Digital Health and Informatics Innovations for Sustainable Health Care Systems
J. Mantas et al. (Eds.)

doi:10.3233/SHTI240745

Development of a CNN for Adult Brain Tumour Characterisation: Implications and Future Directions for Transfer Learning

Teesta MUKHERJEE [a], Shramika GOUR [a], Saadullah Farooq ABBASI [a], Omid POURNIK [a] and Theodoros N. ARVANITIS [a,1]
[a] *Department of Electronic, Electrical and Systems Engineering, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*
ORCiD ID: Teesta Mukherjee https://orcid.org/0009-0003-8023-8317
Shramika Gour https://orcid.org/0009-0006-5585-5040
Saadullah Abbasi https://orcid.org/0000-0001-9814-3023
Omid Pournik https://orcid.org/0000-0001-7938-0269
Theodoros N. Arvanitis https://orcid.org/0000-0001-5473-135X

Abstract. Brain tumours are the most commonly occurring solid tumours in children, albeit with lower incidence rates compared to adults. However, their inherent heterogeneity, ethical considerations regarding paediatric patients, and difficulty in long-term follow-up make it challenging to gather large homogenous datasets for analysis. This study focuses on the development of a Convolutional Neural Network (CNN) for brain tumour characterisation using the adult BraTS 2020 dataset. We propose to transfer knowledge, from models pre-trained on extensive adult brain tumour datasets to smaller cohort datasets (e.g., paediatric brain tumours) in future studies, by leveraging Transfer Learning (TL). This approach aims to extract relevant features from pre-trained models, addressing the limited availability of annotated paediatric datasets and enhancing tumour characterisation in children. The implications and potential applications of this methodology in paediatric neuro-oncology are discussed.

Keywords. Brain Tumour, Image Processing, Feature Extraction, Transfer Learning.

1. Introduction

Brain tumours are the most common solid tumours in children, although their incidence rates differ between age groups. Adults exhibit a higher incidence rate of 29.9 per 100,000 individuals, while within the paediatric cohort, the incidence rate accounts for 5.7 per 100,000 children [1]. These tumours are heterogenous and vary by age, sex, and ethnicity [2]. Their rarity, variability, and low availability of large, annotated datasets pose challenges in diagnosis and treatment planning. Transfer learning emerges as a potential solution to bridge this gap, by transferring knowledge from pre-trained

[1] Corresponding Author: Professor Theodoros N. Arvanitis, Department of Electronic, Electrical and Systems Engineering, School of Engineering, College of Engineering and Physical Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom; E-mail: t.arvanitis@bham.ac.uk.

models on a larger adult brain tumour dataset to enhance the performance of the same model in smaller cohorts of children's brain tumours. By fine-tuning the model, it adapts to accommodate the unique features of paediatric tumours, variations in imaging protocols, and differences in imaging characteristics between the two datasets. This leads to improved characterisation of paediatric brain tumours [3].

In this paper, we outline a research approach to develop a robust CNN for the classification of adult brain tumours. While our current work primarily emphasises the development of the CNN architecture, we acknowledge the significance of future steps involving fine-tuning the model to accommodate the unique features of paediatric tumours. Our aim is to improve diagnostic accuracy and patient outcomes by leveraging advanced computational techniques in model construction and subsequent evaluation.

2. Methods

2.1. Study design and datasets

This study is conducted in accordance with the European Society for Paediatric Oncology (SIOPE) brain tumour imaging protocol [4]. The architecture of the proposed methodology is shown in Figure 1.

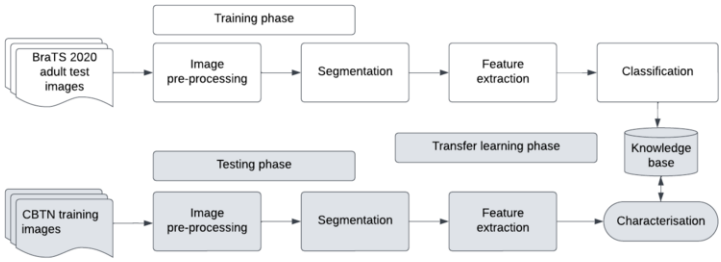

Figure 1. Proposed methodology for CNN development, including proposal for future transfer learning in childhood brain tumour characterisation.

For training the model, the publicly available Multimodal Brain Tumour Segmentation (BraTS) 2020 dataset was acquired [5]. This dataset consists of 660 cases of Glioblastoma (GBM/ HGG) and Lower Grade Glioma (LGG), totalling 2640 multiparametric 3T MRI scans. These scans come with ground truth labels provided by expert neuroradiologists [6]. The scans are stored as NIfTI files (.nii.gz) and describe four different volumes of the same region: Native (T1), Post-contrast T1-weighted (T1CE), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (FLAIR) volumes. They were acquired using various protocols and scanners from 19 institutions.

All images have been manually segmented, approved by experts, and annotated as follows: Label 0 represents unlabelled volume, Label 1 denotes Necrotic and Non-Enhancing Tumour Core (NCR/NET), Label 2 signifies Peritumoral Edema (ED), Label 3 indicates missing data, and Label 4 represents GD-Enhancing Tumour (ET).

Although the testing of the model is outside the scope of this paper, for completeness we discuss the datasets we plan to use, in the future, for this part of the work. During the model testing phase, the Children's Brain Tumor Network (CBTN)

dataset will be used. The CBTN cohort underwent brain MR imaging on 1.5T or 3T Siemens scanners. The imaging sequences include 2D axial T2-weighted turbo spin-echo (with TR/ TE values ranging from 1000 to 7300 ms and 80 to 530 ms, respectively, and section thickness ranging between 0.5 to 5mm), along with 3D axial or sagittal pre-contrast, and 3D axial gadolinium-based contrast agent-enhanced T1-weighted turbo or fast-field echo scans [7].

2.2. Image pre-processing

Images from the BraTS dataset were normalised to the range [0,1], and mask pixel values were adjusted for consistency. Different MRI volumes were combined into single multichannel images, and both images and masks were cropped. Volumes with less than 1% useful information were removed. The dataset was then split into 75:25 training and validation sets using splitfolders for organised data management. These pre-processing steps including intensity normalisation, image registration, and skull stripping optimise the dataset for accurate semantic segmentation model training, facilitating accurate brain tumour analysis [8]. To further improve the quality of MRI scans, SimpleITK will be used for N4 bias field correction along with HD-BET for removing any remaining non-brain tissues.

2.3. Segmentation and feature extraction

Automatic segmentation of tumour regions has been performed using the U-Net algorithm, due to its high accuracy in delineating objects within medical imaging. Additionally, U-Net has feature extraction capabilities, capturing relevant tumour characteristics essential for subsequent analysis. These extracted features serve as a foundation for building the CNN used here, which has been trained to classify brain tumours based on their distinct characteristics. This integrated approach, combining both segmentation and classification using CNN, ensures improved characterisation of brain tumours [9].

2.4. Model evaluation and statistical analysis

The performance and accuracy of the final model will be evaluated by metrics such as sensitivity, specificity, accuracy, and area under the receiver operating characteristic curve. The model will also undergo cross-validation and testing on an independent dataset to assess its generalisability across different children's brain tumour cohorts.

3. Results

The preliminary results of the proposed study as shown in Figure 2, indicates the completion of image pre-processing and segmentation on the BraTS 2020 dataset. A custom data generator for batch processing has been implemented, along with pre-processing steps such as intensity normalisation, rescaling, image registration, and skull stripping.

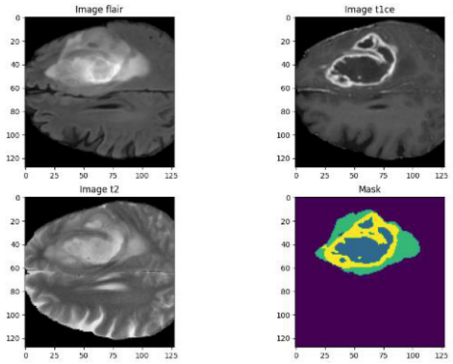

Figure 2. Plotting of FLAIR, T1CE, and T2 volumes with segmentation mask showing NCR, ED, and ET.

4. Discussion

Our future work will involve fine-tuning of the CNN model to accommodate the unique features of paediatric brain tumours, recognizing the significance of this step in extending the applicability of our approach to paediatric neuro-oncology, where the challenges differ from those encountered in adult tumour analysis. By leveraging advanced computational techniques such as TL where the model trained on one task is adapted for a related task [10], we aim to address the limited availability of paediatric datasets and other variations in imaging characteristics between adult and paediatric cohorts.

Childhood brain tumours exhibit distinct biological and clinical characteristics compared to adult tumours, due to their unique molecular signatures and histological features, influencing tumour growth and appearance on imaging scans [11, 12]. The rarity of childhood brain tumour cases poses a challenge in model development, with smaller and less diverse datasets available for training and validation [13]. Specialised imaging protocols may be required for paediatric patients due to factors such as patient cooperation, movement, and high-water content, leading to variations in imaging characteristics between adult and childhood brain tumour datasets [14]. Addressing these technical challenges is crucial to ensure the development of accurate diagnostic models across diverse age groups.

This study emphasises the critical need for comprehensive datasets on childhood brain tumours, highlighting the potential of multi-institutional collaboration to improve disease management. Acknowledging potential limitations is essential. For instance, the limited diversity in existing datasets and concerns regarding the generalisability of the developed model in diverse patient cohorts require further investigation [15]. Future research should explore ways to refine and enhance this approach, potentially through broader data collection efforts and validation studies across various healthcare settings and populations.

5. Conclusions

In conclusion, this research lays the groundwork for the development of a CNN-based approach for brain tumour classification, with implications for both adult and paediatric

applications. The proposed methodology has the potential to significantly improve clinical decision-making in paediatric neuro-oncology, leading to more accurate diagnoses and better patient outcomes.

Acknowledgements

The authors extend their gratitude to the School of Engineering, College of Engineering and Physical Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK for providing Teesta Mukherjee's EPSRC Studentship Scholarship PhD studies.

References

[1] Greuter L, Guzman R, Soleman J. Typical pediatric brain tumors occurring in adults—differences in management and outcome. Biomedicines. 2021 Mar;9(4):356, doi:10.3390/biomedicines9040356

[2] Ostrom QT, Fahmideh MA, Cote DJ, Muskens IS, Schraw JM, Scheurer ME, et al. Risk factors for childhood and adult primary brain tumors. Neuro-Oncology. 2019 Jul;21(11):1357-75, doi:10.1093/neuonc/noz123

[3] Boyd A, Ye Z, Prabhu S, Tjong MC, Zha Y, Zapaishchykova A, et al. Expert-level pediatric brain tumor segmentation in a limited data scenario with stepwise transfer learning. medRxiv [Preprint]. 2023 Sept, doi:10.1101/2023.06.29.23292048

[4] Avula S, Peet A, Morana G, Morgan P, Warmuth-Metz M, Jaspan T. European Society for Paediatric Oncology (SIOPE) MRI guidelines for imaging patients with central nervous system tumours. Child's Nervous System. 2021 May;37(8):2497-508, doi:10.1007/s00381-021-05199-4

[5] Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BraTS). IEEE Transactions on Medical Imaging. 2015 Oct;34(10):1993-2024, doi:10.1109/tmi.2014.2377694

[6] Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby JS, et al. Advancing the cancer genome Atlas Glioma MRI collections with expert segmentation labels and Radiomic features. Nature Scientific Data. 2017 Sept;4(1):170117, doi:10.1038/sdata.2017.117

[7] Lilly JV, Rokita JL, Mason JL, Patton T, Stefankiewiz S, Higgins D, et al. The children's brain tumor network (CBTN) - Accelerating research in pediatric central nervous system tumors through collaboration and open science. Neoplasia. 2023 Jan;35:100846, doi: 10.1016/j.neo.2022.100846

[8] Moradmand H, Aghamiri SM, Ghaderi R. Impact of image preprocessing methods on reproducibility of radiomic features in multimodal magnetic resonance imaging in glioblastoma. Journal of Applied Clinical Medical Physics. 2019 Dec;21(1):179-90, doi:10.1002/acm2.12795

[9] Bhalodiya JM, Lim Choi Keung SN, Arvanitis TN. Magnetic resonance image-based brain tumour segmentation methods: A systematic review. DIGITAL HEALTH. 2022 Jan;8:205520762210741, doi:10.1177/20552076221074122

[10] Valverde JM, Imani V, Abdollahzadeh A, De Feo R, Prakash M, Ciszek R, et al. Transfer learning in Magnetic Resonance Brain Imaging: A systematic review. Journal of Imaging. 2021 Apr;7(4):66, doi:10.3390/jimaging7040066

[11] AlRayahi J, Alwalid O, Mubarak W, Maaz AU, Mifsud W. Pediatric brain tumors in the molecular era: Updates for the radiologist. Seminars in Roentgenology. 2023 Jan;58(1):47-66. doi:10.1053/j.ro.2022.09.004

[12] Pollack IF, Jakacki RI. Childhood brain tumors: Epidemiology, current management and Future Directions. Nature Reviews Neurology. 2011 Jul;7(9):495-506. doi:10.1038/nrneurol.2011.110

[13] Shaari H, Kevrić J, Jukić S, Bešić L, Jokić D, Ahmed N, et al. Deep learning-based studies on pediatric brain tumors imaging: Narrative review of techniques and challenges. Brain Sciences. 2021 May;11(6):716. doi:10.3390/brainsci11060716

[14] Barkovich MJ, Li Y, Desikan RS, Barkovich AJ, Xu D. Challenges in pediatric neuroimaging. NeuroImage. 2019 Jan;185:793-801. doi:10.1016/j.neuroimage.2018.04.044

[15] Familiar AM, Mahtabfar A, Fathi Kazerooni A, Kiani M, Vossough A, Viaene A, et al. Radio-Pathomic Approaches in pediatric neuro-oncology: Opportunities and challenges. Neuro-Oncology Advances. 2023 Jan;5(1):vdad119, doi:10.1093/noajnl/vdad119