Supplemental Materials: Details of Image Pre-Processing, and the CNN Model's Development, Training, and Optimization.

Image Pre-processing for CNN Training

The training set was first visually inspected. Images that had poor quality, including artifact, were removed. The remaining high-quality images underwent extraction of multiple patches (256 × 256 pixels) from within an automatically delineated region of interest that corresponded to the brightest portion of the image (or greater than the 50th percentile of overall image intensity). We increased the number of these patches by performing morphological image post-processing operations to increase the extent of the region of interest and focus the CNN toward relevant portions of the image.

Details of the Model's Development, Training, and Optimization

We modified the original VGG-19 by removing all the layers after layer 16. These subsequent layers are focused on extracting features specific to classifying natural images that are not applicable to the current study. We then added an average pooling layer that groups the feature activations and extracts the most relevant ones. Two fully connected layers were added with a rectified linear unit (*ReLU*) activation function that clips feature activations to stay between 0.0 and 1.0.[1] A *dropout* layer was included to regularize training. Dropout is a commonly used technique in deep learning whereby features are randomly removed during training to increase the robustness of the network to new, unseen images.[2] Finally, an output layer with a *softmax* activation function was used to produce the probabilistic classification of tumor or no tumor.[1]

We trained the model using three-fold crossvalidation, which we have previously shown to be effective for image-based cancer classification.[3] It is beneficial in reducing overfitting, a catastrophic failure of artificial learning modeling. Overfitting occurs when the model fits the training set perfectly, including the unnecessary noise, and it becomes ungeneralizable to other new data sets.[4] In crossvalidation, 1 fold is utilized for evaluation, and the other folds are used for model training. In this way, multiple models are created with each being tested on a subset of the data that was not used for training. The model with the best performance is selected as the final model and evaluated with the testing data set.[5]

The CNN was optimized using Nesterov momentum accelerated adaptive moment optimization with an initial learning rate of 1e-4, and early stopping after a maximum of 40 epochs was utilized with a mini-batch size of 5. A small learning rate, early stopping of the network training, and mini-batch training (random groups of data are presented to the CNN for training) ensured robustness of the CNN to future, unseen data.[6] Additionally, cross-entropy loss was computed to evaluate the probability of misclassification and update the CNN's parameters during training.

REFERENCES

1. Nwankpa CE, Ijomah W, Gachagan A, Marshall S. Activation functions: comparison of trends in practice and research for deep learning 2018. Located at: arXiv. Accessed: June 22, 2021.

2. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res*. 2014;15:1929–1958.

3. Fehr D, Veeraraghavan H, Wibmer A, et al. Automatic classification of prostate cancer Gleason scores from multiparametric magnetic resonance images. *Proc Natl Acad Sci USA*. 2015;112:E6265–E6273.

4. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging*. 2018;9:611–629.

5. Ebigbo A, Palm C, Probst A, et al. A technical review of artificial intelligence as applied to gastrointestinal endoscopy: clarifying the terminology. *Endosc Int Open*. 2019;7:E1616–E1623.

6. Yasaka K, Akai H, Kunimatsu A, Kiryu S, Abe O. Deep learning with convolutional neural network in radiology. *Jpn J Radiol*. 2018;36:257–272.