Image Classification Neural Networks for Horizontal Brain MRI

Proposed by Zhuo Cao, Tse-Wei Lee, Longpeng Xu, Yingqi Zhang

September 22, 2023

1 Topic and Significance

We propose to build multiple neural networks on horizontal brain MRI images (Chatterjee et al., 2022), for brain tumor detection which will be highly accurate. Neural networks like autoencoder, basic CNN, ResNet50, DenseNet and VGG16 are concerned. The outcome will be the best optimized model fully capable of classifying the unseen MRI image data as healthy samples or brain tumor samples.

The project shed light on the global significance on efficient, accurate detection of brain tumors. In Australia, brain tumor is the ninth most common cause of death from cancer (Cancer Australia, 2019). In the world, 308,102 people were diagnosed with brain tumor in 2020. Although the 5-year survival rate for patients aged from 15 to 39 is 72%, the rate drops to 21% for ages over 40 (ASCO, 2023). Therefore, it is important to diagnose the brain tumor as early as possible.

2 Feasibility

2.1 Completed Work

We have completed dataset selection and preprocessing. A dataset by Hamada (2021) is selected for the following reasons. (1) The brains are imaged from the same perspective (one of horizontal, coronal, and sagittal). (2) The images are with binary labels which indicates with-/without tumor(s). (3) The resolutions of the images are generally finer than 150×150, ensuring the quality of model training. (4) The images are of mixed imaging contrasts (T1, T2, or diffusion weighted), ensuring the generalizability of neural networks.

For preprocessing, firstly, we conducted resizing and intensity normalization. Setting a uniform size for each image can improve consistency of the dataset. Intensity carries information about tissue properties, and normalization produces unified images for highly predictive models (Yildirim et al., 2022). For MRI images, Gaussian normalization has subtle advantages over others (Ellingson et al., 2012).

Secondly, we conducted denoising, which will improve predicting accuracy. Many techniques are available, like wavelet transform (Wang et al., 2012), median filter (Ali & Mohammed, 2022), and NN-based approach like autoencoders (Samuel et al., 2023) and DnCNN (Zhang et al., 2022). We selected autoencoders, since it is up-to-date, not an overkill, capable of handling various complex noises, and without side effects like blurring the sharp boundaries. Three stages of using autoencoder include: initially, define an autoencoder by a Python class; then, train the autoencoder; finally, denoise images that are unseen by the autoencoder.

Thirdly, we conducted augmentation, which transforms the existing images. We configured PyTorch module `torchvision.transforms.Compose` so that the augmented images are mixed with original ones automatically. This improves the generalizability of CNN models and expands the original dataset (Fort et al., 2021).

2.2 Planned Work

2.2.1 Model Selection and Architectural Design

We plan to evaluate the following CNN models and compare their performance on the validation set.

Basic CNN Model. Our research aims to design and implement a bespoke CNN architecture featured by: (1) convolutional layers beginning with smaller kernel sizes like 3×3 to capture local patterns; (2) activation function like ReLU due to its nonlinear properties and computational efficiency (Velasco-Forero & Angulo, 2022); (3) pooling Layers like max-pooling to reduce spatial dimensions while retaining critical features (Zhang & Ma, 2020); (4) dense Layers which are positioned towards the end of the network, as these are pivotal for classification decisions.

ResNet. ResNet (He et al., 2015), a.k.a. Residual Network, is crafted to mitigate gradient vanishing and exploding in deep networks. Its salient features include residual blocks incorporating shortcut connections that skip one or more layers, and stacking of Blocks, where depth is achieved by layering these blocks sequentially.

DenseNet. DenseNet (Huang et al., 2018) enhances information flow by establishing direct connections between each layer. Every layer obtains output from all preceding layers as its input. Growth Rate defines the increment in the number of channels between layers within a dense block.

Transfer Learning with Pre-trained Networks. By leveraging extant pre-trained models like VGG16 or ResNet50, the research stands to benefit from features learned on extensive datasets. Optionally, layers of the pre-trained model can be unfrozen and trained on the dataset to further optimize performance.

2.2.2 Training Strategy

We have decided to deploy the Adam optimizer, motivated by the following facets.

Adaptive Learning Rates. Adam offers the advantage of adjusting the learning rate for each parameter.

Momentum Integration. It amalgamates the strengths of Momentum optimization and RMSProp (Kurbiel & Khaleghian, 2017), facilitating quicker convergence. If validation performance does not show enhancement over several epochs, the learning rate will be reduced.

2.2.3 Model Evaluation

We will evaluate the models' performance based on the following criteria.

Accuracy Score. This assesses the model's ability to correctly classify brain tumor images, providing insights into the model's overall performance in this binary classification task.

Confusion Matrix. It is a 2×2 matrix for each model, offering insights based on precision which is calculated as the ratio of true positives to the sum of true and false positives, representing the model's accuracy in predicting positive tumor images.

Recall. It is the ratio of true positives to the sum of true positives and false negatives, important for assessing the model' s capability to correctly identify actual tumor images, crucial in minimizing misclassification in real-world applications.

F1-Score. This is the harmonic mean of precision and recall, balancing the two metrics. A high F1-score denotes high precision and recall achieved by the model.

Computation Cost. In the project, it specifically refers to the number of epochs set during model training will be adjusted to achieve a balance between the training time and the model performance.

3 Expected Outcome

In post-training stages like evaluation, models will bi-classify unseen images into 1 (an MRI image with tumor(s)) and 0 (without tumors). The accuracy score will record correct classifications, and a confusion matrix will illustrate each model's results. Models will be compared based on accuracy, complexity, and computational cost to select the best one. Misclassified images will be analyzed for future performance improvement.

4 Conclusion

Summarily, we have set the background, models and criteria to realize accurate brain tumor detection. The planned work will be elaborated in the final report. Hence, the project proceeds smoothly with fruitful outcomes, thanks to the hard work by every group member.

5 Reference

Ali, A. M., & Mohammed, A. A. (2022). Improving classification accuracy for prostate cancer using noise removal filter and deep learning technique. *Multimedia Tools and Applications, 81*(6), 8653–8669. https://doi.org/10.1007/s11042-022-12102-z

ASCO, Cancer. N. (2023, May 31). *Brain Tumor - statistics.* https://www.cancer.net/cancer-types/brain-tumor/statistics

Cancer Australia, A. G. (2019, December 18). *Brain cancer in Australia statistics.* in Australia statistics | Cancer Australia. https://www.canceraustralia.gov.au/cancer-types/brain-cancer/statistics

Chatterjee, S., Nizamani, F. A., Nürnberger, A., & Speck, O. (2022). Classification of brain tumours in MR images using deep spatiospatial models. *Scientific Reports, 12*(1). https://doi.org/10.1038/s41598-022-05572-6

Ellingson, B. M., Zaw, T., Cloughesy, T. F., Naeini, K. M., Lalezari, S., Mong, S., Lai, A., Nghiemphu, P. L., & Pope, W. B. (2012). Comparison between intensity normalization techniques for dynamic susceptibility contrast (DSC)-MRI estimates of cerebral blood volume (CBV) in human gliomas. *Journal of Magnetic Resonance Imaging, 35*(6), 1472–1477. https://doi.org/10.1002/jmri.23600

Fort, S., Brock, A., Pascanu, R., De, S., & Smith, S. L. (2021). *Drawing Multiple Augmentation Samples Per Image During Training Efficiently Decreases Test Error.* https://doi.org/10.48550/arxiv.2105.13343

Hamada, A. (2021, November 14). *BR35H :: Brain tumor detection 2020.* Kaggle. https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detection

He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition* (arXiv:1512.03385). arXiv. https://doi.org/10.48550/arXiv.1512.03385

Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2018). *Densely Connected Convolutional Networks* (arXiv:1608.06993). arXiv. https://doi.org/10.48550/arXiv.1608.06993

Kurbiel, T., & Khaleghian, S. (2017). *Training of Deep Neural Networks based on Distance Measures using RMSProp* (arXiv:1708.01911). arXiv. https://doi.org/10.48550/arXiv.1708.01911

Samuel, S., Ochawar, R. S., & Rukmini, M. S. S. (2023). Hybrid deep autoencoder network based adaptive cross guided bilateral filter for motion artifacts correction and denoising from MRI. *The Imaging Science Journal, ahead-of-print*(ahead-of-print), 1–16. https://doi.org/10.1080/13682199.2023.2196494

Wang, Y., Che, X., & Ma, S. (2012). Nonlinear filtering based on 3D wavelet transform for MRI denoising. *EURASIP Journal on Advances in Signal Processing, 2012*(1), 1–14. https://doi.org/10.1186/1687-6180-2012-40

Velasco-Forero, S., & Angulo, J. (2022). MorphoActivation: *Generalizing ReLU activation function by mathematical morphology* (arXiv:2207.06413). arXiv. https://doi.org/10.48550/arXiv.2207.06413

Yildirim, Z., Hançer, E., Samet, R., Mali, M.T., & Nemati, N. (2022). Effect of Color Normalization on Nuclei Segmentation Problem in H&E Stained Histopathology Images. *2022 30th Signal Processing and Communications Applications Conference (SIU)*, 1-4.

Zhang, C., Benz, P., Argaw, D. M., Lee, S., Kim, J., Rameau, F., Bazin, J.-C., & Kweon, I. S. (2021). *Resnet or DenseNet? Introducing dense shortcuts to resnet.* 2021 IEEE Winter Conference on Applications of Computer Vision (WACV). https://doi.org/10.1109/wacv48630.2021.00359

Zhang, H., & Ma, J. (2020). *Hartley Spectral Pooling for Deep Learning.* https://doi.org/10.4208/csiam-am.2020

Zhang, G., Huang, C., Jiang, J., Xu, W., Chen, J., & Xu, X. (2022). Denoising of brain magnetic resonance images using a MDB network. *Multimedia Tools and Applications, 81*(29), 41751–41763. https://doi.org/10.1007/s11042-021-11521-8