README

This program contains some models capable of classifying benign and malignant brain tumors based on the MRI image. There are three image classification models that we explore, namely Convolutional Neural Network (CNN), Support Vector Machines (SVM), and Decision Trees. In the end, we compare the performance of those three models. **This program can only be run at Google Colab**.

Before Running the Program

Download the dataset here and save it locally. You can find MRI images of two types of brain tumors: malignant and benign. There are 154 malignant images and 77 benign images. In each folder, the images are saved as JPG files. Each image is in grayscale and has a size of 244 x 244.

[Optional] If you want to run this program using the pre-trained model, download the trained model here and save it locally. You can find one weight for the CNN methods and one for the SVM-CNN model, each containing five folders named train_1 to train_5.

Running the Program

1. Run Section 2.1 to import the libraries and dataset
    - When running the third cell, click the 'Choose Files' button and select all of the Benign images in your local file. Click upload to upload all of them at once. Do not upload a single folder, you have to upload all of the benign images at once. After uploading, you will see a folder named 'Benign' containing all of the images that you have uploaded in the Colab's File.
    - When running the fourth cell, click the 'Choose Files' button and select all of the Malignant images in your local file. Click upload to upload all of them at once. Do not upload a single folder, you have to upload all of the malignant images at once. After uploading, you will see a folder named 'Malignant' containing all of the images that you have uploaded in the Colab's File.

2. Run Section 2.2 to do data augmentation

   After running this section, you will see a folder called ‘augmented-images’ containing augmented benign and malignant images in the Colab’s File. Running the last cell in this section will show you the number of augmented benign and malignant pictures.

3. Run the first cell of Section 2.3 to define the function to crop the images

   Now, we want to see an example of an image before and after being cropped. You have to key in the filename of the image BY YOURSELF following these steps :

   - Go to the Colab’s file, look at the contents of the ‘augmented-images’ folder.
   - You can choose to see the image from either benign or malignant. Open the respective folder and choose one image.
   - In the second cell of Section 2.3, change the content of the filename variable following this format “[label]/[filename.jpg]”
   - For example, if you want to choose the image from the benign folder with the name “aug_10_0_1163.jpg”, then change the content of the filename variable to be “benign/aug_10_0_1163.jpg”.

   Run the last cell of Section 2.3 to see the image before and after being cropped.

4. Run Section 2.4 to load the data into numpy array
5. Run Section 2.5 to convert the 4D data into 2D data
6. Run Section 2.6 to split the data into train, validation, and test data
7. Run Section 3.1 to define the evaluation functions
8. Run the first cell Section 3.2 to create the CNN model.

   If you want to use the pre-trained model

   - Do not run Section 3.2.1.
   - Click the ‘Choose Files’ button. Upload EVERY FILE inside the ‘Train_1’ of the CNN folder AT ONCE. Do not upload a single folder and do not upload each file one by one. After that, another ‘Choose Files’ button will appear.
   - Click the ‘Choose Files’ again, then upload EVERY FILE inside the ‘Train_2’ of the CNN folder AT ONCE. Do not upload a single folder. After that, another ‘Choose Files’ button will appear.
   - Repeat the same steps for ‘Train_3’, ‘Train_4’, ‘Train_5’. After you successfully upload all the files in ‘Train_5’, you can continue to Section 3.2.3

If you don't want to use the pre-trained model

- Run Section 3.2.1 to compile, fit, and do k-fold validation with the model.
- Do not run Section 3.2.2. You can directly go to Section 3.2.3.

Run Section 3.2.3 to see the summary of the model's performance (regardless of having the pretrained model or not).

9. Run the first cell of Section 3.3 to create the hybrid SVM-CNN model.

   If you want to use the pre-trained model

   - Do not run Section 3.3.1.
   - Click the 'Choose Files' button. Upload EVERY FILE inside the 'Train_1' of the SVM-CNN folder AT ONCE. Do not upload a single folder and do not upload each file one by one. After that, another 'Choose Files' button will appear.
   - Click the 'Choose Files' again, then upload EVERY FILE inside the 'Train_2' of the SVM-CNN folder AT ONCE. Do not upload a single folder. After that, another 'Choose Files' button will appear.
   - Repeat the same steps for 'Train_3', 'Train_4', 'Train_5'. After you successfully upload all the files in 'Train_5', you can continue to Section 3.3.3

   If you don't want to use the pre-trained model

   - Run Section 3.3.1 to compile, fit, and do k-fold validation with the model.
   - Do not run Section 3.3.2. You can directly go to Section 3.3.3.

   Run Section 3.3.3 to see the summary of the model's performance (regardless of using the pre-trained model or not).

10. Run Section 3.4 to compile the Decision Tree model. In the end of this section, you can see how well is the performance of this model
11. Run section 4 to see the performance comparison of those three models.