

國立台灣師範大學 資訊工程學系

National Taiwan Normal University

影像處理課程成果報告書

不同影像預處理方式在 U-Net 卷積神經網路

的重疊程度分析-以腦腫瘤影像為例

Analysis of Different Image Preprocessing Methods on
the Overlap Performance of U-Net Convolutional Neural
Networks: A Case Study on Brain Tumor MRI
Segmentation

學生姓名：徐偉翔 (Wei-Hsiang, Hsu)

課程教授：陳世旺 教授 (Prof. Sei-Wang, Chen)

中華民國 一 一 四 年 六 月

目錄

第一章 緒論	1
第一節 研究背景	1
第二節 研究動機	1
第三節 研究目的	2
第四節 研究問題與假設	2
第二章 研究方法	2
第三章 研究結果	4
第四章 討論	6
第五章 未來展望	6
參考文獻.....	8

第一章 緒論

第一節 研究背景

隨著人工智慧與深度學習技術的快速發展，醫學影像分析逐漸朝向自動化與智慧化發展。在臨床上，腦腫瘤的診斷常依賴磁振造影（MRI）影像，但傳統手工標註腫瘤區域不僅耗時，且容易因醫師經驗差異而造成結果不一致（Isensee et al., 2021）。這樣的挑戰使得如何提升影像分割的準確性與一致性，成為醫學影像領域的重要課題。

在眾多模型中，U-Net 以其編碼器—解碼器（encoder-decoder）架構與跳躍連結（skip connections），已被廣泛應用於醫學影像分割任務，尤其在資料有限的情境下仍能展現良好的表現（Ronneberger, Fischer, & Brox, 2015）。多項研究證實，U-Net 在腦腫瘤影像的分割工作中具備高效能，能協助醫師更準確掌握腫瘤的位置、形狀與邊界，進而提升診斷與治療決策的可靠性（Zhou et al., 2019）。

同時，模型的效能不僅取決於架構設計，影像預處理策略亦扮演關鍵角色。過去文獻指出，預處理方法如高斯濾波（Gaussian）、直方圖均衡化（Histogram Equalization, HE）及對比受限自適應直方圖均衡化（Contrast Limited Adaptive Histogram Equalization, CLAHE），會影響輸入影像的對比度、雜訊與紋理特徵，進而影響深度學習模型的學習效果與分割品質（Pizer et al., 1987; Rehman et al., 2023）。

第二節 研究動機

醫學影像分割在腦腫瘤診斷與治療中具有高度臨床價值，但其結果仍受到多種因素影響，其中影像預處理方式常被忽略。臨床上，腦腫瘤 MRI 影像常存在雜訊、對比不足與灰階分布不均等問題，這些因素可能降低模型對腫瘤邊界的辨識能力（Menze et al., 2015）。適當的影像預處理能改善輸入影像品質，強化目標區域特徵，進而提升深度學習模型的準確性與穩定性（Shen, Wu, & Suk, 2017）。

雖然已有許多研究證實 U-Net 及其變形架構在腦腫瘤影像分割上表現優異（Ronneberger et al., 2015; Isensee et al., 2021），但關於不同影像預處理方式對模型學習效果的比較研究仍有限。例如，高斯濾波（Gaussian）能有效降低影像中的隨機雜訊，使邊界更為平滑，有助於模型在訓練初期快速收斂（Gonzalez & Woods, 2018）；直方圖均衡化（HE）能提升影像整體對比度，幫助模型辨識腫瘤與周邊組織（Pizer et al., 1987）；而對比受限自適應直方圖均衡化（CLAHE）則能針對影像中的局部區域調整亮度與對比，讓細節更清晰呈現，但也可能引入額外雜訊（Rehman et al., 2023）。如何在不同預處理策略間取得平衡，仍是提升醫學影像分割表現的重要課題。

因此，本研究以腦腫瘤 MRI 為例，系統性地比較 Gaussian、HE 與 CLAHE 三種常見影像預處理方式對 U-Net 分割效能的影響。此舉不僅能釐清各種方法的優劣勢，亦可為後續醫學人工智慧應用提供實證參考，協助研究人員與臨床醫師在實務中選擇最適合的影像處理流程。

第三節 研究目的

本研究旨在以腦腫瘤磁振造影 (MRI) 影像為例，系統性比較不同影像預處理方式對 U-Net 模型分割效能的影響。研究將選用三種常見的預處理技術，包括高斯濾波 (Gaussian)、直方圖均衡化 (Histogram Equalization, HE) 以及對比受限自適應直方圖均衡化 (Contrast Limited Adaptive Histogram Equalization, CLAHE)，並以同一組資料集進行訓練與驗證，以確保比較的公平性與一致性。

在效能評估方面，本研究將採用多項客觀指標，包括整體準確率 (Accuracy)、重疊係數 (Dice coefficient) 以及平均交並比 (Mean Intersection over Union, Mean IoU)，以全面檢視不同預處理策略對模型分割品質的影響。

透過上述分析，本研究希望能釐清不同預處理方式在 U-Net 腦腫瘤影像分割任務中的優劣差異，找出能夠提升模型表現的關鍵影像處理流程，並為醫學影像人工智慧應用提供更具實證性與參考價值的依據，協助後續相關研究與臨床決策支持系統的發展。

第四節 研究問題與假設

在本研究中，我們著重於比較三種常見影像預處理方式，Gaussian、HE、以及 CLAHE 在 U-Net 模型進行腦腫瘤 MRI 分割任務時的效能表現。由於影像預處理能顯著影響輸入影像的對比度、紋理特徵與雜訊分布，進而改變模型對腫瘤邊界與細節的辨識能力，因此探討不同策略的影響具有實務與研究上的價值。基於文獻回顧與影像特性考量，我們假設 HE 在提升整體對比度與均衡灰階分布上具備優勢，能使模型更有效地區分腫瘤區域與周邊組織，因此預期 HE 在 Dice coefficient 與 Mean IoU 重疊度相關指標上表現優於 Gaussian 與 CLAHE，並在 Accuracy 指標上亦能展現相對較佳的結果。

第二章 研究方法

本研究使用多來源的腦腫瘤磁振造影 (MRI) 影像資料集，總計約 4000 張影像，其中腫瘤影像與非腫瘤影像各占一半。資料來源包含公開可取得的 Kaggle Brain Tumor MRI Dataset (Nickparvar, 2021)、Figshare Brain Tumor Dataset (Cheng, 2017)，以及臺北榮民總醫院所提供的臨床 MRI 影像。影像標記方面，本研究採用手動標記方式建立腫瘤遮罩 (mask)。標記過程使用 LabelMe 軟體進行腫瘤區域描繪，透過 Polygon 工具在灰階 MRI 影像上圈選腫瘤輪廓，並將結果輸出為二值化遮罩 (mask)。資料劃分則依程式實作固定為訓練集 85% 與驗證集 15%，後續所有比較均在一致的分割下進行。

為確保模型輸入的一致性與品質，本研究將所有影像轉換為灰階，並調整大小至 256×256 像素，強度值 rescale 至 $[0,1]$ 區間。影像預處理分別採用高斯濾波（Gaussian）、直方圖均衡化（Histogram Equalization, HE）與對比受限自適應直方圖均衡化（Contrast Limited Adaptive Histogram Equalization, CLAHE）三種方式。處理後的資料以 NumPy 陣列格式載入，確保與模型輸入格式一致。訓練階段額外施加隨機翻轉、90 度旋轉以及亮度與對比度偏移等資料增強策略，以提升模型對多樣化成像條件的泛化能力。

模型骨幹採用 U-Net 架構（圖 1），其設計以編碼器–解碼器對稱結構結合跳躍連結，適合二維醫學影像分割（Ronneberger et al., 2015）。輸入尺寸設定為 $256 \times 256 \times 1$ 的單通道灰階影像，編碼端特徵通道數逐層遞增為 64、128、256、512，解碼端則依序遞減為 512、256、128、64。最終輸出層使用 1×1 卷積與 sigmoid 激活函數，產生單通道機率圖，每個像素值對應於腫瘤區域的預測機率。

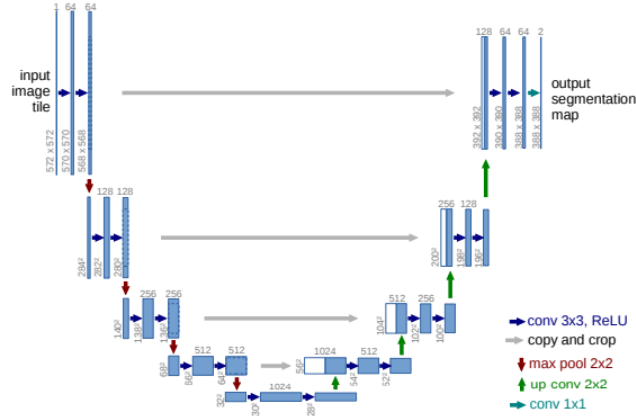


圖 1. U-Net 架構示意圖（改自 Ronneberger et al., 2015）。模型採用對稱式編碼–解碼設計。編碼器透過卷積與池化提取多層次特徵，解碼器則結合跳躍連結與上採樣逐步恢復空間解析度，最終輸出逐像素的分割結果。

在訓練設定上，模型以 batch size = 4 進行，總訓練週期為 50 個 epoch。最佳化器選用 Adam，初始學習率為 1×10^{-4} 。損失函數採用二元交叉熵（Binary Cross Entropy, BCE）與 Tversky Loss 的混合損失以平衡類別不均問題，以減緩腫瘤區域與背景比例不均所造成的學習偏差。其數學形式可表達為：

$$\mathcal{L} = \mathcal{L}_{BCE} + \mathcal{L}_{Tversky} \quad (1)$$

Binary Cross Entropy 的定義為：

$$\mathcal{L}_{BCE} = \frac{1}{N} \sum_{i=1}^N [y_i \log p_i + (1 - y_i) \log(1 - p_i)] \quad (2)$$

其中， y_i 為像素真實標籤， p_i 為模型 sigmoid 輸出機率（Goodfellow, Bengio, & Courville, 2016）。

Tversky Loss 則定義為：

$$\mathcal{L}_{\text{Tversky}} = 1 - \frac{TP}{TP + \alpha FP + \beta FN} \quad (3)$$

其中， TP 、 FP 、 FN 分別為真陽性、假陽性與假陰性， $\alpha = 0.7$, $\beta = 0.3$ 用於調節假陽性與假陰性的懲罰比重 (Salehi, Erdogmus, & Gholipour, 2017)。

為提升模型的泛化能力並避免過擬合，訓練過程中引入 Early Stopping 機制，當驗證集損失在連續數個週期內未顯著下降時即提前終止；同時搭配 ReduceLROnPlateau 策略，當驗證表現停滯時，自動將學習率按比例降低，以確保後期仍具收斂能力。另一方面，整體訓練流程啟用自動混合精度訓練 (automatic mixed precision, AMP) 與梯度縮放 (gradient scaling)，在保持數值穩定性的同時，有效降低記憶體使用量並提升運算效率。

評估指標採用三項常見分割指標，包括準確率 (Accuracy)、重疊係數 (Dice coefficient) 與平均交並比 (Mean Intersection over Union, Mean IoU)。這些指標由 TensorFlow 框架於每個 epoch 自動計算，並同時回報訓練集與驗證集的結果。其中，訓練集指標用於監控模型的學習進展，而最終結果則取自 50 個訓練週期內，驗證集於模型表現最佳回合所達成的指標值，以確保模型效能評估的客觀性與可靠性。三個評估指標的數學公式如下：

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

其中， TP 為真陽性、 TN 為真陰性、 FP 為假陽性、 FN 為假陰性。表示模型正確分類像素的比例。

$$\text{Dice coefficient} = \frac{2 \times |X \cap Y| + \epsilon}{|X| + |Y| + \epsilon} \quad (5)$$

其中，平滑參數以 $\epsilon = 10^{-6}$ 定義， X 為模型預測的腫瘤區域， Y 為真實標註。

$$\text{IoU} = \frac{|X \cap Y|}{|X \cup Y|} \quad (6)$$

$$\text{Mean IoU} = \frac{1}{C} \sum_{i=1}^C \frac{TP_i}{TP_i + FP_i + FN_i}$$

其中， C 為類別數 (本研究為 2：腫瘤/非腫瘤)，最終取兩類平均。

最終，本研究透過在統一的 U-Net 架構與一致的資料劃分設定下，分別對 Gaussian、HE 與 CLAHE 三種預處理方式進行獨立實驗，並以 Accuracy、Dice coefficient 與 Mean IoU 進行系統性比較。此設計使我們能夠直接針對研究問題檢驗不同預處理方式對分割效能的影響，並驗證「HE 在重疊度相關指標上優於 Gaussian 與 CLAHE」之研究假設。

第三章 研究結果

本研究分別比較 Gaussian、HE 與 CLAHE 三種影像預處理方式在 U-Net

模型下的分割效能，所有結果均取自 50 個 epoch 訓練過程中驗證集的最佳回合。整體而言，三種方法在準確率 (Accuracy) 上差異不大，最高值皆約為 0.986 (圖 2)，顯示模型在區分腫瘤與非腫瘤區域上表現穩定。然而，在重疊係數 (Dice coefficient) 指標上，HE 表現最佳，其最大值為 0.1975，高於 CLAHE 的 0.1844 與 Gaussian 的 0.1835 (圖 3)，這意味著 HE 在腫瘤邊界的偵測與區域重疊度上更具優勢；在平均交並比 (Mean IoU) 的比較中，HE 仍然取得最高值 0.5606，相較於 CLAHE 的 0.5572 與 Gaussian 的 0.5551 略有提升 (圖 4)，顯示其在整體空間區域辨識上相對更為穩健。整體而言，三種預處理方式皆能輔助 U-Net 完成腦腫瘤影像分割，但 HE 在各項指標中整體表現最為優異，特別是在 Dice coefficient 與 Mean IoU 等重疊度相關指標上表現最為突出，驗證了研究假設的合理性。

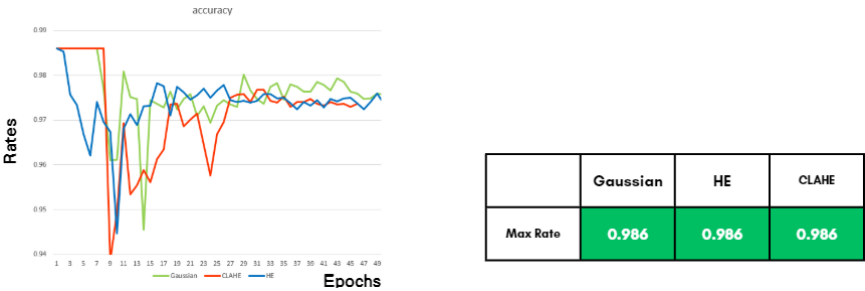


圖 2. 三種影像預處理方式在 U-Net 模型下的準確率表現。左圖為 Gaussian、HE 與 CLAHE 在 50 個訓練週期中的準確率變化，初期波動明顯，但隨著訓練進行，三者逐漸趨於穩定，最終皆收斂至約 0.97。右表為各方法於驗證集的最高準確率，Gaussian、HE 與 CLAHE 皆達到 0.986，差異不大，顯示不同預處理方式對最終準確率影響有限。

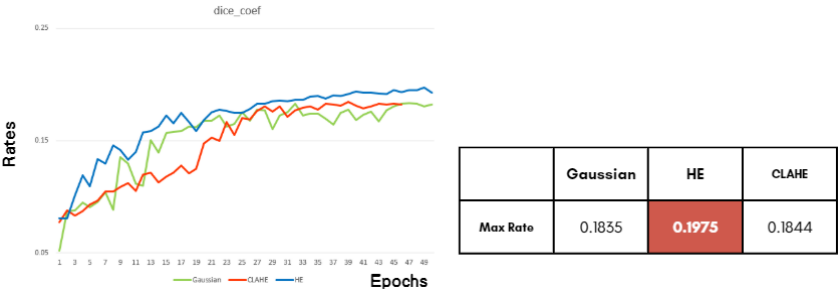


圖 3. 三種影像預處理方式在 U-Net 模型下的 Dice coefficient 表現。左圖顯示 Gaussian、HE 與 CLAHE 在 50 個訓練週期中的 Dice coefficient 變化，三者皆隨訓練逐步提升，HE 全程表現最佳。右表為各方法的最高值，HE 為 0.1975，高於 Gaussian (0.1835) 與 CLAHE (0.1844)，顯示 HE 在重疊度指標上最具優勢。

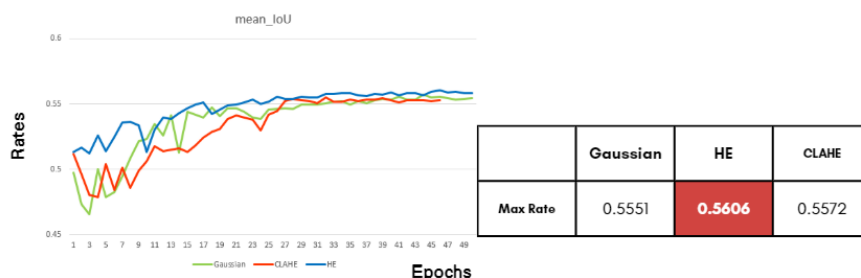


圖 4. 三種影像預處理方式在 U-Net 模型下的 Mean IoU 表現。左圖顯示 Gaussian、HE 與 CLAHE 在 50 個訓練週期中的 Mean IoU 變化，三者皆隨訓練逐步提升並趨於穩定。右表為各方法的最高值，HE 為 0.5606，略高於 CLAHE（0.5572）與 Gaussian（0.5551），顯示 HE 在空間區域重疊度上的整體表現最佳。

第四章 討論

本研究結果顯示，三種預處理方式在 U-Net 模型下皆能穩定完成腦腫瘤 MRI 影像的分割任務，整體準確率相近，代表模型在腫瘤與非腫瘤的區別上表現一致。然而，Dice coefficient 與 Mean IoU 等重疊度相關指標則凸顯了預處理策略的差異。Gaussian 在早期訓練中收斂較快，顯示降噪處理有助於初期特徵學習，但中後期成長趨緩，最終效果不及其他方法。CLAHE 雖能改善局部對比，但可能放大區域性雜訊，使模型對邊界的辨識受限。相較之下，HE 在整體訓練過程中表現平穩，最終在 Dice coefficient 與 Mean IoU 均取得最佳結果，顯示其增強整體對比度的特性更有助於腫瘤邊界的學習與分割。綜合來看，HE 是三種方法中對 U-Net 分割效能影響最顯著的策略，支持研究假設的合理性。

儘管本研究驗證了不同預處理方式對分割效能的影響，但仍存在若干限制。首先，標記方式雖以手動標註完成，但部分樣本的邊界標記並未能完全貼合腫瘤實際輪廓，可能導致模型學習偏差。其次，資料集中腫瘤樣本數有限且來源異質，影像分布存在差異，可能影響模型在真實臨床場景的泛化能力。此外，最終的 Dice coefficient 與 Mean IoU 仍偏低，顯示模型對腫瘤邊界的捕捉能力不足，無法達到臨床實用標準。

第五章 未來展望

本研究針對不同影像預處理方式在 U-Net 模型腦腫瘤分割中的影響進行比較，結果顯示 Histogram Equalization (HE) 整體效能最佳。然而，現有結果仍有改進空間，未來研究可從幾個方向延伸，首先，可擴充資料來源並增加腫瘤樣本的多樣性，特別是不同病理類型與影像品質的 MRI 資料，以提升模型在實際臨床情境下的泛化能力。其次，可嘗試導入更先進的模型架構，例如 Attention U-Net、UNet++ 或結合 Transformer 的 SwinUNETR，以補足傳統 U-Net 在捕捉細微邊界特徵上的不足。第三，未來可進一步評估多重預處理方式的組合效果，或探索自適應式影像增強策略，建立更靈活的前處理流程，提升模型對不同成像

條件的適應性。最後，隨著臨床應用需求提升，未來研究亦可將模型與臨床決策支援系統結合，驗證其在實際診斷與治療規劃中的效益，推進腦腫瘤智慧化影像分析的臨床落地。

參考文獻

- Cheng, J. (2017). Brain tumor dataset. Figshare. <https://doi.org/10.6084/m9.figshare.1512427.v5>
- Gonzalez, R. C., & Woods, R. E. (2018). Digital image processing (4th ed.). Pearson.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
- Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: A self-adapting framework for U-Net-based medical image segmentation. Nature Methods, 18(2), 203–211. <https://doi.org/10.1038/s41592-020-01008-z>
- Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner, E., Weber, M. A., Arbel, T., Avants, B. B., Ayache, N., Buendia, P., Collins, D. L., Cordier, N., Corso, J. J., ... Van Leemput, K. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). IEEE transactions on medical imaging, 34(10), 1993–2024. <https://doi.org/10.1109/TMI.2014.2377694>
- Nickparvar, M. (2021). Brain tumor MRI dataset. Kaggle. <https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection>
- Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., ... & Zimmerman, J. B. (1987). Adaptive histogram equalization and its variations. Computer Vision, Graphics, and Image Processing, 39(3), 355–368. [https://doi.org/10.1016/S0734-189X\(87\)80186-X](https://doi.org/10.1016/S0734-189X(87)80186-X)
- Maragatham, G., & Roomi, S. M. M. (2015). A Review of Image Contrast Enhancement Methods and Techniques. Research Journal of Applied Sciences, Engineering and Technology, 9(5), 309–326. <https://www.airitilibrary.com/Article/Detail?DocID=20407467-201502-201506220034-201506220034-309-326>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 (pp. 234–241). Springer. https://doi.org/10.1007/978-3-319-24574-4_28
- Salehi, S. S. M., Erdogmus, D., & Gholipour, A. (2017). Tversky loss function for image segmentation using 3D fully convolutional deep networks. In Proceedings of the International Workshop on Machine Learning in Medical Imaging (MLMI 2017) (pp. 379–387). Springer. https://doi.org/10.1007/978-3-319-67389-9_44
- Chan, H. P., Samala, R. K., Hadjiiski, L. M., & Zhou, C. (2020). Deep Learning in Medical Image Analysis. Advances in experimental medicine and biology, 1213, 3–21. https://doi.org/10.1007/978-3-030-33128-3_1

Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2019). UNet++: A nested U-Net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (pp. 3–11). Springer. https://doi.org/10.1007/978-3-030-00889-5_1