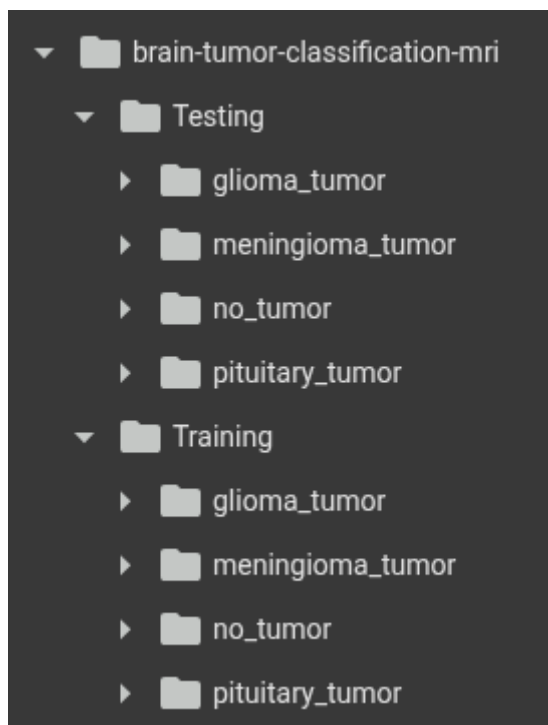


## Bases de dados

### 1. Base de dados utilizada para treinamento (referência: Bhuvaji et al. 2020)

- **Quantas imagens o artigo diz que utilizou?** 3264 imagens. Desse total, sendo:
  - Meningioma: 937 imagens.
  - Glioma: 926 imagens.
  - Hipofisário: 901 imagens.
  - Sem tumor: 500 imagens.
- **Link informado nas referências do artigo para baixar a base:** <https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri/data>
- **Após baixar a base, é possível identificar a seguinte estrutura:**



- **Após baixar a base, existem imagens faltando?**

Para cada categoria de tumor, vamos contar quantas imagens existem dentro das pastas testing e training:

- Meningioma: 115 (testing) + 822 (training) = 937 imagens (bate com a quantidade informada no artigo).

- Glioma: 100 (testing) + 826 (training) = 926 imagens (bate com a quantidade informada no artigo).
- Hipofisário: 74 (testing) + 827 (training) = 901 imagens (bate com a quantidade informada no artigo).
- Sem tumor: 105 (testing) + 395 (training) = 500 imagens (bate com a quantidade informada no artigo).

Logo, não existem imagens faltando.

- **As imagens estão em que formato?**

Todas as imagens estão em formato .jpg.













- **As imagens estão em tons de cinza?**

Sim.

## 2. Primeira base de dados utilizada para validação (referência: Cheng 2017)

- **Quantas imagens o artigo diz que utilizou?** 3064 imagens. Desse total, sendo:
  - Meningioma: 708 imagens.
  - Glioma: 1426 imagens.
  - Hipofisário: 930 imagens.
- **Link informado nas referências do artigo para baixar a base:**  
[https://figshare.com/articles/dataset/brain\\_tumor\\_dataset/1512427](https://figshare.com/articles/dataset/brain_tumor_dataset/1512427)
- **Observação sobre o link acima:**

Ao clicar no link acima, você perceberá que existem vários arquivos:

brainTumorDataPublic_1-766.zip 204.47 MB		
brainTumorDataPublic_1533-22... zip 205.58 MB		
brainTumorDataPublic_767-1532.zip 207.76 MB		
brainTumorDataPublic_2299-30... zip 220.95 MB		
cvind.mat 5.6 kB		
README 2024.txt 3.23 kB		

Para essa base, as imagens foram divididas nos quatro zips acima. Cada zip contém 766 imagens. Logo, será necessário baixar cada um.

- **Após baixar cada zip e extrair, existem imagens faltando?**

Não, cada pasta possui 766 imagens.  $766 \times 4 = 3064$  imagens (bate com a quantidade informada no artigo).

- **As imagens estão em que formato?**

Todas as imagens estão em formato .mat. Talvez seja necessário converter para .jpg (ou outro formato desejado). No arquivo README 2024.txt (mostrado na imagem acima) tem informações de como cada arquivo .mat está estruturado.

- **As imagens estão em tons de cinza?**

Sim.

### **3. Segunda base de dados utilizada para validação (referência: Chakrabarty 2017)**

- **Quantas imagens o artigo diz que utilizou?** 1500 imagens, sendo todas do tipo sem tumor.

- **Link informado nas referências do artigo para baixar a base:**  
<https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection>

- **Após baixar a base, existem imagens faltando?**

Sim, existem e muitas. O artigo menciona que utiliza 1500 imagens (do tipo sem tumor) dessa base, porém ao clicar no link acima e baixar a base, no total existem apenas 253 imagens (98 do tipo sem tumor, 155 do tipo com tumor). Logo, é necessário avaliar o que será feito para essa base.

- **As imagens estão em que formato?**

Todas as imagens estão em formato .jpg.

- **As imagens estão em tons de cinza?**

Sim.