

Supplementary Materials: LLaVA-Ultra: Large Chinese Language and Vision Assistant for Ultrasound

Anonymous Authors

The supplementary material is structured as follows:

- Our US-Hospital dataset.
- Detailed description of data redundancy.
- More qualitative comparison results.

1 OUR US-HOSPITAL DATASET

Shown in Fig. A1, our ultrasound dataset from the hospital (US-Hospital) consists of over 170k sets of text instances and 2M images from multiple departments for over 20 examination site categories. These include heart, thyroid, kidney, liver, gallbladder, spleen and pancreas, male reproductive related and female reproductive related organs, etc. Such a wide range of body parts involved is rarely existed in previous medical multimodal datasets. It also provides a strong foundation for the model to learn the extensive and diverse medical semantics.

2 DETAILED DESCRIPTION OF DATA REDUNDANCY

Shown in Fig. B2, in the left figure, we show a complete example of data redundancy. Here an instance of text corresponds to 6 different images from several frames of a video of the same thyroid ultrasound of the same person. However, only figures (b) and (e) show the nodule described by the text. These two images are close to similar cuts showing the same nodule. Only figure (b) represents gray scale ultrasound and figure (e) represents color Doppler ultrasound, and the red and blue colors in it represent the signals of blood flow. The other figures do not show the lesion described by the text. Thus unlike the case where the text is one-to-one with the valid images, data redundancy occurs here. This is the superiority of our dataset over existing datasets to more closely match actual medical scenarios. In the right figure, we give the counting statistics for data redundancy, showing the number of images corresponding to a single text instance. It can be seen that most text instances correspond to between 2 and 10 images.

3 MORE QUALITATIVE COMPARISON RESULTS

3.1 More Results on SLAKE Dataset

On Chinese language dataset SLAKE-zh. Shown in Fig. C3, SLAKE dataset contains medical quizzes that tend to be concise. In these Q&A scenarios, LLaVA-Ultra can produce more robust results than LLaVA and LLaVA-Med.

3.2 More Results on OpenI Dataset

On Chinese language dataset OpenI-zh. Shown in Fig. C4, Open-I dataset contains medical questions and answers that tend to be more detailed. The examples we give contain both relatively healthy images and ultrasound with significant abnormalities. LLaVA tends to give incorrect judgments, indicating a lack of true understanding of the medical semantics. LLaVA-Med also sometimes misjudges due to the presence of hallucinations. In contrast, our LLaVA-Ultra gives more correct and robust diagnoses in both these normal and abnormal images.

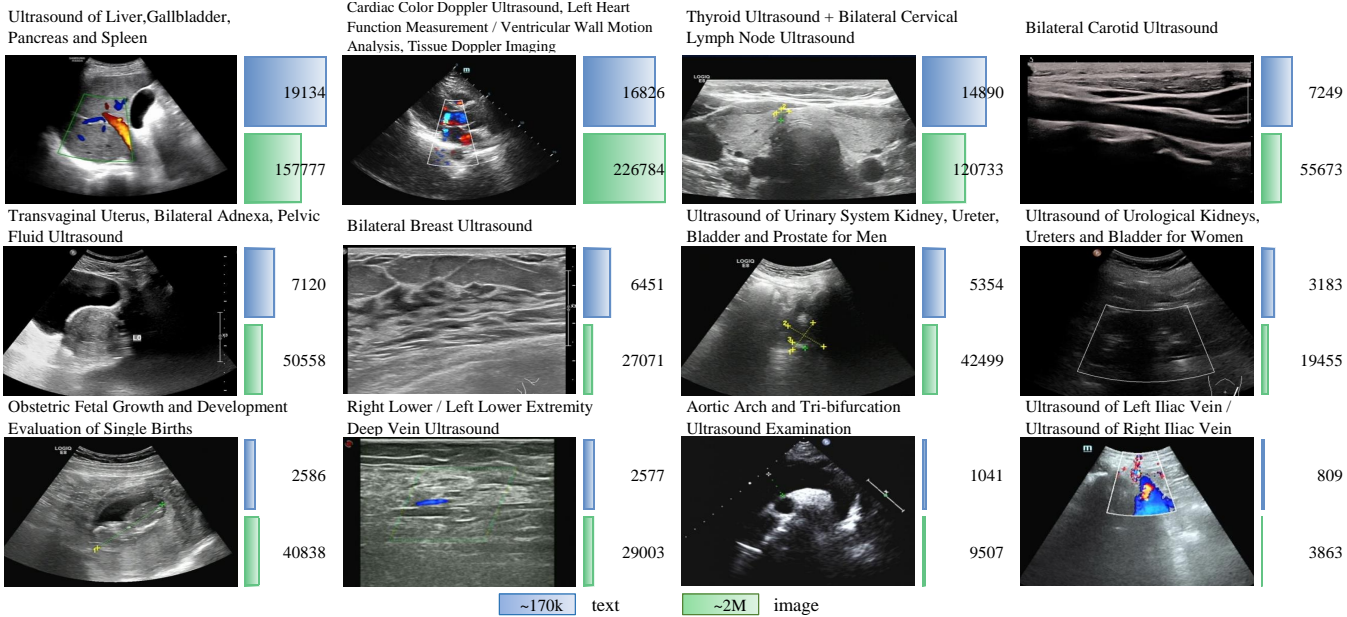


Figure A1: Some instances of our ultrasound dataset from the hospital (US-Hospital), where the numbers represent the counts of items of text and images.

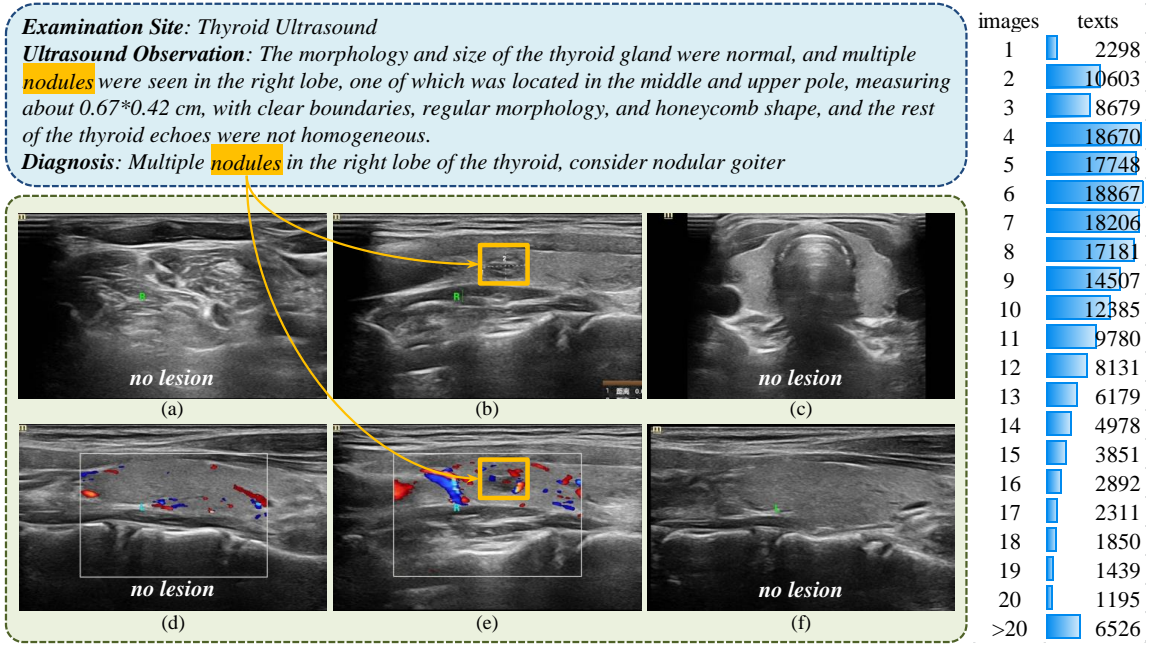


Figure B2: Detailed Description of Data Redundancy: (Left) An example of thyroid ultrasound from the same scanning examination of the same person. (Right) Statistics on the number of images matched by a single text instance.

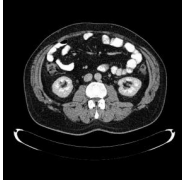

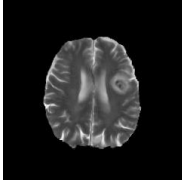

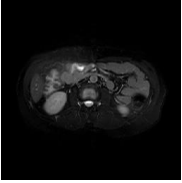
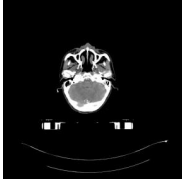

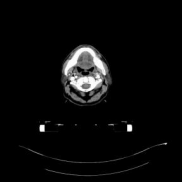

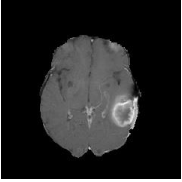
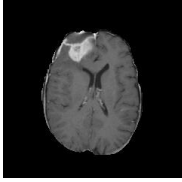



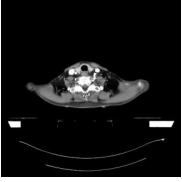

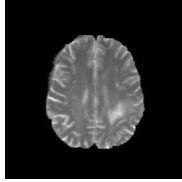
					
	User Does the picture include a liver?	User Is pneumonia in right upper lung?	User Which part in the image belong to?	User What diseases are included in the picture?	User What weighting was used in this MRI?
	LLaVA Included	LLaVA No, it's not.	LLaVA Head	LLaVA Lung Cancer	LLaVA T1
	LLaVA-Med No.	LLaVA-Med Yes, it is.	LLaVA-Med Parotid gland	LLaVA-Med Lung mass	LLaVA-Med T2
	LLaVA-Ultra Not included	LLaVA-Ultra No.	LLaVA-Ultra Head	LLaVA-Ultra Lung mass	LLaVA-Ultra T2
	Ground-truth Not included	Ground-truth No.	Ground-truth brain	Ground-truth Lung mass	Ground-truth T2
					
	User What color does the brain stem appear?	User How many hearts are present?	User What is the imaging method of this picture?	User What the images represent in the body?	User What type of imaging is shown in this image?
	LLaVA White	LLaVA 1	LLaVA X-Ray	LLaVA Abdomen	LLaVA CT
	LLaVA-Med Gray	LLaVA-Med 0	LLaVA-Med CT	LLaVA-Med Abdomen	LLaVA-Med CT
	LLaVA-Ultra Gray	LLaVA-Ultra 1	LLaVA-Ultra CT	LLaVA-Ultra Pelvis	LLaVA-Ultra Magnetic resonance
	Ground-truth Gray	Ground-truth 1	Ground-truth CT	Ground-truth Pelvis	Ground-truth Magnetic resonance
					
	User What are the major organ systems?	User What are symptoms of heart disease?	User What color does the right lung shown?	User Is pleural effusion in lower right of thorax?	User What is the shape of this patient's trachea?
	LLaVA Spinal cord	LLaVA Chest tightness, arrhythmia.	LLaVA black	LLaVA No way	LLaVA Irregularity
	LLaVA-Med Head	LLaVA-Med Pneumonia	LLaVA-Med gray	LLaVA-Med No way	LLaVA-Med ellipse
	LLaVA-Ultra Head	LLaVA-Ultra Chest tightness, arrhythmia.	LLaVA-Ultra black	LLaVA-Ultra Okay	LLaVA-Ultra Round shape
	Ground-truth Head	Ground-truth Chest tightness, arrhythmia.	Ground-truth black	Ground-truth Yes.	Ground-truth Round shape
		User How was this image imaged?		User What diseases are included?	
		LLaVA X-Ray		LLaVA Brain edema, brain tumor	
		LLaVA-Med CT		LLaVA-Med Cerebral edema	
		LLaVA-Ultra CT		LLaVA-Ultra Cerebral edema	
		Ground-truth CT		Ground-truth Cerebral edema	
		User What part of the body is it?		User Where is the anomaly in the picture?	
		LLaVA The thoracic cavity		LLaVA Right lower lobe	
		LLaVA-Med The thoracic cavity		LLaVA-Med Bottom left	
		LLaVA-Ultra The thoracic cavity		LLaVA-Ultra Bottom left	
		Ground-truth The thoracic cavity		Ground-truth Left lower lobe	
		User Are the lungs abnormal?		User Whether the abnormality is high or low density?	
		LLaVA Yes.		LLaVA Low density	
		LLaVA-Med No		LLaVA-Med High density	
		LLaVA-Ultra Yes.		LLaVA-Ultra High density	
		Ground-truth Yes.		Ground-truth High density	

Figure C3: More comparison results of our LLaVA-Ultra, prior SOTA LLaVA and LLaVA-Med on the SLAKE-zh dataset. The green and red parts highlight correct and incorrect results respectively.

User: Introduce this chest X-ray image.



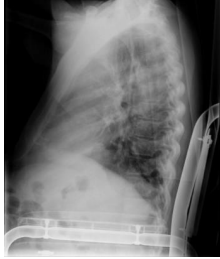
LLaVA: Chest X-ray showed mild cardiac enlargement, pulmonary vascular congestion, and increased pulmonary interstitial markers. However, there was no evidence of acute pulmonary complications, pleural effusion, or pneumothorax. In addition, there were mild degenerative changes in the thoracic spine. On the basis of these findings, impression is of cardiomegaly, increased pulmonary interstitial edema, the absence of any acute pulmonary disease.

LLaVA-Med: The chest X-ray showed that the outline of the heart and mediastinum was within normal limits, the lungs were clear, and the bony structure was intact. There was no evidence of acute cardiopulmonary disease, suggesting good health.

LLaVA-Ultra: Chest X-ray showed a normal cardiac contour and clear lungs without any evidence of pneumothorax or pleural effusion. The chest X-ray was normal.

GT: Chest X-ray showed that the heart, lungs, and mediastinae were within normal limits. There was no pleural effusion or pneumothorax. There was no focal opacification of the void suggestive of pneumonia. In addition, the patient was shielded during X-ray procedure. In summary, no acute cardiopulmonary disease was found on the X-ray.

User: Introduce this chest X-ray image.



LLaVA: Chest X-ray showed enlarged cardiac contours, low lung volumes, and crowded bronchi. No acute cardiopulmonary dysfunction was observed. However, there is evidence of chronic lung disease, possibly emphysema or Oxford emphysema, and bronchial crowding. Degenerative changes in the thoracic spine are also visible. Consultation with a healthcare professional is recommended for further evaluation and management of these findings.

LLaVA-Med: Chest X-ray showed clear lungs without any focal lacunar disease, pleural effusion, or pneumothorax. The contours of the heart and mediastinum were within normal limits, and there were degenerative changes in the spine. Based on these findings, there was no evidence of acute cardiopulmonary disease.

LLaVA-Ultra: Chest X-ray showed that the heart and mediastinum were normally contoured, the lungs were clear, and the bone structure was intact. There was no evidence of acute cardiopulmonary disease. This indicates that the patient's chest X-ray was normal.

GT: Chest X-ray showed no evidence of acute cardiopulmonary abnormalities. Specifically, there was no focal consolidation, pneumothorax, or pleural effusion, and the cardio-thoracic shape was stable and normal. In addition, no acute bone abnormalities were found. Overall, these findings suggest that the chest radiograph is normal and there is no immediate concern about acute illness.

User: Introduce this chest x-ray image.



LLaVA: The chest x-ray shows a normal heart size and clear lungs without any focal airspace disease, pleural effusion, or pneumothorax. The impression is that there is no acute cardiopulmonary disease present in the patient.

LLaVA-Med: The chest X-ray examination shows no evidence of acute cardiopulmonary process. The cardiomeastinal contours are normal and there is no focal consolidation, pleural effusion or pneumothorax identified. However, degenerative changes of the thoracic spine are observed. The patient also has right greater than left biapical bullous emphysema.

LLaVA-Ultra: The x-ray shows that the heart and mediastinum are normal in size and shape. The lungs are clear and there is no evidence of any acute disease. The bony structures are intact, except for some degenerative changes in the thoracic spine. The radiologist found no signs of any abnormalities or illnesses in the chest area.

GT: The x-ray shows that the heart and mediastinum are normal in size and shape. The lungs are clear without any signs of infiltrate, pneumothorax, or effusion. There is no acute cardiopulmonary disease. The bony structures are intact except for some degenerative changes in the thoracic spine.

User: Introduce this chest X-ray image.



LLaVA: Chest X-ray showed clear lungs with no evidence of pneumothorax or pleural effusion. The size and contour of the heart and mediastinum were normal. However, there were degenerative changes in the spine, while one vertebral body in the middle of the thoracic spine had a mild wedge deformity. Overall, no acute cardiopulmonary abnormalities were found on radiographs.

LLaVA-Med: Chest X-ray shows that the outline of the heart and mediastinum is normal. The lungs were clear without any focal lacunar disease, pleural effusion, or pneumothorax. However, an 8-mm nodule was found in the left lower lobe and was found to be a calcified granuloma. No bone abnormalities were found. The impression indicated that there was no imaging manifestation of an acute cardiopulmonary process.

LLaVA-Ultra: Chest X-ray showed a normal-sized heart, clear lungs, and no evidence of pneumothorax or pleural effusion. The chest X-ray was normal.

GT: The X-ray image showed a normal size and blood vessels of the heart, while the lungs appeared to be clear without any evidence of hydrops or pneumothorax. No abnormalities were found in the visible bony structures. Overall, the chest appeared to be normal, with no significant findings.

Figure C4: More comparison results of our LLaVA-Ultra, prior SOTA LLaVA and LLaVA-Med on the OpenI-zh dataset. The green and red parts highlight correct and incorrect results respectively.