

SingleStrip: 从单个标记实例中学习头骨剥离

Bella Spektor-Fadida¹ and Malte Hoffmann²⁻⁴

¹ Department of Medical Imaging Sciences, University of Haifa, Israel

² Athinoula A. Martinos Center for Biomedical Imaging, Boston, MA, USA

³ Department of Radiology, Harvard Medical School, Boston, MA, USA

⁴ Department of Radiology, Massachusetts General Hospital, Boston, MA, USA

Abstract. 深度学习分割严重依赖于标记数据，但手动标记既费力又费时，尤其是对像脑部磁共振成像（MRI）这样的体积图像。虽然最近的领域随机化技术通过从标签图中合成多样的训练图像缓解了对标记数据的依赖，但当可用标签图非常少时，它们提供的解剖变异性有限。半监督自我训练通过反复将模型预测纳入训练集来解决标签稀缺问题，使网络能够学习未标记的数据。在本研究中，我们将领域随机化与自我训练相结合，使用尽可能少的单个标记示例来训练三维去颅骨网络。首先，我们自动对体素强度进行分箱，生成用于合成图像的标签，这些图像用于训练初始的去颅骨模型。其次，我们在标记示例上训练一个卷积自编码器（AE），并利用其重建误差评估未标记数据预测的脑掩膜的质量。第三，我们选择排名靠前的伪标签来微调网络，在超出分布的数据上实现接近使用更多标记图像训练的模型的去颅骨性能。我们将基于 AE 的排名与测试时增强下的基于一致性的排名进行比较，发现 AE 方法与分割精度有更强的相关性。我们的结果突出了结合领域随机化和基于 AE 的质量控制的潜力，以实现从极其有限的标记数据中进行有效的半监督分割。这一策略可能会减轻标记负担，从而加速涉及新解剖结构或新兴成像技术研究的进展。

Keywords: segmentation · deep learning · one-shot learning · synthetic data · self-training · quality control

在医学图像中对解剖结构进行分割对于许多临床任务是至关重要的，例如评估肿瘤的大小和进展。然而，手动标注既费力、又需要专业知识，并且结果因评价者而异，这推动了自动化技术的发展。例如，脑部磁共振成像（MRI）中的颅骨剥离 [20, 23] 是一种在神经影像分析流程中自动化的技术，可以去除令后续分析分心的非脑部图像内容 [5, 6, 13]。现代方法利用深度学习进行快速而准确的颅骨剥离，但通常依赖于带有注释的高质量训练数据集，这些数据集的创建成本高昂 [27]。因此，使用尽可能少的数据进行训练的策略（可能仅需要单个标注的示例图像）对于开发在没有大量标注工作的新解剖结构或成像模式的分割模型非常有价值。

Domain randomization: 最近的合成策略或领域随机化策略 [2, 9–11]，在通过从少量标签图生成多样的训练数据这一方式上，展示了在多种模式下的稳健脑部分割性能，这减少了对大规模数据集的需求。这些方法通过为每个脑部结构和背景结构的体素分配随机强度来合成图像。然后，它们施加一系列强烈的增强步骤，从而在解剖结构中形成复杂的强度模式。这种策略导致了一些工具的产生，这些工具能够可靠地处理那些未见过的图像类型而无需重新训练。然而，尽管有几何增强，但生成的训练图像并不一定能覆盖训练稳健、可推广网络所需的真实世界解剖变异性，特别是在仅使用单一标签图的情况下。

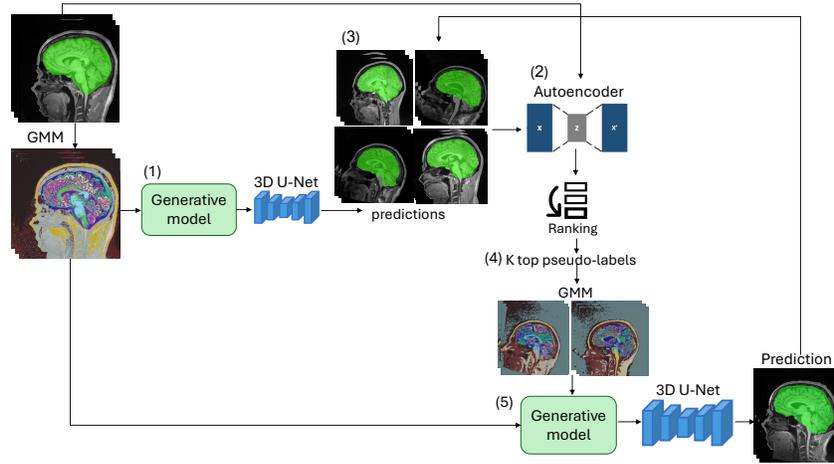


Fig. 1. 方法概述。为了从单个标记图像中学习去颅，我们首先拟合一个高斯混合模型 (GMM)，将体素强度分配给 c 个类别。然后，我们根据这些类别合成多样化的图像来训练去颅的 U-Net。同时，我们训练一个自动编码器 (AE) 来重建大脑掩膜以进行质量控制。假设高质量掩膜在 AE 重建中变化最小，我们使用这两个网络从未标记的数据集中去颅，并保留变化最小的预测，以通过 GMM 合成微调 U-Net。根据需要，该方案可以重复多次。

Semi-supervised learning: 在语义分割中，有多种其他方法解决标注稀缺的问题 [27]。一个显著的例子是带有自我训练 (ST) 的半监督学习 (SSL)，其创建伪标签。在 ST 中，模型通过首先预测一个未标记数据集的标签，然后将伪标签添加到训练集中以从网络自身的预测中学习，从而提高自身性能 [4]。这种方法使得无需人工努力即可显著增加训练数据的多样性，因为未标记的数据集通常比已标记的数据集更普遍可用。

关于 ST 的研究经常以不确定性感知的方式训练两个或多个网络 [4, 21, 27]。一种方法 [16] 利用卷积对抗框架，将分割网络与置信网络相结合。后者预测一个注意力图，用于选择从未标记数据中预测出的可靠伪标签，以进一步进行半监督学习。另一种方法通过共同训练多个网络来生成伪标签 [28, 30]。它将输入旋转和翻转成多个视图，并用每个视图训练独立的网络，通过不确定性加权对未标记视图保持一致性。相比之下，一致性正则化 [3, 17] 通过对同一输入的不同版本应用扰动来生成分歧，通过强化未标记数据的预测一致性来训练单个模型。最近，一种离线 ST 方法涉及教师和学生网络，已经取得了有前景的分割性能 [25, 26, 29, 32]。首先，教师网络在标记数据上训练，然后为未标记数据集预测伪标签。其次，学生网络利用手工标注的真实标签和生成的伪标签进行训练。

为了减少人为负担，ST 采用自动化质量控制 (QC) 技术来区分可靠和不太准确的伪标签。用于分割 QC 的监督方法通常借助回归网络估计一个单一指标 [1, 7, 12, 18, 22]。然而，用单个标记示例训练回归网络具有挑战性。另一种策略是使用一个二维 (2D) 卷积自编码器 (AE) 来估计分割质量，该自编码器在地面真实标签图上训练，以从潜在空间中重建感兴趣结构的形状。然后，QC 指标计算为输入标签图和其重建之间的 Dice 得分 [8]。

无监督质量控制技术旨在没有标签的情况下评估分割质量。例如，测试时增强 (TTA) 对测试图像应用随机变换，并在撤销这些变换后比较每个版本的预测。TTA 假设模型对变换间变异性最小的预测最有信心，并已在估测分割质量方面显示出前景 [15, 24]。最近，一种两阶段方法 [31] 将真实的肿瘤合成与伪标签结合用于无监督分割。类似于师生方法 [25, 26, 29, 32]，该方法启动了第一个网络并使用其预测来训练第二个，重点在于缩小真实图像与合成图像之间的差距。

在这项工作中，我们结合了合成技术和伪标记，并利用自动质量估计在半监督环境下引导高性能分割网络。我们的生成模型不像通常那样追求真实感，而是合成训练数据超出典型医学影像的范围，以优先让模型接触到强度变化。我们展示了鲁棒的性能，在训练去颅骨网络时仅需一个标记的图像作为人为输入。

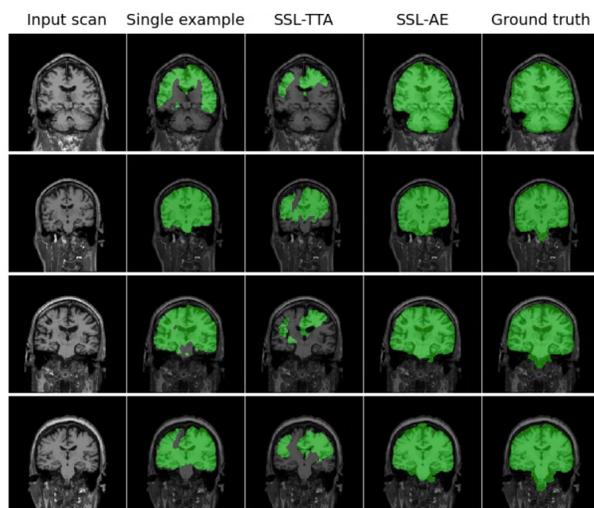


Fig. 2. 去骨示例。方法 SSL-TTA 和 SSL-AE 分别通过使用测试时增强和自编码器重构选择的伪标签，微调用“单个示例”训练的 U-Net。每行显示不同的 ASL 受试者。

1 方法

我们的半监督方法通过五个步骤（见图 1）从单个标记图像中学习头骨剥离。首先，我们对图像拟合高斯混合模型 (GMM)，将其体素强度分配到 c 类。这些用于合成高变动性的图像流和对应的标签图，以训练初始分割网络。其次，我们训练一个卷积自编码器 (AE)，将标签图映射到潜在空间，然后重建它们——仅使用原始标签图，并进行增强。第三，我们将分割网络应用于未标记的数据，并生成一个具有伪标签图的辅助训练集。第四，我们根据伪标签图与其 AE 重建的 Dice 重叠选择 k 个最高质量的伪标签图。最后，我们为辅助数据集拟合 GMM 标签，并在从辅助训练集生成的合成图像上重新训练分割网络，同时结合原始标记的例子。需要时步骤 3-5 可以重复多次。

Domain-randomized skull-stripping: 基于最近的颅骨剥离方法 [11, 14]，我们训练了一个三维 (3D) 卷积网络来预测大脑掩膜，使用合成的图像和标签图对。简而言之，一个生成模型通过应用平滑的随机形变（包括平移、旋转、缩放和剪切），空间增强了 c GMM 拟合类的输入标签图。从新的标签图中，模型通过为标签图中的每个类分配一个均匀采样的强度值来生成图像。最后，一系列随机的干扰包括加性噪声、平滑、裁剪、强度调制和指数运算，使图像产生高度复杂的强度模式。

尽管先前的研究使用包含多个解剖结构的真实标签图 [11, 14]，但我们的重点是以尽可能少的人力输入启动一个分割网络。因此，我们从一个单一的真实脑掩膜开始。我们分别在脑内和脑外拟合非解剖的高斯混合模型标签，并从中生成输入图像进行训练。训练优化网络预测与空间增强的真实脑掩膜之间的软 Dice 指标。我们没有开发新的网络架构，而是采用了现有的 U-Net [11, 19]。

为了在无需人工干预的情况下区分可靠的和较不准确的伪标签图，我们采用了一种基于重构的方法 [8]。我们训练一个二维卷积自动编码器 (AE) 将输入的脑掩膜映射到潜在空间并进行重构，结合使用软 Dice 和均方误差准则。假设高质量掩膜的重构是准确的，而未观察到的较低质量掩膜导致不准确的重构，我们将 AE 输入和输出之间的 Dice 分数视为脑掩膜质量的代理。我们根据体积 Dice 分数对三维脑掩膜进行排序。

对于 AE，我们采用了现有的网络架构 [8]，选择了一个 100 元素的潜向量，批大小为 8，学习率为 0.001。我们使用 Adam 优化器对 AE 进行训练 500 个 epoch，权重衰减为 10^{-5} ，应用旋转在 -30° 和 30° 之间，缩放因子在 0.7 到 1.4 之间，随机中心偏移，并进行翻转以增强数据。

Semi-supervised learning: 我们使用 SSL 通过伪标签来改进初始网络。具体来说，我们应用训练好的网络来预测未标记的辅助数据集的脑掩膜。我们期望这些掩膜的质量是可变的，并通过 AE 重构来评估它们。我们选择具有最低重构误差的 k 个脑掩膜并将其添加到训练集中。以与初始的真实脑掩膜相同的概率对每个伪脑掩膜进行采样，我们重新训练分割网络，以去除由 c 个 GMM 拟合类生成的合成图像的颅骨，并对原始的单样本 (SE) 网络进行微调。我们可以根据需要重复此步骤多次。

2 实验与结果

我们在颅骨去除任务中评估我们的 SSL 策略，使用来自 SynthStrip 数据集的 81 个具有真实脑掩膜的 T1 加权结构性脑 MRI 扫描。我们选择 38 个 FreeSurfer 维护 (FSM) 图像用于训练和分布内 (ID) 测试，以及 43 个来自动脉自旋标记 (ASL) 子集的图像用于分布外 (OOD) 测试。我们将 FSM 子集分成训练、验证和测试集，大小分别为 16、2 和 20。对于训练集中的每张图像，我们提取 $c = 16$ 个 GMM 类进行图像合成——脑内 10 个，脑外 6 个。

我们通过在单个随机选择的 FSM 图像和相关的大脑掩膜上训练我们的网络，并使用剩余的 15 个 FSM 和 43 个 ASL 图像作为辅助的未标记数据集，测试 5 折。对于每一折，我们使用网络为辅助数据集生成伪标签，并选择排名最高的 $k = 10$ 个大脑掩膜进行微调。为了评估基于 AE 的排名效用，我们与基于 TTA 的 QC 进行比较，并将选择与真实大脑掩膜最接近的预测作为上限进行对比。对于 TTA，我们通过应用随机强度缩放（最高到 $\pm 20\%$ ）和旋转（在 -30° 和 30° 之间）创建 10 个大脑掩膜副本，使用其体素中位数作为伪标签。作为基线，

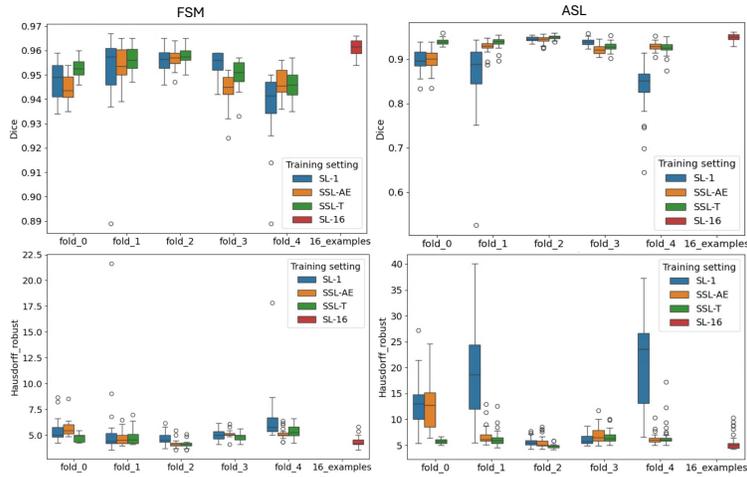


Fig. 3. 头骨剥离准确性在分布内 FSM 和分布外 ASL 图像的表现。我们比较了使用单个样本 (SL-1) 进行训练、进一步使用基于自动编码器重构选择的伪标签 (SSL-AE) 或与真实脑面具最相似的伪标签 (SSL-T) 进行微调, 以及使用 16 个样本 (SL-16) 进行训练的效果。较高的 Dice 值和较低的 Hausdorff 值更好。

我们还在所有 16 个标记的 FSM 示例上训练一个网络, 并使用经过微调的伪标签与从头开始的重新训练进行比较。对于每一折, 我们根据 Dice 得分和 95th 百分位 Hausdorff 距离 (HD95) 评估性能。

2.1 结果

Semi-supervised learning: 图 2 显示了有代表性的分割示例。图 3 比较了使用单一示例 (SL-1) 进行训练、使用真实标签选择伪标签的半监督学习 (SSL-T) 以及基于 AE 的 QC 自动选择伪标签的半监督学习 (SSL-AE)。与使用单一示例进行训练的网络相比, SSL-T 在 FSM 上的准确率从平均 Dice 得分的 0.950 提高到 0.953, HD95 从 5.5 减少到 4.8 mm。在 ASL 中, 差异更为显著: Dice 得分从 0.897 提高到 0.936, HD95 从 12.8 减少到 6.0 mm。我们发现对于 FSM, SL-1 和 SSL-AE 之间几乎没有区别——Dice 得分从 0.949 提高到 0.950, HD95 分别从 5.0 增加到 5.5 mm——但在 ASL 数据上的性能从 0.897 提高到 0.924 Dice 和从 12.8 减少到 7.0 mm HD95。在大多数分析中, SSL 提高了颅骨去除的准确性。然而, 对于第 3 分折, 即使使用真实标签选择伪标签 (SSL-T), 相比于 SL-1, 性能有所下降。

图 4 比较了我们的 SSL-AE 方法与使用 TTA 排序伪标签 (SSL-TTA) 以及使用伪标签来微调初始网络与从头训练的精度。SSL-TTA 的表现最不精准, 与 SL-1 相比, Dice 从 0.950 降至 FSM 的 0.863 和从 ASL 的 0.897 降至 0.624。从头训练 SSL-T 的精度比微调降低, FSM 数据上的 Dice 为 0.943, 而微调后的为 0.953, ASL 数据上的 Dice 为 0.895, 而微调后的为 0.936。

表 1 比较了通过 AE 重建和 TTA 进行伪标签选择的 Dice 分数与真实值分数的平均绝对误差 (MAE) 和皮尔逊相关系数。在基于 TTA 的 QC 中, 我们使

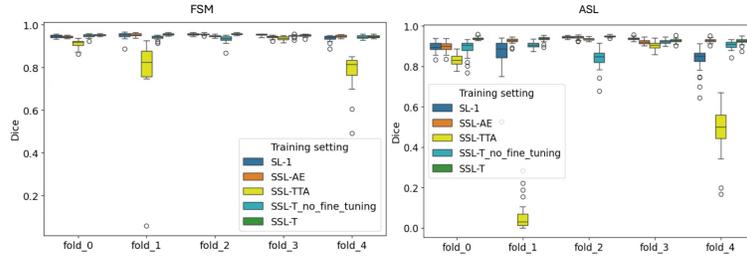


Fig. 4. 消融实验对比了使用单个示例 (SL-1) 进行训练, 与使用通过自动编码器重构选择的伪标签 (SSL-AE)、测试时增强 (SSL-TTA) 或通过与地面实况脑掩模比较 (SSL-T) 进行进一步微调的效果。此外, 我们测试了从头开始的重新训练 (SSL-T 无微调)。

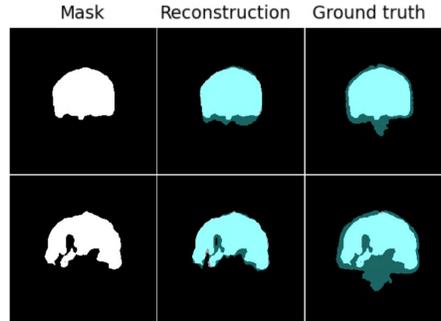


Fig. 5. 自编码器 (AE) 重建。左图: 输入大脑掩膜。中间: AE 输出与输入叠加。右图: 真实掩膜与输入叠加。

用了一个在单一标签图 (SL-1) 上训练的模型。虽然基于 AE 的 QC 与真实值分数达到了 0.708 的较高平均相关性, 但基于 TTA 的 QC 则呈现出一个弱的、负的相关性 -0.1 , 这表明其在估计大脑掩膜质量方面的不足。同样地, 我们发现 AE 的 MAE 显著低至 0.067, 而 TTA 的 MAE 为 0.107。图 5 展示了对一个中部冠状切片的代表性 AE 重建结果。AE 输出与输入在高质量和低质量的大脑掩膜下均有显著差异。然而我们发现低质量输入在重建中发生了更多变化, 这样基于 AE 的伪标签排序可以相较于仅使用单个例子进行训练, 提升 SSL 分割性能。

我们提出了一个新颖的 SSL 框架, 用于训练分割网络, 所需的人为输入仅限于一个有标签的例子。我们的方法结合了最近的图像合成技术和完全自动化的伪标签选择。在一个 3D 颅骨去除任务中, 该方法在异质数据上的性能表现出稳健性, 并且相比于在一个较大有标签数据集上训练的网络, 表现良好。

从单个训练实例开发一个可靠的自动分割质量控制系统本质上是具有挑战性的。在这项工作中, 我们将基于自编码器的质量控制与一种在早期研究中显示出希望的无监督测试时间调整 (TTA) 质量控制方法进行了比较。在只有一个标记示例的情况下, 我们发现基于 TTA 的质量控制不可靠, 而自编码器重建产生的伪标签排名与基于真实标签计算的 Dice 分数强烈相关。这个结果有些令人惊讶, 因为二维自编码器是在单个脑部掩膜切片上操作的。

Table 1. 伪标签 Dice 评估用于测试时增强 (TTA) 和自编码器重建 (AE) 的准确性。我们通过平均绝对误差 (MAE) 和皮尔逊相关系数评估伪标签与真实标签之间的 Dice 评分的一致性。

Fold	MAE ↓		Pearson ↑	
	TTA	AE	TTA	AE
0	0.079	0.048	-0.239	0.427
1	0.214	0.130	-0.170	0.975
2	0.036	0.010	-0.145	0.771
3	0.043	0.025	0.111	0.460
4	0.161	0.122	-0.056	0.909
Mean	0.107	0.067	-0.100	0.708

未来, 我们将研究低质量和高质量脑掩膜的自编码器重建之间的差异, 并探索 3D 自编码器是否可以在 SSL 设置中进一步提高排序精度。我们还计划评估在更大训练集下的性能扩展, 并将我们的方法扩展到更具挑战性的多类分割任务。

References

1. Arbelle, A., Elul, E., Raviv, T.R.: QANet—quality assurance network for image segmentation. arXiv preprint arXiv:1904.08503 (2019)
2. Billot, B., Greve, D.N., Puonti, O., Thielscher, A., Van Leemput, K., Fischl, B., Dalca, A.V., et al.: SynthSeg: segmentation of brain MRI scans of any contrast and resolution without retraining. *Medical Image Analysis* **86**, 102789 (2023)
3. Chen, X., Yuan, Y., Zeng, G., Wang, J.: Semi-supervised semantic segmentation with cross pseudo supervision. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 2613–2622 (2021)
4. Cheplygina, V., De Bruijne, M., Pluim, J.P.: Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. *Medical Image Analysis* **54**, 280–296 (2019)
5. Cox, R.W.: AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research* **29**(3), 162–173 (1996)
6. Fischl, B.: FreeSurfer. *NeuroImage* **62**(2), 774–781 (2012)
7. Fournel, J., Bartoli, A., Bendahan, D., Guye, M., Bernard, M., Rausero, E., Khanji, M.Y., Petersen, S.E., et al.: Medical image segmentation automatic quality control: a multi-dimensional approach. *Medical Image Analysis* **74**, 102213 (2021)
8. Galati, F., Zuluaga, M.A.: Efficient model monitoring for quality control in cardiac image segmentation. In: *International Conference on Functional Imaging and Modeling of the Heart*. pp. 101–111. Springer (2021)
9. Gopinath, K., Hoopes, A., Alexander, D.C., Arnold, S.E., Balbastre, Y., Billot, B., Casamitjana, A., Cheng, Y., Chua, R.Y.Z., Edlow, B.L., et al.: Synthetic data in generalizable, learning-based neuroimaging. *Imaging Neuroscience* **2**, 1–22 (2024)
10. Hoffmann, M.: Domain-randomized deep learning for neuroimage analysis. arXiv preprint arXiv:2507.13458 (2025)

11. Hoopes, A., Mora, J.S., Dalca, A.V., Fischl, B., Hoffmann, M.: SynthStrip: skull-stripping for any brain image. *NeuroImage* **260**, 119474 (2022)
12. Huang, C., Wu, Q., Meng, F.: Qualitynet: segmentation quality evaluation with deep convolutional networks. In: *Visual Communications and Image Processing*. pp. 1–4. IEEE (2016)
13. Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M.: FSL. *NeuroImage* **62**(2), 782–790 (2012)
14. Kelley, W., Ngo, N., Dalca, A.V., Fischl, B., Zöllei, L., Hoffmann, M.: Boosting skull-stripping performance for pediatric brain images. In: *IEEE International Symposium on Biomedical Imaging*. pp. 1–5. IEEE (2024)
15. Lambert, B., Forbes, F., Doyle, S., Dehaene, H., Dojat, M.: Trustworthy clinical AI solutions: a unified review of uncertainty quantification in deep learning models for medical image analysis. *Artificial Intelligence in Medicine* **150**, 102830 (2024)
16. Nie, D., Gao, Y., Wang, L., Shen, D.: ASDNet: Attention based semi-supervised deep networks for medical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention*. pp. 370–378. Springer (2018)
17. Ouali, Y., Hudelot, C., Tami, M.: Semi-supervised semantic segmentation with cross-consistency training. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 12674–12684 (2020)
18. Robinson, R., Oktay, O., Bai, W., Valindria, V.V., Sanghvi, M.M., Aung, N., Paiva, J.M., Zemrak, F., Fung, K., Lukaschuk, E., et al.: Real-time prediction of segmentation quality. In: *Medical Image Computing and Computer-Assisted Intervention*. pp. 578–585. Springer (2018)
19. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention*. pp. 234–241. Springer (2015)
20. Ségonne, F., Dale, A.M., Busa, E., Glessner, M., Salat, D., et al.: A hybrid approach to the skull stripping problem in MRI. *NeuroImage* **22**(3), 1060–1075 (2004)
21. Shen, W., Peng, Z., Wang, X., Wang, H., Cen, J., Jiang, D., Xie, L., Yang, X., Tian, Q.: A survey on label-efficient deep image segmentation: Bridging the gap between weak supervision and dense prediction. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **45**(8), 9284–9305 (2023)
22. Shi, W., Meng, F., Wu, Q.: Segmentation quality evaluation based on multi-scale convolutional neural networks. In: *IEEE Visual Communications and Image Processing*. pp. 1–4. IEEE (2017)
23. Smith, S.M.: Fast robust automated brain extraction. *Human Brain Mapping* **17**(3), 143–155 (2002)
24. Specktor-Fadida, B., Ben-Sira, L., Ben-Bashat, D., Joskowicz, L.: SegQC: a segmentation network-based framework for multi-metric segmentation quality control and segmentation error detection in volumetric medical images. *Medical Image Analysis* p. 103638 (2025)
25. Specktor-Fadida, B., Levchakov, A., Schonberger, D., Ben-Sira, L., Ben Bashat, D., Joskowicz, L.: Test-time augmentation-based active learning and self-training for label-efficient segmentation. In: *Workshop on Medical Image Learning with Limited and Noisy Data*. pp. 213–223. Springer (2023)
26. Specktor-Fadida, B., Link-Sourani, D., Ferster-Kveller, S., Ben-Sira, L., Miller, E., et al.: A bootstrap self-training method for sequence transfer: state-of-the-art placenta segmentation in fetal MRI. In: *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging, and Perinatal Imaging, Placental and Preterm Image*

- Analysis: 3rd International Workshop, UNSURE 2021, and 6th International Workshop, PIPPI 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, October 1, 2021, Proceedings 3. pp. 189–199. Springer (2021)
27. Tajbakhsh, N., Jeyaseelan, L., Li, Q., Chiang, J.N., Wu, Z., Ding, X.: Embracing imperfect datasets: a review of deep learning solutions for medical image segmentation. *Medical Image Analysis* **63**, 101693 (2020)
 28. Xia, Y., Yang, D., Yu, Z., Liu, F., Cai, J., Yu, L., Zhu, Z., Xu, D., Yuille, A., Roth, H.: Uncertainty-aware multi-view co-training for semi-supervised medical image segmentation and domain adaptation. *Medical Image Analysis* **65**, 101766 (2020)
 29. Yang, L., Zhuo, W., Qi, L., Shi, Y., Gao, Y.: St++: make self-training work better for semi-supervised semantic segmentation. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 4268–4277 (2022)
 30. Zeng, X., Puonti, O., Sayeed, A., Herisse, R., Mora, J., Evancic, K., Varadarajan, D., Balbastre, Y., Costantini, I., Scardigli, M., et al.: Segmentation of supragranular and infragranular layers in ultra-high-resolution 7T ex vivo MRI of the human cerebral cortex. *Cerebral Cortex* **34**(9), bhac362 (2024)
 31. Zhang, X., Ou, N., Liu, C., Zhuo, Z., Matthews, P.M., Liu, Y., Ye, C., Bai, W.: Un-supervised brain MRI tumour segmentation via two-stage image synthesis. *Medical Image Analysis* **102**, 103568 (2025)
 32. Zoph, B., Ghiasi, G., Lin, T.Y., Cui, Y., Liu, H., Cubuk, E.D., Le, Q.: Rethinking pre-training and self-training. *Advances in Neural Information Processing Systems* **33**, 3833–3845 (2020)