Image Quality-aware Diagnosis via Meta-knowledge Co-embedding

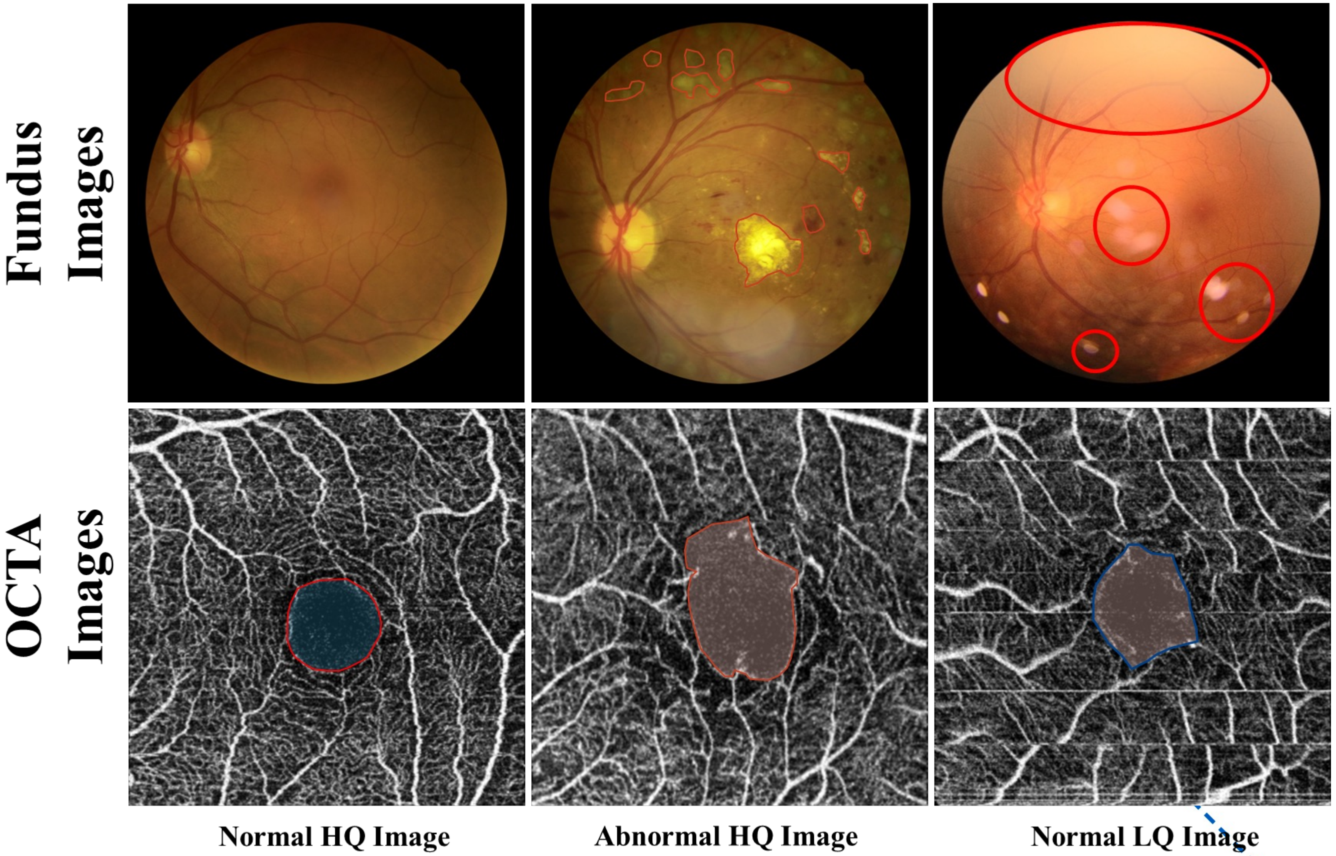
Haoxuan Che, Siyu Chen, Hao Chen
Department of Computer Science and Engineering
hche@ust.hk

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

計算機科學及工程學系
THE DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

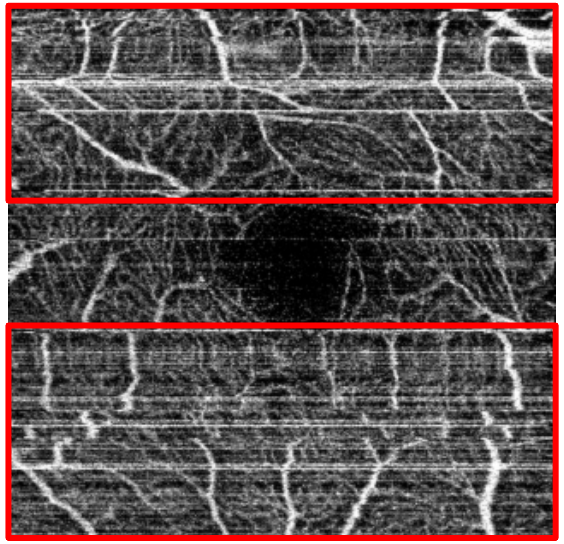
Background

- AI shown desirable performance for high-quality (HQ) medical images, yet fail to generalize on low-quality (LQ) images [Nagendran et al., 2020].
  - Real-world deployment example from Google [Beede et al., 2020].
  - Research study of LQ fundus images [Liu et al., 2022].

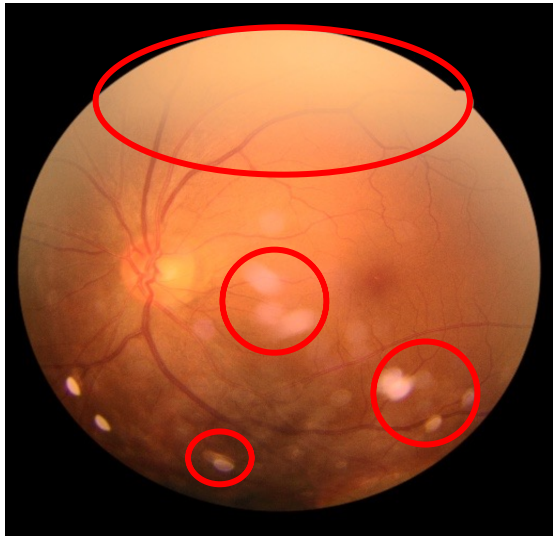


Background

- Medical image degradations can significantly affect diagnostic semantics of images:
  - It can affect diagnostic measurements.
  - It can raise potential false abnormalities.



Influence vessel area calculation.



Shade anatomical retinal structures, generate lesion-like artifacts.

IMAGE QUALITY MATTERS

Motivation

Image Quality Assessment

- Reject "Bad"/ungradable/LQ samples [Fu et al., 2019], but sub-optimal [Yii et al., 2022].

Image Enhancement

- Improve image quality, but costive [Shen et al., 2020] and boost marginally [Liu et al., 2022].

Multi-task Learning

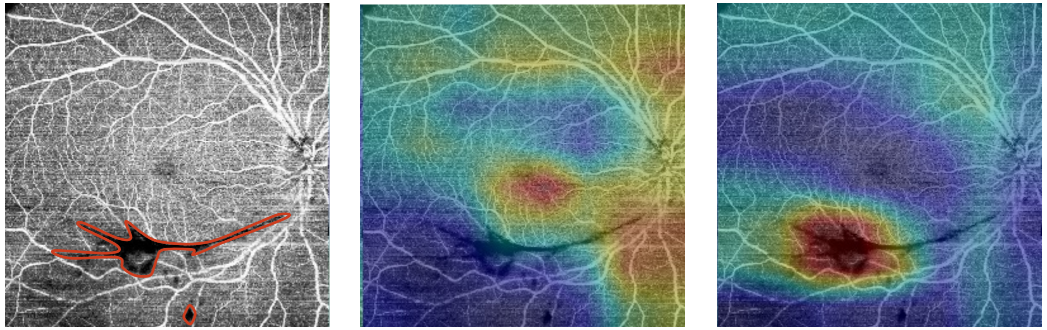
- Explore correlations between diseases and tasks [Che et al., 2022], but ignore the potential benefits from quality information.

We raise the concept of Image Quality-aware Diagnosis (IQAD), aiming to leverage LQ images and corresponding image quality labels to boost disease diagnosis performance.

Challenges

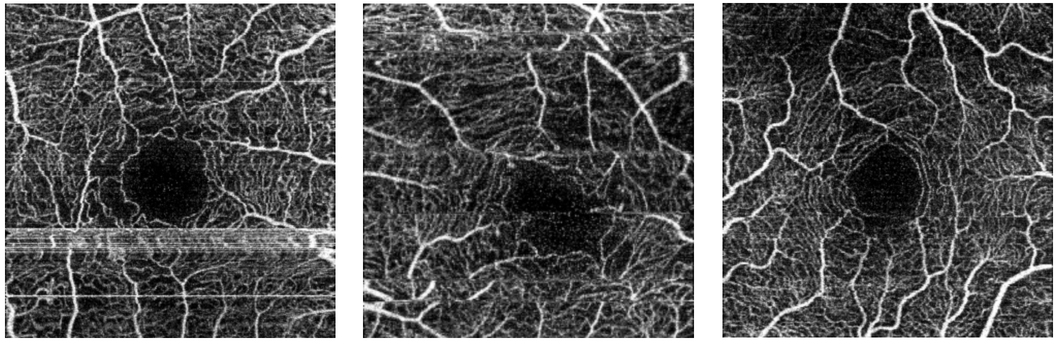
- **Non-direct relationship between image quality and diagnosis.**

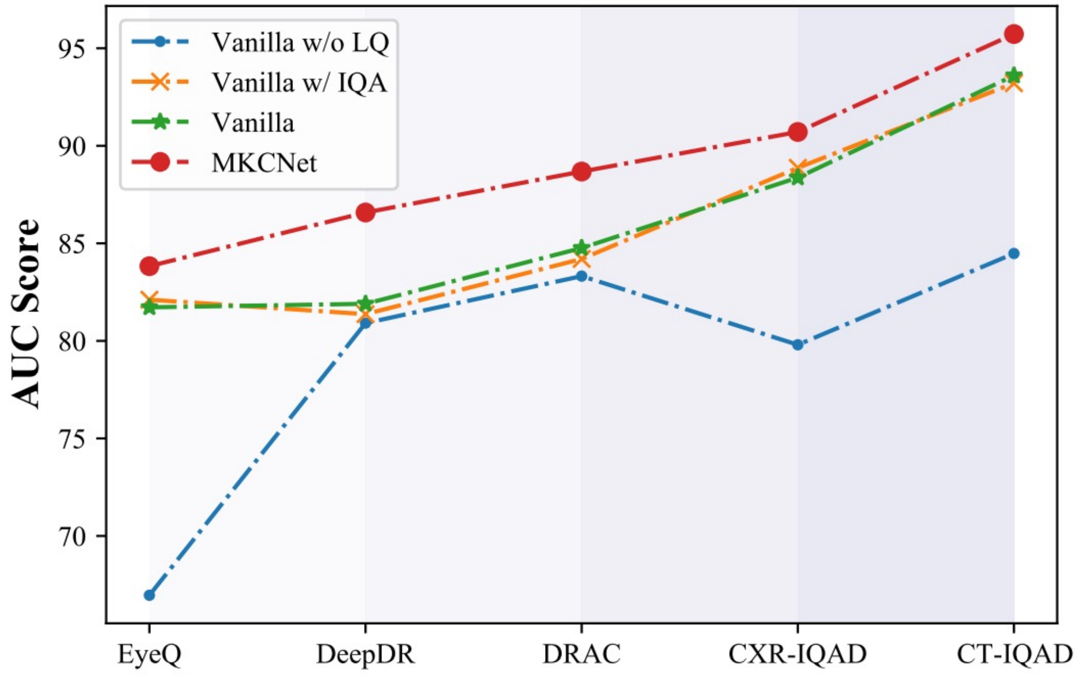
Original Images | Diagnosis Focus | Degradation Area



- **Limited granularity of binary or multi-class quality annotations.**

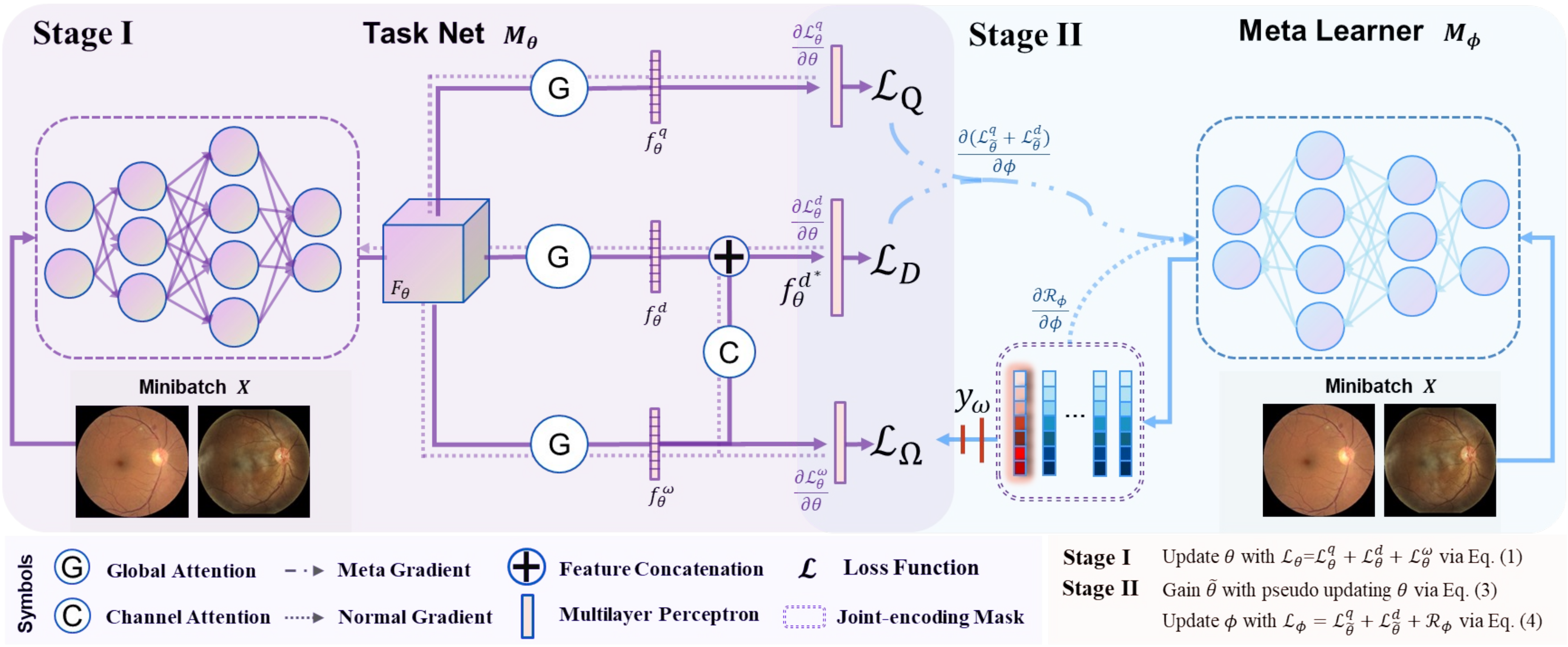
Segmentation | Movement | Shadow





Proposed Method - Overview

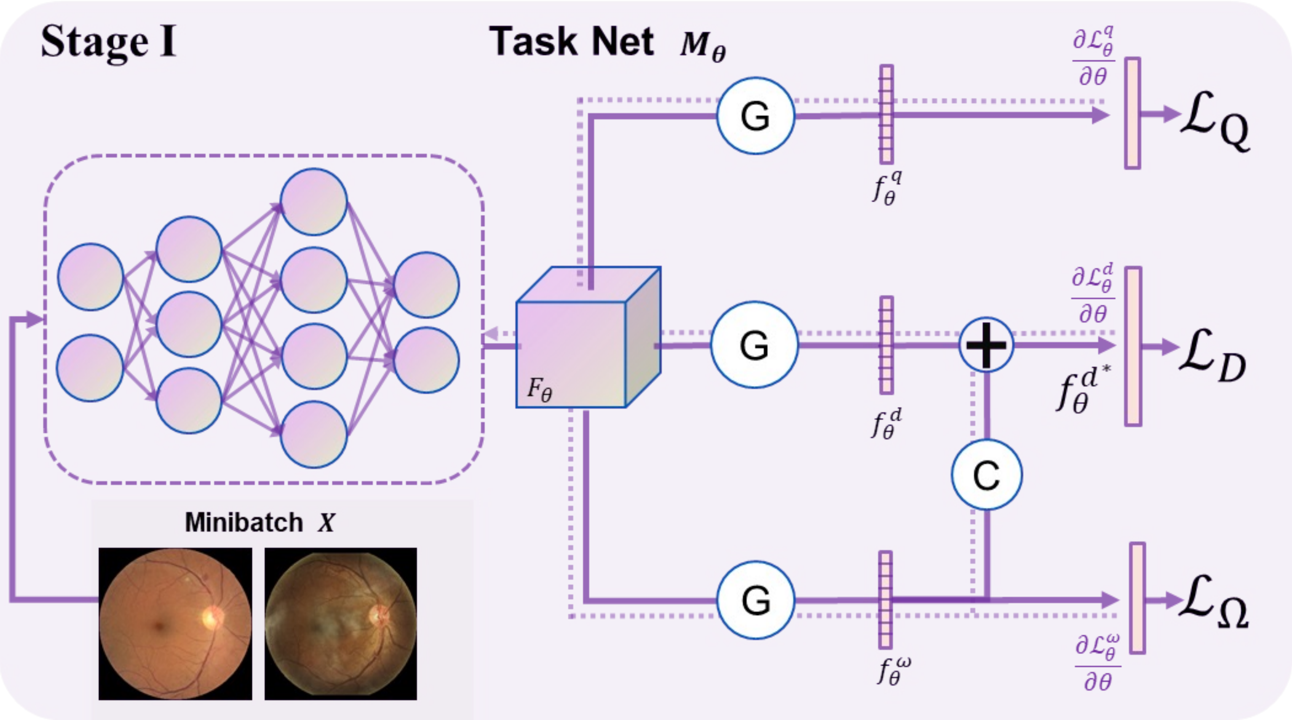
- We propose a meta-knowledge co-embedding network (MKCNet) to tackle IQAD.



Proposed Method – Stage I

❑ Task Net M_θ

- Three task branches for learning the image quality label y_q, diagnosis label y_d, and auxiliary label embedding y_ω.
- y_ω contains desired information related to quality and diagnosis, being optimized by Meta Learner.
- Global attention block (G) helps to instruct a generalizable F_θ.
- Meta-knowledge assistance block (C) explicitly leverage f_θ^ω via extracting useful information to diagnose.



Learning objective: $\mathcal{L}_\theta = \mathcal{L}_D\left(M_\theta^d(x), y_d\right) + \mathcal{L}_Q\left(M_\theta^q(x), y_q\right) + \mathcal{L}_\Omega\left(M_\theta^\omega(x), y_\omega\right)$ **Eq. (1)**

Proposed Method – Stage II

❑ Meta Learner M_ϕ

- Design the joint-encoding masking to ensure the semantics of y_ω.

$$y_\omega = \mathcal{B}y_{d,q}(M_\phi(x))$$

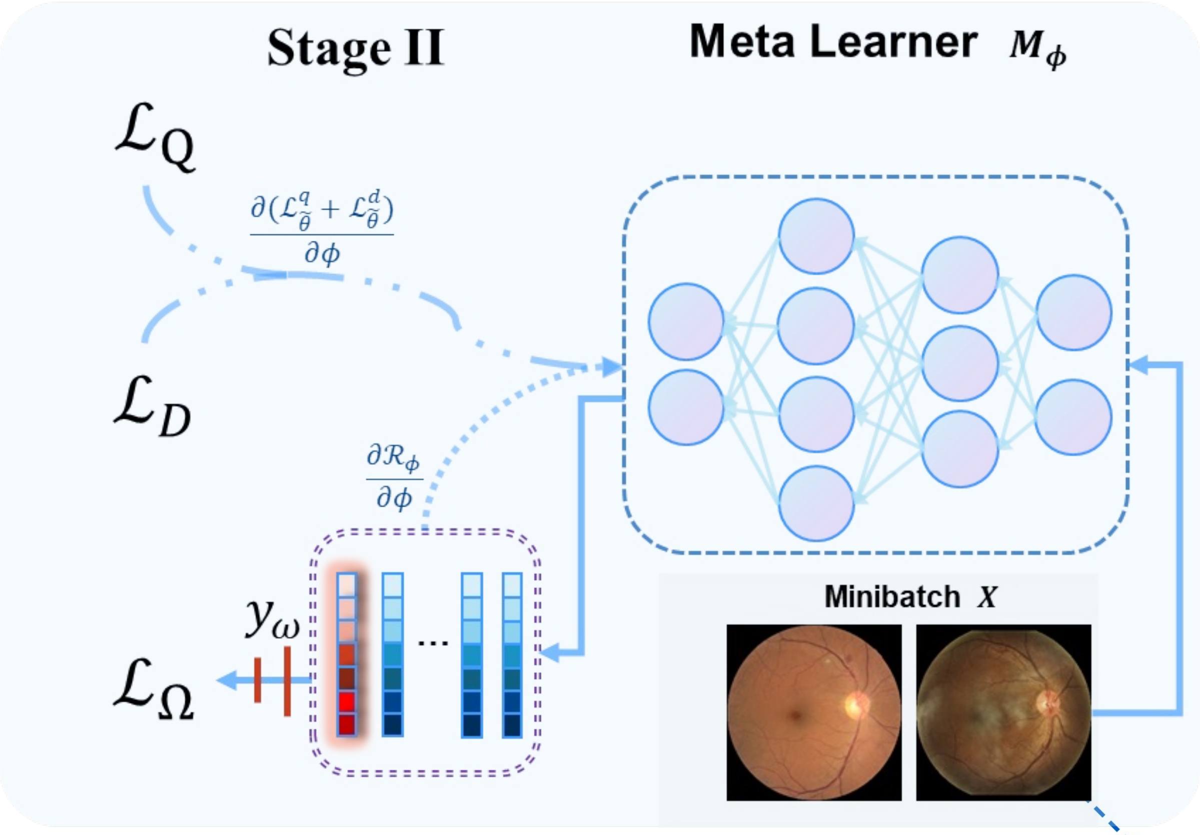
- Adopt meta-auxiliary learning to enable meta learner optimize Task Net:

1. Pseudo update (Eq. (3)):

$$\tilde{\theta} \coloneqq \theta - \alpha\nabla_\theta[\mathcal{L}_D(M_\theta^d(x), y_d) + \mathcal{L}_Q(M_\theta^q(x), y_q) + \mathcal{L}_\Omega(M_\theta^\omega(x), y_\omega)]$$

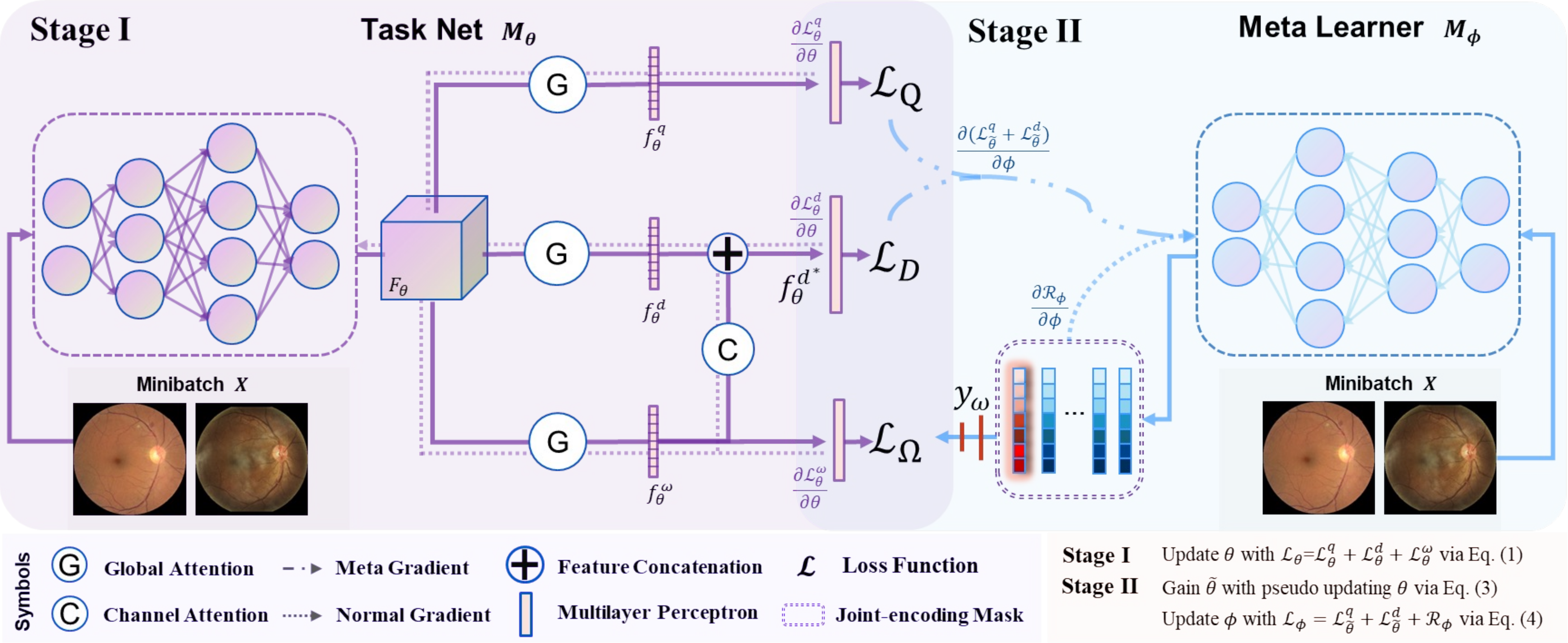
2. Meta-optimization (Eq. (4)):

$$\phi \coloneqq \phi - \beta\nabla_\theta[\mathcal{L}_D(M_\theta^d(x), y_d) + \mathcal{L}_Q(M_\theta^q(x), y_q) + \mathcal{R}(\mathcal{B}y_{d,q}(M_\phi(x))]$$



Proposed Method

- 1) IQAD, 2) explicit utilization mechanism, 3) semantics constraints.



Experiment

- **Datasets**: DRAC (OCTA), EyeQ (Fundus), DeepDR (Fundus), CT-IQAD (CT), CXR-IQAD (CXR).
- **Test paradigms:** train-test for DRAC, train-val-test for others.
- **Extensive ablation study, evaluation metrics**: AUC, Macro F1 score, Accuracy.

Table 1. Statistics for lung disease diagnosis datasets (left group) and ophthalmic disease diagnosis datasets (right group).

Dataset	CT-IQAD			CXR-IQAD					DRAC			DeepDR			EyeQ			
Label	ALL	HQ	LQ	ALL	HQ	LQ-C	LQ-A	Label	ALL	HQ	LQ	ALL	HQ	LQ	ALL	HQ	LQ-U	LQ-P
Normal	1,149	800	349	2,582	792	791	999	NoDR	545	444	101	914	532	382	20,680	12,308	4,553	3,747
Covid-19	1,197	800	397	-	-	-	-	NPDR	344	303	41	974	506	468	7,533	4,405	1,690	1,438
Pneumonia	-	-	-	5,704	2,137	2,136	1,431	PDR	108	93	15	112	50	62	558	104	191	353
ALL	2,346	1,600	746	8,286	2,929	2,927	2,430	ALL	997	840	157	2,000	1,088	912	28,789	16,817	6,434	5,538

Experiment – In-distribution Test

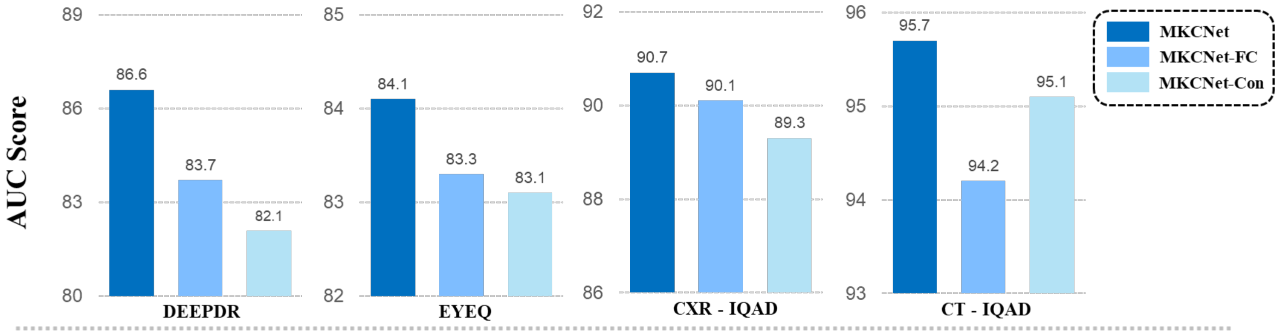
❑ **Comparison with the state-of-the-art methods.**

Table 2. Comparison with state-of-the-art approaches in disease diagnosis.

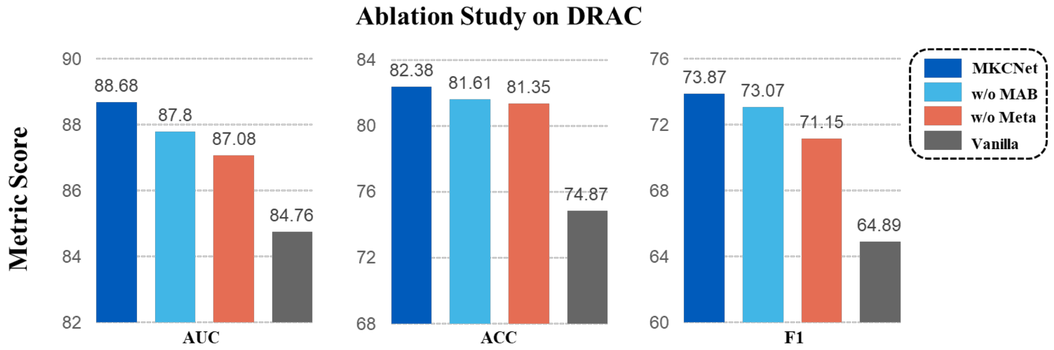
	Dataset	CT-IQAD			CXR-IQAD			DRAC			DeepDR			EyeQ			Avg.		
	Metrics	AUC	ACC	F1	AUC	ACC	F1	AUC	ACC	F1	AUC	ACC	F1	AUC	ACC	F1	AUC	ACC	F1
ODS	MMCNN [53]	86.57	86.54	86.51	89.13	89.07	91.82	82.41	77.20	66.18	76.39	79.50	54.40	69.36	73.22	35.48	80.77	81.11	66.88
ODS	BIRA-Net [54]	88.46	88.46	88.21	89.88	90.76	93.23	84.00	77.46	67.00	80.67	78.50	64.10	80.13	72.50	56.42	84.63	81.54	73.79
ODS	GREEN [20]	92.54	92.52	92.44	90.06	91.30	93.67	84.37	73.83	63.03	81.14	76.50	**64.88**	81.23	77.21	59.08	85.87	82.27	74.64
ODS	CABNet [19]	91.01	91.03	90.79	88.38	89.73	92.51	81.94	72.28	59.07	82.39	77.75	55.97	79.21	76.95	53.48	84.59	81.55	70.36
Other	Mixstyle [55]	91.06	91.03	90.99	88.85	90.52	93.14	85.03	73.32	60.91	81.89	73.50	60.34	82.10	74.92	61.25	85.79	80.66	73.33
Other	AugMix [56]	89.17	89.10	89.26	88.21	89.79	92.58	84.94	70.47	57.66	84.79	80.00	54.74	77.38	77.00	44.26	84.90	81.27	67.70
Other	DDAIG [57]	92.72	92.74	92.54	88.19	90.28	93.00	84.95	77.72	62.91	79.40	75.75	51.66	70.71	74.64	37.84	83.19	82.23	67.59
Other	Mixup [58]	92.12	92.10	92.04	88.98	90.70	93.28	82.41	72.80	62.84	81.69	76.00	63.57	82.35	77.74	**63.07**	85.51	81.87	74.89
MTAL	QGNet [16]	92.71	92.74	92.50	89.51	90.70	93.22	84.38	63.73	47.80	84.21	73.50	64.66	82.40	74.73	60.36	86.64	79.08	71.71
MTAL	MAXL [45]	93.35	93.38	93.16	87.99	89.86	92.66	87.24	77.72	69.20	81.59	71.25	57.58	80.64	73.10	58.81	86.16	81.06	74.28
MTAL	CANet [36]	93.36	93.38	93.19	90.51	91.49	93.78	86.70	75.13	64.38	81.83	77.50	62.74	81.89	77.20	60.26	86.86	82.94	74.87
MTAL	MT-Net [39]	92.56	92.52	92.51	90.07	90.58	93.05	84.63	73.06	61.89	82.40	76.75	64.51	82.66	75.97	59.92	86.46	81.78	74.38
MTAL	MTMR-Net [38]	94.46	94.44	94.37	89.69	91.61	93.96	85.79	76.42	65.12	82.30	63.75	53.61	79.71	76.57	60.50	86.39	80.56	73.51
MTAL	DETACH [37]	92.53	92.52	92.41	89.52	90.64	93.16	84.81	76.94	65.24	83.61	71.75	62.14	81.68	74.80	59.64	86.43	81.33	74.52
	MKCNet (Ours)	**95.73**	**95.73**	**95.65**	**90.71**	**91.91**	**94.12**	**88.68**	**82.38**	**73.87**	**86.58**	**80.00**	54.73	**83.83**	**78.01**	59.68	**89.11**	**85.61**	**75.61**

Experiment – Ablation Study

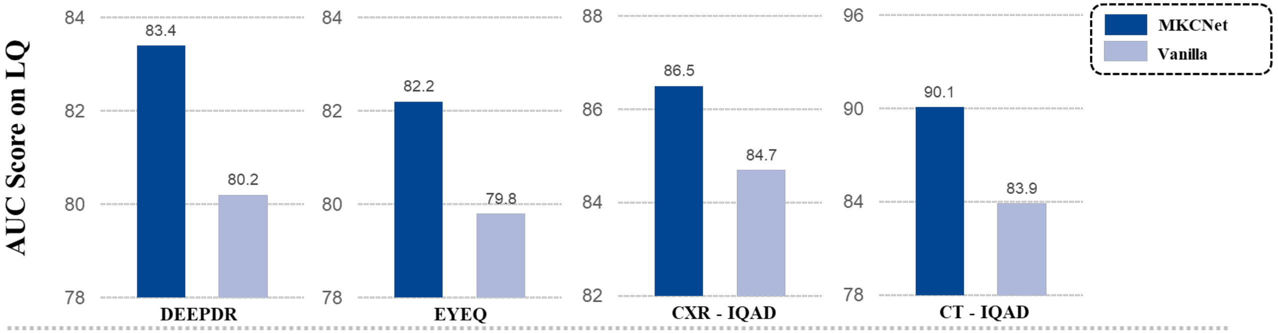
(2) Task Net Design



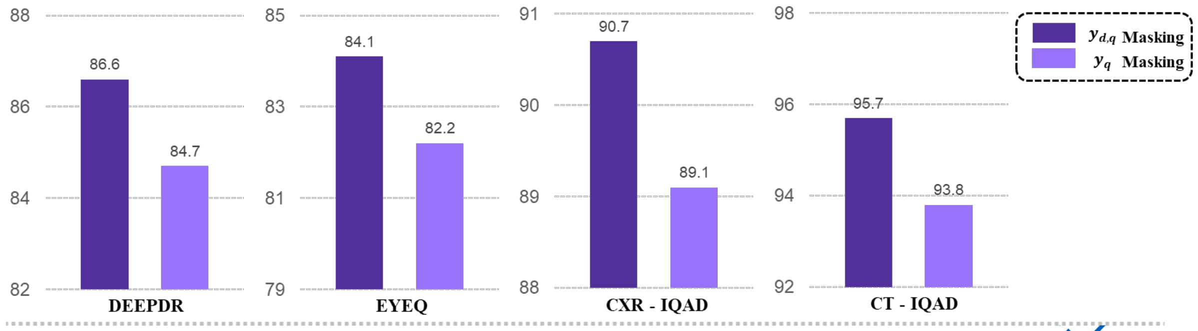
(1) Components



(3) Performance on LQ images

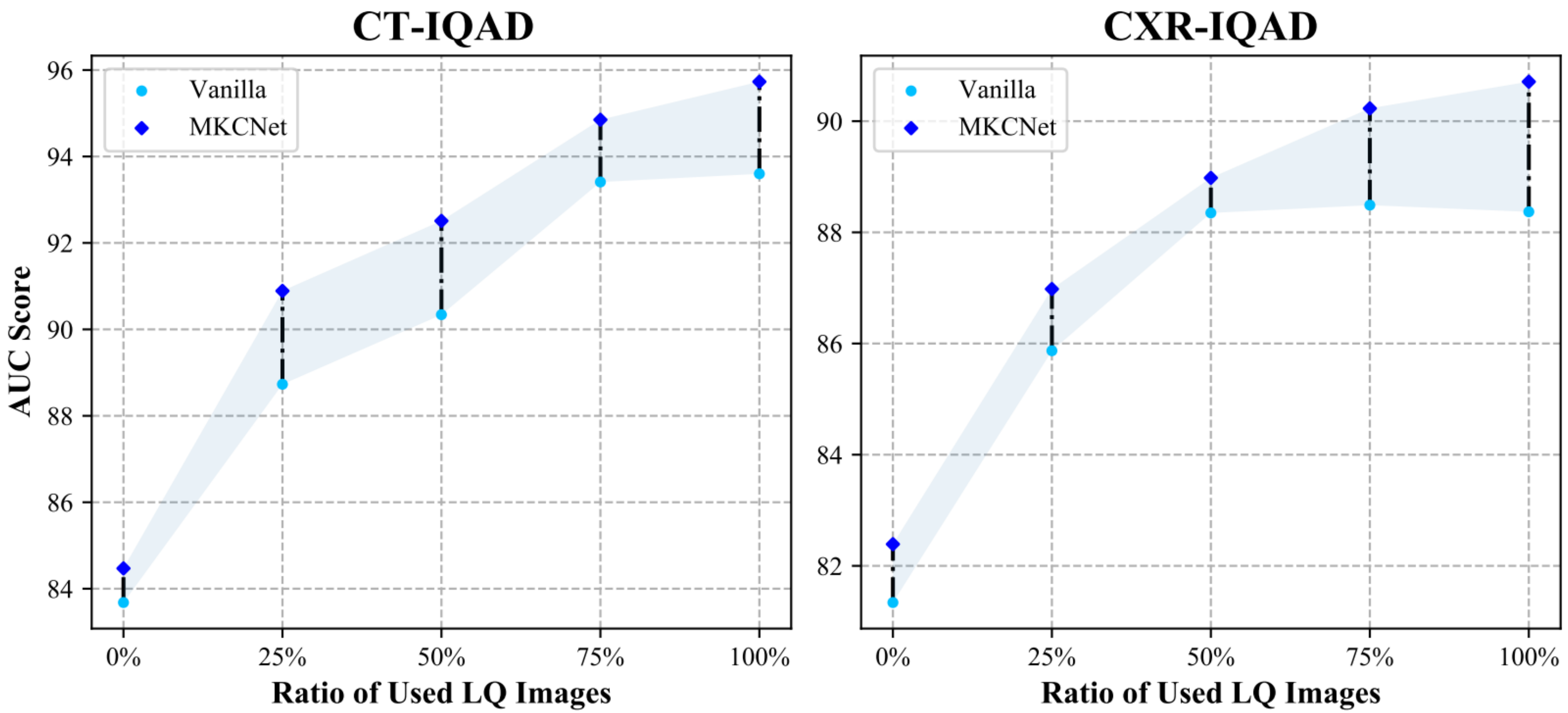


(4) Masking Design



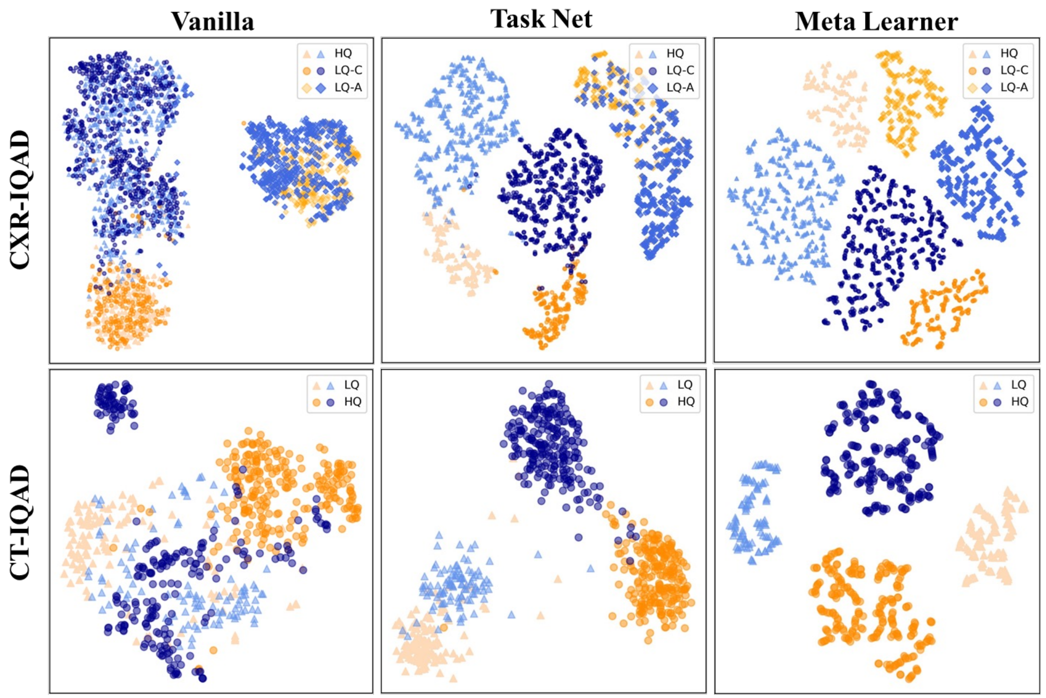
Experiment – Out-of-distribution Test

❑ **Simulate different ratio of LQ images to gain distribution gaps.**

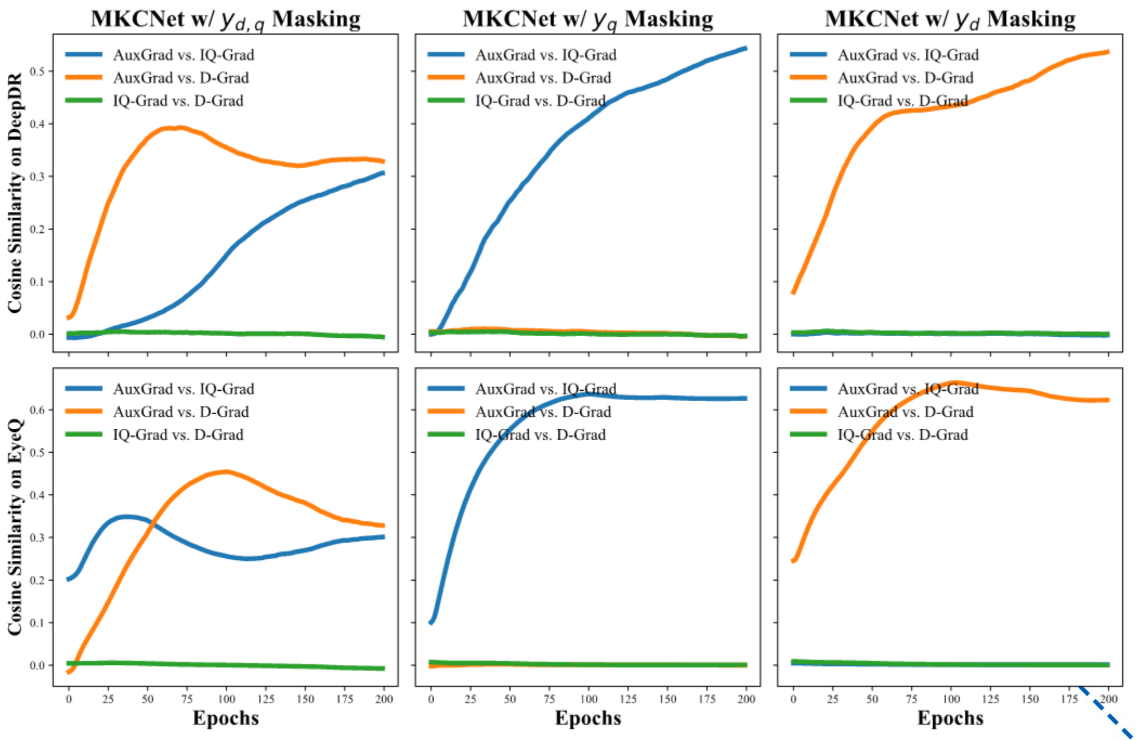


Experiment – Feature Analysis

t-SNE Analysis

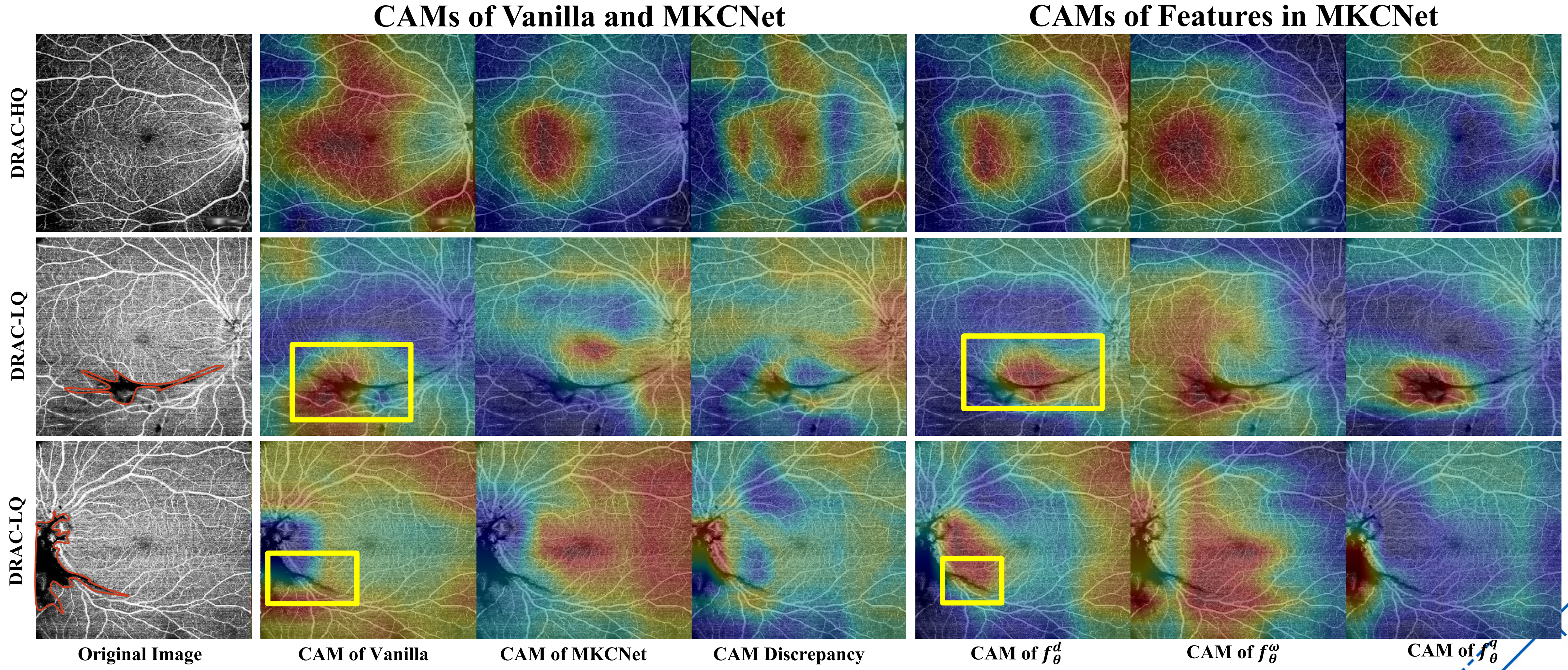


Gradient Analysis



香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Experiment – Qualitative Analysis



Conclusion

- ❑ First time to raise the concept of **image-quality aware diagnosis.**
- ❑ Explore to design image quality label **utilization mechanism**.
- ❑ Enable **semantics constraints** in the meta-auxiliary learning.
- ❑ Extensive experiments on **five** datasets with **four** widely-used imaging modalities.