Kim KW
Cardiff University
School of Engineering

Duman A
Cardiff University
School of Engineering

Spezi E
Cardiff University
School of Engineering

AI AND DEEP LEARNING

RGU-Net: Computationally Efficient U-Net for Automated Brain Extraction of mpMRI with Presence of Glioblastoma

Brain extraction refers to the process of removing non-brain tissues in brain scans and is one of the initial pre-processing procedures in neuroimage analysis. Since errors produced during this process can be challenging to amend in subsequent analyses, accurate brain extraction is crucial. Most deep learning-based brain extraction models are optimised on performance, leading to computationally expensive models. Such models may be ideal for research; however, they are not ideal in a clinical setting. In this work, we propose a new computationally efficient 2D brain extraction model, named RGU-Net. RGU-Net incorporates Ghost modules and residual paths to accurately extract features and reduce computational cost. Our results show that RGU-Net has 98.26% fewer parameters compared to the original U-Net model, whilst yielding state-of-the-art performance of 97.97 ± 0.84% Dice similarity coefficient. Faster run time was also observed in CPUs which illustrates the model's practicality in real-world applications.

Keywords:
Brain extraction, skull stripping, U-Net, lightweight, deep learning, magnetic resonance imaging.

Corresponding author:
KimKW1@cardiff.ac.uk

K.W. Kim, A. Duman, and E. Spezi, 'RGU-Net: Computationally Efficient U-Net for Automated Brain Extraction of mpMRI with Presence of Glioblastoma', *Proceedings of the Cardiff University School of Engineering Research Conference 2024*, Cardiff, UK, 2024, pp. 28-32.

doi.org/10.18573/conf3.h

INTRODUCTION

Brain extraction, also known as brain segmentation or skull stripping, is one of the initial pre-processing stages for neuroimage analysis and refers to the process of separating brain and non-brain (e.g. eyes, neck, skull) tissues in Magnetic Resonance Imaging (MRI) scans through medical image segmentation [1, 2]. The quality of brain extraction has a significant impact on subsequent analysis as errors produced during this process will be present in further stages of the analysis. To this day, manual segmentation is still considered the "gold standard", however, such a method is prone to introducing inter and intra-rater variability between radiologists and the process is extremely time-consuming [3].

In recent years, there has been an increase in the demand for neuroimage analysis and the size of datasets, resulting in the difficulty of performing manual segmentation. To tackle this problem, several deep learning models have been developed [4 - 9], as the implementation of the deep learning approach allows the decrease of the cost associated with the radiologist and the time taken to produce the segmentation. However, most studies to date focused on increasing the performance, often resulting in computationally inefficient models which are not ideal for deployment in a clinical setting.

To address this problem, we developed a computationally efficient 2D brain extraction model which has similar performance as the original U-Net [10] but requires significantly less computational power. The proposed model, Residual Ghost U-Net (RGU-Net), integrates Ghost modules and residual paths to enhance feature extraction whilst decreasing the number of trainable parameters to achieve computational efficiency. Our objective is not to claim the proposed model's optimality for the task or to compare the model to existing brain extraction algorithms, but instead focus on potential lightweight architecture to increase the practicality of deep learning-based approaches in clinical settings.

MATERIALS AND METHODS

Dataset: UPenn-GBM
The proposed model was developed and validated using the publicly available dataset, UPenn-GBM[1]. The UPenn-GBM dataset contains multi-parametric MRI (mpMRI) scans of 630 patients diagnosed with glioblastoma and a brain segmentation produced using a deep learning model [11]. The dataset includes 611 MRI scans prior to initial surgery and 60 follow-up scans for the second resection surgery. For this study, four structural MRI scans were used: T1-weighted (T1; n=671), gadolinium-based agent-enhanced T1 (T1Gd; n=671), T2-weighted (T2; n=671), and Fluid Attenuated Inversion Recovery (FLAIR; n=671) scans. The dataset was split into training (n=1612), validation (n=536), and testing (n=536) datasets.

Data Pre-processing and Augmentation
To develop a computationally efficient model, a 2D segmentation approach was proposed as 3D convolutional layers are computationally expensive to operate. Due to this, the MRI scans were converted from NIfTI imaging format to a series of 2D JPEG images.

Additionally, data augmentation and transformation techniques were applied to increase the variety of the training and validation data. Five transformations were applied to the original scans: horizontal flip, vertical flip, rotation, random brightness, and elastic transformation. Each transformation had a 50% chance of being applied, allowing random combinations of transformations, creating a unique dataset with a wide variety of data.

Network Architecture and Implementation
RGU-Net (Fig. 1) is a modification of the U-Net architecture [10] optimised for automated brain extraction, with a focus on decreasing computational cost through the incorporation of Ghost modules and residual paths. Ghost modules, which were first proposed by Han et al. [12], were implemented to produce feature maps at a lower computational cost compared to the conventional convolutional layer approach. In addition to this, residual paths [13] were also integrated into the architecture to allow the deeper layers of the model to capture appropriate features during the training.

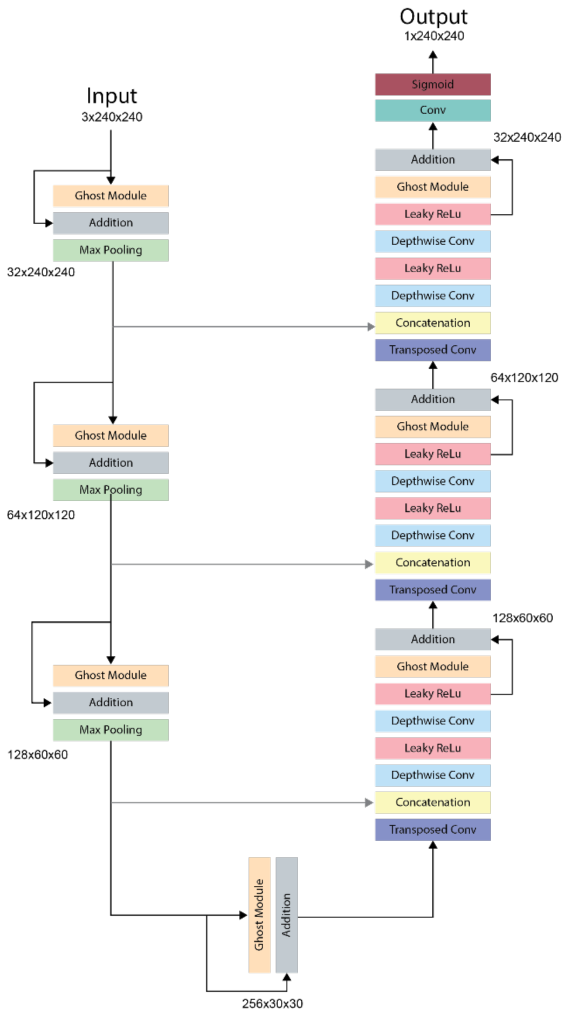

Fig. 1. Overview of RGU-Net architecture.

The ablation study showed that the combination of Ghost modules and residual paths alone reduces the performance of the model. Therefore, asymmetrical encoder-decoder architecture was introduced to further enhance the performance. The decoder block incorporates two additional depthwise convolutional layers followed by Leaky ReLU (α=0.01). This approach further reduces noise

1 doi.org/10.7937/TCIA.709X-DN49

and increases performance, by reconstructing accurate segmentation masks of the brain. The combination of Ghost modules, residual paths, and asymmetrical encoder-decoder has shown success in producing accurate predicted masks and requires significantly fewer trainable parameters.

RGU-Net was implemented using the PyTorch framework [14]. During the training process, a Dice loss function was implemented to monitor the performance of the model. The Dice loss function ranges from 0 to 1, where 0 means the predicted mask (PM) is identical to the ground truth (GT). The loss function is expressed as:

$$L_{Dice} = 1 - \frac{2 \cdot |PM \cap GT|}{|PM| + |GT|} \tag{1}$$

Furthermore, to prevent overfitting of the model, a validation loop was implemented after every training epoch, and the model parameters were saved when the validation loss improved. The model was trained with the Adam optimiser algorithm [15] with an initial learning rate of 0.0001. If the validation loss does not improve for three epochs, the learning rate decays by a factor of 0.1. The decaying learning rate aids the model to learn complex patterns and increases the performance [16]. Lastly, to decrease training time, an early stopper function was implemented to stop the training when the model converges.

Post-processing
3D segmentation was produced through the concatenation of a series of 2D PMs. Thresholding and a post-processing technique introduced by Lucena et al. [4] were applied to the concatenated 3D masks. This allowed the reduction of floating segments and noise, producing a clean brain segmentation mask.

RESULTS

The proposed model was evaluated using the Dice similarity coefficient (DSC), which compares the similarity between the PM and the GT [17]. The DSC can be calculated as follows:

$$DSC = \frac{2 \cdot |PM \cap GT|}{|PM| + |GT|} \times 100 \tag{2}$$

Additionally, the number of model parameters were calculated using the model summary function from PyTorch.

Figure 2 shows the segmentation results in mpMRI scans of three random slices from each MRI sequence to illustrate the general segmentation generated by RGU-Net. The red arrow indicates areas of over/under-segmentation of the brain.

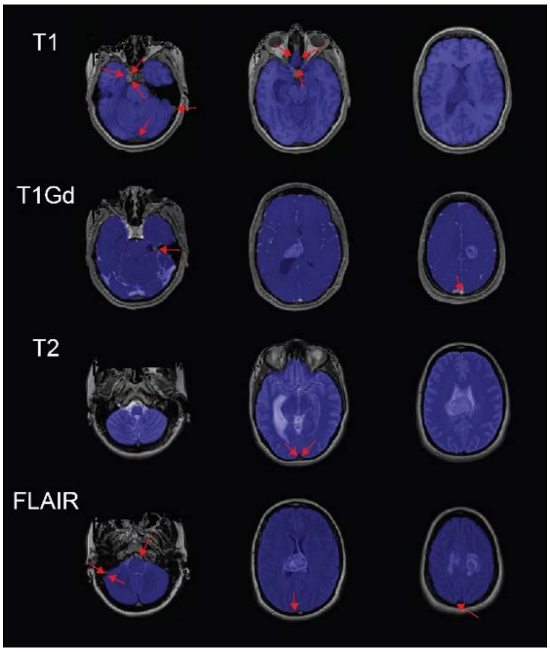

Fig. 2. The segmentation results of RGU-Net in different MRI sequences. Red arrows indicate regions of over/under-segmentation.

Ablation study
The ablation study was conducted to investigate the effect of encoder and decoder blocks using a baseline model and two additional variations of RGU-Net: U-Net (baseline), RGU-Net without residual paths, RGU-Net without asymmetrical encoder-decoder, and RGU-Net. All models were implemented as specified and hyper-parameters were kept constant. Table 1 shows the comparison of model performances. Furthermore, the average run time to produce a 3D segmentation on a CPU device (Intel Core i7-1165G7) was measured and showed that the proposed RGU-Net is 80 seconds faster than the U-Net model.

Model	No. Param	Run time (s)	DSC (%)			
			T1	T1Gd	T2	FLAIR
U-Net (baseline)	31.03M	113	98.10.9	98.01.0	97.80.9	97.61.0
RGU-Net w/o residual paths	0.56M	19	97.71.0	97.51.1	97.40.9	97.11.1
RGU-Net w/o asymmetrical encoder-decoder	0.65M	24	97.61.0	97.51.1	97.31.0	97.11.4
RGU-Net	0.54M	33	98.00.8	97.81.0	97.70.9	97.51.1

Table 1. Performance comparison for ablation study.

DISCUSSION

RGU-Net demonstrated state-of-the-art performance in automated brain extraction, whilst maintaining significantly less trainable parameters, as shown in Table 1. Despite slightly lower performance compared to the U-Net, all differences are considered to be clinically insignificant since segmentations produced by trained radiologists have greater variance between segmentation masks [18].

All models in the ablation study exhibited the highest performance on T1 scans, followed by T1Gd, T2, and FLAIR scans, suggesting sensitivity to the different contrast settings of MRI scans. This could be mitigated by introducing contrast transformation to the training data during data augmentation. While RGU-Net demonstrates high performance. In some MRI slices, signs of over/under-segmentation were observed (Fig. 2), suggesting further optimisation of the model is needed. The use of a single dataset during the development of the model reduced the variety of image acquisition within the training data. Leading to poorer performance when MRI scans with different imaging parameters are inputted. To address this, an increase in the number of datasets and input from radiology consultants is highly recommended.

Furthermore, to increase the accuracy of the model whilst decreasing the number of parameters, an asymmetrical encoder-decoder architecture was introduced. However, the method increased the number of operations, and Table 1 suggests a correlation between the model complexity and the run time of the model. Therefore, further investigation on the model's complexity and run time is required to further optimise the model to improve computational efficiency. Additionally, to decrease computational cost of the model, 2D convolutional layers were implemented, limiting the model's ability to capture 3D contextual and spatial information. Although the high DSC implies that this loss is not significant, it is worth exploring a potential implementation of the 3D convolutional layer for optimised performance.

One of the major strengths of RGU-Net is the computational efficiency of the model, achieved through the implementation of Ghost modules and residual paths which make the model ready for in real-world applications in environments where access to powerful computing devices is not warranted. This is crucial for the integration of automated brain extraction tools into clinical workflow, as the run time of the model will accumulate when inputted with a large dataset.

CONCLUSION

Overall, RGU-Net showed promising results for automated brain extraction. Its high performance, computational efficiency and compatibility with CPUs make the model a valuable tool for neuroimage analysis. Further research and optimisation of the model could further enhance the reliability, performance and computational efficiency of the model.

Conflicts of Interest
The authors declare no conflict of interest.

Data Availability Statement
The UPenn-GBM dataset is publicly available at The Cancer Imaging Archive (TCIA): doi.org/10.7937/TCIA.709X-DN49. Source code of this study has been made publicly available at: https://github.com/KWKIM128/Brain_Extraction

REFERENCES

[1] H. Z. U. Rehman, H. Hwang, and S. Lee, ‘Conventional and Deep Learning Methods for Skull Stripping in Brain MRI’, *Applied Sciences*, vol. 10, no. 5, p. 1773, Mar. 2020. doi.org/10.3390/app10051773

[2] P. Kalavathi and V. B. S. Prasath, ‘Methods on Skull Stripping of MRI Head Scan Images—a Review’, *J. Digit. Imaging*, vol. 29, no. 3, pp. 365–379, Jun. 2016. doi.org/10.1007/s10278-015-9847-8

[3] F. Yepes-Calderon and J. Gordon McComb, ‘Manual Segmentation Errors in Medical Imaging. Proposing a Reliable Gold Standard’, in *Applied Informatics*, vol. 1051, H. Florez, M. Leon, J. M. Diaz-Nafria, and S. Belli, Eds., Cham: Springer International Publishing, 2019, pp. 230–241. doi.org/10.1007/978-3-030-32475-9_17

[4] O. Lucena, R. Souza, L. Rittner, R. Frayne, and R. Lotufo, ‘Convolutional neural networks for skull-stripping in brain MR imaging using silver standard masks’, *Artificial Intelligence in Medicine*, vol. 98, pp. 48–58, Jul. 2019. doi.org/10.1016/j.artmed.2019.06.008

[5] S. Thakur *et al*., ‘Brain extraction on MRI scans in presence of diffuse glioma: Multi-institutional performance evaluation of deep learning methods and robust modality-agnostic training’, *NeuroImage*, vol. 220, p. 117081, Oct. 2020. doi.org/10.1016/j.neuroimage.2020.117081

[6] F. Isensee *et al*., ‘Automated brain extraction of multisequence MRI using artificial neural networks’, Human Brain Mapping, vol. 40, no. 17, pp. 4952–4964, Dec. 2019. doi.org/10.1002/hbm.24750

[7] L. Pei *et al*., ‘A general skull stripping of multiparametric brain MRIs using 3D convolutional neural network’, *Sci. Rep.*, vol. 12, no. 1, p. 10826, Jun. 2022. doi.org/10.1038/s41598-022-14983-4

[8] A. Hoopes, J. S. Mora, A. V. Dalca, B. Fischl, and M. Hoffmann, ‘SynthStrip: skull-stripping for any brain image’, *NeuroImage*, vol. 260, p. 119474, Oct. 2022. doi.org/10.1016/j.neuroimage.2022.119474

[9] J. S. Park, S. Fadnavis, and E. Garyfallidis, ‘Multi-scale V-net architecture with deep feature CRF layers for brain extraction’, *Commun. Med.*, vol. 4, no. 1, p. 29, Feb. 2024. doi.org/10.1038/s43856-024-00452-8

[10] O. Ronneberger, P. Fischer, and T. Brox, ‘U-Net: Convolutional Networks for Biomedical Image Segmentation’, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, vol. 9351, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., Cham: Springer International Publishing, 2015, pp. 234–241. doi.org/10.1007/978-3-319-24574-4_28

[11] S. Bakas *et al.*, ‘The University of Pennsylvania glioblastoma (UPenn-GBM) cohort: advanced MRI, clinical, genomics, & radiomics’, *Sci. Data*, vol. 9, no. 1, p. 453, Jul. 2022. doi.org/10.1038/s41597-022-01560-7

[12] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, ‘GhostNet: More Features From Cheap Operations’, in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA: IEEE, Jun. 2020, pp. 1577–1586. doi.org/10.1109/CVPR42600.2020.00165

[13] K. He, X. Zhang, S. Ren, and J. Sun, ‘Deep Residual Learning for Image Recognition’, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR*), Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 770–778. doi.org/10.1109/CVPR.2016.90

[14] A. Paszke *et al*.,’Pytorch: An imperative style, high-performance deep learning library’. *NeurIPS* 2019, 32. doi.org/10.48550/ARXIV.1912.01703

[15] D. P. Kingma and J. Ba, ‘Adam: A Method for Stochastic Optimization’, *arXiv*, Jan. 29, 2017. doi.org/10.48550/arXiv.1412.6980

[16] K. You, M. Long, J. Wang, and M. I. Jordan, ‘How Does Learning Rate Decay Help Modern Neural Networks?’, *arXiv*, 26 Sep. 2019. doi.org/10.48550/arXiv.1908.01878

[17] L. R. Dice, ‘Measures of the Amount of Ecologic Association Between Species’, *Ecology*, vol. 26, no. 3, pp. 297–302, Jul. 1945. doi.org/10.2307/1932409

[18] M. A. Deeley *et al*., ‘Comparison of manual and automatic segmentation methods for brain structures in the presence of space-occupying lesions: a multi-expert study’, *Phys. Med. Biol.*, vol. 56, no. 14, pp. 4557–4577, Jul. 2011. doi.org/10.1088/0031-9155/56/14/021

E. Spezi and M. Bray (eds.) 2024. *Proceedings of the Cardiff University School of Engineering Research Conference 2024.* Cardiff: Cardiff University Press.
doi.org/10.18573/conf3

Cardiff University School of Engineering Research Conference 2024 was held from 12 to 14 June 2024 at Cardiff University.

The work presented in these proceedings has been peer reviewed and approved by the conference organisers and associated scientific committee to ensure high academic standards have been met.

First published 2024

Cardiff University Press
Cardiff University, Trevithick Library
First Floor, Trevithick Building, Newport Road
Cardiff CF24 3AA

cardiffuniversitypress.org

Editorial design and layout by
Academic Visual Communication

ISBN: 978-1-9116-5351-6 (PDF)