brat: Aligned Multi-View Embeddings for Brain MRI Analysis

Supplementary Material

6. Dataset Details

In this section, we discuss MSKBrain, the largest existing dataset of brain MRIs and radiology reports.

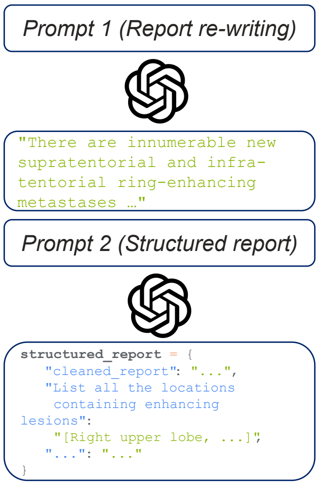
6.1. Raw Dataset

We collected the brain MRI scans and their corresponding clinical reports from a cancer center, covering the period from 2012 to 2017. These scans were primarily obtained to monitor brain metastases and tumours in cancer patients, resulting in a dataset rich in positive findings (89.7% of scans show abnormalities, an average of 134 words or 8 sentences per report) and representative of a diverse patient population. We also collected the clinical reports corresponding to these images, as well as extensive demographic data, primary diagnosis, ongoing chemotherapy and radiotherapy treatments, and survival information, which will be utilized in future studies.

MRI sessions typically consist of multiple MRI modality scans (e.g. FLAIR, T2, etc.); however, in this first iteration, we focus on T1-post contrast MRIs, the most informative type of scan. To extract T1 post-contrast scans, we generated a long, clinician-validated list of keywords typically used to refer to these scans. The list contained over 50 expressions such as "Axial T1 post SENSE" or "Ax T1 POST". We removed around 3,000 sessions that did not include T1 post-contrast imaging. DICOM medical images were converted and saved as 3D NIFTI files. We first sorted and loaded the DICOM slices into a 3D array, and then, for each image stack, preserved relevant metadata, such as pixel spacing and slice thickness, in the affine matrix. To standardize the intensity values, we thresholded the images at the 99th percentile and rescaled them to a range of 0-800, converting the final values to 16-bit integers. We saved the processed 3D volumes as compressed NIFTI files (.nii.gz). We then connect the MRI sessions to patient data stored in a REDCap database. For each imaging session, we attached the patient's demographic information and added calculated fields like age and time to death. We then mapped treatments and diagnoses to imaging sessions by finding the closest diagnosis date and checking which medications and radiation therapies were active at the time of imaging. We provide an overview of key patient data in Figure 11. The final dataset includes 77,228 brain MRI image-report pairs from 24,262 unique patients. To develop our model, we performed a patient-wise split of the data into 75,142 examples for training, 945 for development, and 1,141 for the test set.

> Again status post left-sided craniotomy with stable post-operative changes and with slight increase in the heterogeneously enhancing mass lesion centred in the left temporal lobe which now measures 7.5 x 4.8 cm on image 13 series 14 from 6.7 x 4.7 cm , though the enhancement within it is more irregular and less intense than before . The mass is not completely imaged on the perfusion sequence but there is hyperperfusion inferiorly within the nodular enhancing component which is incompletely demonstrated . The surrounding hyperintense T2/FLAIR infiltrating nonenhancing signal abnormality is stable consistent with nonenhancing tumor/edema. No new discontinuous suspiciously enhancing brain lesions. There is slightly increased dilatation of the ventricles with slightly increased hyperintense T2/FLAIR signal in the periventricular white matter particularly about the frontal horns and atrium, suggesting transependymal flow of CSF from a communicating hydrocephalus. Stable mild midline shift to the right without significant downward herniation. No acute intracranial hemorrhage, infarct, or new extra-axial collections.

Figure 9. An example report showing references to prior scans in blue and descriptions of findings not visible on T1 post-contrast scans in yellow .



```
"There are innumerable new
supratentorial and infra-
tentorial ring-enhancing
metastases …"
```

```
structured_report = {
   "cleaned_report": "...",
   "List all the locations
    containing enhancing
lesions":
    "[Right upper lobe, ...]",
   "...": "..."
}
```

Figure 10. The two-step GPT-4 based report processing pipeline. Prompt 1 and 2 are in Figure 12 and 17, respectively.

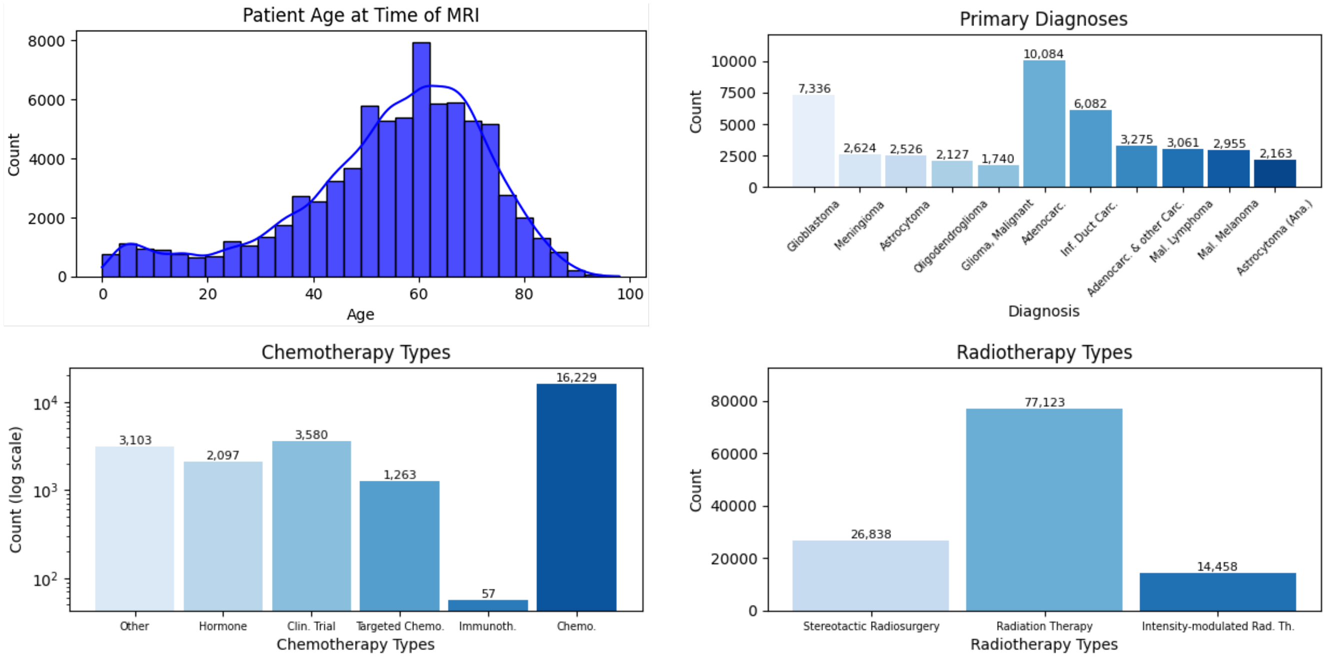


Figure 11. Participant demographics. Primary diagnoses refers to the primary cancer diagnosis for the patients for whomst the scan was ordered. Chemotherapy and radiotherapy types show a count of all the types of chemo/radio sessions assigned to the patients in this dataset.

6.2. Text Processing

As this paper focuses on learning image representations from brain MRIs, we had to ensure that all report content was visually grounded in the corresponding images. For example, keyword filtering revealed that 94% of reports make references to prior scans. Figure 9 contains an example report showing the two main types of information not detectable from the image: references to prior scans and description of findings not visible on T1 post-contrast images.

> You are a highly experienced radiologist. Re-write the given brain MRI report and only modify the following:
>
> (a) Leave out any details not visible on T1-weighted post-contrast images. Note that T2/FLAIR hyperintensities can often be seen on T1 Images. Observations related to e.g. perfusion, plasma volume or K trans cannot be seen and should be excluded.
> (b) Leave out any terms that suggest temporal change or progression (e.g. dates, “new”, “increased”, “previous”, “now”, “compared to”, “since last”, “more”, “less”, etc.)
> (c) Remove any PHI.

Figure 12. The final prompt that was used to re-write the reports and remove PHI and information not visible from the T1 post-contrast images.

Inspired by recent work demonstrating GPT-4 performs well on radiology report processing [28], we developed an on-premise GPT-4-based report processing pipeline. This pipeline enabled us to achieve three objectives without using expensive human annotators: the data was anonymized by removing protected health information (PHI), the reports were re-written to remove the aforementioned references, and we extracted structured information. Through iterative prompt engineering and radiologist feedback, we arrived at a 2-stage approach that extracted information from the reports with a 96% accuracy on a gold standard set of 50 manually annotated reports. This accuracy was calculated by counting the share of exact matches for all structured data points. By evaluating the second step of the two-step approach, we also implicitly ensure that the intermediate re-written report captures all the relevant information from the original report.

Annotating all the reports costs approximately $1,600, which is significantly lower than the cost of expert annotation. Figure 12 shows the report re-writing prompt and Figure 17 shows the information extraction prompt. Example of this two-step processing are shown in Figures 20, 18, and 19.

As there were around 80,000 long radiology reports, we used Python’s asyncio framework to process multiple reports in parallel with GPT-4. Each report underwent two API calls: one for re-writing and one for extracting structured information. We managed the API rate limits by processing reports in batches and including sleep time between batches. We also added a logit bias to avoid certain temporal medical terms (e.g., "increase", "new") in the answer. The temperature was set to 0.0 and top_p to 1.0 for deterministic outputs. With parallelization, processing 80,000 reports took around 48 hours. GPT-4 performed significantly better than GPT-3.5 during preliminary comparisons.

7. Motivation for DPPs

Figure 13 and 14 in the Appendix illustrate how DPPs promote a more desirable feature diversity than pairwise repulsion.

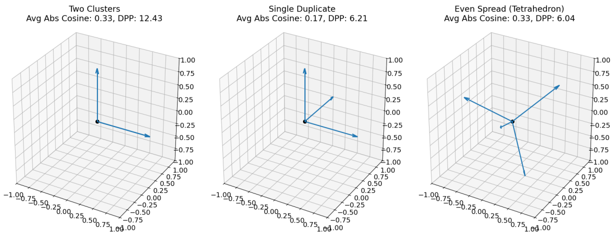


Figure 13. Visualization of the volume spanned by four 3D vectors under different configurations. In the left figure, both vectors are duplicates. In the middle one, one vector is a duplicate. The average pairwise cosine dissimilarity is the same for the left and right figures, even though the right figure represents a much more desirable spread of the vectors. The DPP is much lower for that figure thus minimizing the DPP is superior.

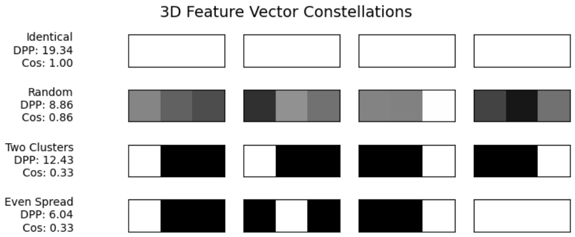


Figure 14. Visualization of four 3D vectors under different configurations, representing attention maps over a 1 × 3 image. It shows that pairwise cosine dissimilarity ("Cos") equally rewards an even spread of attention maps (desired) and the embeddings grouping into two clusters (not desired). In contrast, the DPP rewards the even spread the most.

8. Implementation and Training Details

In this section, we discuss the implementation details of our pre-training and downstream evaluation. All code and models will be made public.

8.1. Pre-Training

Training parameters were determined empirically, with the final set provided in Table 6. Contrary to the general assumption that big batches lead to improved performance for image-text contrastive learning, our results consistently improved for comparitvely small batch sizes, in the range of 25-32. This allowed us to train each model on a single A/H100 GPU. Our experiments also found that the Q-Former's language modeling loss consistently improved performance across nearly all configurations, while image-text matching did not yield benefits, leading us to omit the image-text matching loss. We also found that using a biomedically pre-trained BERT outperformed the standard BERT pre-training version in all evaluated scenarios. All model weights were selected based on the best average metrics on the development set. For MSKBrain, all models were trained with the same image processing: $1mm \times 1mm \times 1mm$ voxel spacing, intensity normalization, and resizing to $32 \times 256 \times 256$. Preliminary analysis on MSKBrain also showed lowered performance with standard data augmentation such as Gaussian noise, image rotation and translation, as well as random view cropping, and was removed from subsequent analyses. We also explored various text augmentation techniques. In "text dropout", a given share of text tokens is randomly masked during the forward pass. In "sentence dropout", different sentences, i.e., clinical features, are randomly removed from the reports. In "PVA dropout," we randomly drop a given share of paired multi-view embeddings and clinical features from the overall similarity matrix. None of these techniques have led to performance gains.

8.2. Downstream Tasks

We evaluated our pre-training methods by fine-tuning on several downstream tasks. When feasible, our hyperparameters were selected via grid search. The ADNI hyperparameter are given in Table 5. For ADNI, image preprocessing was performed using Clinica's t1-volume-tissue-segmentation pipeline. For report generation, we follow the parameters chooses in the Llama paper [12]. We used a batch size of 1024 and learning rate of 0.0002. We use an AdamW optimizer with a cosine decay and a warm-up ratio of 0.3. For segmentation, we used nnUNet as the baseline model, fine-tuning it with an initial learning rate of 1e-2 and a weight decay of 3e-5. The training pipeline included standard nnUNet pre-processing, data augmentation was not used. Result model weights were selected based on the highest mean Dice score for BraTS-2021 and the best Lesion-Wise metrics for BraTS-2023-METS on the validation set. All model weights were selected based on the best average metrics on the development set.

Parameter	1%	10%	100%
Batch Size	16	32	32
Learning Rate	1.00E-06	1.00E-05*	1.00E-05*
Training Precision		Bfloat16	
Augmentation	Yes	No	No
Trained Layers	MLP Only	All	All
MLP size		2 layers	

Table 5. Implementation details of our Alzheimer classification downstream task. *For ViT we used 1.00E-06 across all data amounts. Augmentation consisted of: random flipping, random intensity scaling, random intensity shifting, adding gaussian noise, gaussian smoothing, random contrast adjustment, and random low resolution simulation. More details can be found in our code.

Parameter	all models on MSKBrain	brat on BIMCV-R
Batch Size	32	25
3D Vision Model M	Densenet-121/ViT/ResNet-50	Densenet-121
Weights Init. of M	None	
Architecture of $E_{I/R}$	BERT-base[2]	
Weights Init. of $E_{I/R}$	BiomedBERT[3]	
Learning Rate M	5.00E-04*	
Learning Rate $E_{I/R}$	5.00E-05	
Max. Text Length E_R	256	
N_Q (# of Query Tokens)	32	
Cross-Attention Frequency	2	
Max. Number of Sentences	20	
Training Precision	Bfloat16	
Augmentation	None	

Table 6. Implementation details of our pre-training. Except for BIMCV-R , the batch size was chosen to be maximal given compute resources. *For ViT, we used a lower learning rate of 1.00E-07.

8.3. BIMCV-R Dataset

We found quality issues with the BIMCV-R dataset that may explain the overall lower performance obtained on this dataset compared to MSKBrain. Figure 15 shows how for some images the middle slice (depicted) is already no longer in the lung, suggesting that the scan mainly depicts other body parts. Several images also seem to depict localizer scans, which makes it difficult to connect them to radiology reports. Appropriate processing of these images would likely lead to significant performance improvements.

9. Additional Results

In this section, we provide more detailed results and examples.

9.1. Alzheimer's Classification

More detailed results for Alzheimer's classification are provided in Table 7.

9.2. Report Generation

Results in Table 4 show that, for the GREEN metric, we see sizeable gains from general vision-language pre-training but no big difference between traditional Q-Former training versus the brat framework. The GREEN metric provides a structured clinical assessment by identifying key radiology report errors derived from expert evaluations. However, it was developed mainly on chest X-ray reports, and therefore, its applicability to out-of-distribution modalities is limited. The authors evaluated their metric on an abdomen CT dataset and found a high absolute error (5.31). I provide this metric as it can offer a rough assessment of the clinical correctness, however, it is likely not well suited to assess minor performance differences in brain MRI reports. An inspection of the generated evaluations by GREEN confirms that they contain many errors. A more dependable approach would involve direct human evaluation or leveraging a stronger LLM such as GPT-4 for assessment.

Figure 16 shows examples of generated reports.

9.3. Segmentation

Figure 8 and Figure 9 provide more detailed results on segmentation.

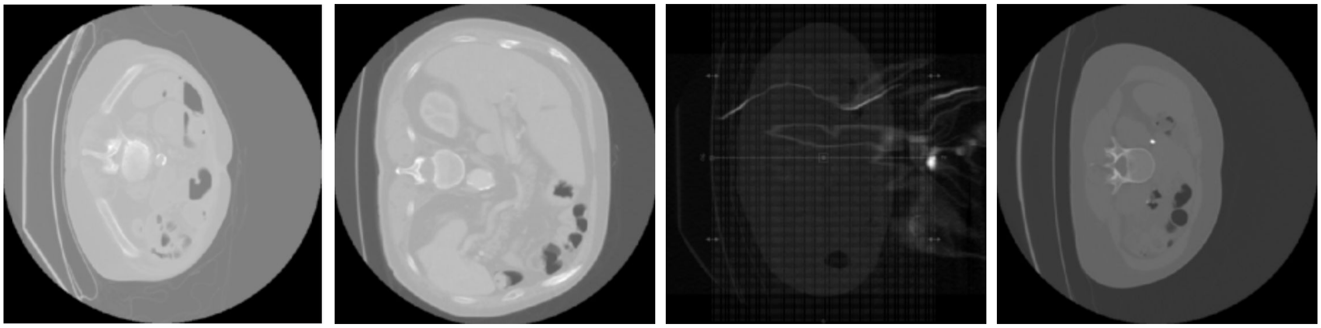


Figure 15. BIMCV-R example images of localizer scans or where the middle slice is already in the abdomen or pelvis.

Pre-training Approaches		1% Training Data (n=19)				10% Training Data (n=193)				100% Training Data (n=1,932)			
Vision Model M	Weight Init.	Alz.	Normal	MCI	μ	Alz.	Normal	MCI	μ	Alz.	Normal	MCI	μ
Densenet-121	Random	0.523	0.513	0.527	0.521 [0.495, 0.547]	0.640	0.560	0.498	0.567 [0.535, 0.596]	0.724	0.629	0.535	0.629 [0.608, 0.649]
	CLS	0.514	0.517	0.511	0.514 [0.487, 0.538]	0.614	0.598	0.523	0.578 [0.555, 0.602]	0.720	0.628	0.556	0.635 [0.612, 0.650]
	Q-Former	0.565	0.525	0.486	0.526 [0.506, 0.547]	0.688	0.627	0.550	0.623 [0.604, 0.640]	0.747	0.662	0.581	0.663 [0.651, 0.681]
	brat	0.560	0.559	0.505	0.543 [0.497, 0.579]	0.720	0.644	0.518	0.628 [0.606, 0.653]	0.793	0.687	0.505	0.661 [0.650, 0.672]
ResNet-50	Random	0.497	0.566	0.541	0.535 [0.497, 0.569]	0.516	0.529	0.541	0.530 [0.498, 0.561]	0.590	0.525	0.528	0.548 [0.514, 0.586]
	brat	0.527	0.531	0.532	0.530 [0.500, 0.556]	0.621	0.456	0.452	0.510 [0.490, 0.532]	0.636	0.542	0.533	0.569 [0.519, 0.612]
ViT	Random	0.517	0.485	0.473	0.492 [0.471, 0.512]	0.554	0.491	0.502	0.515 [0.495, 0.532]	0.528	0.473	0.515	0.505 [0.476, 0.533]
	brat	0.518	0.498	0.460	0.491 [0.458, 0.523]	0.607	0.555	0.467	0.543 [0.522, 0.561]	0.622	0.521	0.450	0.531 [0.513, 0.551]

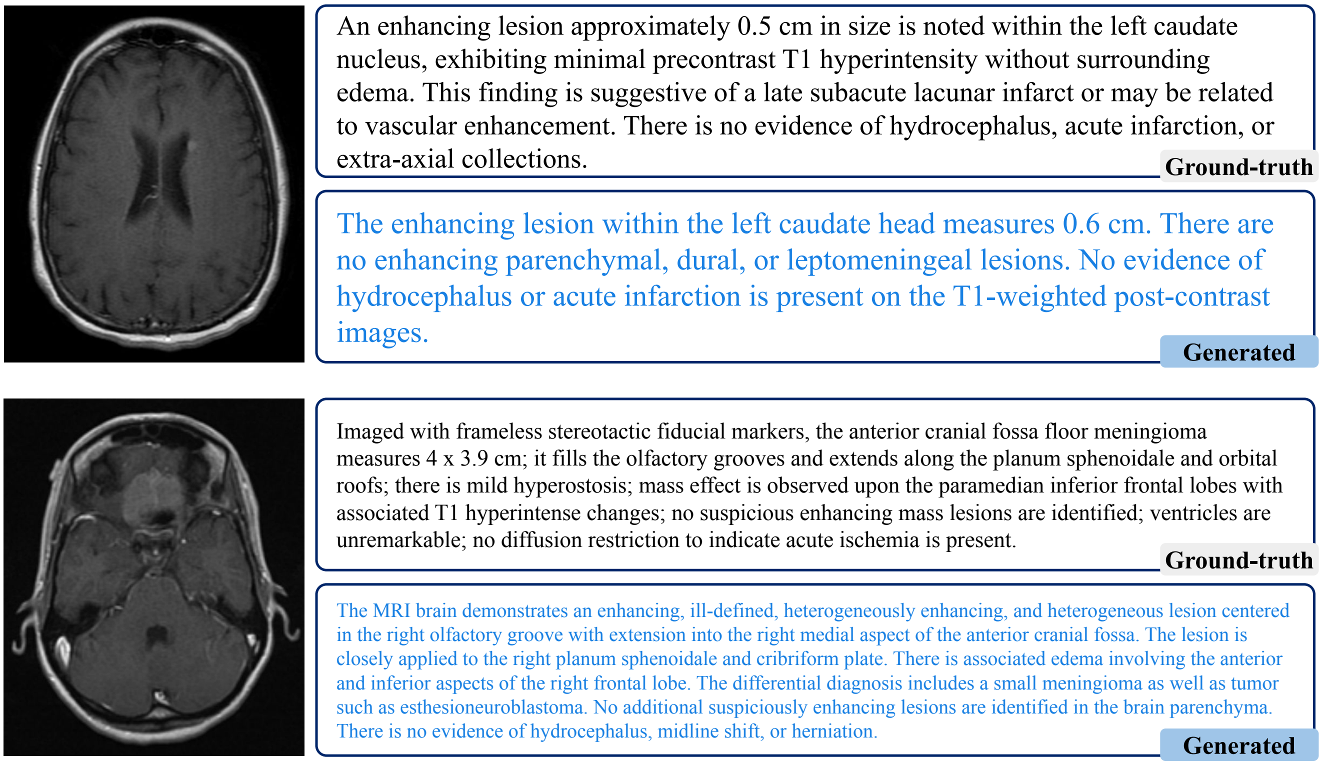
Table 7. Evaluation results (AUC scores) for different initialisations using 1%, 10%, and 100% of training data. "Alz." stands for Alzheimer's disease and "MCI" for mild cognitive impairment. The column μ is the average of the per-class AUC scores computed on the balanced test set; only this column displays the confidence interval.

Pre-training Approaches		1% Training Data (n=12)			10% Training Data (n=120)			100% Training Data (n=1200)		
Vision Model M	Weight Init.	Whole Tumor	Tumor Core	Enhanced Tumor	Whole Tumor	Tumor Core	Enhanced Tumor	Whole Tumor	Tumor Core	Enhanced Tumor
Densenet-121	Random	0.780	0.646	0.585	0.875	0.791	0.710	0.903	0.865	0.779
Densenet-121	brat	0.796	0.633	0.580	0.870	0.785	0.707	0.903	0.864	0.776

Table 8. Segmentation performance (Dice scores) for different pre-training initialisations using 1%, 10%, and 100% of the training data. The values correspond to the Dice scores for the Whole Tumor, Tumor Core, and Enhanced Tumor regions.

Pre-training Approaches		1% Training Data (n=12)			10% Training Data (n=120)			100% Training Data (n=1200)		
Vision Model M	Weight Init.	Whole Tumor	Tumor Core	Enhanced Tumor	Whole Tumor	Tumor Core	Enhanced Tumor	Whole Tumor	Tumor Core	Enhanced Tumor
Densenet-121	Random	0.780	0.646	0.585	0.875	0.791	0.710	0.922	0.854	0.761
brat	brat	0.796	0.633	0.580	0.870	0.785	0.707	0.925	0.867	0.762

Table 9. Segmentation performance (Lesion-wise Dice scores) for different pre-training initialisations using 1%, 10%, and 100% of the training data. The values correspond to the lesion-wise Dice scores for the Whole Tumor, Tumor Core, and Enhanced Tumor regions.



An enhancing lesion approximately 0.5 cm in size is noted within the left caudate nucleus, exhibiting minimal precontrast T1 hyperintensity without surrounding edema. This finding is suggestive of a late subacute lacunar infarct or may be related to vascular enhancement. There is no evidence of hydrocephalus, acute infarction, or extra-axial collections.

The enhancing lesion within the left caudate head measures 0.6 cm. There are no enhancing parenchymal, dural, or leptomeningeal lesions. No evidence of hydrocephalus or acute infarction is present on the T1-weighted post-contrast images.

Imaged with frameless stereotactic fiducial markers, the anterior cranial fossa floor meningioma measures 4 x 3.9 cm; it fills the olfactory grooves and extends along the planum sphenoidale and orbital roofs; there is mild hyperostosis; mass effect is observed upon the paramedian inferior frontal lobes with associated T1 hyperintense changes; no suspicious enhancing mass lesions are identified; ventricles are unremarkable; no diffusion restriction to indicate acute ischemia is present.

The MRI brain demonstrates an enhancing, ill-defined, heterogeneously enhancing, and heterogeneous lesion centered in the right olfactory groove with extension into the right medial aspect of the anterior cranial fossa. The lesion is closely applied to the right planum sphenoidale and cribriform plate. There is associated edema involving the anterior and inferior aspects of the right frontal lobe. The differential diagnosis includes a small meningioma as well as tumor such as esthesioneuroblastoma. No additional suspiciously enhancing lesions are identified in the brain parenchyma. There is no evidence of hydrocephalus, midline shift, or herniation.

Figure 16. Two example MRI reports generated by our VLM with a brat backbone. The reports largely capture the correct image findings.

You are a highly experienced radiologist. Accurately answer the questions below based on the given brain MRI report. Your output must be in json format.

(a) For each question, choose the appropriate answer (wording must match exactly). If answers are mutually exclusive, choose one. If multiple answers can apply, list all that are true, separated by semicolons (";").
(b) If the MRI report does not contain information to answer a specific question, use the default answer indicating a normal status.
(c) Note the following assumptions: meningiomas are considered enhancing lesions; burr holes and ventriculostomy and Ommaya catheters are considered prior surgery; punctate lesions are less than 1cm.

Questions (Answer options):
Is there evidence of prior surgery? (Yes / No)
What kind of surgery was performed? (NA / left frontal craniotomy; right frontal craniotomy; left parietal craniotomy; right parietal craniotomy; left temporal or pterional craniotomy; right temporal or pterional craniotomy ; left occipital craniotomy; right occipital craniotomy)
Are there any enhancing lesions? (Yes / No)
What is the length of the biggest mass lesion? (NA / Less than 1cm / 1 to 2cm / More than 2cm)
Which side of the brain has more enhancing lesions? (NA / Left / Right)
List all the locations that contain enhancing lesions. (NA / Left frontal lobe; Right frontal lobe; Left parietal lobe; Right parietal lobe; Left temporal lobe; Right temporal lobe; Left occipital lobe; Right occipital lobe; Left thalamus or basal ganglia; Right thalamus or basal ganglia; Cerebellum; brainstem; cervical spinal cord)
How many enhancing lesions are there? (NA / One / Between 2 and 6 / Between 7 and 15 / More than 15)
Is there a herniation or midline shift? (Yes / No)
Are there any signs of white matter disease (e.g., leukoaraiosis or leukoencephalopathy)? (Yes / No)
Is the pituitary gland normal in appearance? (Yes / No)
Are there abnormalities in the sella or parasellar regions? (Yes / No)
Where is there evidence of invasion into or compression of adjacent structures? (Nowhere / ventricles; brainstem; cranial nerves)

Brain MRI report: ${Insert processed report}$

Figure 17. The final prompt that was used to extract structured information from the reports. Note that the structured information is mainly used to gain more understanding of our dataset and evaluate models. The only model in this paper trained on these labels is the Classification baseline which we compare against in Table 4 and Figure 7.

Original Report

FINDINGS: BRAIN: Unchanged plaque-like left parietal convexity enhancing meningioma at 0.9 x 0.3 cm with mild overlying hyperostosis; No suspicious parenchymal enhancing mass lesions, increase in Ktrans or increase in plasma volume; Normal ventricles; no diffusion restriction to indicate acute ischemia.
IMPRESSION: Since February 2015, no brain metastasis. Unchanged subcm left parietal convexity meningioma. Unchanged mild enlargement of the pituitary gland. Unchanged benign left orbital lesion, possibly a hemangioma or dermoid/epidermoid.

↓

Rewritten Report

Plaque-like left parietal convexity enhancing meningioma at 0.9 x 0.3 cm with mild overlying hyperostosis is present. No suspicious parenchymal enhancing mass lesions are observed. Ventricles appear normal. There is no diffusion restriction indicating acute ischemia.

↓

Structured Report

- Evidence of prior surgery? No
- Kind of surgery performed? NA
- Enhancing lesions? Yes
- Non-enhancing lesions? No
- Biggest lesion length? ¡1cm
- Side with more enhancing lesions? Left
- Locations with enhancing lesions? Left parietal lobe
- Number of enhancing lesions? One
- Herniation/midline shift? No
- Hydrocephalus? No
- Type of hydrocephalus? NA
- White matter changes? No
- Grey matter abnormalities? No
- Pituitary gland normal? Yes
- Sella/parasellar abnormalities? No
- Evidence of invasion/compression? Nowhere

Figure 18. An example of our automated report processing. Here, the intermediate re-written report omits the enlargement of the pituitary gland and left orbital lesion.

Original Report

FINDINGS: BRAIN: The peripherally enhancing centrally cystic/necrotic metastasis within inferior right frontal lobe bulging into frontal horn of the right lateral ventricle has not significantly changed in size when allowing for differences in image acquisition. Extensive right anterior hemispheric vasogenic edema which extends into the basal ganglia and anterior thalamus appears mildly improved. Mass effect resulting in right-to-left midline shift has mildly improved as well. The Ommaya catheter remains in place with its tip within necrotic core of the right inferior frontal lobe mass. The nodular dural based thickening along the anteroinferior margin of the right frontal burr hole has increased now measuring 1.8 x 1.3 cm on the axial images, having previously measured 1.1 x 0.8 cm. On the coronal images, it measures 1.5 x 0.8 cm and likely represent focal dural based tumor. The previously seen superficial focus of enhancement along the right inferior frontal gyrus appears less conspicuous. Small enhancing lesion adjacent to upper aspect of the Ommaya catheter has increased in size now measuring 4 to 5 mm. Right high parafalcine parietal hemorrhagic lesion has not significantly changed. Background white matter changes are stable.
IMPRESSION: Compared to December 21, 2011; 1. Mild progression of dural based disease along the anteroinferior margin of the right frontal burr hole. 2. Mild progression of the metastasis adjacent to upper aspect of the right Ommaya catheter. 3. Stable hemorrhagic lesion within high right parafalcine parietal lobe. 4. The heterogeneous right inferior frontal lobe mass its stable in size however, right hemispheric edema and mass effect have mildly improved.

↓

Rewritten Report

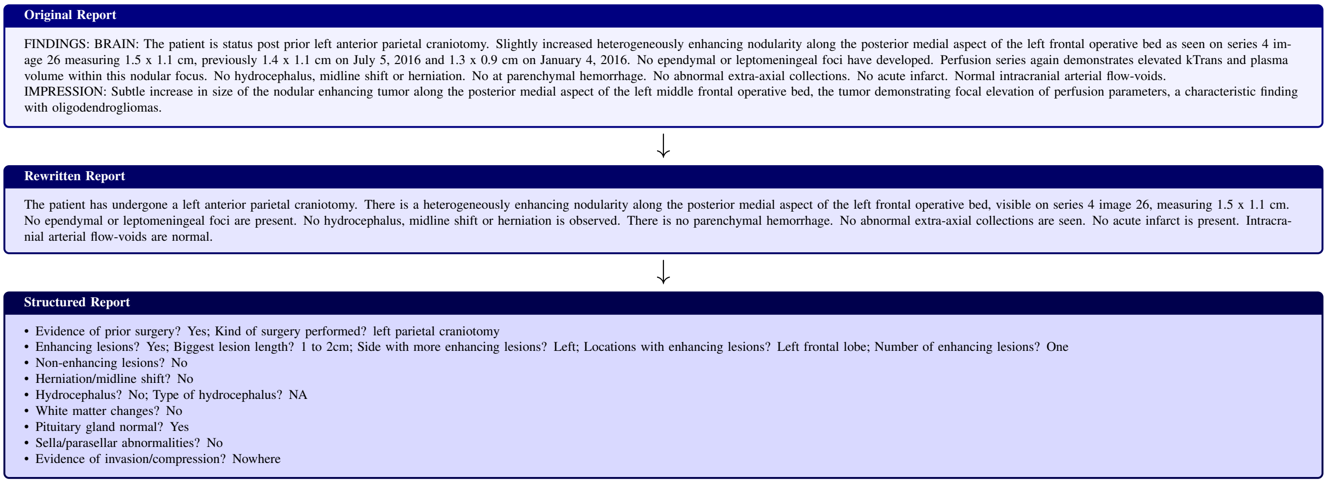
The peripherally enhancing centrally cystic/necotic metastasis is located within the inferior right frontal lobe, bulging into the frontal horn of the right lateral ventricle. There is extensive right anterior hemispheric vasogenic edema extending into the basal ganglia and anterior thalamus. Mass effect results in right-to-left midline shift. The Ommaya catheter is in place with its tip within the necrotic core of the right inferior frontal lobe mass. There is nodular dural based thickening along the anteroinferior margin of the right frontal burr hole, measuring 1.8 x 1.3 cm on the axial images. On the coronal images, it measures 1.5 x 0.8 cm and likely represents a focal dural based tumor. There is a superficial focus of enhancement along the right inferior frontal gyrus. A small enhancing lesion is adjacent to the upper aspect of the Ommaya catheter, measuring 4 to 5 mm. There is a right high parafalcine parietal hemorrhagic lesion. Background white matter changes are present.

↓

Structured Report

- Is there evidence of prior surgery? Yes
- What kind of surgery was performed? Right frontal craniotomy
- Are there any enhancing lesions? Yes
- Are there any non-enhancing lesions? Yes
- What is the length of the biggest lesion? 1 to 2 cm
- Which side of the brain has more enhancing lesions? Right
- List all the locations that contain enhancing lesions. Right frontal lobe; Right thalamus or basal ganglia
- How many enhancing lesions are there? Between 2 and 6
- Is there a herniation or midline shift? Yes
- Is there hydrocephalus? No
- What kind of hydrocephalus is there? NA
- Are there white matter changes suggestive of other pathologies? Yes
- Are there abnormalities in the cortical or subcortical grey matter suggestive of other pathologies? No
- Is the pituitary gland normal in appearance? Yes
- Are there abnormalities in the sella or parasellar regions? No
- Where is there evidence of invasion into or compression of adjacent structures? Ventricles

Figure 19. An example of a long and complex report that was processed successfully.



Original Report

FINDINGS: BRAIN: The patient is status post prior left anterior parietal craniotomy. Slightly increased heterogeneously enhancing nodularity along the posterior medial aspect of the left frontal operative bed as seen on series 4 image 26 measuring 1.5 x 1.1 cm, previously 1.4 x 1.1 cm on July 5, 2016 and 1.3 x 0.9 cm on January 4, 2016. No ependymal or leptomeningeal foci have developed. Perfusion series again demonstrates elevated kTrans and plasma volume within this nodular focus. No hydrocephalus, midline shift or herniation. No at parenchymal hemorrhage. No abnormal extra-axial collections. No acute infarct. Normal intracranial arterial flow-voids.
IMPRESSION: Subtle increase in size of the nodular enhancing tumor along the posterior medial aspect of the left middle frontal operative bed, the tumor demonstrating focal elevation of perfusion parameters, a characteristic finding with oligodendrogliomas.

↓

Rewritten Report

The patient has undergone a left anterior parietal craniotomy. There is a heterogeneously enhancing nodularity along the posterior medial aspect of the left frontal operative bed, visible on series 4 image 26, measuring 1.5 x 1.1 cm. No ependymal or leptomeningeal foci are present. No hydrocephalus, midline shift or herniation is observed. There is no parenchymal hemorrhage. No abnormal extra-axial collections are seen. No acute infarct is present. Intracranial arterial flow-voids are normal.

↓

Structured Report

- Evidence of prior surgery? Yes; Kind of surgery performed? left parietal craniotomy
- Enhancing lesions? Yes; Biggest lesion length? 1 to 2cm; Side with more enhancing lesions? Left; Locations with enhancing lesions? Left frontal lobe; Number of enhancing lesions? One
- Non-enhancing lesions? No
- Herniation/midline shift? No
- Hydrocephalus? No; Type of hydrocephalus? NA
- White matter changes? No
- Pituitary gland normal? Yes
- Sella/parasellar abnormalities? No
- Evidence of invasion/compression? Nowhere

Figure 20. An example of our automated report processing. The re-writing correctly rephrases the sentences, making references to changes in findings and removes references to, e.g. kTrans findings. The structured information extraction correctly answers all our instructions.