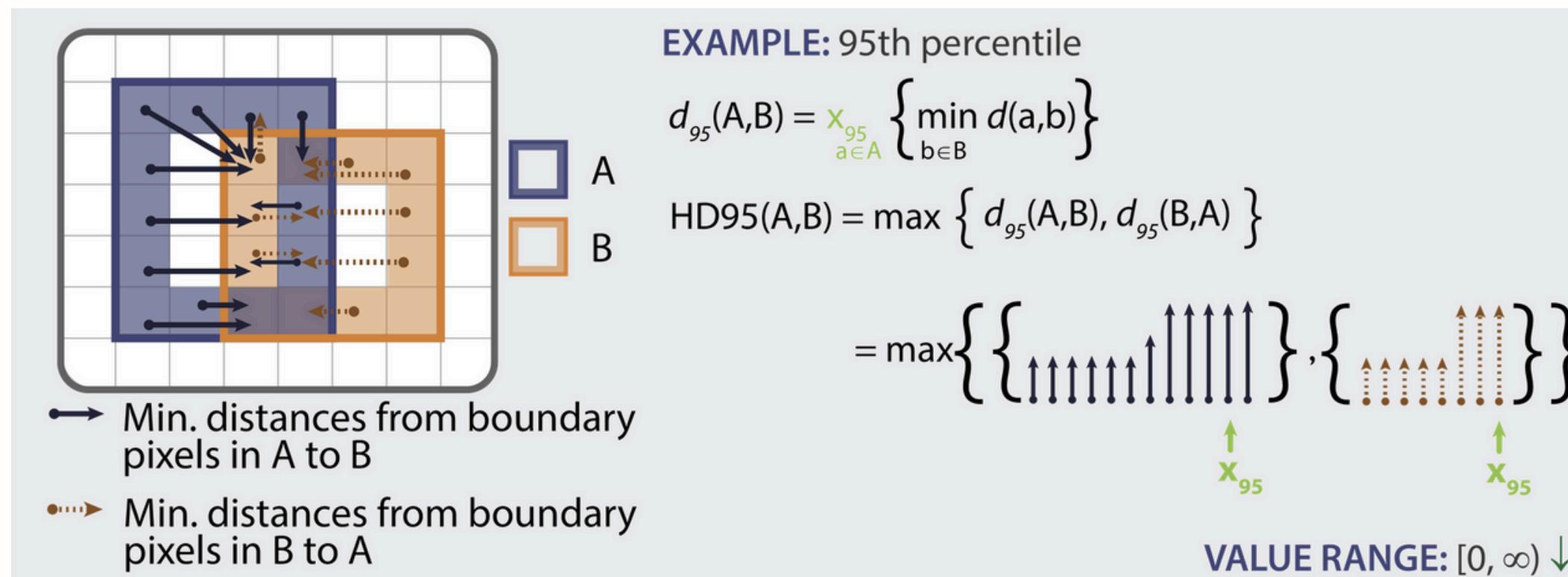
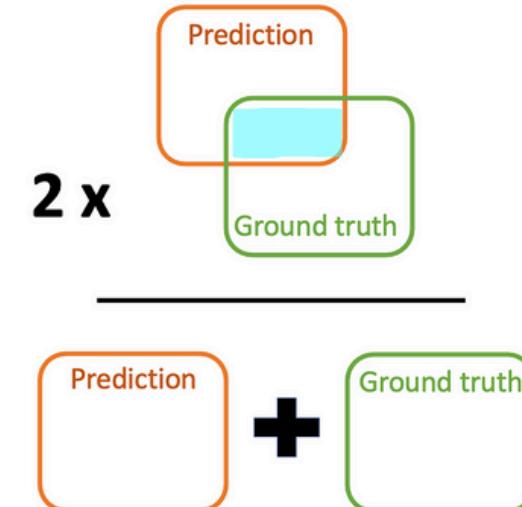


Evaluation

Evaluation Metrics Used

- **Dice score**
- **Hausdorff Distance**

$$\text{Dice} = \frac{2 \times \text{Area of overlap}}{\text{Total area}} =$$



Evaluation

Evaluation Metrics Used

- Precision
- Recall
- F1 score

		POSITIVE	NEGATIVE
ACTUAL VALUES	POSITIVE	TP	FN
	NEGATIVE	FP	TN

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F1 Score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Evaluation

Results

Metrics	SA model	LA model
Dice score	0.5568	0.8964
Hausdorff distance	76.8162	62.4874
Precision	0.4190	0.5264
Recall	0.4986	0.9197
F1 score	0.4327	0.6547



Future Work - Combine SA& LA (Multiview/2.5D)

Current approach: Two separate SegFormer3D models for SA and LA views.

Limitation: No cross-view context , each model only sees one orientation.

Option A (2.5D): stack neighbouring slices and the orthogonal view as channels for more spatial context without changing input size

Option B (late fusion): keep separate SA/LA encoders; add a small fusion head to combine their logits before thresholding.



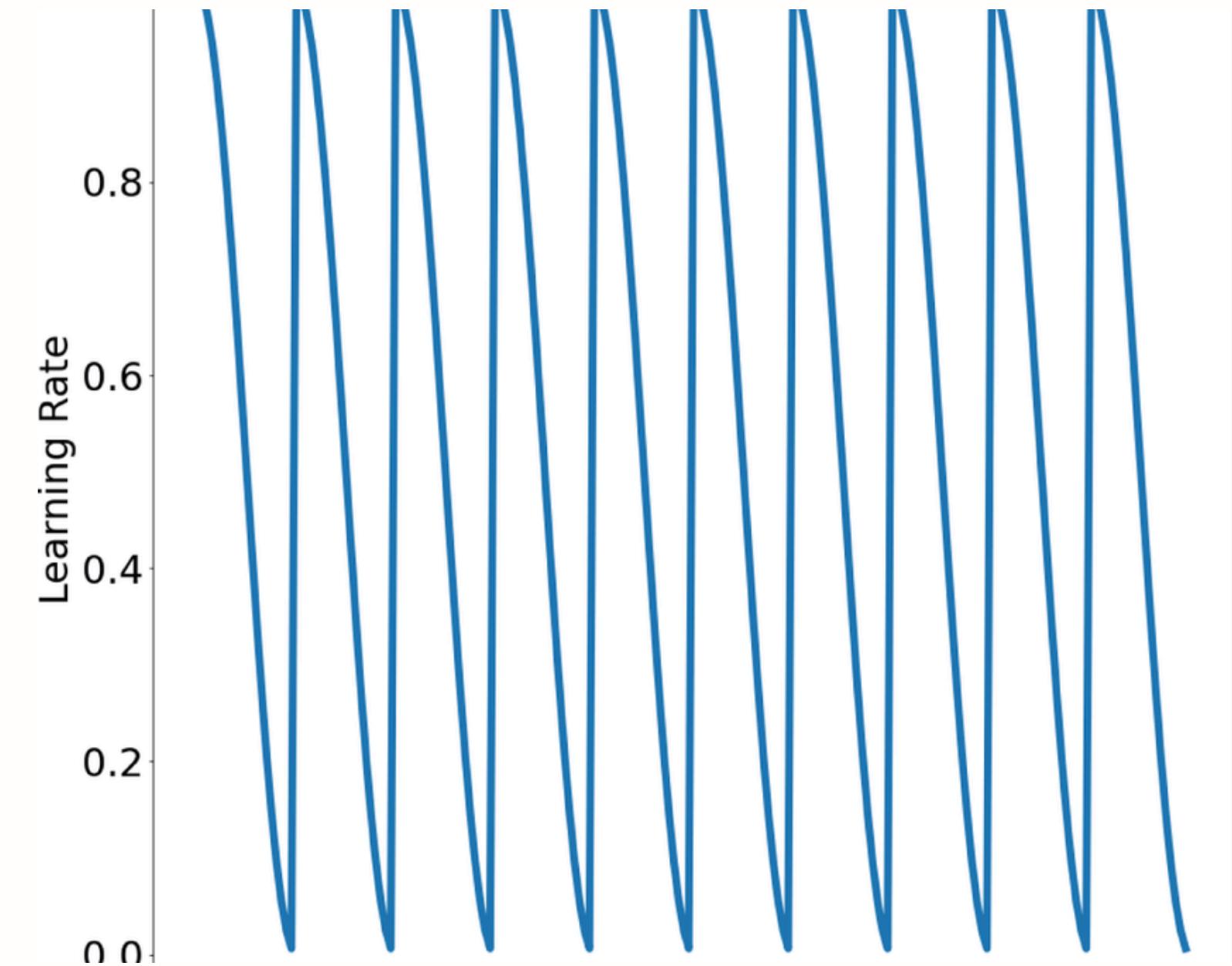
Future Work- Training: Cosine Annealing + Restarts

Current setup: AdamW + ReduceLROnPlateau may lead to reactive learning rate drops.

Limitation: May converge too quickly or get stuck in local minima for non-pretrained models.

Cosine annealing: Smooth, cyclic decay of learning rate, acts as warm restarts.

With restarts: Periodically reset LR to a higher value and encourages escaping local minima and better generalisation.





Future Work - Boundary Aware Loss

Current loss = Dice + Focal which optimises region overlap, not boundary precision.

Limitation: high Dice can still occur with poor contour alignment (large HD95).

Boundary loss: Uses distance transform of ground truth to penalise errors near edges more heavily.

Alternative: Hausdorff distance loss can explicitly minimise largest boundary deviations.

Expected benefit: Reduces sharp misalignments and improves contour accuracy, important for clinical measurements.



Future Work - Interpretability (MC Dropout + Grad-CAM)

Monte-Carlo dropout: Keep dropout active at inference, run K forward passes and generate uncertainty maps highlighting low-confidence regions.

Grad-CAM: Visualises the image regions that most influence the model's predictions, allowing us to verify whether it is focusing on the right ventricle borders or being distracted by cavities or background, and to compare attention patterns between SA and LA views.

References

- Tseng, W. Y. I., Su, M. Y. M., & Tseng, Y. H. E. (2016). Introduction to Cardiovascular Magnetic Resonance: Technical Principles and Clinical Applications. *Acta Cardiologica Sinica*, 32(2), 129. <https://doi.org/10.6515/ACS20150616A>
- Hamosh, P., & Cohn, J. N. (1971). Left ventricular function in acute myocardial infarction. *The Journal of Clinical Investigation*, 50(3), 523–533. <https://doi.org/10.1172/JCI106521>
- Tretter, J. T., & Redington, A. N. (2018). The Forgotten Ventricle?: The Left Ventricle in Right-Sided Congenital Heart Disease. *Circulation: Cardiovascular Imaging*, 11(3). https://doi.org/10.1161/CIRCIMAGING.117.007410/SUPPL_FILE/CIRCCVIM_CIRCCVIM-2017-007410-T_SUPP4.PDF
- Perera, S., Navard, P. and Yilmaz, A., 2024. SegFormer3D: an Efficient Transformer for 3D Medical Image Segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW 2024), 16–22 June 2024, Seattle, WA, USA. Los Alamitos, CA, USA: IEEE Computer Society, pp. 4981–4988. doi:10.1109/CVPRW63382.2024.00503.
- Zhan, X., Zhang, C., Wang, Z., Han, Y., Xiong, P. and He, L. (2024) FasterPest: a multi-task classification model for rice pest recognition. *IEEE Access*, 12, pp. 167845–167855. doi:10.1109/ACCESS.2024.3485616.
- Kervadec, H., Bouchtiba, J., Desrosiers, C., Granger, E., Dolz, J. and Ayed, I.B., 2019, May. Boundary loss for highly unbalanced segmentation. In *International conference on medical imaging with deep learning* (pp. 285-296). PMLR.
- Gal, Y. and Ghahramani, Z., 2016, June. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning* (pp. 1050-1059). PMLR.
- Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D., 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision* (pp. 618-626).
- Kim, Y., Kang, D., Mok, Y., Kwon, S. and Paik, J., 2023. Bidirectional meta-Kronecker factored optimizer and Hausdorff distance loss for few-shot medical image segmentation. *Scientific reports*, 13(1), p.8088.
- DeGrave, A.J., Janizek, J.D. and Lee, S.I., 2021. Course Corrections for Clinical AI. *Kidney360*, 2(12), pp.2019-2023.
- Müller, D., Soto-Rey, I. and Kramer, F., 2022. Towards a guideline for evaluation metrics in medical image segmentation. *BMC Research Notes*, 15(1), p.210.
- Jungo, A., Meier, R., Ermis, E., Herrmann, E. and Reyes, M., 2018. Uncertainty-driven sanity check: Application to postoperative brain tumor cavity segmentation. *arXiv preprint arXiv:1806.03106*.
- Jungo, A. and Reyes, M., 2019, October. Assessing reliability and challenges of uncertainty estimations for medical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 48-56). Cham: Springer International Publishing.

References

- Martín-Isla, C., Campello, V.M., Izquierdo, C., Kushibar, K., Sendra-Balcells, C., Gkontra, P., Sojoudi, A., Fulton, M.J., Arega, T.W., Punithakumar, K. and Li, L., 2023. Deep learning segmentation of the right ventricle in cardiac MRI: the M&Ms challenge. *IEEE Journal of Biomedical and Health Informatics*, 27(7), pp.3302-3313.
- Liu, D., Gao, Y., Zhangli, Q., Han, L., He, X., Xia, Z., Wen, S., Chang, Q., Yan, Z., Zhou, M. and Metaxas, D., 2022, September. Transfusion: multi-view divergent fusion for medical image segmentation with transformers. In *International conference on medical image computing and computer-assisted intervention* (pp. 485-495). Cham: Springer Nature Switzerland.
- Meng, Q., Qin, C., Bai, W., Liu, T., De Marvao, A., O'Regan, D.P. and Rueckert, D., 2022. MulViMotion: Shape-aware 3D myocardial motion tracking from multi-view cardiac MRI. *IEEE transactions on medical imaging*, 41(8), pp.1961-1974.
- Isensee, F., Jaeger, P.F., Kohl, S.A.A. et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods* 18, 203–211 (2021).
- Loshchilov, I. and Hutter, F., 2016. Sgdr: Stochastic gradient descent with warm restarts. arXiv preprint arXiv:1608.03983.
- Loshchilov, I. and Hutter, F., 2017. Decoupled weight decay regularization. arXiv preprint arXiv:1711.05101.
- Karimzadeh, M., Seyedarabi, H., Jodeiri, A. and Afrouzian, R., 2025. Enhanced Brain Stroke Lesion Segmentation in MRI Using a 2.5 D Transformer Backbone U-Net Model. *Brain Sciences*, 15(8), p.778.
- Chen, C., Qin, C., Qiu, H., Tarroni, G., Duan, J., Bai, W. and Rueckert, D., 2020. Deep learning for cardiac image segmentation: a review. *Frontiers in cardiovascular medicine*, 7, p.25.
- Foret, P., Kleiner, A., Mobahi, H. and Neyshabur, B., 2020. Sharpness-aware minimization for efficiently improving generalization. arXiv preprint arXiv:2010.01412.