

基于 UNETR++ 的新生儿缺氧缺血性脑病病灶分割

摘要

本文复现了一种基于 Vision Transformer 架构的 U 型 3D 医学图像分割网络——UNETR++。该网络的创新之处在于引入了高效成对注意模块 (Efficient Pairwise Attention Module)，该模块采用双分支结构分别处理空间和通道注意力机制，从而有效捕获丰富的空间和通道特征。特别地，通过共享查询 (Query) 和键 (Key) 映射函数的权重，该模块在保持模型性能的同时显著降低了模型复杂度和参数量。实验表明，UNETR++ 在多个基准数据集上均展现出优秀的分割精度和推理效率。基于其优异的性能表现，本文将该网络架构应用于新生儿缺血缺氧性脑病 (Hypoxic-Ischemic Encephalopathy, HIE) 病灶分割任务，取得了一定的性能提升，为复杂医学图像分割任务提供了一种解决方案。

关键词：医学图像分割；空间注意力；通道注意力；新生儿缺血缺氧性脑病

1 引言

新生儿缺氧缺血性脑病 (Hypoxic Ischemic Encephalopathy, HIE) 是围生期因窒息缺氧引发的脑部缺氧缺血性损伤，其病因可能发生于产前或产时的子宫内缺氧阶段，或产后窒息导致有效呼吸建立延迟的情况。患儿出生后通常表现出脑病症状，严重者可能发展为脑瘫 (发生率约为 10% ~ 20%)、智力低下、生长发育迟缓或癫痫等问题，对儿童健康造成严重威胁，并可能导致持续至成年期的神经发育异常 [12]。HIE 的诊断目前主要依赖于临床表现的综合评估，但医生间诊断水平的差异、医疗资源的区域性不平衡等问题使得 HIE 的脑部磁共振成像 (MRI) 诊断存在较大不确定性，尤其在医疗条件薄弱地区，这给 HIE 的早期预警和个性化临床决策带来了巨大挑战 [3]。因此，开发客观、高效的方法来精准识别新生儿脑部损伤，并进行科学的预后评估和治疗效果监测，已成为亟待解决的关键问题。

脑部磁共振成像 (MRI) 被公认为评估 HIE 最敏感且最特异的成像手段，在分析新生儿大脑病理改变及预测发育结局方面具有显著优势，目前已成为 HIE 早期诊断和预后评估的主要辅助工具和首选筛查方式 [4]。然而，由于 HIE 病灶通常表现为体积小且弥散性分布，传统影像分析方法难以有效捕捉这些特征。而 Abdelrahman 等人 [8] 提出了一种用于 3D 医学分割的 UNETR++ 模型。该模型引入了一种高效的成对注意力 (EPA) 块，通过使用空间和通道注意力来编码丰富的相互依赖的空间和通道特征。在 EPA 块内，共享查询和键映射函数的权重，以更好地在空间和通道分支之间进行通信，提供互补优势并减少参数。与现有的最佳方法相比，UNETR++ 在腹部器官、心脏、脑肿瘤、肺癌上取得了良好的分割结果，同时显

著降低了模型复杂度。因此，本文探索将先进的 UNETR++ 模型迁移至 HIE 的病灶分割任务中，旨在实现精准分割。

2 相关工作

近年来，深度学习技术在医学图像分割领域取得了显著进展，尤其是在脑部病变分割任务中。然而，HIE 的病变分割仍然是一个极具挑战性的任务，主要原因在于 HIE 病变通常较小且弥散，占据脑部体积的比例极低（通常小于 1%），这使得传统的分割方法难以取得理想的效果。针对这一问题，研究者们提出了多种基于深度学习的分割方法。例如，Elham Soltani Kazemi 等人 [9] 提出了一种基于 Swin-UNETR 的集成策略，通过自注意力机制和多层次编码捕捉病变的多尺度特征，显著提升了分割性能，并在 BONBID-HIE 挑战赛中取得了前三名的成绩。此外，M. Arda Aydin 等人 [2] 提出了一种新颖的互逆变换 (Reciprocal Transformation)，结合 SegResNet 和 Attention U-Net，进一步优化了 HIE 病变的分割效果，尤其是在处理小病变时表现出色。

在模型架构方面，Imad Eddine Toubal 等人 [5] 提出了一种结合 Swin-UNETR 和随机森林分类器的混合方法，通过全局上下文信息和局部特征的融合，显著提升了分割精度，尤其是在表面距离度量 (MASD 和 NSD) 方面表现优异。Nazanin Tahmasebi 和 Kumaradevan Punithakumar [10] 则采用了 nnU-Net 进行 HIE 病变分割，利用其自配置能力减少了手动调参的需求，尽管在验证集上表现良好，但在测试集上的泛化能力仍有待提升。此外，Marek Wodzinski 等人 [11] 通过重数据增强策略，结合 3D ResUNet 架构，显著提升了模型在处理大病变时的分割性能，但在小病变分割上仍有改进空间。

集成学习方法也在 HIE 病变分割中展现了其优势。Chiranjeev Prasad Koirala 等人 [7] 提出了一种结合六种 U-Net 变体和三种损失函数的集成方法，通过多模型输出的融合，显著提升了分割的准确性和鲁棒性。

总之，上述方法在 HIE 病变分割中展现出了其有效性，但存在分割精度低，模型复杂度高，因此本文探索将先进的 UNETR++ 模型迁移至 HIE 的病灶分割任务中，旨在实现快速而精准的分割。

3 方法

3.1 整体架构

图 1 展示了 UNETR++ 架构，它包含一个分层的编解码器结构。UNETR++ 框架基于最近提出的 UNETR [6]，在编解码器之间有跳跃连接 (skip connections)，之后是卷积块 (ConvBlocks) 用于生成预测掩码。UNETR++ 在编码器中没有使用固定的特征分辨率，而是采用了分层设计，在每个阶段特征的分辨率逐渐降低为原来的二分之一。在 UNETR++ 框架中，编码器有四个阶段。四个阶段的通道数为 $[C1, C2, C3, C4]$ 。第一阶段包括补丁嵌入 (patch embedding)，将体积输入划分为 3D 补丁，之后是高效成对注意力 (EPA) 块。在补丁嵌入中，将每个 3D 输入 (体积) $x \in R^{H \times W \times D}$ 划分为非重叠的补丁 $x_u \in R^{N \times (P1, P2, P3)}$ ，其中 $(P1, P2, P3)$ 是每个补丁的分辨率， $N = (\frac{H}{P1} \times \frac{W}{P2} \times \frac{D}{P3})$ 表示序列的长度。然后，这些补丁被投影到 C 个通道维度，产生大小为 $\frac{H}{P1} \times \frac{W}{P2} \times \frac{D}{P3} \times C$ 的特征图。对于其余的每个编码器阶

段，使用了非重叠卷积的下采样层将分辨率降低为原来的二分之一，之后是高效成对注意力（EPA）块。

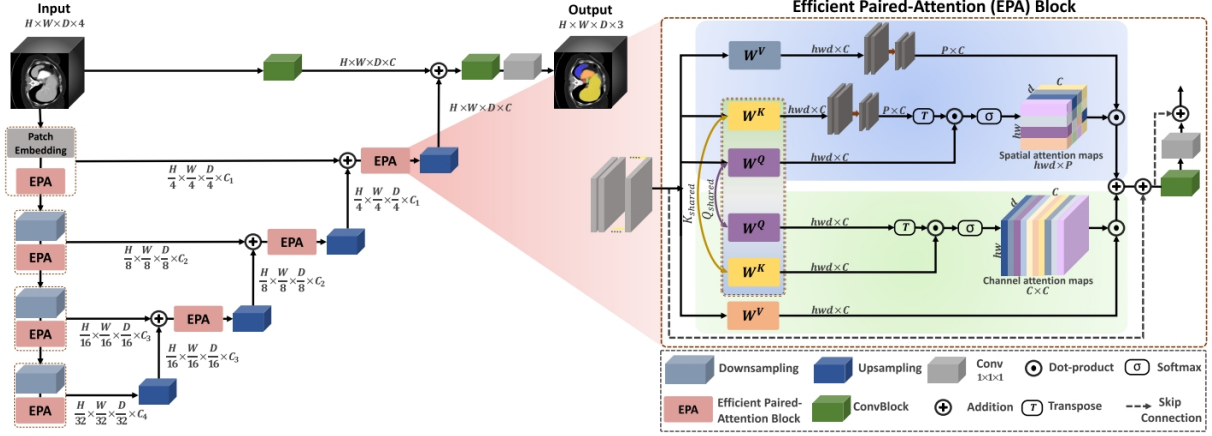


图 1. UNETR++ 架构

3.2 高效成对注意力模块（Efficient Paired - Attention Block）

大多数现有的混合方法采用的自注意力操作的复杂度是 token 数量的二次方。在体分割（volumetric segmentation）的情况下，尤其在混合设计中交错窗口注意力（interleaving window attention）和卷积组件时计算复杂度更加严重。此外，通过将键（keys）和值（values）的空间矩阵投影到低维空间，可以有效地学习空间注意力信息。有效地结合空间维度上的相互作用以及通道特征之间的相互依赖关系，能够提供丰富的上下文空间 - 通道特征表示，从而改进掩模预测（mask predictions）。

高效成对注意力（EPA）模块执行高效的全局注意力操作，并有效地捕获丰富的空间 - 通道特征表示。EPA 模块包含空间注意力（spatial attention）和通道注意力（channel attention）模块。空间注意力模块将自注意力的复杂度从二次方降低到线性。另一方面，通道注意力模块有效地学习通道特征图之间的相互依赖关系。EPA 模块基于在两个注意力模块之间共享键 - 查询（keys - queries），以便相互提供信息，从而生成更好的特征表示。这可能是由于通过共享键和查询但使用不同的值层来学习互补特征。为了增强训练稳定性，在每个 EPA 模块的开头采用层归一化（Layer Normalization, LayerNorm）。

如图 1（右侧）所示，输入特征图 x 被送入 EPA 模块的通道注意力和空间注意力模块。 Q 和 K 线性层的权重在两个注意力模块之间共享，并且每个注意力模块使用不同的 V 层。两个注意力模块的计算如下：

$$\begin{aligned}\hat{X}_s &= SA(Q_{shared}, K_{shared}, V_{spatial}) \\ \hat{X}_c &= CA(Q_{shared}, K_{shared}, V_{channel})\end{aligned}$$

其中， \hat{X}_s 、 \hat{X}_c 、 SA 和 CA 分别表示空间注意力图、通道注意力图、空间注意力模块和通道注意力模块。 Q_{shared} 、 K_{shared} 、 $V_{spatial}$ 和 $V_{channel}$ 分别是共享查询、共享键、空间值层和通道值层的矩阵。

3.2.1 空间注意力 (Spatial Attention)

在本模块中, 通过将复杂度从 $O(n^2)$ 降低到 $O(np)$ (其中 n 是 token 的数量, p 是投影向量的维度, 且 $p \ll n$) 来有效地学习空间信息。

给定形状为 $hwd \times C$ 的归一化张量 X , 使用三个线性层计算 Q_{shared} 、 K_{shared} 和 $V_{spatial}$ 投影, 得到 $Q_{shared} = W_Q X$, $K_{shared} = W_K X$, $V_{spatial} = W_V X$, 其维度为 $hwd \times C$, 其中 W_Q 、 W_K 和 W_V 分别是 Q_{shared} 、 K_{shared} 和 $V_{spatial}$ 的投影权重。然后, 执行三个步骤。首先, 将 K_{shared} 和 $V_{spatial}$ 层从 $hwd \times C$ 投影到形状为 $p \times C$ 的低维矩阵。其次, 通过将 Q_{shared} 层乘以投影后的 K_{shared} 的转置, 然后进行 $softmax$ 操作来计算空间注意力图, 以测量每个特征与其余空间特征之间的相似性。第三, 将这些相似性乘以投影后的 $V_{spatial}$ 层, 以生成形状为 $hwd \times C$ 的最终空间注意力图。空间注意力定义如下:

$$\hat{X}_s = \text{Softmax} \left(\frac{Q_{shared} K_{proj}^\top}{\sqrt{d}} \right) \cdot \tilde{V}_{spatial}$$

其中, Q_{shared} 、 K_{proj} 、 $\tilde{V}_{spatial}$ 分别表示共享查询、投影共享键和投影空间值层, d 是每个向量的大小。

3.2.2 通道注意力 (Channel Attention)

本模块通过在通道值层和通道注意力图之间的通道维度上应用点积操作来捕获特征通道之间的相互依赖关系。使用与空间注意力模块相同的 Q_{shared} 和 K_{shared} , 使用线性层计算通道的值层以学习互补特征, 得到 $V_{channel} = W_V X$, 其维度为 $hwd \times C$, 其中 W_V 是 $V_{channel}$ 的投影权重。通道注意力定义如下:

$$\hat{X}_c = V_{channel} \cdot \text{Softmax} \left(\frac{Q_{shared}^\top K_{shared}}{\sqrt{d}} \right)$$

其中, $V_{channel}$ 、 Q_{shared} 、 K_{shared} 分别表示通道值层、共享查询和共享键。

最后, 执行求和融合 (sum fusion), 并通过卷积块对来自两个注意力模块的输出进行变换, 以获得丰富的特征表示。EPA 模块的最终输出 \hat{X} 如下获得:

$$\hat{X} = \text{Conv1}(\text{Conv3}(\hat{X}_s + \hat{X}_c))$$

其中, \hat{X}_s 和 \hat{X}_c 分别表示空间注意力图和通道注意力图, Conv1 和 Conv3 分别是 $1 \times 1 \times 1$ 和 $3 \times 3 \times 3$ 的卷积块。

3.3 损失函数

损失函数基于常用的 soft dice 损失和交叉熵损失之和, 可以利用这两个互补损失函数的优势。其定义如下:

$$L(Y, P) = 1 - \sum_{i=1}^I \frac{2 \cdot \sum_{v=1}^V Y_{v,i} \cdot P_{v,i}}{\sum_{v=1}^V Y_{v,i}^2 + \sum_{v=1}^V P_{v,i}^2} + \sum_{v=1}^V Y_{v,i} \log P_{v,i}$$

其中, I 表示类别数量, V 表示体素数量, $Y_{v,i}$ 和 $P_{v,i}$ 分别表示体素 v 对于类别 i 的真实值和输出概率。

4 复现细节

4.1 与已有开源代码对比

本文在 UNERT++ 开源代码的基础上 [1], 针对 HIE (缺氧缺血性脑病) 患者的 MRI 影像, 特别是 ADC (表观扩散系数) 模态, 增加了图像预处理的功能。具体包括以下三个部分:

1. 采用 N4ITK 偏置场校正算法对 MRI 图像进行亮度不均匀性校正。该算法基于 B 样条曲线拟合, 通过迭代优化过程估计并消除由磁场不均匀性引起的低频强度变化。具体而言, 算法首先对图像进行对数变换, 将乘性偏置场转换为加性场, 然后通过 B 样条平滑估计偏置场, 最后进行校正。这一步骤改善了图像质量, 确保了不同组织区域间亮度的一致性, 为后续分析提供了可靠的基础。

2. 通过仿射变换实现图像的空间标准化。该过程使用线性代数变换矩阵, 将每个患者的 MRI 图像配准到标准空间。算法采用互信息作为相似性度量, 通过优化变换参数 (包括平移、旋转和缩放) 实现最佳配准。这一步骤确保了所有图像具有一致的空间分辨率和覆盖范围, 便于跨病例的定量比较和分析。

3. 使用基于 B 样条的自由形变 (FFD) 配准算法, 将每个患者的图像精确映射到标准解剖空间。该算法通过控制点网格的弹性变形实现局部配准, 采用归一化互相关 (NCC) 作为相似性度量, 通过梯度下降法优化变形场。这一步骤确保了不同病例的解剖结构在相同坐标系下对齐, 为病变检测和分割提供了统一的参考框架。

以上增加的算法的可以提升图像的亮度一致性和空间对齐性, 保证了后续处理病灶分割的精度

4.2 实验环境搭建

本实验在 Ubuntu 18.04 操作系统环境下进行, 采用 Python 3.8 作为主要编程语言环境。为满足 3D 医学图像分割任务对计算性能的需求, 实验环境配置了 CUDA 11.3 并行计算平台, 并安装了与之兼容的 PyTorch 1.11.0 深度学习框架及 torchvision 0.12.0 计算机视觉库。硬件平台采用两块 NVIDIA Tesla V100 GPU 进行并行加速计算, 以支持大规模 3D 医学图像数据的高效处理与模型训练。

4.3 创新点

UNETR++ 是一种高效的 3D 医学图像分割模型, 在参数量、浮点运算次数 (FLOPs)、GPU 内存消耗和推理速度等方面实现了优异的分割精度与效率提升。其核心创新是引入新颖的高效成对注意力 (EPA) 模块, 该模块通过在两个分支中分别应用空间和通道注意力, 能够有效捕捉丰富的空间和通道特征, 同时通过共享查询和键映射函数的权重降低了模型复杂度和参数量。UNETR++ 在多个基准数据集 (如 Synapse、BTCV、ACDC、BraTS 和 Decathlon - Lung) 上的分割精度和推理效率上均表现突出。鉴于 HIE 病变病灶体积小、弥散且占脑部位比例不足 1%, 本文创新性地将 UNETR++ 迁移至 HIE 病灶分割任务中, 充分利用其优异性能来应对该任务的挑战。

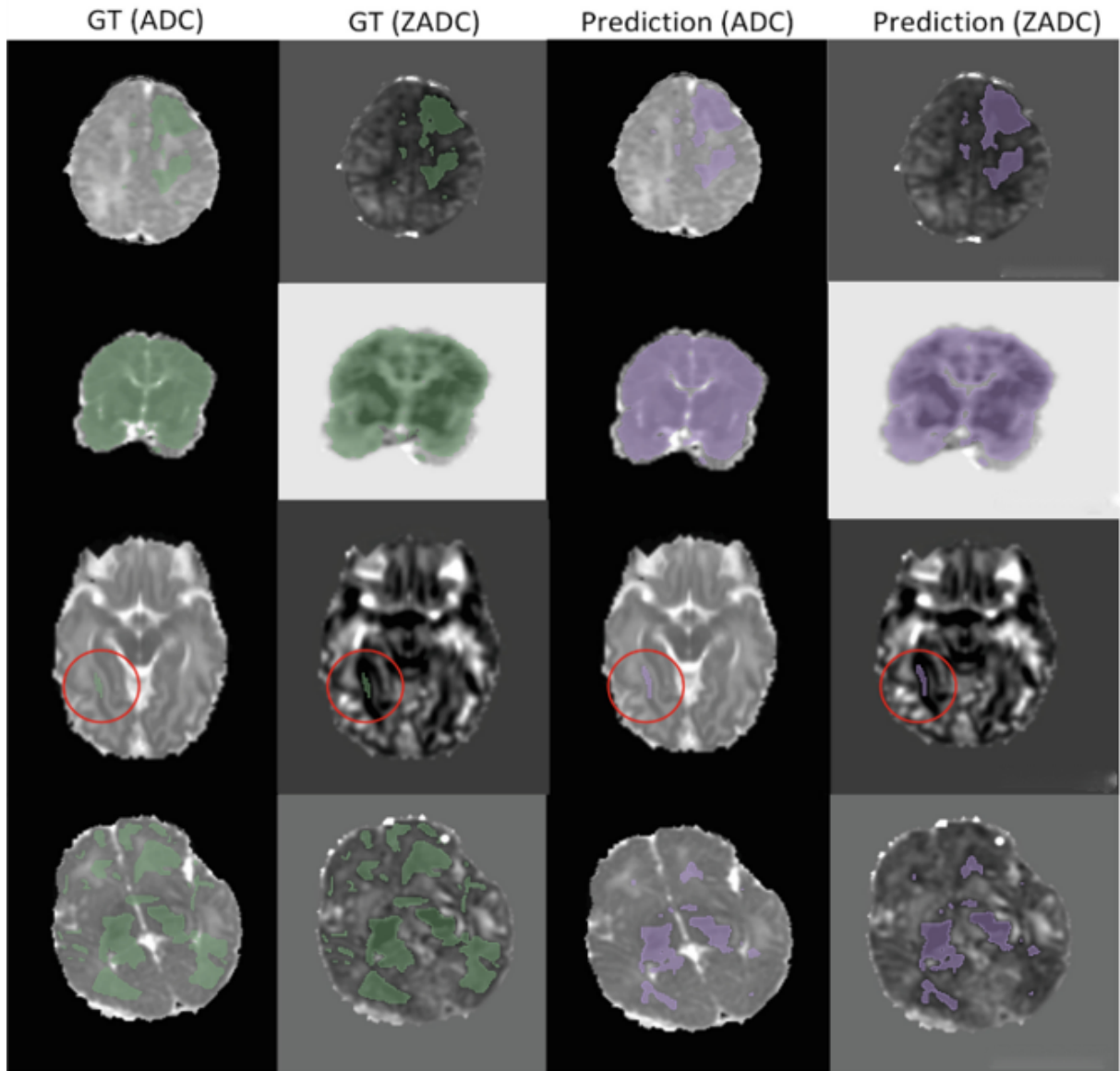


图 2. UNETR++ 分割结果可视化展示（绿色区域表示为 GT，紫色区域表示为预测结果）

5 实验结果分析

表 1 展示了不同方法在 HIE 病灶分割任务中的性能对比。从结果可以看出，UNETR++ 模型随着交叉验证折数的增加呈现出明显的性能提升趋势，其中五折交叉验证取得了最佳性能，Dice 值达到 0.5741。然而，这一结果仍较当前最优方法低 0.049，且低于 Swin-UNETR Ensemble 和 3-DCNN 等对比方法。

图 2 通过四个典型病例的可视化结果，图中第一列和第二列分别展示了 ADC 和 ZADC 模态的 Ground Truth，第三列和第四列则是对应的分割结果。在第四行的可视化结果，尽管原始图像中存在多个病灶区域，但 UNETR++ 仅能检测到面积较大的显著病灶，而对较小病灶的识别能力明显不足。这表明，现有 UNETR++ 架构在处理多病灶场景时难以充分利用跨病灶的全局上下文信息，导致对小病灶的识别灵敏度不足。其现有结构对病灶尺寸变化的适应性较差，无法同时准确捕捉大病灶的整体特征和小病灶的细节特征。

	Dice	MASD	NSD
UNETR++ Fold1	0.5265	3.115	0.7036
UNETR++ Fold2	0.532	3.1606	0.7171
UNETR++ Fold3	0.5378	2.9548	0.7151
UNETR++ Fold5	0.5741	2.6143	0.7491
Swin-UNETR Ensemble [9]	0.5742	2.6243	0.7491
3-DCNN [11]	0.5770	2.9419	0.7379
Deep Random Forest Fusion [5]	0.6215	2.2556	0.7678

表 1. 不同分割方法的性能比较

6 总结与展望

本文将 UNETR++ 模型应用于 HIE 病灶分割任务，五折交叉验证结果显示其 Dice 值为 0.5741，较最优方法低 0.049，尤其在多病灶切面的分割性能较差。这表明，现有 UNETR++ 在上下文信息整合及尺度适应性方面存在不足。因此未来研究将在 UNETR++ 基础上引入多尺度卷积模块、跨层特征融合机制及空间-通道双重注意力机制，并优化损失函数设计，以增强模型对多尺度病灶特征的提取能力和边界定位精度。

参考文献

- [1] Amshaker. Unetr++: A github repository for unetr++ implementation. https://github.com/Amshaker/unetr_plus_plus, 2023. Accessed: 2023-10-01.
- [2] M Arda Aydın, Elvin Abdinli, and Gozde Unal. Segresnet based reciprocal transformation for bonbid-hie lesion segmentation. In *Trauma Thompson Challenge*, pages 39–44. Springer, 2023.
- [3] Sonia Lomeli Bonifacio and Shandee Hutson. The term newborn: evaluation for hypoxic-ischemic encephalopathy. *Clinics in perinatology*, 48(3):681–695, 2021.
- [4] Deepak Chawla. Biomarkers for prognostication in hypoxic-ischemic encephalopathy. *The Indian Journal of Pediatrics*, 87(10):777–778, 2020.
- [5] Imad Eddine Toubal, Elham Soltani Kazemi, Gani Rahmon, Taci Kucukpinar, Mohamed Almansour, Mai-Lan Ho, and Kannappan Palaniappan. Fusion of deep and local features using random forests for neonatal hie segmentation. In *Trauma Thompson Challenge*, pages 3–13. Springer, 2023.
- [6] Ali Hatamizadeh, Yucheng Tang, Vishwesh Nath, Dong Yang, Andriy Myronenko, Bennett Landman, Holger R Roth, and Daguang Xu. Unetr: Transformers for 3d medical image segmentation. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 574–584, 2022.

- [7] Chiranjeev Prasad Koirala, Sovesh Mohapatra, and Gottfried Schlaug. An ensemble approach for segmentation of neonatal hie lesions. In *Trauma Thompson Challenge*, pages 23–27. Springer, 2023.
- [8] Abdelrahman M Shaker, Muhammad Maaz, Hanoona Rasheed, Salman Khan, Ming-Hsuan Yang, and Fahad Shahbaz Khan. Unetr++: delving into efficient and accurate 3d medical image segmentation. *IEEE Transactions on Medical Imaging*, 2024.
- [9] Elham Soltani Kazemi, Imad Eddine Toubal, Gani Rahmon, Taci Kucukpinar, Mohamed Almansour, Mai-Lan Ho, and Kannappan Palaniappan. Enhancing lesion segmentation in the bonbid-hie challenge: An ensemble strategy. In *Trauma Thompson Challenge*, pages 14–22. Springer, 2023.
- [10] Nazanin Tahmasebi and Kumaradevan Punithakumar. A deep neural network approach for the lesion segmentation from neonatal brain magnetic resonance imaging. In *Trauma Thompson Challenge*, pages 34–38. Springer, 2023.
- [11] Marek Wodzinski and Henning Müller. Improving segmentation of hypoxic ischemic encephalopathy lesions by heavy data augmentation: Contribution to the bonbid challenge. In *Trauma Thompson Challenge*, pages 28–33. Springer, 2023.
- [12] Edibe Pembegül Yıldız, Barış Ekici, and Burak Tatlı. Neonatal hypoxic ischemic encephalopathy: an update on disease pathogenesis and treatment. *Expert review of neurotherapeutics*, 17(5):449–459, 2017.