Enhancing Hypoxic-Ischemic Encephalopathy Classification through Region-Specific Brain MRI Segmentation

Elias J. Faro Gómez[1], Alejandra Aguado-Del-Hoyo[2], Belén Bernardo-Atienza[3], Dorotea Blanco-Bravo[3], Juan J. Gómez-Valverde[1,4] and Lucilio Cordero Grande[1,4]

[1]Biomedical Image Technologies, ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid,Spain, [2]Radiology Service, Hospital General Universitario Gregorio Marañón, Madrid, Spain, [3]Neonatology Service, Hospital General Universitario Gregorio Marañón, Madrid, Spain, [4]CIBER-BBN, ISCIII, Madrid, Spain

Abstract

INTRODUCTION: Accurate classification of neonatal HIE severity is essential for clinical decision-making and outcome prediction [1]. While existing scoring systems have shown strong predictive potential, they are time-consuming and prone to observer variability [2, 3]. Automating scoring through advanced imaging could lead to more consistent assessments. Namely, this study investigates whether refined brain region segmentations enhance HIE classification accuracy. Using a public neonatal dataset and a clinical cohort, we study the added value of region-specific segmentation in distinguishing mild from severe cases.

METHODS: Automatic segmentation. The nnU-Net [4] was trained on 679 annotated T2- and T1-weighted MRI volumes from the developing Human Connectome Project dataset [5] to segment 12 neonatal brain regions, including 6 clinically relevant for HIE, White Matter (WM), deep Gray Matter (dGM), Posterior Limb of the Internal Capsule (PLIC), Basal Ganglia (BG), and Thalamus with high (Th-H) and low (Th-L) intensity [1]. The model was trained for 500 epochs using default nnU-Net 2D settings and applied to a clinical dataset of 71 neonates with mild to severe HIE. **Classification of severity.** The clinical dataset was slightly imbalanced, so a 3-fold stratified cross-validation was used. Various classification experiments were conducted: (I) Deep Learning (DL) Classification. A DenseNet model was trained using two aproaches: (a) 2D, pretrained on ImageNet and (b) 3D, no pretraining available. In both cases, extensive tests were performed to optimize the hyperparameters for maximized performance. (II) Radiomics classification. Features were extracted using Pyradiomics [6] removing features with correlation > 0.9 from: (a) A whole-brain mask and (b) The 6 HIE-relevant regions. We conducted two experiments: one with feature selection and another without. For the former, mutual information analysis was applied, retaining features with values greater than 0.2 for T2- and 0.1 for T1-weighted images. Additionally, ridge (α=0.01) and Lasso (α=0.1) regularization were respectively applied to T2 and T1 features. Furthermore, ANOVA-based feature selection was used to identify the five most important features. An ensemble approach combining, including Random Forest, Logistic Regression, Gradient Boosting, and Multiayer Perceptron, was adopted to improve robustness and reduce model bias.

RESULTS & DISCUSSION: Segmenting the six key regions improves prediction accuracy, AUC, and G-Mean compared to using only a whole-brain mask (Fig. 1, left), emphasizing the relevance of region-specific analysis. Feature selection is not strictly necessary, as models without it perform similarly, though selecting five key features slightly enhances T1-based predictions, warranting further study. Supervised DL results were comparable to radiomics when using a whole-brain mask, suggesting that DL does not inherently outperform radiomics under these conditions. Incorporating segmentations in DL models may improve performance, though a larger cohort could be needed. Directionality of key features identified by SHAP (Fig. 1, right) appears potentially consistent with trend towards lower image conspicuity from PLIC myelin and abnormal thalamic intensities in the presence of lesion [7]. Similarly promising results have been reported by other approaches [8]. However, data and pipeline standardization for automatic lesion detection in HIE is an on-going work, demanding further statistical analysis and validation, to which we expect to contribute with our future research.

Modality	Experiment	Reduction	Accuracy	AUC	G-Mean
T2	Rad Segs	Yes	**0.915±0.034**	0.999±0.000	**0.911±0.042**
	Rad Segs	No	0.887±0.021	**0.910±0.035**	0.875±0.028
	DL 2D	No	0.845±0.053	0.960±0.041	0.829±0.050
	Rad	No	0.831±0.034	0.867±0.028	0.805±0.022
	DL 3D	No	0.830±0.038	0.892±0.054	0.811±0.039
	Rad	Yes	0.732±0.051	0.743±0.030	0.712±0.045
T1	Rad Segs	Yes	**0.888±0.038**	**0.921±0.021**	**0.870±0.045**
	Rad Segs	No	0.831±0.008	0.871±0.006	0.818±0.002
	Rad	No	0.803±0.016	0.841±0.036	0.801±0.026
	DL 3D	No	0.803±0.070	0.785±0.036	0.795±0.065
	DL 2D	No	0.802±0.056	0.793±0.115	0.796±0.058
	Rad	Yes	0.691±0.076	0.739±0.051	0.627±0.089

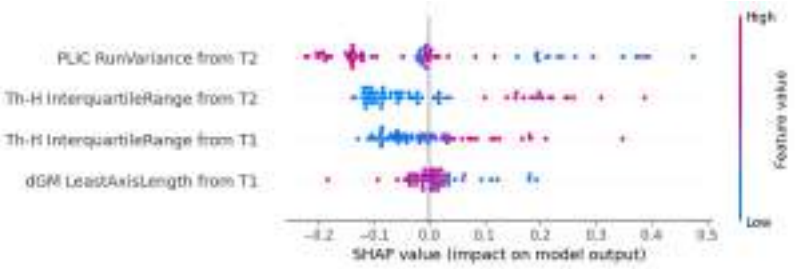

Figure 1 – (Left) Summary of DL and radiomic results using T2 and T1. 'Rad': radiomic experiment with a whole brain mask. 'Rad Segs': radiomic experiment with the 6 clinically relevant regions. (Right) Beeswarm plot of SHAP values showing most important features for both modalities. Higher SHAP values (shift to the right) are associated with more severe lesions.

Acknowledgements. The authors gratefully acknowledge the Universidad Politécnica de Madrid for providing computing resources on Magerit Supercomputer and funding from the MCIN / AEI/ 10.13039/501100011033 / FEDER, UE, under Project PID2021-129022OA-I00.

References

[1] M Martinez-Biarge et al. "Predicting motor outcome and death in term hypoxic-ischemic encephalopathy". Neurology 76.24 (2011), pp. 2055–2061. [2] S. Shankaran et al. "Brain injury following trial of hypothermia for neonatal hypoxic–ischaemic encephalopathy". Archives of Disease in Childhood - Fetal and Neonatal Edition 97.6 (2012), F398–F404. [3] LC. Weeke et al. "A Novel Magnetic Resonance Imaging Score Predicts Neurodevelopmental Outcome After Perinatal Asphyxia and Therapeutic Hypothermia". The Journal of Pediatrics 192 (2018), 33–40.e2. [4] F Isensee et al. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation." Nature Methods 18.2 (2021). [5] A. Makropoulos et al. "The developing human connectome project: A minimal processing pipeline for neonatal cortical surface reconstruction". NeuroImage 173 (2018), pp. 88–112. [6] Joost J.M. van Griethuysen et al. "Computational Radiomics System to Decode the Radiographic Phenotype". Cancer Research 77.21 (2017), e104–e107. [7] JJM et al. "Neuroimaging in the term newborn with neonatal encephalopathy". Seminars in Fetal and Neonatal Medicine 26.5 (2021). [8] T. Brochet et al. "Neonatal Hypoxic Ischemic Encephalopathy Severity Grading Using Multimodal Swin Transformer". 2024 IEEE International Symposium on Biomedical Imaging. 2024, pp. 1–4.