

# Research Internship Notes

Chang

me

her email

my email

This is more of a log. Will probably show results of experiments , screenshots .. of the stuff that i would be working

on like a journal

## 1. Results from ROCO dataset

- Dataset at : <https://huggingface.co/datasets/adishourya/ROCO-QA-Train>
- Sample Generation from our new VQA model; model card at [adishourya/resultsrocogoogle\\_paligemma-3b-mix-448\\_\\_fullrun\\_\\_0111-165804:](#)



Figure 1: ROCO\_44266

```
In [8]: show_sample()
ID : ROCO_44266
=====question=====
Does the lesion described in the coronal CT image cross the midline?
=====generation=====
Coronal CT image shows the lesion crossing midline.
=====answer=====
Coronal CT after intravenous contrast injection: expansive cervical process showing
that the medial border of the mass crosses the middle line.
```

txt

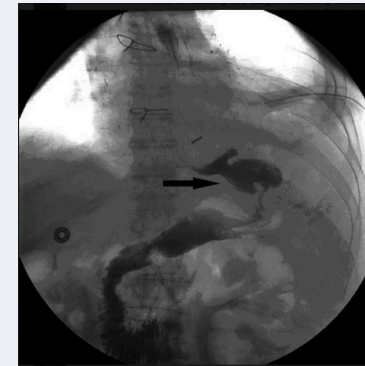


Figure 2: ROCO\_29769

```
In [9]: show_sample()
ID : ROCO_29769
=====question=====
What is causing the slight narrowing observed in the postoperative upper
gastrointestinal tract image?
=====generation=====
Postoperative upper gastrointestinal tract image showing slight narrowing of the
gastroepigastrium.
=====answer=====
Postoperative UGI showing slight narrowing at mid-body of stomach (arrow).
```

txt

- Generation Samples not cherry picked
- I havent made a full dataset of the entire test results yet. I'll do it this week

## 2. Training

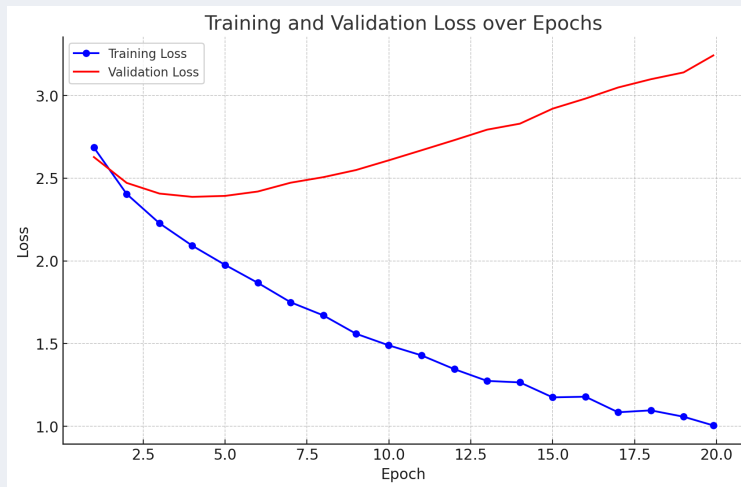


Figure 3: overfitting in medpix at 2nd epoch

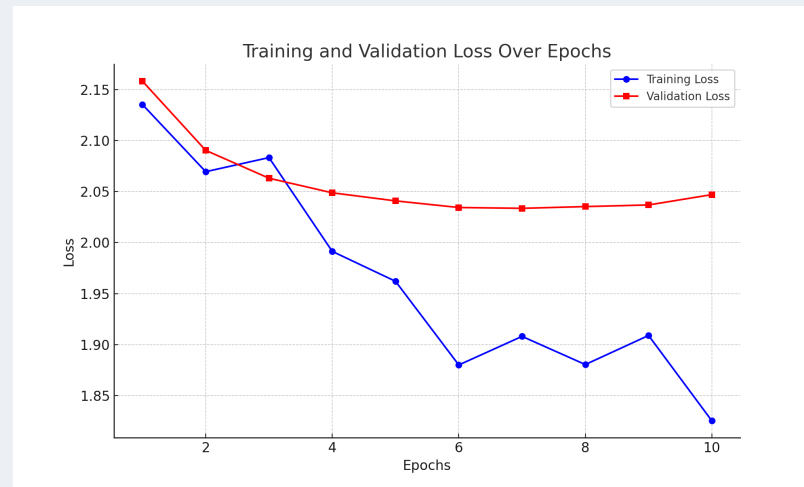


Figure 4: Early stop at 7th epoch on ROCO dataset

### 3. Evalutation

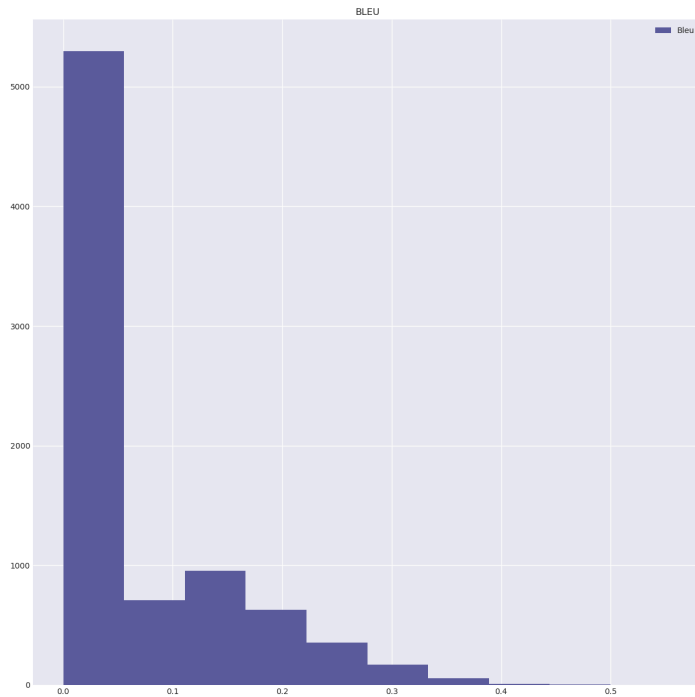


Figure 5: Bleu Score

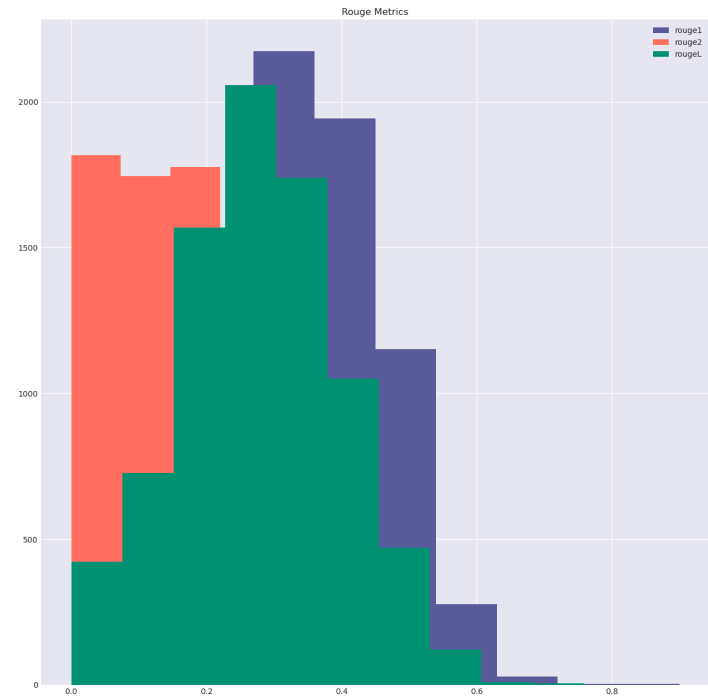


Figure 6: Rouge Score

- achieves good Rouge score (fluency in answer)
- Bleu score is weak compared to some of the other vqa models i looked at.

## 4. About metrics

### 1. Rouge score

- ROUGE-1: Measures the overlap of unigrams (individual words) between the generated and reference texts. It helps assess general relevance. **[0.4-0.6 is treated as good]**
- ROUGE-2: Measures the overlap of bigrams (pairs of words), capturing more contextual accuracy than ROUGE-1 and indicating how well the model maintains meaningful word pairs. **[0.2-0.4 is treated as good]**
- ROUGE-L: Based on the longest common subsequence (LCS), it identifies the longest sequence of words in the correct order, which is particularly useful for capturing fluency and syntactic structure. **[0.3-0.5 is treated as good]**

### 2. Bleu scores

- Bleu score is more apt for translations tasks
- as : In medical VQA, a lower BLEU score doesn't necessarily indicate poor performance, as multiple correct answers could be phrased differently but remain clinically valid. Therefore, BLEU is often used in combination with other metrics, like ROUGE, accuracy, or domain-specific human evaluation, to give a more complete performance assessment. **[0.2-0.4 is seen as average]**

<|END of Notes|>

## 5. Appendix

```
[
{"image_id": "ROCO_00001", "question": "What type of imaging modality is being used to display the coronal view?", "predicted_answer": "answer en What type of imaging modality is being used to display the coronal view?\nCoronal view.", "true_answer": "Axial MRI (coronal view).", "bleu_score": 0, "rouge1": 0.19047619047619047, "rouge2": 0.10526315789473684, "rougeL": 0.19047619047619047},
{"image_id": "ROCO_00006", "question": "What type of cancer is most likely causing the multiple large tumor masses seen in this coronal plain computed tomography image?", "predicted_answer": "answer en What type of cancer is most likely causing the multiple large tumor masses seen in this coronal plain computed tomography image?\nCoronal plain computed tomography image showing multiple large tumor masses in a patient with multiple primary tumors.", "true_answer": "Coronal plain computed tomography image showing multiple large tumor masses with edge enhancement inside the abdominal cavity and liver.", "bleu_score": 0.22915295704820032, "rouge1": 0.4067796610169491, "rouge2": 0.3157894736842105, "rougeL": 0.37288135593220345},
{"image_id": "ROCO_00016", "question": "What is the diagnosis indicated by the abnormal arterial signal elevation in the left more than right cavernous sinuses on this intracranial magnetic resonance angiogram?", "predicted_answer": "answer en What is the diagnosis indicated by the abnormal arterial signal elevation in the left more than right cavernous sinuses on this intracranial magnetic resonance angiogram?\nAbnormal arterial signal elevation in the left more than right cavernous sinuses on intracranial magnetic resonance angiogram.", "true_answer": "Axial source image from an intracranial magnetic resonance angiogram reveals abnormal arterial signal elevation in the left more than right cavernous sinuses consistent with a carotid cavernous fistula, as indicated by the arrow.", "bleu_score": 0.3117658802075856, "rouge1": 0.5194805194805195, "rouge2": 0.42666666666666664, "rougeL": 0.4155844155844156},
{"image_id": "ROCO_00025", "question": "What are the characteristics being evaluated in the medical image of the mineral trioxide aggregate plug?", "predicted_answer": "answer en What are the characteristics being evaluated in the medical image of the mineral trioxide aggregate plug?\nThe characteristics of the mineral trioxide aggregate plug evaluated in the medical image.", "true_answer": "The apical height, homogeneity, and the thickness of mineral trioxide aggregate plug", "bleu_score": 0.08635800047213178, "rouge1": 0.3255813953488372, "rouge2": 0.14634146341463417, "rougeL": 0.3255813953488372},
]
```

```

{"image_id": "ROCO_00031", "question": "What procedure was performed on the Right Coronary Artery?", "predicted_answer": "answer en What procedure was performed on the Right Coronary Artery?\nRight Coronary Artery Occlusion Dissection", "true_answer": "CTO of RCA (closure in the 2nd segment)", "bleu_score": 9.109159947227211e-232, "rouge1": 0.08333333333333333, "rouge2": 0.0, "rougeL": 0.08333333333333333},
{"image_id": "ROCO_00036", "question": "What was shown on the venogram before thrombolysis?", "predicted_answer": "answer en What was shown on the venogram before thrombolysis?\nVenogram before thrombolysis.", "true_answer": "Venography shows a large amount of thrombogenesis in the femoropoplitea vein before thrombolysis.", "bleu_score": 5.554837769749797e-155, "rouge1": 0.23076923076923078, "rouge2": 0.08333333333333333, "rougeL": 0.23076923076923078},
{"image_id": "ROCO_00061", "question": "What was the effect on the size of the cyst after surgery?", "predicted_answer": "answer en What was the effect on the size of the cyst after surgery?\nPostoperative cyst size was reduced.", "true_answer": "Postoperative decreased size of the cyst", "bleu_score": 0.13400826781778894, "rouge1": 0.39999999999999997, "rouge2": 0.2608695652173913, "rougeL": 0.32},
{"image_id": "ROCO_00084", "question": "What is causing the turbulence in the fetal ductus arteriosus, shown by aliasing and visible through its arch?", "predicted_answer": "answer en What is causing the turbulence in the fetal ductus arteriosus, shown by aliasing and visible through its arch?\nTurbulence in the fetal ductus arteriosus, shown by aliasing and visible through its arch.", "true_answer": "Fetal aortic arch. The solid arrow represents the ascending aorta while the dotted arrow represents the aortic arch distal to the ductus arteriosus. Note the aliasing in the ductal arch reflecting turbulent blood flow as a result of external compression due to mass-effect.", "bleu_score": 3.2172751716810823e-155, "rouge1": 0.2820512820512821, "rouge2": 0.052631578947368425, "rougeL": 0.20512820512820512},
{"image_id": "ROCO_00138", "question": "What is indicated by the arrow in this abdominal computed tomography scan?", "predicted_answer": "answer en What is indicated by the arrow in this abdominal computed tomography scan?\nAbdominal computed tomography scan showing the arrow pointing to the presence of the tumor.", "true_answer": "Abdominal computed tomography scan. Arrow points to the likely cholecystocutaneous fistulous track.", "bleu_score": 2.0299490458960478e-78, "rouge1": 0.4, "rouge2": 0.3157894736842105, "rougeL": 0.4},
{"image_id": "ROCO_00153", "question": "What is the cause of distension in the digestive system with levels of hydro-air colics and pelvic emptiness?", "predicted_answer": "answer en What is the cause of distension in the digestive system with levels of hydro-air colics and pelvic emptiness?\nDistension of the digestive system with levels of hydro-air colics and pelvic emptiness.", "true_answer": "Distension digestive avec niveaux hydro-a\u00e9riques coliques et vacuit\u00e9 pelvienne", "bleu_score": 9.039352811507815e-232, "rouge1": 0.13043478260869562, "rouge2": 0.0, "rougeL": 0.13043478260869562},

```



```

{"image_id": "ROCO_00176", "question": "What is the radiographic appearance of the left mandible four years after initial examination?", "predicted_answer": "answer en What is the radiographic appearance of the left mandible four years after initial examination?\nFour years later radiograph of the left mandible shows normal radiolucency.", "true_answer": "Panoramic X ray taken four years later showing a unilocular radiolucent area in the left ramus.", "bleu_score": 5.7030969235351194e-155, "rouge1": 0.3255813953488372, "rouge2": 0.14634146341463417, "rougeL": 0.2325581395348837},
{"image_id": "ROCO_00185", "question": "What does the anteroposterior radiograph show about the patient's pelvis?", "predicted_answer": "answer en What does the anteroposterior radiograph show about the patient's pelvis?\nAnteroposterior radiograph of the patient's pelvis showing the anterior acetabular dysplasia.", "true_answer": "Anteroposterior radiograph of pelvis at most recent follow up.", "bleu_score": 2.0230944175374798e-78, "rouge1": 0.23529411764705882, "rouge2": 0.125, "rougeL": 0.23529411764705882},
{"image_id": "ROCO_00190", "question": "Is there cerebrospinal fluid signal present around the spinal cord on this magnetic resonance image?", "predicted_answer": "answer en Is there cerebrospinal fluid signal present around the spinal cord on this magnetic resonance image?\nNo", "true_answer": "Immediate postinjury magnetic resonance image demonstrating intact spinal cord canal with cerebrospinal fluid signal surrounding the spinal cord at all levels.", "bleu_score": 3.6962883090160686e-78, "rouge1": 0.4615384615384615, "rouge2": 0.3243243243243243, "rougeL": 0.30769230769230765},
{"image_id": "ROCO_00206", "question": "What is the appearance of the mass in the arterial phase dynamic CT scan?", "predicted_answer": "answer en What is the appearance of the mass in the arterial phase dynamic CT scan?\nAxial phase dynamic CT scan shows a well-defined mass with a smooth contour and a low attenuation.", "true_answer": "Arterial phase dynamic CT shows homogenous enhancement of the mass (arrow). In the other level, normal right side adrenal gland was visualized.", "bleu_score": 2.568175498735308e-78, "rouge1": 0.35714285714285715, "rouge2": 0.22222222222222222, "rougeL": 0.21428571428571427},
{"image_id": "ROCO_00218", "question": "What is being shown by the arrow on the angiogram?", "predicted_answer": "answer en What is being shown by the arrow on the angiogram?\nAngiogram showing the site of the aneurysm (arrow).", "true_answer": "Angiogram showing pseudoaneurysm (arrow) formation in proximal left superficial femoral artery.", "bleu_score": 4.017758378857155e-155, "rouge1": 0.19354838709677416, "rouge2": 0.06896551724137931, "rougeL": 0.19354838709677416},
{"image_id": "ROCO_00251", "question": "What procedure is being performed on the patient as shown by the image?", "predicted_answer": "answer en What procedure is being performed on the patient as shown by the image?\nPatient receiving IV contrast injection with the needle in the left upper quadrant.", "true_answer": "Image from CT

```

```

fluoroscopic guidance rhizotomy demonstrates the radiofrequency cannular tip in the junction of superior articular process and transverse process.", "bleu_score":

3.743881382418666e-155, "rouge1": 0.16326530612244897, "rouge2": 0.0425531914893617, "rougeL": 0.16326530612244897},
{"image_id": "ROCO_00258", "question": "What is visible on the panoramic radiography?", "predicted_answer": "answer en What is visible on the panoramic radiography?\nPanoramic

radiography showing no abnormality.", "true_answer": "Bilateral styloid process elongation in a subject on a panoramic radiography.", "bleu_score": 1.1200407237786664e-231,

"rouge1": 0.23999999999999996, "rouge2": 0.08695652173913043, "rougeL": 0.23999999999999996},
{"image_id": "ROCO_00261", "question": "What is the nature of the lesion shown in the ultrasound image with thickened hyperechoic wall and tortuous arterial vessels seen on

color Doppler?", "predicted_answer": "answer en What is the nature of the lesion shown in the ultrasound image with thickened hyperechoic wall and tortuous arterial vessels

seen on color Doppler?\nUltrasound image showing a thickened hyperechoic wall with tortuous arterial vessels on color Doppler.", "true_answer": "Solid-cystic lesion with

thickened hyperechoic wall. In the region of the polycyclic solid structure and in the wall of the lesion, color Doppler (CDUS) shows tortuous arterial vessels", "bleu_score":

2.8669966493343836e-78, "rouge1": 0.5217391304347826, "rouge2": 0.2686567164179105, "rougeL": 0.34782608695652173},
{"image_id": "ROCO_00264", "question": "What is the likely diagnosis based on the chest x-ray showing a left tension pneumothorax with shift of the mediastinum to the right,

pleural effusion left, basal dorsolateral rib fractures and air visible under the right diaphragm?", "predicted_answer": "answer en What is the likely diagnosis based on the

chest x-ray showing a left tension pneumothorax with shift of the mediastinum to the right, pleural effusion left, basal dorsolateral rib fractures and air visible under the

right diaphragm?\nChest x-ray showing left tension pneumothorax with shift of the mediastinum to the right, pleural effusion left, basal dorsolateral rib fractures and air

visible under the right diaphragm.", "true_answer": "Initial chest x-ray showing a left tension pneumothorax with shift of the mediastinum to the right, pleural effusion left,

basal dorsolateral rib fractures. There's also air visible under the right diaphragm (arrow).", "bleu_score": 0.3495274586585512, "rouge1": 0.5631067961165048, "rouge2":

0.5346534653465347, "rougeL": 0.5631067961165048},
{"image_id": "ROCO_00271", "question": "What is the name of the vein that this tributary joins?", "predicted_answer": "answer en What is the name of the vein that this

tributary joins?\nThe left radial vein of the forearm.", "true_answer": "A large subcutaneous tributary pierces the superficial fascia to join the LSV in the thigh region.",

"bleu_score": 1.2183324802375697e-231, "rouge1": 0.2777777777777778, "rouge2": 0.058823529411764705, "rougeL": 0.22222222222222224},
{"image_id": "ROCO_00300", "question": "What are the characteristics of the lesions visible in the frontal lobes on the axial FLAIR MRI sequence?", "predicted_answer": "answer

en What are the characteristics of the lesions visible in the frontal lobes on the axial FLAIR MRI sequence?\nAxial FLAIR MRI sequence. Lesions in the frontal lobes are

```

```
shown.", "true_answer": "Brain magnetic resonance imaging (MRI) after gadolinium injection showed an image in the axial (FLAIR sequence) showing hyper-intensity lesions in the  
white matter of the frontal lobes. There is no signal abnormality of the cortex. Note that there is no mass effect on the ventricular cavities or midline structures.",  
  
"bleu_score": 4.402644441718793e-155, "rouge1": 0.3950617283950617, "rouge2": 0.22784810126582278, "rougeL": 0.24691358024691357},  
]
```