



WavePolyp: Video Polyp Segmentation via Hierarchical Wavelet-Based Feature Aggregation and Inter-Frame Divergence Perception Class

Yuhua Zhang¹ Guilian Chen¹ Yuanqin He¹ Huisi Wu¹ Jing Qin²

¹Shenzhen University

²The Hong Kong Polytechnic University

<https://github.com/FishballZhang/WavePolyp>

Challenges

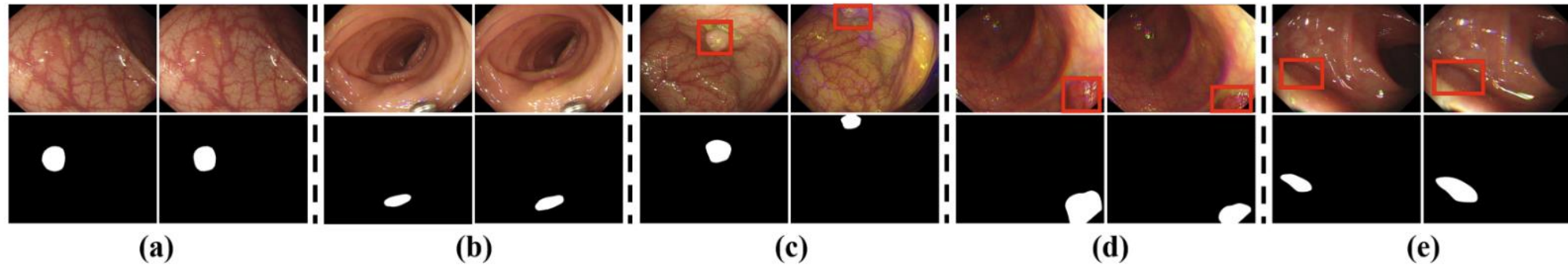
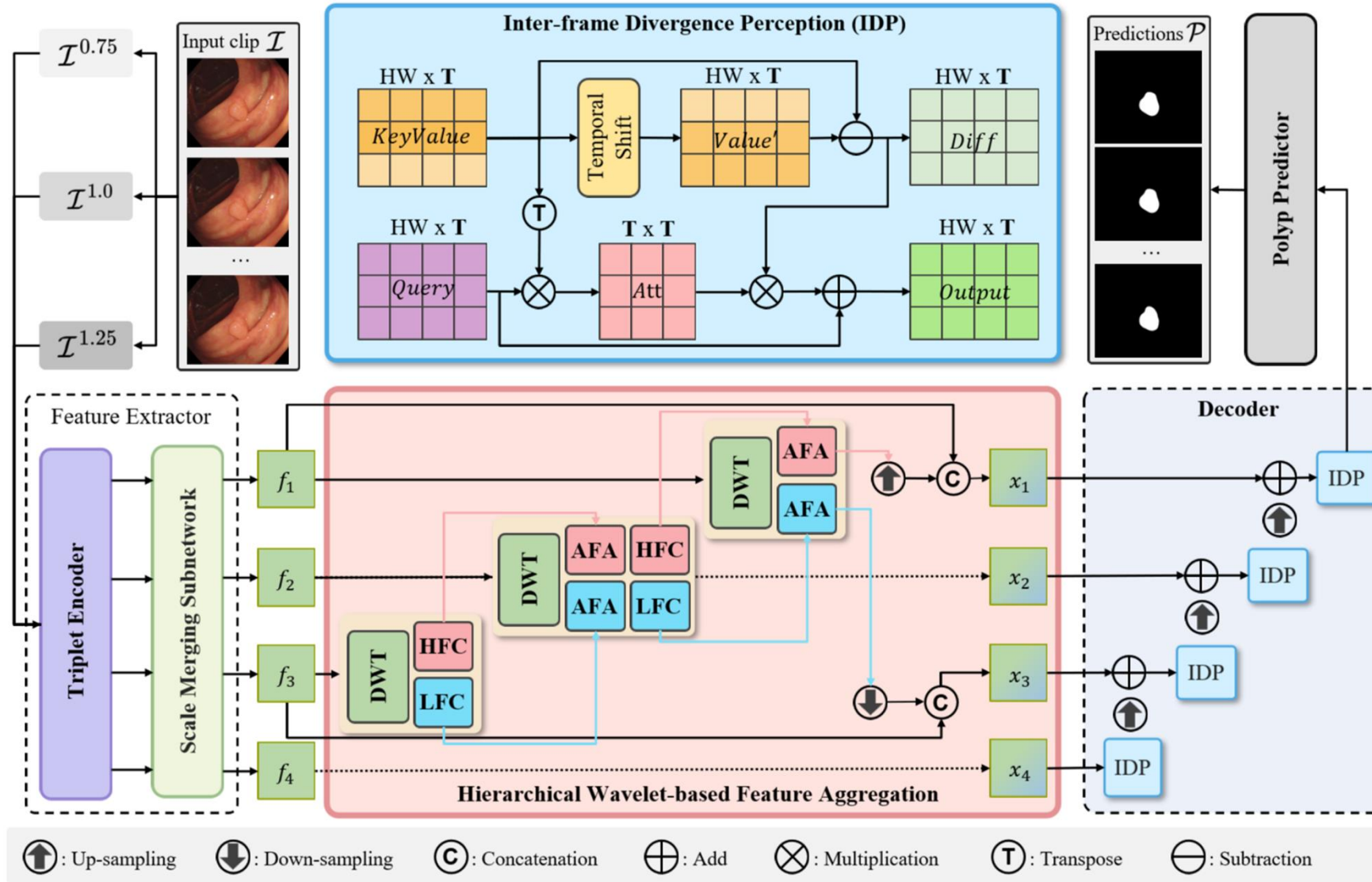


Figure 1: Challenges in VPS: (a) and (b) the high camouflage of polyps, (c) position variation, (d) shape variation, (e) size variation. Note that (a)-(e) represent two strictly adjacent frames.

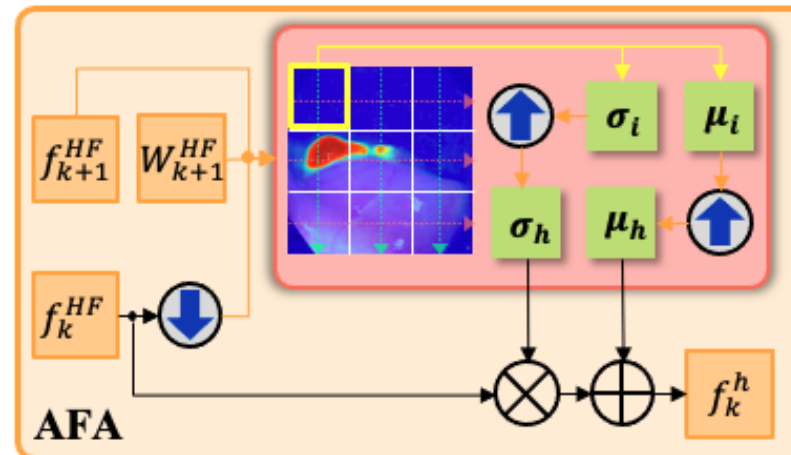
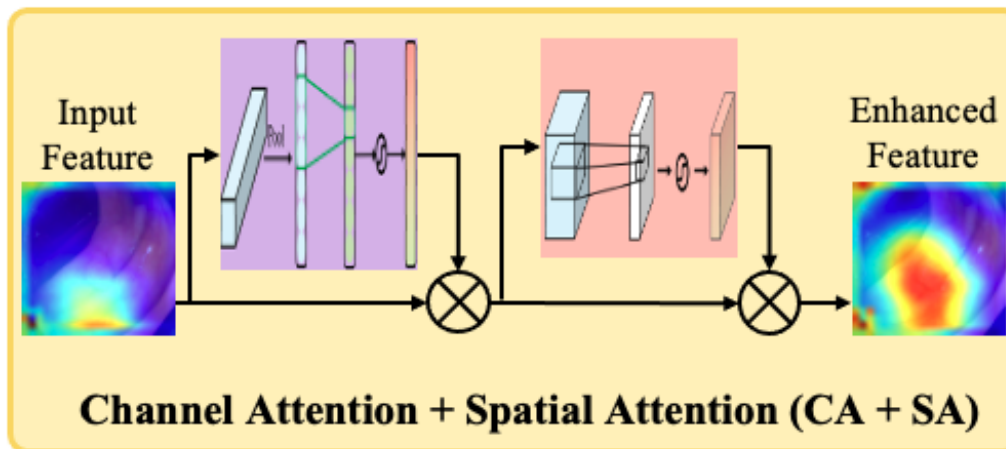
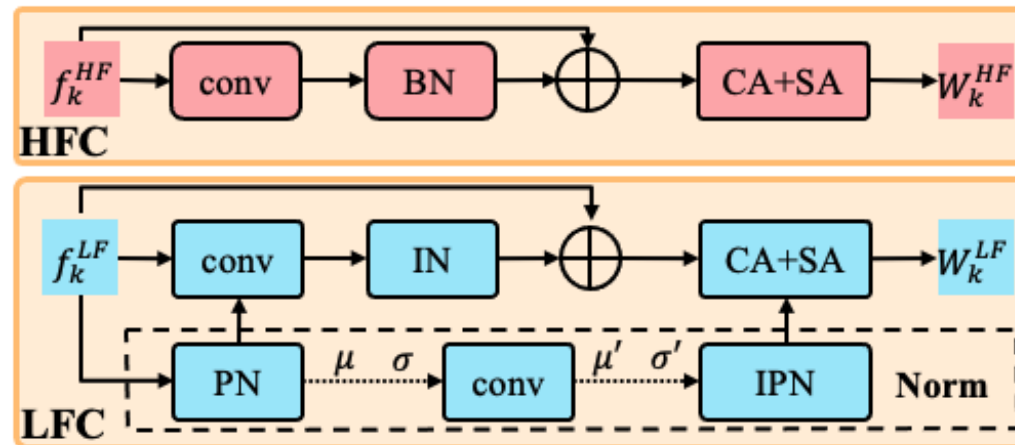
- Significant inter-frame divergence in videos
- High camouflage of polyps in normal colon structures
- Clinical requirement of real-time performance

Method



Method

- Details of the proposed HFC, LFC and AFA in HWFA.



Experiments

- Dataset

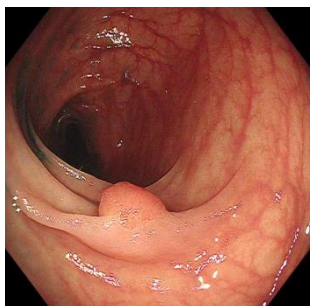
1. SUN-SEG

- training set (112 clips / 19, 544 frames), test set SUN-SEG-Easy (119 clips / 17,070 frames) and test set SUN-SEG-Hard (54 clips / 12,544 frames).

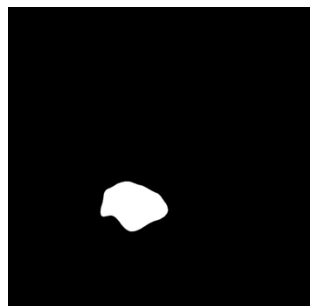
2. CVC-612

- Containing 612 frames from 31 colonoscopy clips.

Resizing each frame into 352×352 size



Image

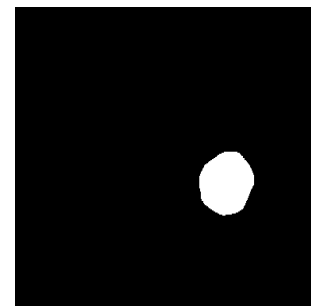


Label

SUN-SEG Dataset



Image

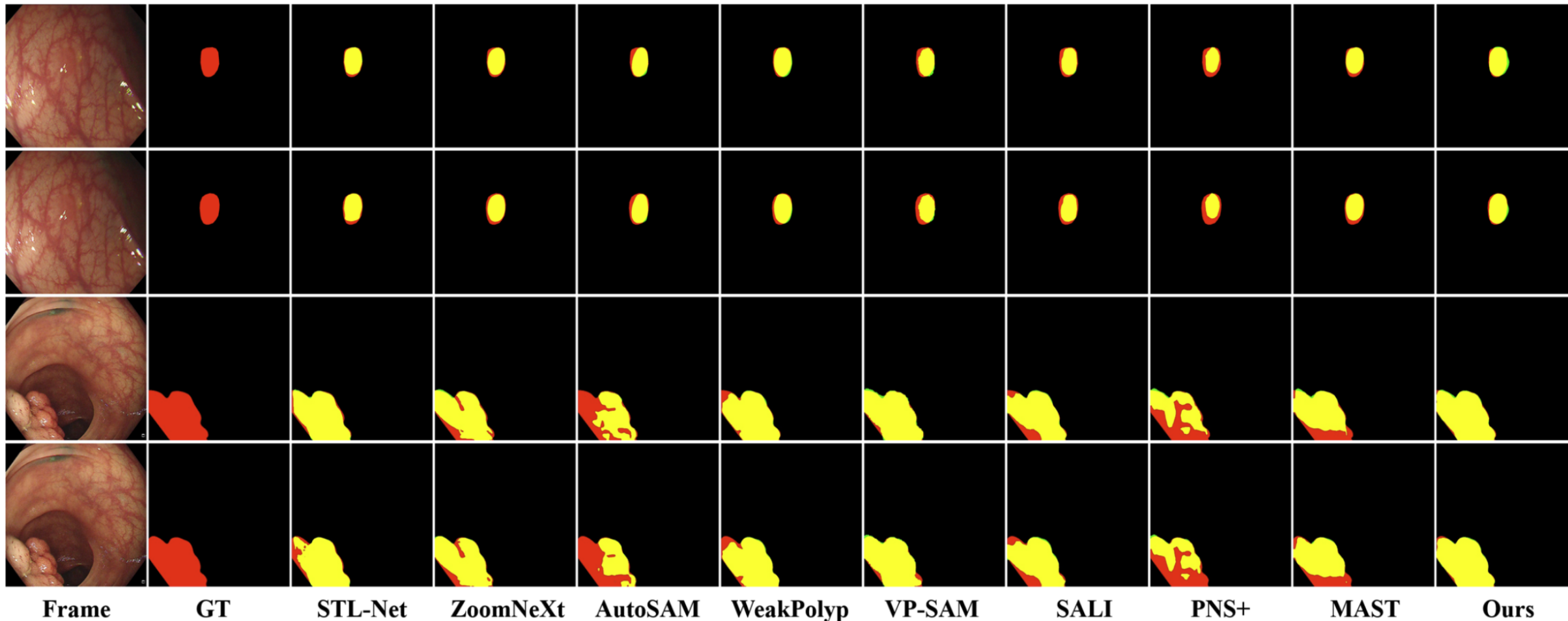


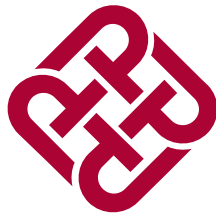
Label

CVC-612 Dataset

Experiments

- Visual Comparison





Experiments

- Comparison with SOTA Methods

Model	Backbone	SUN-SEG-Easy				SUN-SEG-Hard				CVC-612			
		S_α	E_ϕ^{mn}	F_β^w	$Dice$	S_α	E_ϕ^{mn}	F_β^w	$Dice$	S_α	E_ϕ^{mn}	F_β^w	$Dice$
SLT-Net	PVTv2-B5	90.39	<u>93.75</u>	84.35	87.15	88.06	92.05	80.31	83.55	<u>94.61</u>	97.70	92.06	92.96
ZoomNeXt	PVTv2-B5	89.81	92.25	84.64	87.55	88.51	91.34	82.21	85.22	94.54	97.53	90.91	92.73
AutoSAM	ViT-B	86.28	91.69	78.25	81.27	83.59	89.91	73.08	77.25	90.56	96.19	87.68	88.34
WeakPolyp	PVTv2-B5	<u>90.51</u>	93.72	<u>84.89</u>	<u>87.57</u>	<u>90.19</u>	<u>93.77</u>	<u>83.74</u>	<u>86.73</u>	91.51	95.18	88.74	89.07
PNS+	Res2Net-50	85.75	86.11	76.14	81.91	83.98	85.68	72.75	79.32	94.49	96.44	89.05	92.54
MAST	PVTv2-B2	87.91	92.87	81.40	84.44	87.44	92.82	80.27	83.79	93.54	96.07	89.93	90.12
SALI	PVTv2-B5	89.45	93.01	83.73	86.07	89.19	93.21	83.05	85.54	94.41	97.12	<u>92.16</u>	<u>93.01</u>
VP-SAM	ViT-B	90.49	93.68	85.64	88.19	90.03	93.74	83.30	86.94	94.68	97.83	92.51	93.45
Ours	PVTv2-B5	90.93	94.15	86.74	88.96	90.28	93.96	85.17	87.55	95.24	98.61	93.57	94.36



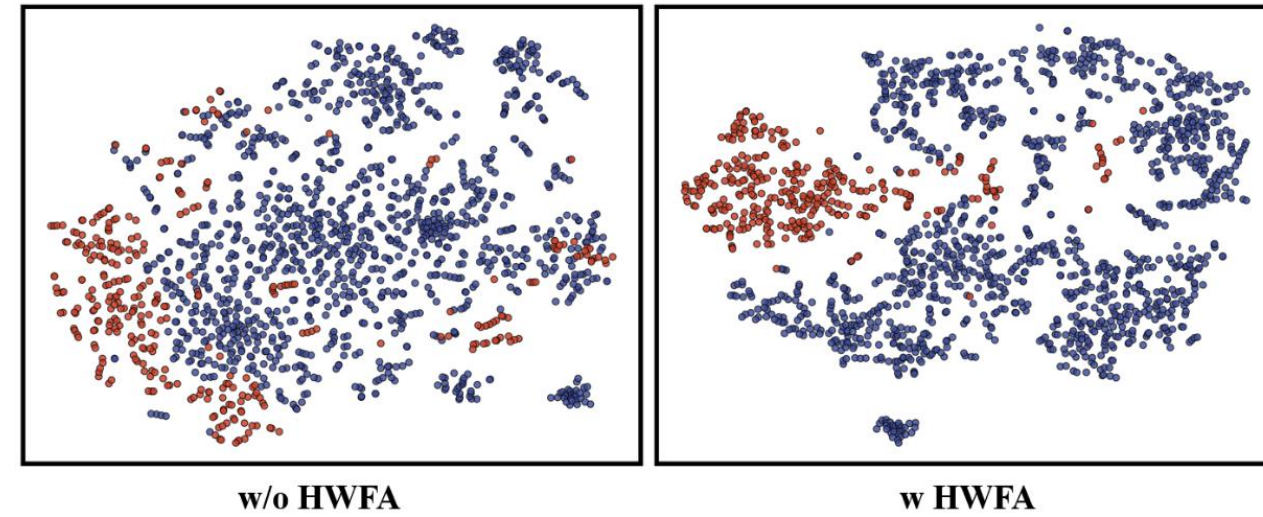
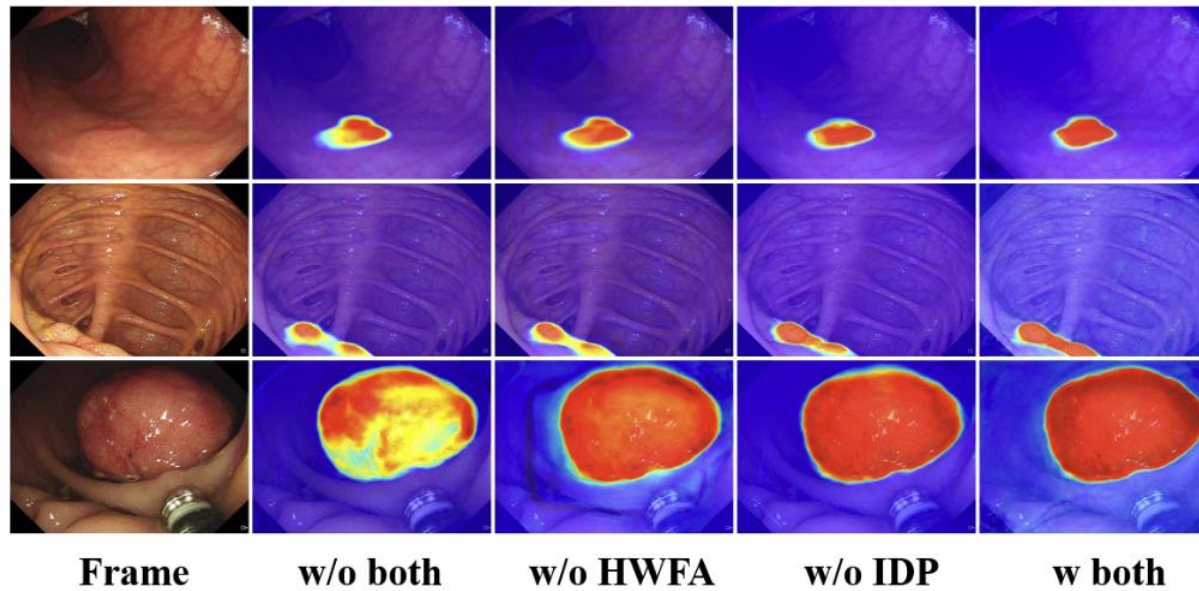
Experiments

- Ablations

HWFA				IDP			SUN-SEG-Easy		SUN-SEG-Hard	
HFC	LFC	Norm	AFA	Shift	IDA	IDD	S_α	Dice	S_α	Dice
							89.12	87.51	88.84	86.12
				✓	✓	✓	90.07	87.69	90.09	86.65
	✓	✓	✓	✓	✓	✓	90.31	88.06	89.75	87.12
✓			✓	✓	✓	✓	89.81	87.29	89.79	86.85
✓	✓		✓	✓	✓	✓	90.21	87.76	89.96	87.01
✓	✓	✓		✓	✓	✓	87.92	87.39	87.75	85.55
✓	✓	✓	✓				90.03	87.62	89.45	86.76
✓	✓	✓	✓		✓	✓	90.26	88.01	89.74	87.11
✓	✓	✓	✓	✓	✓		90.57	88.36	90.06	87.12
✓	✓	✓	✓	✓	✓	✓	90.93	88.96	90.28	87.55

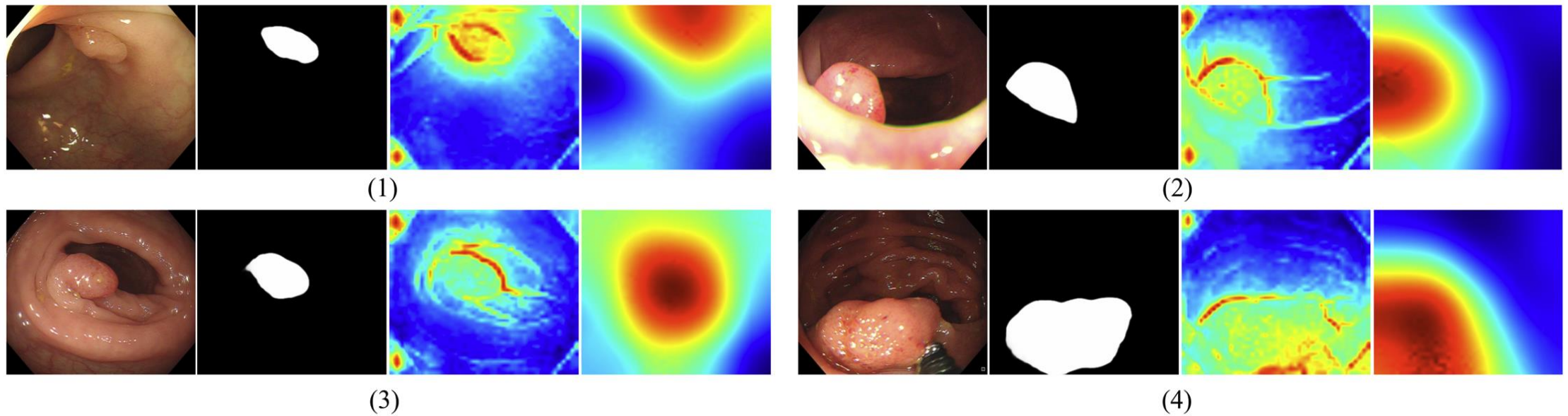
Experiments

- Ablations



Experiments

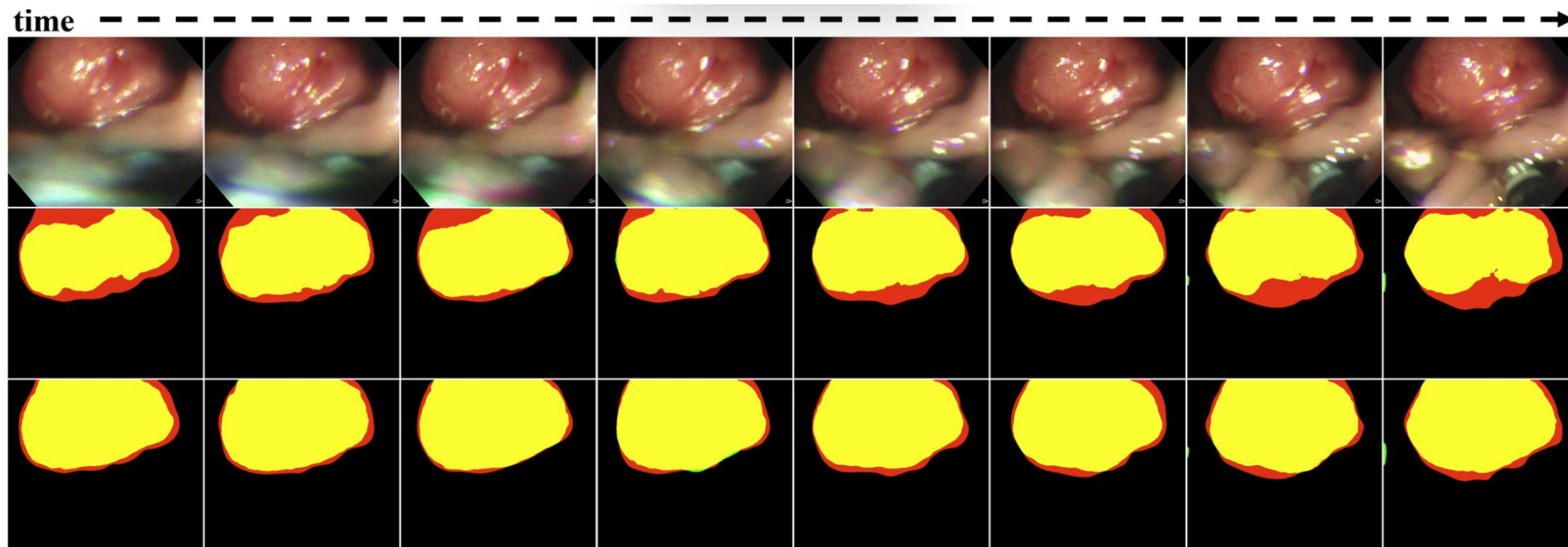
- Ablations



From left to right: (a) Original Frame, (b) Prediction Mask, (c) High-Frequency Features (f_1^{HF}), and (d) Low-Frequency Features (f_1^{LF}).

Experiments

- Ablations

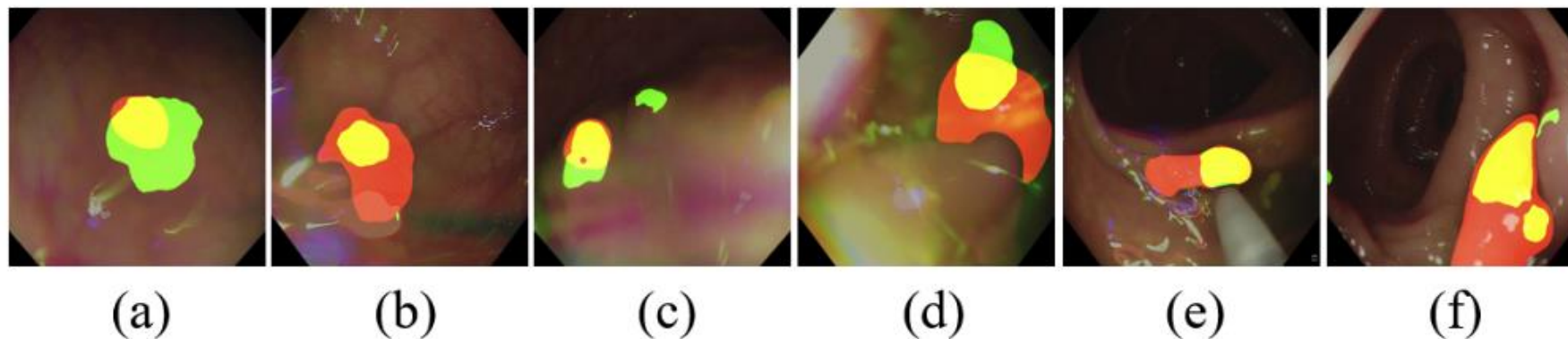


Row 2: w/o IDP; Row 3: Ours.

Conclusion

- WavePolyp
 - Excavation of intra-frame discriminative features
 - Perception of inter-frame divergence
 - State-of-the-art performance on two polyp video datasets

Limitation





THANKS FOR YOUR LISTENING

Please Contact Us with:

- Email: 2400101085@mails.szu.edu.cn
- Email: hswu@szu.edu.cn

See video in YouTube: <https://youtu.be/5R4hQIx86v4?si=xdiiRJSIhYVcFhZV>