Multimodal representation learning of brain pathology

Information Retrieval - 27 June 2023

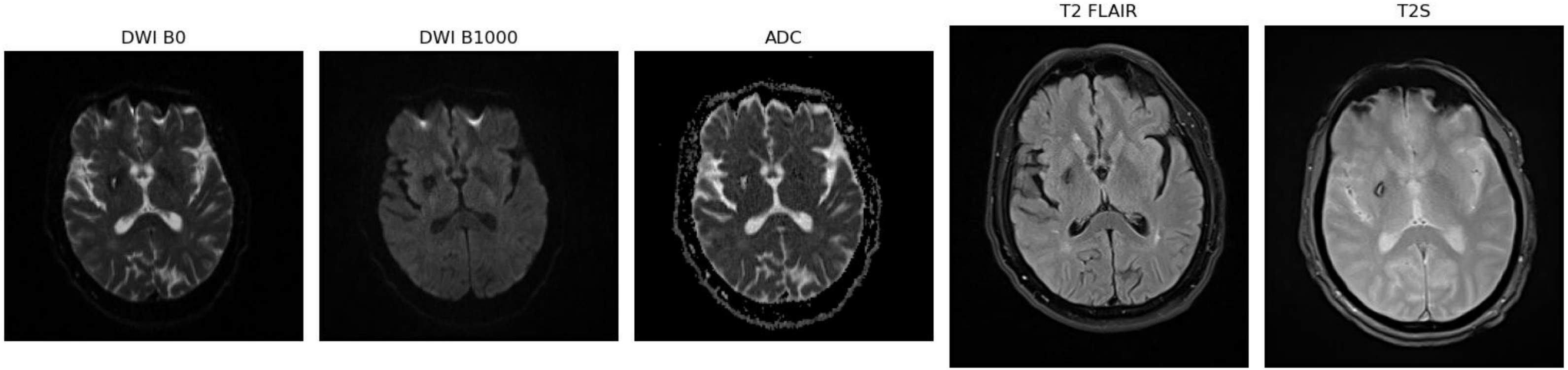
Alice Schiavone

Multimodal representation learning of brain pathology

1. The input is made by multiple data formats
2. Features that represent a data point
3. Machine Learning (Deep Learning)
4. Data domain and task definition

Brain pathology

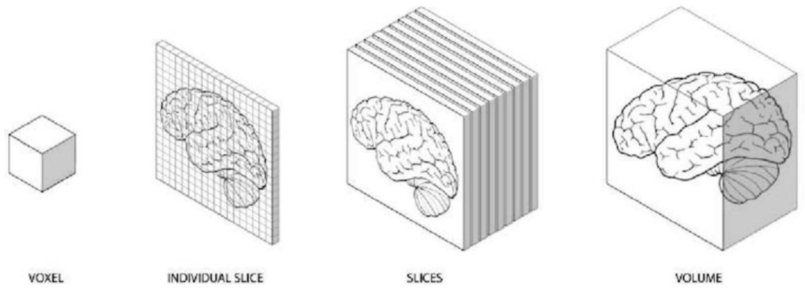
1. Neurologist visits patient and might require brain imaging
2. Technologist captures brain anatomy and pathology via multiple MRI sequences
3. Radiologists annotate findings on the image and on a text report



"Chronic infarcts in the right lentiform nucleus, right corona radiata, body of right caudate nucleus posterior limb of right internal capsule. Hemosiderin deposits in the posterior limb of right internal capsule Age related cerebral atrophy with Fazekas Grade I small vessel ischemic changes. Left maxillary sinusitis."

Data

- Image acquisition results in a 3-d volume (atomic unit: voxel)
- Images are usually not isotropic
    - MRI tends to lose information along the third dimension
- Huge domain shift when testing on new data



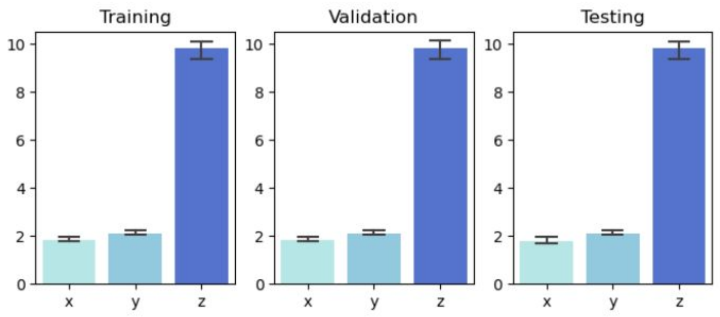


Fig. 2 Voxel size distribution per image dimension.

Data

- More data samples are available (~1100 cases) but around half of them belong to a single class (normal brain).
- Text data is made by simple annotations and is often repetitive
- Data samples classified automatically by a NLP algorithm

	total	infarct	tumor	hemorrhage	normal	others
training	211	40	47	29	48	47
validation	24	5	6	5	3	5
testing	27	5	6	4	6	6

Table 1 Distribution of samples into training splits by class.

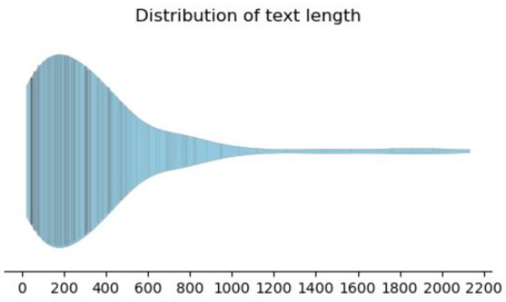


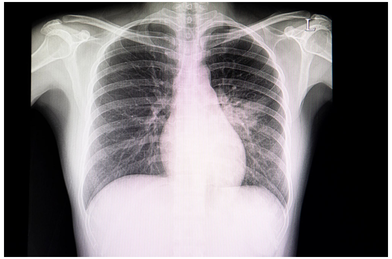
Fig. 3 Distribution of length of medical reports in number of characters.

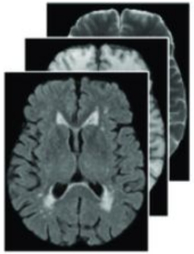
Preprocessing

- No preprocessing for text, except for:
    - Tokenization
    - Clipping to 512 tokens (required by pre-trained language model)
- Image preprocessing:
    - Stack 3 MRI sequences along a new dimension
        - Will be fed as image channels (analogously to RGB 2-d images)
    - Resample and crop to common size: 112x112x16x3
    - Apply data augmentation on-the-fly
        - Affine transform (scaling, rotation, translation)
        - Gaussian noise
        - Flipping (left-right)
        - Change of contrast

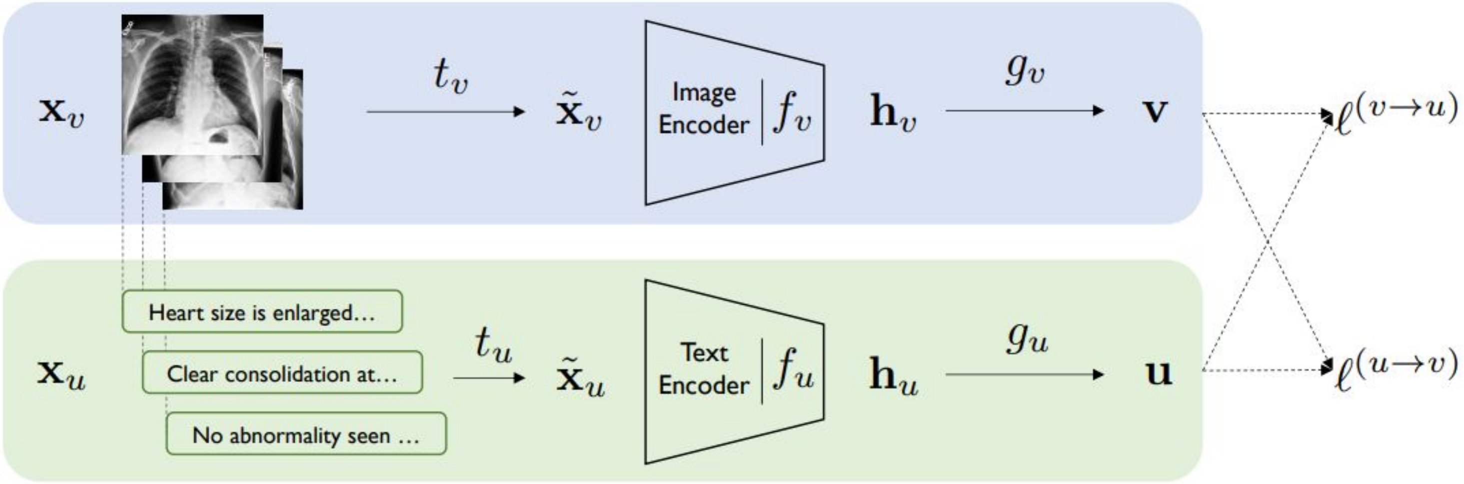
Previous work

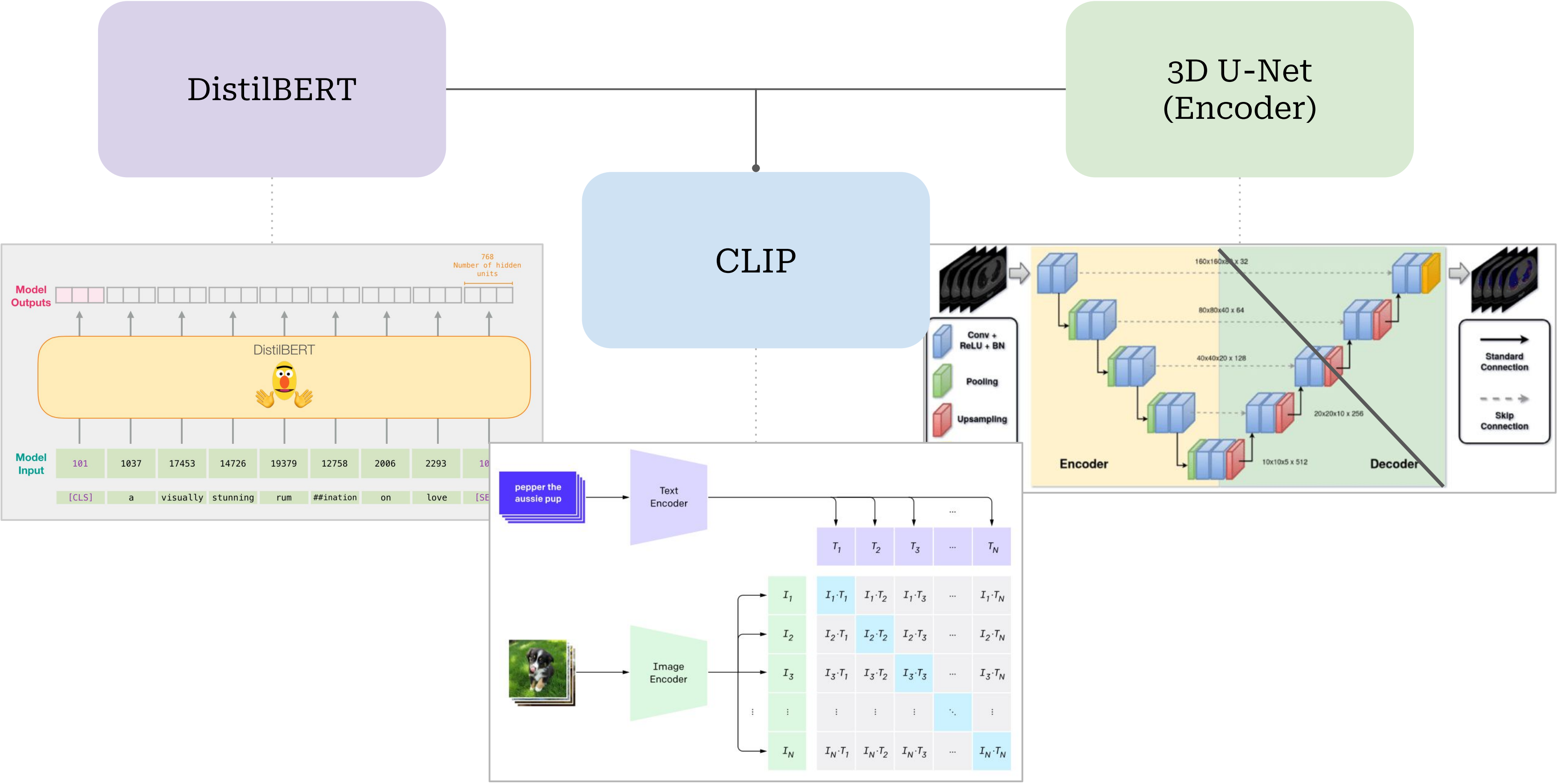
Previous tasks	Our task
2D	4D
in medical imaging, binary classification	multiple classes
much bigger training sets (thousands or millions)	262



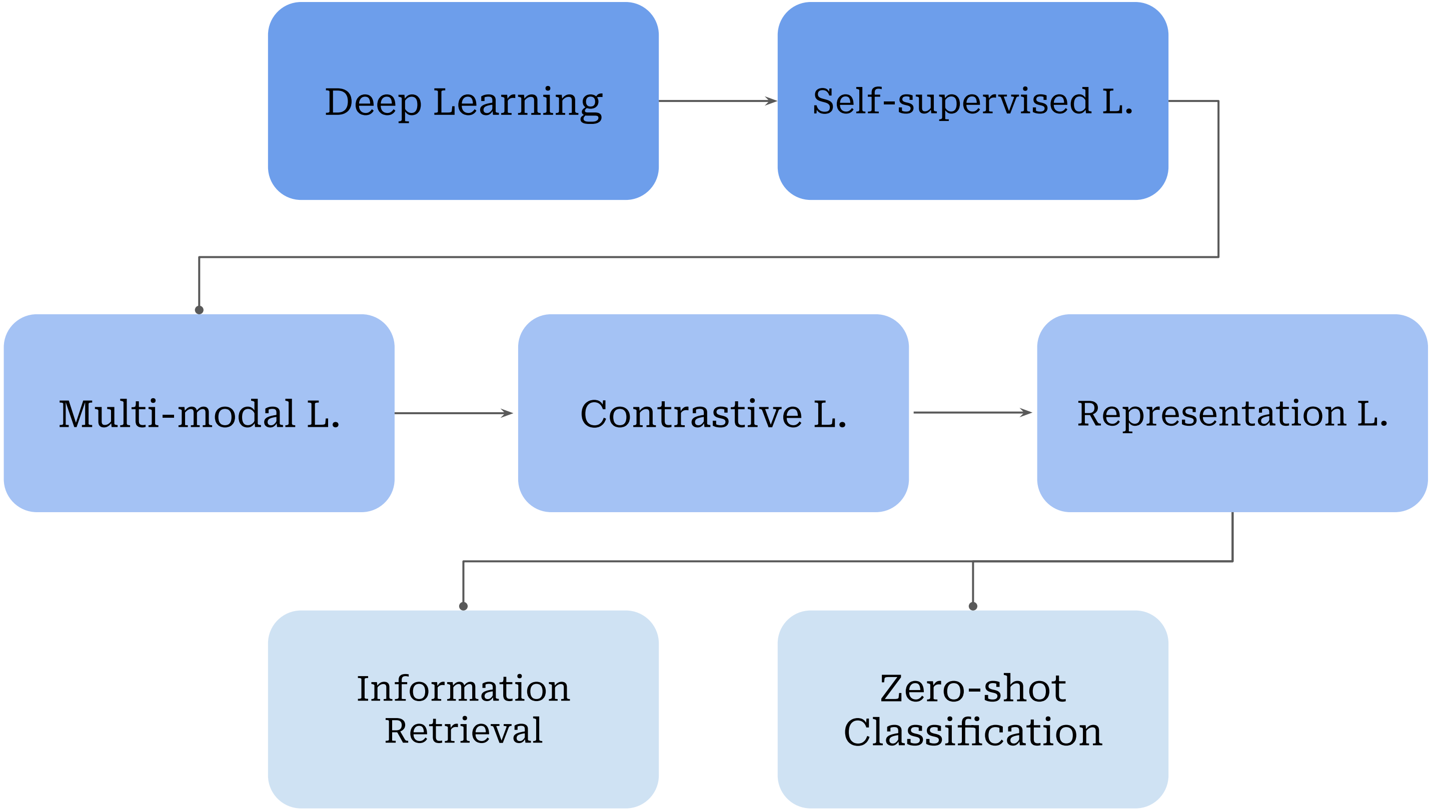


Proposed method

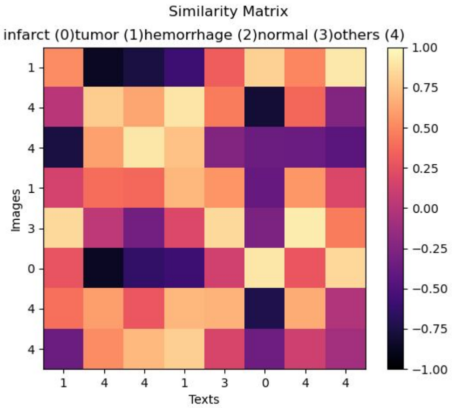


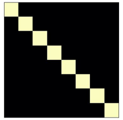

DistilBERT
3D U-Net
(Encoder)
CLIP
768
Number of hidden units
Model Outputs
DistilBERT
Model Input
101
1037
17453
14726
19379
12758
2006
2293
[CLS]
a
visually
stunning
rum
##ination
on
love
pepper the aussie pup
Text Encoder
Image Encoder
T_1
T_2
T_3
T_N
I_1
I_2
I_3
I_N
$I_1 \cdot T_1$
$I_1 \cdot T_2$
$I_1 \cdot T_3$
$I_1 \cdot T_N$
$I_2 \cdot T_1$
$I_2 \cdot T_2$
$I_2 \cdot T_3$
$I_2 \cdot T_N$
$I_3 \cdot T_1$
$I_3 \cdot T_2$
$I_3 \cdot T_3$
$I_3 \cdot T_N$
$I_N \cdot T_1$
$I_N \cdot T_2$
$I_N \cdot T_3$
$I_N \cdot T_N$
Conv + ReLU + BN
Pooling
Upsampling
80x80x40 x 64
40x40x20 x 128
20x20x10 x 256
10x10x5 x 512
Encoder
Decoder
Standard Connection
Skip Connection

Proposed method



Contrastive Loss





Listing 1 The forward function for the proposed method.

```
def forward(image, input_id_report, attention_mask_report, label):
        I_f = image_encoder(image)
        T_f = text_encoder(input_id_report, attention_mask_report)

        W_i = image_projection(I_f)
        W_t = text_projection(T_f)

        I_e, T_e = map(lambda t: F.normalize(t, p = 2, dim = -1), (W_i, W_t))
        sim = torch.einsum('i d, j d -> i j', I_e, T_e)
        sim *= temperature.exp()

        labels = torch.arange(I_f.size(0))
        I_loss = F.cross_entropy(sim, labels)
        T_loss = F.cross_entropy(sim.T, labels)

        loss = (I_loss + T_loss) / 2

        return loss
```

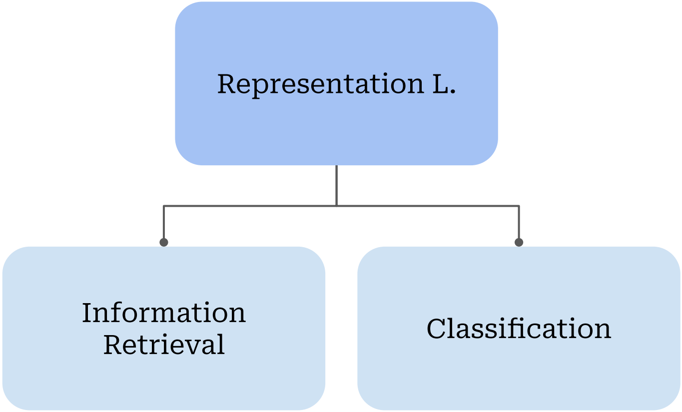
Training set up

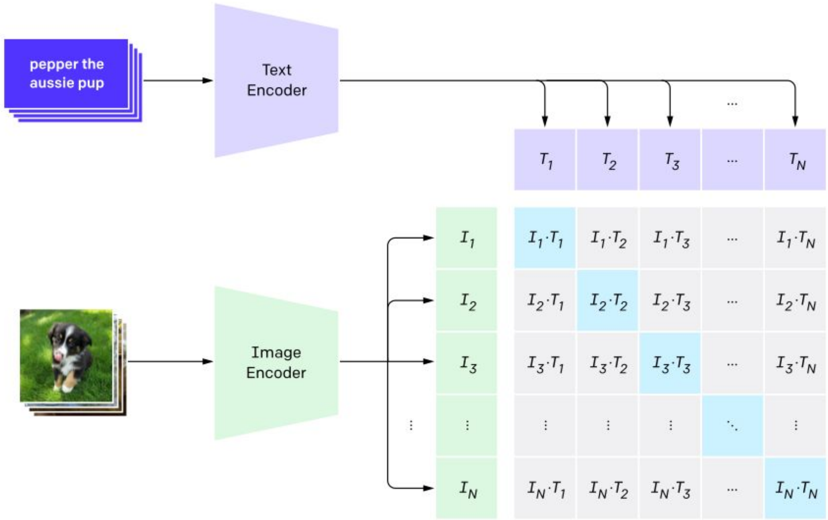
- Parameter choice: trial and error (no tuning)
- Adam optimizer with initial learning rate of 0.01
  - reducing on plateau with patience of 10 epochs on the validation loss by a factor of 0.5.
- The mini-batch size of 8 pairs (0.04% |TR|)
  - CLIP |TR|=400million+ and mini-batch size == 32,768 (0.0008% |TR|)
- 300 epochs
- Better with
  - Data augmentation
  - More data by including more classes and MRI sequences

- Classification training was inconclusive

Results

- Test set of 27 image-text pairs
- Test ability to match a text query to other images





Results

	Score	Report	Findings
1	0.813	Normal study of brain. No acute infarcts / haemorrhage / focal lesions. Vascular loop from right superior cerebellar artery is seen indenting the cisternal segment of right trigeminal nerve - suggested clinical correlation to r/o trigeminal neuralgia.	Normal
2	0.770	Few non-specific T2 FLAIR white matter hyperintensities are seen in bilateral frontal lobes. No significant abnormality in the brain, MR Angiogram and MR Venogram.	Normal
3	0.767	No significant abnormality is seen in the brain.	Normal
4	0.760	Grade I global cortical atrophy with scattered punctate white matter ischaemic changes. No acute infarcts / haemorrhage / focal lesions. Page 1 of 1	Normal
5	0.700	Normal traumatic neuroparenchymal injury. No acute infarcts / haemorrhage / focal lesions. Subgaleal hematoma in the left frontal, parietal and temporal region with contusions in the left temporalis muscle.	Normal

Table 2 Highest ranking medical reports from matched image-text pair for the query "No significant abnormality is seen in the brain."

NLP ground truth

Tumor

Hemorrhage

Results

For each report in the test set (n = 27), we compute its similarity between each other sample (n − 1). Here we show the medical reports for the images with the highest similarity (rank = 1) that also matched the query ground truth class and the matched pair class. In bold, the images that matched between one another. This is the desired behaviour, indicating that if two images in the test set display the same features, then they should be matched together when the corresponding reports are queried by the model. Three images with the ”Normal (brain)” class all matched to the same image. In general, it seems that the ”Normal” class was the easiest to learn and match.

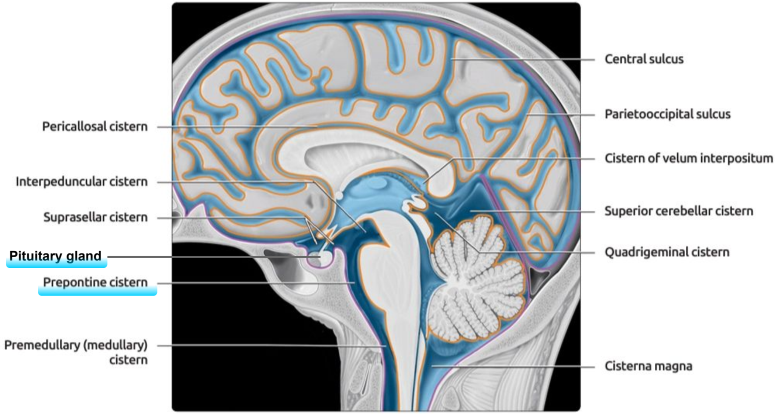


Image	Text (query)	Matched Image	Matched Text	Class
image_236	No significant abnormality is seen in the brain.	image_474	No significant abnormality is seen in the brain.	Normal
image_576	No significant abnormality is seen in the brain.	image_474	No significant abnormality is seen in the brain.	Normal
image_1069	Sellar and suprasellar lesion indenting on inferior aspect of optic chiasma, likely pituitary adenoma. Suggest: Post contrast imaging. Mild small vessel ischemic changes.	**image_812**	Mildly enhancing, well-defined, oval, T2 heterogenous hypointense to isointense lesion at left prepontine cistern with a wide angle to the dura as described above, likely representing meningioma. Homogenous strongly enhancing, well-defined, rounded T1/T2 isointense lesion in left parafalcine region as described above. It may likely represent meningioma. Adv clinopathological correlation and review with relevant clinical details.	Tumor
image_387	No significant abnormality is seen in the brain.	image_474	No significant abnormality is seen in the brain.	Normal
image_812	Mildly enhancing, well-defined, oval, T2 heterogenous hypointense to isointense lesion at left prepontine cistern with a wide angle to the dura as described above, likely representing meningioma. Homogenous strongly enhancing, well-defined, rounded T1/T2 isointense lesion in left parafalcine region as described above. It may likely represent meningioma. Adv clinopathological correlation and review with relevant clinical details.	**image_1069**	Sellar and suprasellar lesion indenting on inferior aspect of optic chiasma, likely pituitary adenoma. Suggest: Post contrast imaging. Mild small vessel ischemic changes.	Tumor
image_989	No significant abnormality in the brain. Hypoplastic intracranial right vertebral artery. Rest of MRA normal. Hypoplastic left transverse and sigmoid sinus. Rest of MRV normal. Kindly correlate clinically.	image_110	Grade I global cortical atrophy with scattered punctate white matter ischaemic changes. No acute infarcts / hemorrhage / focal lesions. Page 1 of 1	Others
image_411	No significant abnormality is seen in the brain. Page 2 of 2	image_576	No significant abnormality is seen in the brain.	Normal

Conclusion

- No meaningful results, but the model seems to learn something
  - Easier to distinguish healthy brains from the others
- Previous work is built on much simpler premises
  - 2D data, binary classification, much bigger training sets
- In future work:
  - increase the training set / apply more aggressive data augmentation to improve performance
  - train a classifier with more reliable ground truth labels

Thank you! 😄