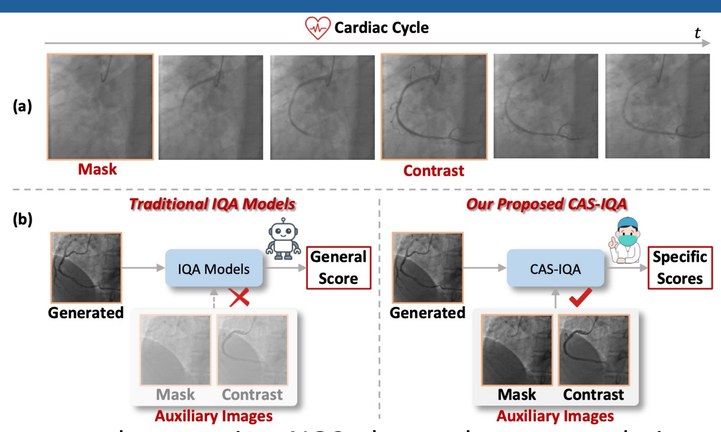
ICONIP2025
November 20-24 Okinawa

CAS-IQA: Teaching Vision-Language Models for Synthetic Angiography Quality Assessment

Bo Wang*, De-Xing Huang*, Xiao-Hu Zhou†, Mei-Jiang Gui,
Nu-Fang Xiao, Jian-Long Hao, Ming-Yuan Liu†, and Zeng-Guang Hou

⋆ Equal Contributions † Corresponding Author

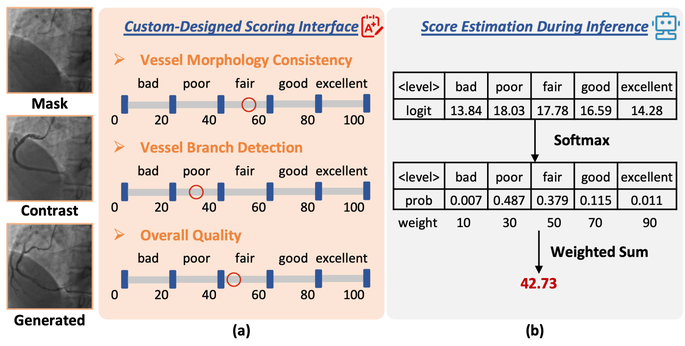
Motivation



- Recent advances in AIGC have demonstrated impressive capabilities in generating photo-realistic angiographies from non-contrast X-rays.
- However, ***the fidelity of such synthetic angiographies to essential clinical constraints remains unquantified***, underscoring the need for rigorous quality assessment.
- Existing IQA models **fail to leverage auxiliary images** as references during evaluation and **lack fine-grained, task-specific metrics** necessary for clinical relevance.

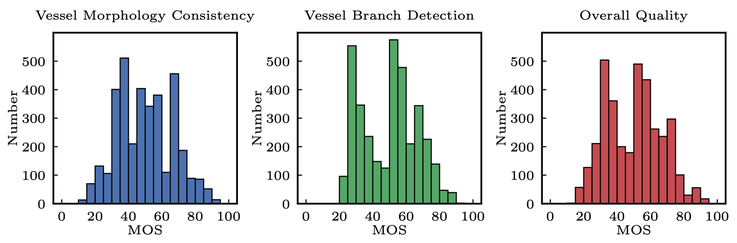
CAS-3K Dataset

Image Collection and Scoring Rules



- ***Image Collection.***
  - We construct CAS-3K by manually selecting 713 high-quality mask–contrast angiography pairs from **XCAD** and **CADICA**.
  - Using the masks as input, we generate **3,565** synthetic angiograms with **five** state-of-the-art generative models, forming paired Mask–Generated–Contrast samples.
- ***Task-specific Evaluation Metrics.***
  1. **Vessel Morphology Consistency (VMC)**
  2. **Vessel Branch Detection (VBD)**
  3. **Overall Quality (OQ)**

Subjective Quality Assessment and Data Processing

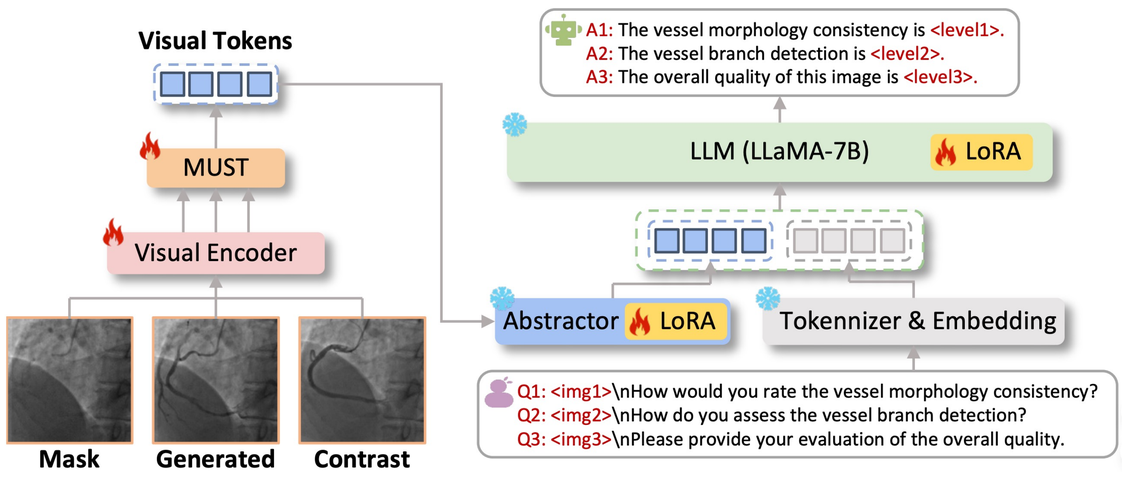


- **Three** graduate students majoring in biomedical engineering are recruited to participate in subjective evaluation experiments.
- The **MOS** is calculated using the following formulas.

$$MOS_j = \frac{1}{N}\sum_{i=1}^{N} \hat{Z}_{ij} \qquad Z_{ij} = \frac{S_{ij} - \mu_i}{\sigma_i}$$

Method

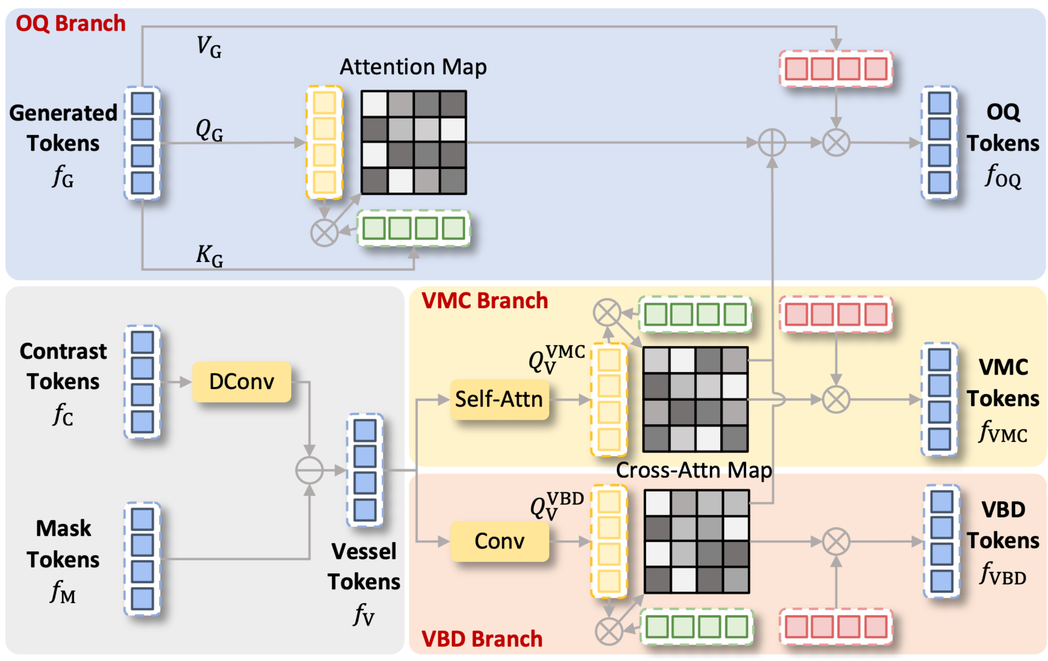
Overall Framework



- ***VLM-based.***
  - Our method is built upon the **mPLUG-Owl2**.
- ***Instruction Tuning.***
  - We design a series of prompts framed as **multi-turn question–answer interactions**, as illustrated in the figure on the right.
  - Both the Visual Abstractor and the LLM are fine-tuned using **LoRA**.
- ***Score Estimation.***
  - We apply softmax over five level logits and take a weighted sum using interval midpoints to obtain the final quality score.

Vessel Morphology Consistency:
#User: <img1>As an experienced interventional radiologist, how would you rate the vessel morphology consistency of this image?
#Assistant: The vessel morphology consistency of this image is <level1>.
Vessel Branch Detection:
#User: <img2>From your perspective as an interventional radiologist, how do you assess the vessel branch detection in this image?
#Assistant: The vessel branch detection of this image is <level2>.
Overall Quality:
#User: <img3>With your experience in interventional radiology, please provide your evaluation of the overall quality of this image.
#Assistant: The overall quality of this image is <level3>.

Multi-Path FeatUre FuSion and RouTing (MUST) Module

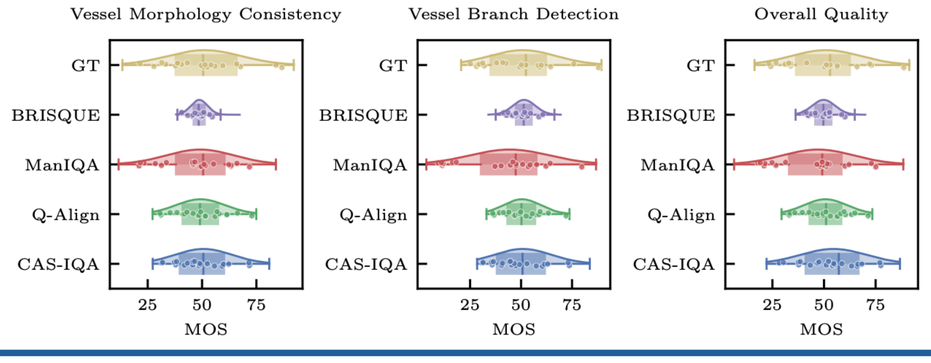


- A MUST module is presented, which ***fuses features of Mask, Contrast, and Generated***, and ***routes them to metric-specific branches for optimized assessment***.
- We use **different attention maps** to focus on the synthetic angiography from different perspectives.

Experiments

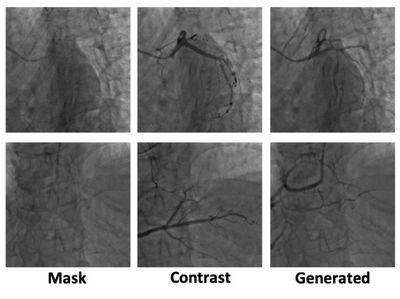
Comparisons with SOTA Methods

Category	Methods	VMC		VBD		OQ	
		PLCC	SRCC	PLCC	SRCC	PLCC	SRCC
Handcrafted	NIQE [16] [SPL'12]	0.0520	0.0438	0.0289	0.0256	0.0337	0.0202
	BRISQUE [15] [TIP'12]	0.1876	0.1901	0.2359	0.2391	0.2224	0.2139
DNN-based	DBCNN [31] [TCSVT'20]	0.4990	0.4853	0.5856	0.5647	0.4745	0.4666
	HyperIQA [21] [CVPR'20]	0.6657	0.6356	0.5942	0.5873	0.6339	0.6216
	AHIQ [8] [CVPR'22]	0.6251	0.6434	0.6050	0.6010	0.6654	0.6621
	ManIQA [28] [CVPR'22]	0.6823	0.6759	0.6089	0.6060	0.6778	0.6689
	QCN [20] [CVPR'24]	0.6750	0.6651	0.6308	0.6235	0.6796	0.6765
	LoDa [27] [CVPR'24]	0.6859	0.6664	0.6353	0.6279	0.6899	0.6737
VLM-based	Q-Align [25] [ICLR'24]	0.6030	0.5907	0.5683	0.5651	0.6096	0.5999
	MA-AGIQA [24] [MM'24]	0.6892	0.6625	0.6507	0.6314	0.6696	0.6598
	CAS-IQA (Ours)	0.6925	0.6855	0.6639	0.6605	0.6985	0.6986



Ablation Study

MUST	VMC		VBD		OQ	
	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC
✗	0.6517	0.6636	0.6451	0.6491	0.6830	0.6750
✓	**0.6925**	**0.6855**	**0.6639**	**0.6605**	**0.6985**	**0.6986**



Metric	w/o MUST	w/ MUST	GT
VMC	50.16	47.31	41.45
VBD	50.19	50.87	43.30
OQ	62.91	53.19	44.26

Metric	w/o MUST	w/ MUST	GT
VMC	24.75	19.98	17.89
VBD	29.79	28.34	28.93
OQ	16.48	24.02	23.30

- We compare it against a comprehensive set of baselines, including ***two handcrafted IQA methods***, ***six DNN-based approaches***, and ***two VLM-based methods***.
- Extensive experimental results indicate that CAS-IQA significantly ***outperforms state-of-the-art IQA models*** on the CAS-3K dataset.
- Extensive ablation studies ***validate the effectiveness*** of our MUST module.

Conclusion

- **The first high-quality dataset for angiography IQA (CAS-3K)** is constructed, consisting of 3,565 synthetic angiographies with metric-based annotations.
- **A novel IQA framework based on vision-language models** is proposed to comprehensively evaluate the quality of synthetic angiographies. Moreover, the MUST module is designed to effectively fuse visual cues from multiple image sources and adapt them to different evaluation metrics.
- Future work will focus on improving the interpretability of CAS-IQA to facilitate its deployment in real-world clinical applications.

Contact

Bo Wang **De-Xing Huang** **HEART Group**