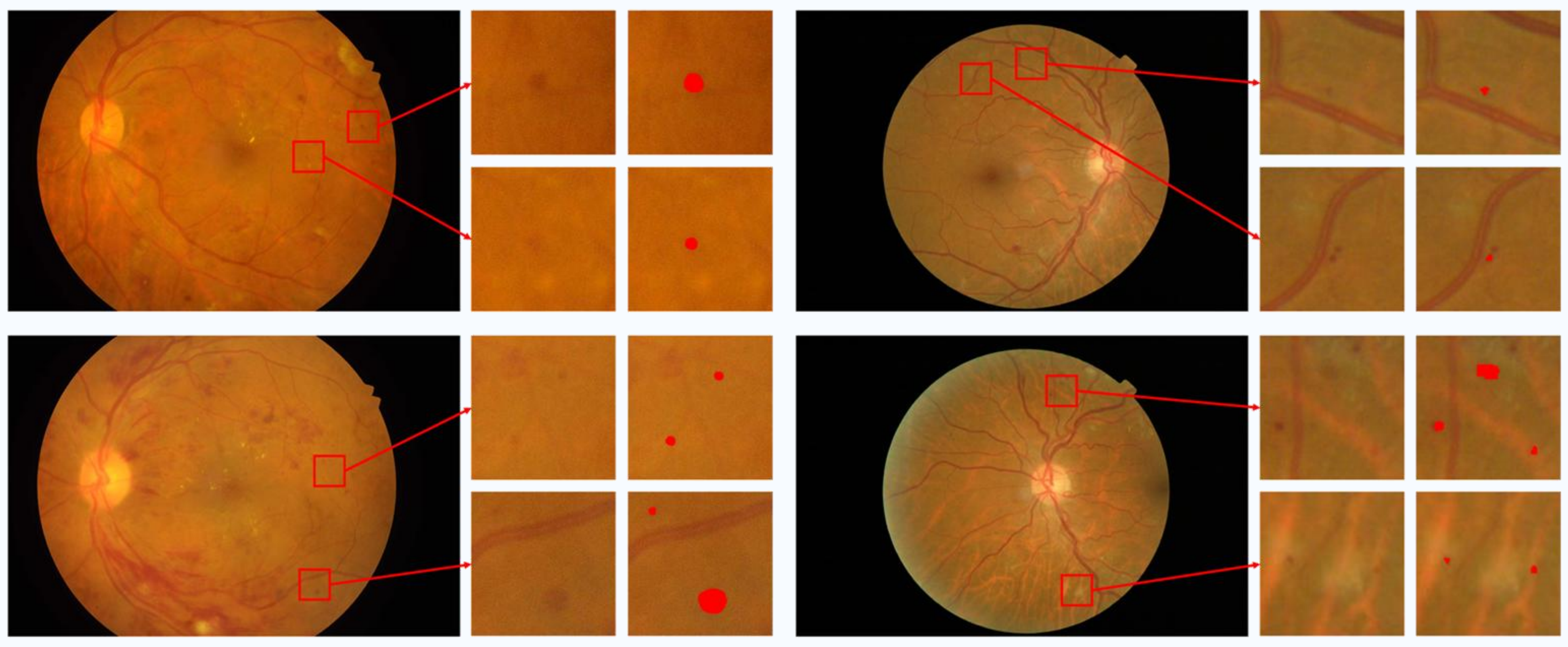
WDT-MD: Wavelet Diffusion Transformers for Microaneurysm Detection in Fundus Images

AAAI-26 / IAAI-26 / EAAI-26
JANUARY 20–27, 2026 | SINGAPORE

Yifei Sun, Yuzhi He, Junhao Jia, Jinhong Wang, Ruiquan Ge✉, Changmiao Wang✉, Hongxia Xu✉

INTRODUCTION

Microaneurysms (MAs), the earliest pathognomonic signs of **Diabetic Retinopathy (DR)**, present as sub-60 μm lesions in fundus images with highly variable photometric and morphological characteristics, rendering manual screening not only labor-intensive but also error-prone.



QUANTITATIVE RESULTS

Table 1. Comparison of different methods on IDRiD.

Method	Source	Pixel-level					Image-level				
		AUC	ACC	F1	SEN	SPE	AUC	ACC	F1	SEN	SPE
AnoDDPM	$CVPR_{22}$	81.76	99.91	53.34	63.62	99.93	71.90	76.92	62.50	55.56	88.24
CPC	$WACV_{23}$	76.77	99.93	49.75	53.63	99.96	77.45	80.77	70.59	66.67	88.24
DAE	$MedIA_{23}$	71.52	99.69	35.64	43.23	99.72	56.86	53.85	50.00	66.67	47.06
HACDR-Net	$AAAI_{24}$	56.38	95.07	4.03	18.82	95.12	63.07	65.38	52.63	55.56	70.59
AE[$d_{optimal}$]	$MICCAI_{24}$	75.06	99.24	19.52	50.88	99.27	62.75	61.54	54.55	66.67	58.82
Dif-fuse	TMI_{24}	81.82	99.95	69.55	63.65	99.97	71.57	73.08	63.16	66.67	76.47
GatingAno	PR_{24}	78.73	92.07	11.49	63.04	92.09	54.25	53.85	45.45	55.56	52.94
DTU-Net	$WACV_{25}$	75.70	99.95	58.68	51.44	99.97	68.63	69.23	60.00	66.67	70.59
WDT-MD	Ours	**82.80**	**99.96**	**74.43**	**65.61**	**99.98**	**85.95**	**88.46**	**82.35**	**77.78**	**94.12**

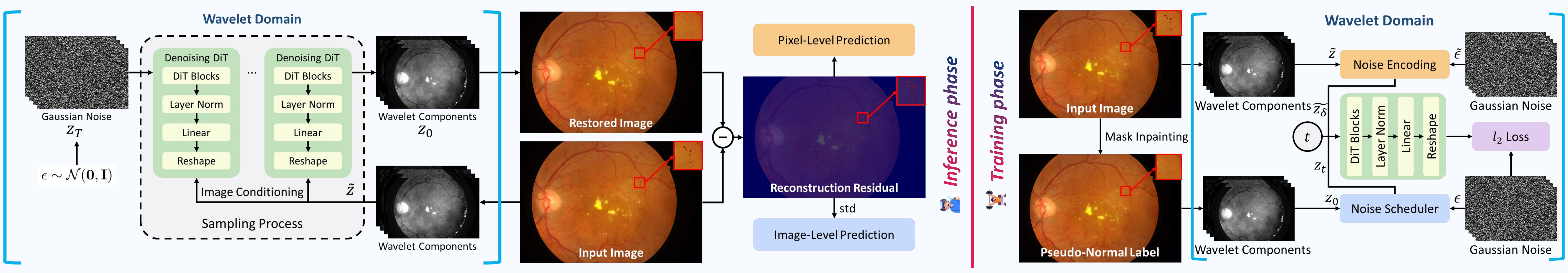
Table 2. Comparison of different methods on e-ophtha MA.

Method	Source	Pixel-level					Image-level				
		AUC	ACC	F1	SEN	SPE	AUC	ACC	F1	SEN	SPE
AnoDDPM	$CVPR_{22}$	79.26	99.96	32.09	58.56	99.97	60.00	61.54	51.61	53.33	66.67
CPC	$WACV_{23}$	76.28	99.98	35.34	52.57	99.98	65.42	66.67	58.06	60.00	70.83
DAE	$MedIA_{23}$	72.64	99.95	21.96	45.31	99.96	57.08	56.41	51.43	60.00	54.17
HACDR-Net	$AAAI_{24}$	54.13	98.03	3.53	9.56	98.04	41.67	43.59	31.25	33.33	50.00
AE[$d_{optimal}$]	$MICCAI_{24}$	78.86	99.98	32.21	57.75	99.98	62.08	64.10	53.33	53.33	70.83
Dif-fuse	TMI_{24}	80.82	99.96	32.48	61.67	99.96	61.25	61.54	54.55	60.00	62.50
GatingAno	PR_{24}	78.27	98.83	2.253	58.45	98.84	63.33	64.10	56.25	60.00	66.67
DTU-Net	$WACV_{25}$	80.72	99.98	42.99	61.46	99.98	49.58	48.72	44.44	53.33	45.83
WDT-MD	Ours	**81.08**	**99.99**	**57.70**	**62.16**	**99.99**	**70.83**	**71.79**	**64.52**	**66.67**	**75.00**

CHALLENGES

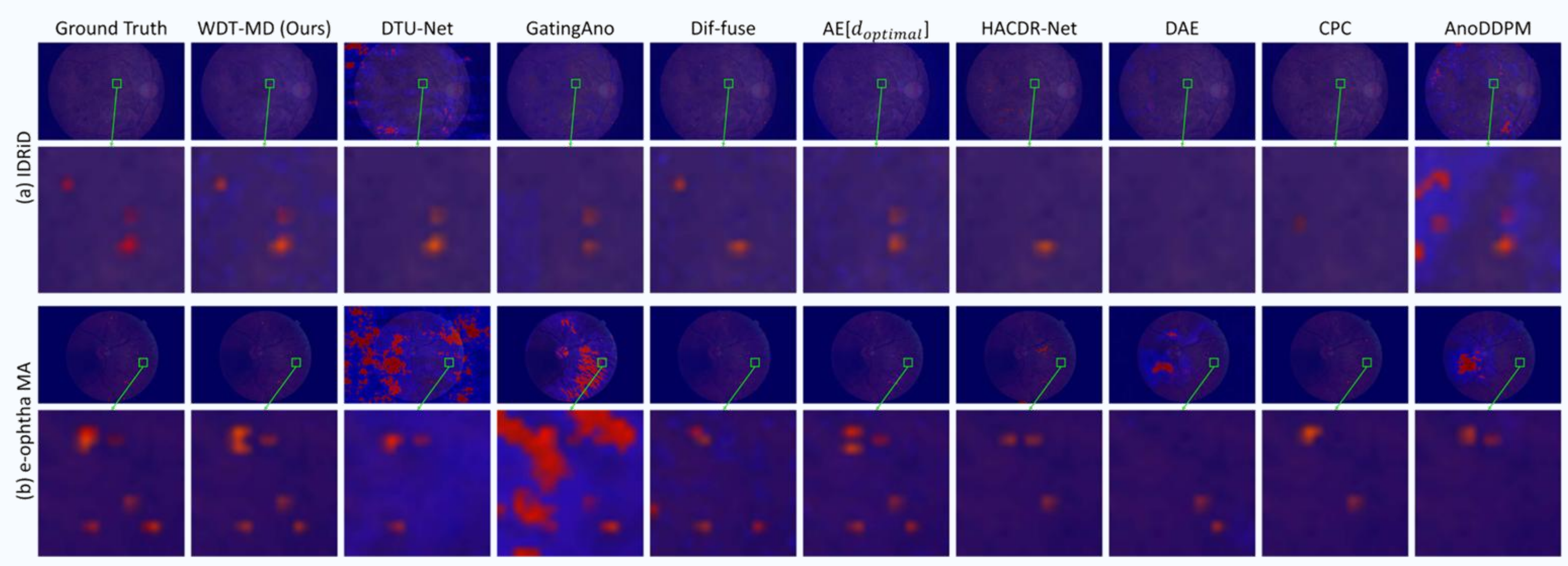
1. The inherent risk of learning "**identity mapping**" still persists in existing frameworks based on diffusion models.
2. The inability to distinguish MAs from other anomalies leads to **high false positives**, undermining clinical utility.
3. The **suboptimal reconstruction quality** of normal features hampers the performance of anomaly detection.

METHOD

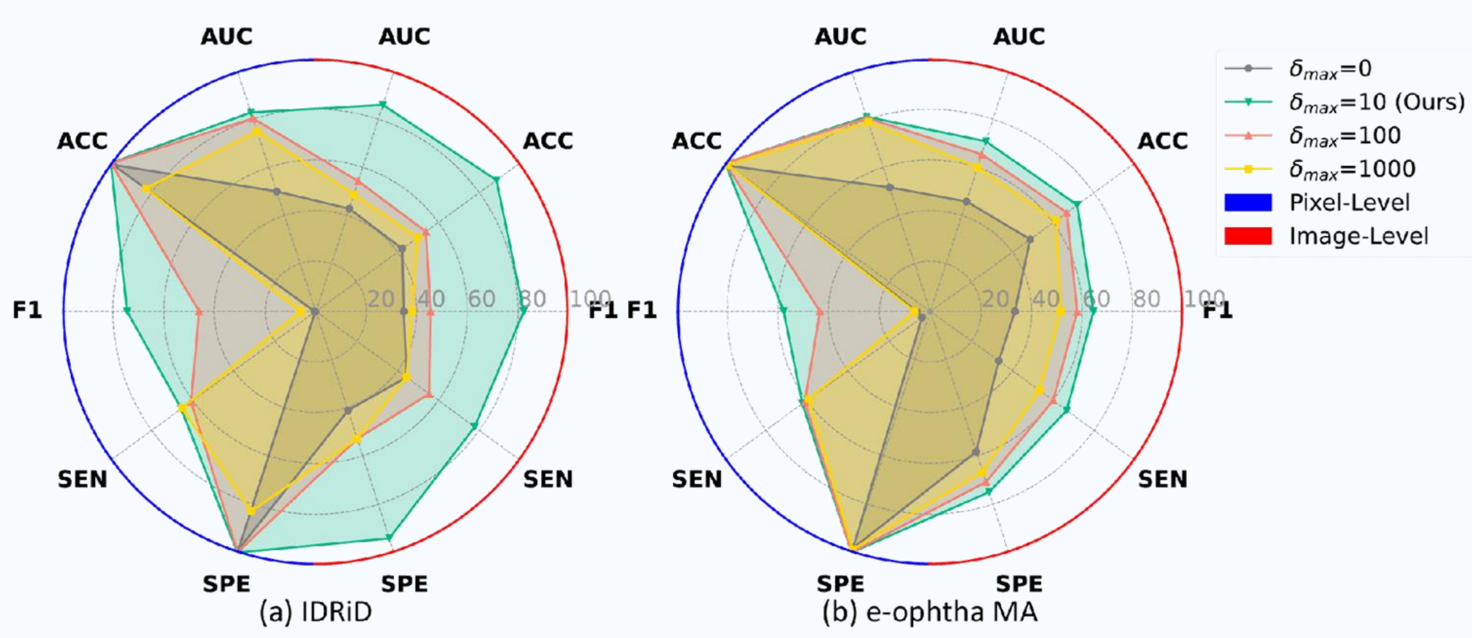


1. **Identity mapping** -> We propose a **noise-encoded image conditioning** mechanism for diffusion-based MA detection.
2. **High false positives** -> We introduce pixel-level supervision signals in training through **pseudo-normal pattern synthesis**.
3. **Suboptimal reconstruction quality** -> We propose a **wavelet diffusion Transformer** architecture.

QUALITATIVE RESULTS



HYPERPARAMETER ANALYSIS



Yifei Sun | Email: diaoquesang@gmail.com
Yuzhi He | Email: 23012100059@stu.xidian.edu.cn
Junhao Jia | Email: 23080631@hdu.edu.cn

arXiv | Code | Project Page | Survey | WeChat