Real-Time Polyp Detection in Colonoscopy using Lightweight Transformer: Supplementary Material

Youngbeom Yoo[1,*], Jae Young Lee[1,*], Dong-Jae Lee[1], Jiwoon Jeon[2], and Junmo Kim[1]
[1]School of Electrical Engineering, KAIST, South Korea
bum8552@hufs.ac.kr, {mcneato, jhtwosun, junmo.kim}@kaist.ac.kr
[2]AINex, South Korea jiwoon.jeon@ainex.io

1. Detection Results

Fig. 1 presents a comparison of inference result examples at an image size of 320 using the YOLOv5m [2] and the proposed YOLOv5m-TST models on the SUN dataset [6]. Fig. 2 presents a comparison of inference result examples at an image size of 320 using the YOLOv5m and the proposed YOLOv5m-TST models on the KUMC dataset [5]. Fig. 3 presents a comparison of inference result examples at an image size of 320 using the YOLOv5m and the proposed YOLOv5m-TST models on the Kvasir dataset [1]. We can observe that our proposed method demonstrates more robust performance compared

[*]equal contribution

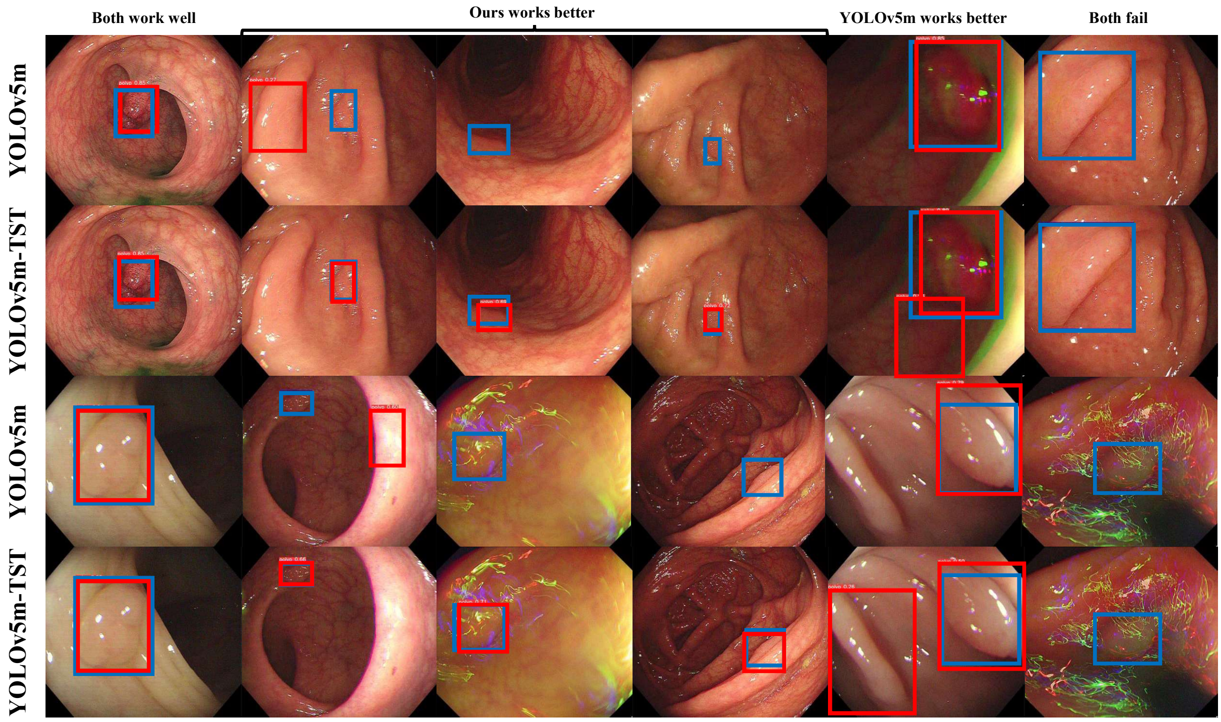

Figure 1. Detection results on the SUN dataset [6] of YOLOv5m [2] and YOLOv5m-TST (Ours) with an image size of 320. The red and blue boxes signify the prediction results and ground truth, respectively.

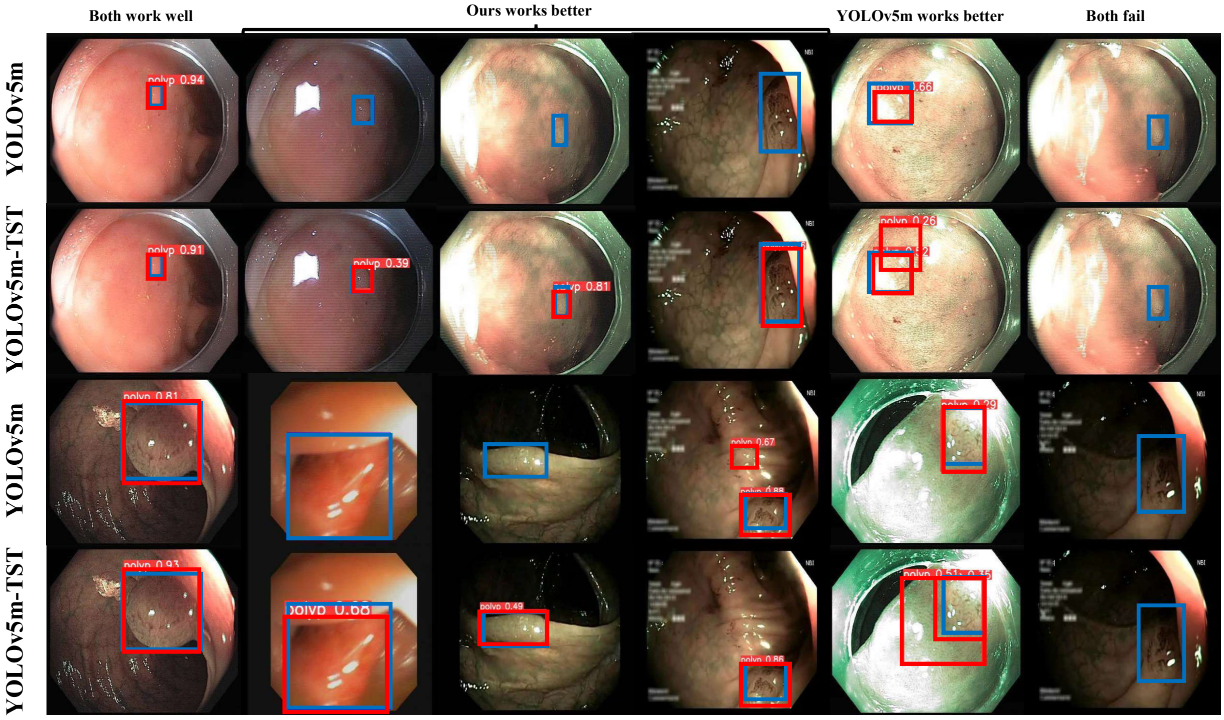

Figure 2. Detection results on the KUMC dataset [1] of YOLOv5m [2] and YOLOv5m-TST (Ours) with an image size of 320. The red and blue boxes signify the prediction results and ground truth, respectively.

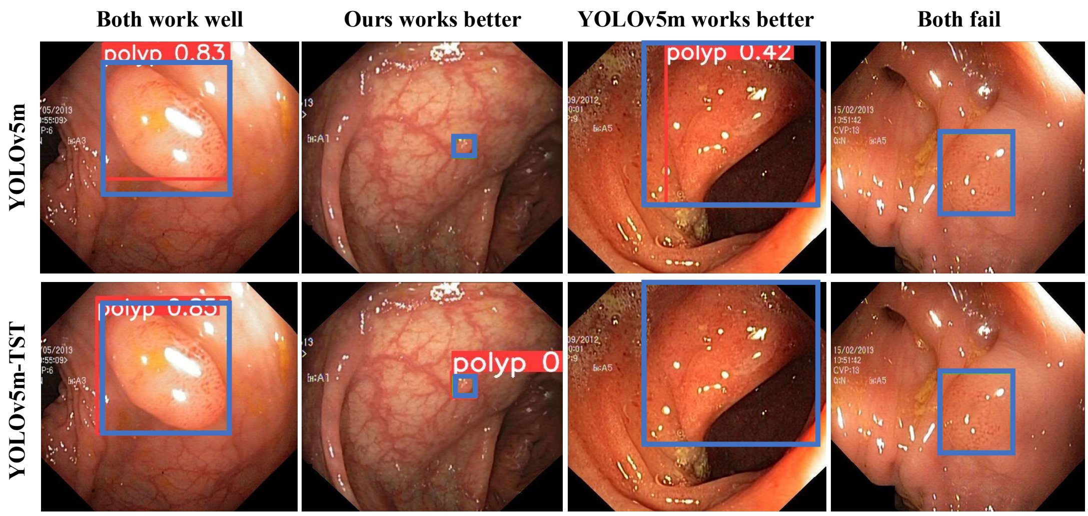

Figure 3. Detection results on the Kvasir dataset [5] of YOLOv5m [2] and YOLOv5m-TST (Ours) with an image size of 320. The red and blue boxes signify the prediction results and ground truth, respectively.

to the YOLOv5m in cases of small polyps, polyps with shadows, situations where the frame is dark, or when the polyps are partially occluded.

# Fold	Model	320						384						480					
		Precision	Recall	F1	mAP25	mAP50	mAP50:95	Precision	Recall	F1	mAP25	mAP50	mAP50:95	Precision	Recall	F1	mAP25	mAP50	mAP50:95
1	YOLOv5s	0.9580	**0.9028**	0.9295	**0.9518**	**0.9398**	**0.6001**	**0.9597**	**0.9124**	**0.9354**	**0.9545**	**0.9430**	0.6131	**0.9674**	**0.9080**	**0.9367**	**0.9555**	**0.9454**	**0.7300**
	YOLOv5s-TST	**0.9640**	0.9010	**0.9314**	0.9512	0.9381	0.5842	0.9502	0.8997	0.9243	0.9479	0.9347	**0.6766**	0.9425	0.8931	0.9171	0.9458	0.9294	0.6794
	YOLOv5m	**0.9636**	0.8941	0.9275	0.9457	0.9318	0.6010	**0.9686**	0.9028	0.9345	0.9494	0.9409	**0.6246**	**0.9715**	0.8972	0.9329	0.9513	0.9403	**0.6177**
	YOLOv5m-TST	0.9623	**0.9135**	**0.9373**	**0.9590**	**0.9488**	**0.6124**	0.9553	**0.9153**	**0.9349**	**0.9580**	**0.9486**	0.6130	0.9647	**0.9170**	**0.9403**	**0.9598**	**0.9512**	0.6164
	YOLOv5l	**0.9664**	0.8993	0.9316	0.9480	0.9384	**0.6220**	**0.9667**	0.9057	0.9352	0.9532	0.9442	**0.6368**	0.9561	**0.9216**	0.9386	**0.9634**	**0.9552**	**0.6351**
	YOLOv5l-TST	0.9663	**0.9100**	**0.9373**	**0.9613**	**0.9506**	0.6157	0.9659	**0.9223**	**0.9436**	**0.9653**	**0.9572**	0.6287	**0.9754**	0.9089	**0.9410**	0.9547	0.9451	0.6346
2	YOLOv5s	**0.9245**	0.8441	**0.8825**	**0.9203**	**0.8785**	**0.5361**	**0.9128**	0.8572	**0.8842**	**0.9242**	**0.8816**	**0.5493**	**0.9238**	0.8489	**0.8847**	**0.9245**	**0.8890**	**0.5505**
	YOLOv5s-TST	0.9092	**0.8546**	0.8810	0.9198	0.8710	0.5183	0.9109	**0.8579**	0.8836	0.9181	0.8716	0.5158	0.9016	**0.8569**	0.8787	0.9191	0.8673	0.5190
	YOLOv5m	**0.9316**	0.8309	0.8784	0.9069	0.8672	**0.5334**	**0.9334**	0.8222	0.8743	0.9064	0.8670	**0.5415**	**0.9341**	0.8590	**0.8950**	**0.9301**	**0.8897**	**0.5569**
	YOLOv5m-TST	0.9244	**0.8607**	**0.8914**	**0.9256**	**0.8777**	0.5262	0.9130	**0.8686**	**0.8902**	**0.9285**	**0.8897**	0.5406	0.9246	**0.8622**	0.8923	0.9280	0.8877	0.5482
	YOLOv5l	**0.9193**	0.8443	0.8802	0.9117	0.8770	**0.5500**	0.9284	0.8356	0.8795	0.9101	0.8683	**0.5543**	**0.9337**	0.8359	0.8821	0.9176	0.8786	**0.5608**
	YOLOv5l-TST	0.9180	**0.8658**	**0.8911**	**0.9290**	**0.8852**	0.5398	**0.9331**	**0.8670**	**0.8989**	**0.9333**	**0.8891**	0.5456	0.9126	**0.8780**	**0.8950**	**0.9389**	**0.8941**	0.5467
3	YOLOv5s	**0.9157**	0.8508	0.8821	**0.9245**	**0.8660**	**0.5654**	**0.9342**	**0.8603**	**0.8957**	**0.9348**	**0.8832**	**0.5351**	**0.9198**	**0.8608**	**0.8894**	**0.9279**	**0.8806**	**0.5411**
	YOLOv5s-TST	0.9075	**0.8615**	**0.8839**	0.9221	0.8646	0.5015	0.9022	0.8475	0.8740	0.9178	0.8690	0.5136	0.8844	0.8524	0.8681	0.9150	0.8661	0.5044
	YOLOv5m	0.9158	0.8538	0.8837	0.9226	0.8757	**0.5386**	**0.9229**	0.8542	**0.8872**	0.9173	0.8688	**0.5316**	**0.9326**	0.8491	**0.8889**	0.9261	0.8715	0.5407
	YOLOv5m-TST	**0.9257**	**0.8736**	**0.8989**	**0.9330**	**0.8866**	0.5288	0.8993	**0.8657**	0.8822	**0.9241**	**0.8768**	0.5310	0.9137	**0.8634**	0.8878	**0.9297**	**0.8918**	**0.5477**
	YOLOv5l	**0.9361**	0.8600	**0.8965**	0.9331	**0.8872**	**0.5551**	**0.9352**	0.8667	0.8997	0.9352	0.8774	**0.5440**	**0.9349**	0.8628	**0.8974**	**0.9310**	0.8687	**0.5500**
	YOLOv5l-TST	0.9299	**0.8645**	0.8960	**0.9351**	0.8806	0.5355	0.9242	**0.8806**	**0.9019**	**0.9412**	**0.8876**	0.5390	0.8974	**0.8692**	0.8831	0.9290	**0.8854**	0.5446
4	YOLOv5s	**0.9103**	0.8181	0.8617	**0.8956**	**0.8527**	**0.5085**	**0.9280**	**0.8194**	**0.8703**	**0.9109**	**0.8680**	**0.5247**	**0.9278**	**0.8302**	**0.8763**	**0.9148**	**0.8739**	**0.5329**
	YOLOv5s-TST	0.9079	**0.8249**	**0.8644**	0.8938	0.8457	0.4918	0.9041	0.8005	0.8492	0.8830	0.8387	0.4912	0.9004	0.8042	0.8495	0.8927	0.8436	0.4895
	YOLOv5m	0.9170	0.7878	0.8475	0.8811	0.8372	0.5166	**0.9240**	0.8017	0.8586	0.8928	0.8554	0.5197	0.9136	0.8048	0.8558	0.8952	0.8605	0.5328
	YOLOv5m-TST	**0.9193**	**0.8294**	**0.8720**	**0.9108**	**0.8717**	**0.5206**	0.9238	**0.8529**	**0.8869**	**0.9200**	**0.8822**	**0.5341**	**0.9186**	**0.8420**	**0.8786**	**0.9192**	**0.8810**	**0.5349**
	YOLOv5l	0.9023	0.7981	0.8470	0.8847	0.8472	0.5207	0.9201	0.8120	0.8627	0.8954	0.8552	0.5315	**0.9309**	0.8215	0.8728	0.9022	0.8642	**0.5420**
	YOLOv5l-TST	**0.9173**	**0.8379**	**0.8758**	**0.9209**	**0.8734**	**0.5268**	**0.9290**	**0.8560**	**0.8910**	**0.9250**	**0.8825**	**0.5356**	0.9225	**0.8399**	**0.8793**	**0.9159**	**0.8774**	0.5358
5	YOLOv5s	**0.9389**	0.8202	0.8756	**0.9112**	0.8630	0.5220	**0.9438**	0.8171	**0.8759**	**0.9109**	**0.8687**	**0.5437**	**0.9304**	**0.8233**	**0.8736**	**0.9103**	**0.8701**	**0.5474**
	YOLOv5s-TST	0.9282	**0.8363**	**0.8799**	0.9061	**0.8673**	**0.5256**	0.9343	**0.8181**	0.8724	0.9007	0.8587	0.5253	0.9203	0.8126	0.8631	0.8962	0.8493	0.5149
	YOLOv5m	0.9231	0.7978	0.8559	0.8846	0.8443	0.5303	**0.9533**	0.8065	0.8738	0.8929	0.8572	**0.5424**	0.9388	0.8135	0.8717	**0.9050**	**0.8625**	0.5471
	YOLOv5m-TST	**0.9364**	**0.8365**	**0.8837**	**0.9076**	**0.8661**	**0.5348**	0.9366	**0.8285**	**0.8792**	**0.9062**	**0.8610**	0.5354	**0.9544**	**0.8168**	**0.8802**	0.8957	0.8567	**0.5538**
	YOLOv5l	**0.9419**	**0.8228**	**0.8783**	**0.9005**	**0.8639**	**0.5468**	0.9289	0.8167	0.8692	0.8959	0.8668	0.5477	**0.9434**	0.8116	0.8725	0.8959	0.8642	**0.5572**
	YOLOv5l-TST	0.9404	0.8151	0.8733	0.9002	0.8610	0.5413	**0.9296**	**0.8295**	**0.8767**	**0.9125**	**0.8695**	**0.5504**	0.9388	**0.8561**	**0.8956**	**0.9235**	**0.8771**	0.5543
Avg.	YOLOv5s	**0.9295**	0.8472	0.8863	**0.9207**	**0.8800**	**0.5464**	**0.9357**	**0.8533**	**0.8923**	**0.9271**	**0.8889**	**0.5532**	**0.9338**	**0.8542**	**0.8921**	**0.9266**	**0.8918**	**0.5804**
	YOLOv5s-TST	0.9234	**0.8557**	**0.8881**	0.9186	0.8773	0.5243	0.9203	0.8447	0.8807	0.9135	0.8745	0.5445	0.9098	0.8438	0.8753	0.9138	0.8711	0.5414
	YOLOv5m	0.9302	0.8329	0.8786	0.9082	0.8712	0.5440	**0.9404**	0.8375	0.8857	0.9118	0.8779	**0.5520**	**0.9381**	0.8447	0.8889	0.9215	0.8849	0.5590
	YOLOv5m-TST	**0.9336**	**0.8627**	**0.8967**	**0.9272**	**0.8902**	**0.5446**	0.9256	**0.8662**	**0.8947**	**0.9274**	**0.8917**	0.5508	0.9352	**0.8603**	**0.8958**	**0.9265**	**0.8937**	**0.5602**
	YOLOv5l	0.9332	0.8449	0.8867	0.9156	0.8827	**0.5589**	0.9359	0.8473	0.8893	0.9180	0.8824	**0.5629**	**0.9398**	0.8507	0.8927	0.9220	0.8862	**0.5690**
	YOLOv5l-TST	**0.9344**	**0.8587**	**0.8947**	**0.9293**	**0.8902**	0.5518	**0.9364**	**0.8711**	**0.9024**	**0.9355**	**0.8972**	0.5599	0.9293	**0.8704**	**0.8988**	**0.9324**	**0.8958**	0.5632

Table 1. Comparison of the performance between our proposed method and various sizes of the YOLOv5 [2] models. Each model is evaluated on the combined datasets of SUN [6], KUMC [5], and Kvasir [1]. The best scores are bold-faced.

# Fold	Model	SUN						KUMC						Kvasir					
		Precision	Recall	F1	mAP25	mAP50	mAP50:95	Precision	Recall	F1	mAP25	mAP50	mAP50:95	Precision	Recall	F1	mAP25	mAP50	mAP50:95
1	YOLOv5m	0.9684	0.9030	0.9346	0.9516	0.9416	**0.6043**	0.9543	**0.8786**	0.9149	0.9337	0.9185	0.6319	0.9520	0.8894	0.9197	0.9430	0.9182	0.6748
	ENDOMIND	0.9679	0.8846	0.9244	0.9496	0.9358	0.5849	0.9380	0.8621	0.8985	0.9240	0.9077	0.6194	**0.9651**	0.8894	**0.9257**	**0.9557**	0.9213	0.6835
	YOLOv5m-Tiny	0.9476	0.8852	0.9153	0.9470	0.9351	0.5658	0.9538	0.8700	0.9100	0.9282	0.9121	0.6247	0.9440	0.8918	0.9172	0.9500	**0.9227**	**0.6883**
	YOLOv5m-TST	**0.9712**	**0.9286**	**0.9494**	**0.9737**	**0.9659**	0.5986	**0.9607**	0.8785	**0.9178**	**0.9349**	**0.9219**	**0.6366**	0.9038	**0.8990**	0.9014	0.9485	0.9185	0.6819
2	YOLOv5m	**0.9124**	0.8381	0.8737	0.9149	0.8779	**0.5160**	0.8914	0.8298	0.8595	0.9132	0.8713	0.5697	0.9119	0.8738	0.8925	0.9282	0.8862	0.6233
	ENDOMIND	0.8999	0.8605	**0.8798**	0.9149	0.8726	0.4981	0.9013	0.8013	0.8484	0.8900	0.8360	0.5298	0.9053	**0.8937**	0.8995	0.9328	0.9054	0.6562
	YOLOv5m-Tiny	0.9193	0.8644	0.8910	0.9233	0.8994	0.5382	0.8821	**0.8544**	**0.8680**	**0.9288**	**0.8720**	**0.5813**	0.8962	0.8875	0.8918	0.9322	0.8879	0.6469
	YOLOv5m-TST	0.9120	**0.8402**	0.8746	**0.9200**	**0.8827**	0.4982	**0.9038**	0.8207	0.8603	0.9152	0.8425	0.5373	**0.9589**	0.8726	**0.9137**	**0.9338**	**0.9157**	**0.6739**
3	YOLOv5m	0.9036	0.8096	0.8540	0.9071	0.8617	**0.5241**	0.9238	0.8318	0.8754	0.9149	0.8677	0.5561	0.9386	0.8784	0.9075	0.9264	0.8889	0.6596
	ENDOMIND	0.8888	0.8186	0.8522	0.8897	0.8178	0.4864	**0.9514**	0.7950	0.8662	0.8869	0.8466	0.5346	**0.9511**	0.8770	0.9126	0.9316	0.8885	0.6659
	YOLOv5m-Tiny	0.8642	0.8132	0.8379	0.8950	0.8364	0.5153	0.9254	0.8028	0.8598	0.8986	0.8643	**0.5688**	0.9088	0.8829	0.8957	0.9314	0.8906	0.6556
	YOLOv5m-TST	**0.9247**	**0.8646**	**0.8936**	**0.9356**	**0.8783**	0.5158	0.9447	**0.8606**	**0.9007**	**0.9367**	**0.8983**	0.5672	0.9413	**0.8874**	**0.9135**	**0.9426**	**0.9175**	**0.7037**
4	YOLOv5m	**0.9238**	0.7762	0.8436	0.8597	0.8231	0.5116	0.8882	**0.8156**	0.8503	0.8916	0.8408	**0.5015**	0.9410	**0.9245**	**0.9327**	0.9506	0.9004	0.6661
	ENDOMIND	0.9051	0.7702	0.8322	0.8611	0.8158	0.5054	0.8972	0.7811	0.8351	0.8716	0.7984	0.4823	0.9303	0.8868	0.9080	0.9480	0.9053	0.6694
	YOLOv5m-Tiny	0.9228	**0.8139**	**0.8649**	**0.9042**	**0.8593**	**0.5339**	0.8609	0.7991	0.8289	0.8808	0.8188	0.5075	0.9214	0.8849	0.9028	0.9452	0.9018	0.6640
	YOLOv5m-TST	0.9189	0.8032	0.8572	0.8908	0.8398	0.4982	**0.9181**	0.8124	**0.8621**	**0.9019**	**0.8512**	0.4947	**0.9490**	0.8915	0.9194	**0.9510**	**0.9112**	**0.6723**
5	YOLOv5m	0.8831	0.7751	0.8256	0.8574	0.8124	0.4422	0.9606	0.8779	0.9174	0.9403	0.9194	0.6308	0.9181	0.8868	0.9022	0.9370	0.9005	0.6916
	ENDOMIND	0.8795	0.7631	0.8172	0.8473	0.7871	0.4124	0.9549	0.8790	0.9154	0.9403	0.9158	0.6169	0.9140	0.8896	0.9016	0.9367	0.9033	0.6897
	YOLOv5m-Tiny	0.8779	**0.7913**	**0.8324**	**0.8727**	**0.8262**	**0.4469**	**0.9732**	**0.9055**	**0.9381**	**0.9579**	**0.9370**	**0.6573**	0.8654	0.8975	0.8812	0.9280	0.9011	0.6812
	YOLOv5m-TST	**0.8849**	0.7678	0.8222	0.8587	0.8013	0.4337	0.9663	0.9048	0.9345	0.9543	0.9270	0.6283	**0.9315**	**0.9070**	**0.9191**	**0.9383**	**0.9121**	**0.6957**
Avg.	YOLOv5m	0.9183	0.8204	0.8663	0.8981	0.8633	0.5196	0.9236	0.8467	0.8835	0.9187	0.8835	0.5780	0.9323	0.8906	0.9109	0.9370	0.8988	0.6631
	ENDOMIND	0.9082	0.8194	0.8612	0.8925	0.8458	0.4974	0.9286	0.8237	0.8727	0.9026	0.8609	0.5566	0.9332	0.8873	0.9095	0.9410	0.9048	0.6729
	YOLOv5m-Tiny	0.9064	0.8336	0.8683	0.9084	0.8713	**0.5200**	0.9191	0.8464	0.8809	0.9188	0.8808	**0.5879**	0.9072	0.8889	0.8977	0.9374	0.9008	0.6672
	YOLOv5m-TST	**0.9223**	**0.8409**	**0.8794**	**0.9157**	**0.8736**	0.5089	**0.9387**	**0.8554**	**0.8951**	**0.9286**	**0.8881**	0.5728	**0.9369**	**0.8915**	**0.9134**	**0.9428**	**0.9150**	**0.6855**

Table 2. Comparison on SUN [6], KUMC [5], and Kvasir [1] datasets with 320 $\times$ 320. For each dataset, comparisons among YOLOv5m, YOLOv5m-TST, YOLOv5m-ENDOMIND (ENDOMIND) [4], and modified Polyp-YOLOv5-Tiny (YOLOv5m-Tiny) [7] are presented. The best scores are bold-faced.

2. 5-Fold Cross-Validation Results

Tab. 1 shows comparison between our proposed method and various sizes of the YOLOv5 models for all experiments of 5-fold cross-validation on combined datasets. Tab. 2 shows comparison between YOLOv5m, YOLOv5m-TST, YOLOv5m-ENDOMIND (shortly ENDOMIND) [4], and modified Polyp-YOLOv5-Tiny (shortly YOLOv5m-Tiny) [7] for all experiments of 5-fold cross-validation on SUN, KUMC, and Kvasir datasets. The original Polyp-YOLOv5-Tiny utilized YOLOv5s and their source code is not publicly available. However, for the purpose of conducting a fair comparison in our experiments, we re-implemented it to align with the YOLOv5m model and conducted our experiments accordingly. Tab. 3 shows comparison between our proposed method applied to YOLOv3 [8] and YOLOv8m [3] for all experiments of 5-fold cross-validation on combined datasets. Tab. 4 shows comparison between our proposed method with varying key dimension sizes and the

# Fold	Model	320						384						480					
		Precision	Recall	F1	mAP25	mAP50	mAP50:95	Precision	Recall	F1	mAP25	mAP50	mAP50:95	Precision	Recall	F1	mAP25	mAP50	mAP50:95
1	YOLOv3	0.9652	0.9316	0.9481	0.9743	0.9665	**0.6782**	**0.9756**	0.9354	**0.9551**	0.9767	**0.9701**	**0.6928**	**0.9705**	0.9365	**0.9532**	**0.9775**	**0.9704**	**0.6944**
	YOLOv3-TST	**0.9667**	**0.9383**	**0.9523**	**0.9761**	**0.9684**	0.6750	0.9665	**0.9439**	0.9551	**0.9778**	0.9693	0.6768	0.9664	**0.9371**	0.9515	0.9733	0.9667	0.6701
	YOLOv8m	0.9616	**0.9348**	**0.9480**	**0.9728**	**0.9661**	**0.6774**	0.9650	0.9232	0.9437	0.9717	0.9642	**0.6845**	**0.9689**	0.9167	0.9421	0.9664	0.9606	0.6820
	YOLOv8m-TST	**0.9691**	0.9238	0.9459	0.9706	0.9636	0.6664	**0.9666**	**0.9374**	**0.9518**	**0.9760**	**0.9669**	0.6742	0.9666	**0.9361**	**0.9511**	**0.9760**	**0.9680**	**0.6824**
	YOLOv5m	**0.9636**	0.8941	0.9275	0.9457	0.9318	0.6010	**0.9686**	0.9028	0.9345	0.9494	0.9409	**0.6246**	**0.9715**	0.8972	0.9329	0.9513	0.9403	**0.6177**
	YOLOv5m-TST	0.9623	**0.9135**	**0.9373**	**0.9590**	**0.9488**	**0.6124**	0.9553	**0.9153**	**0.9349**	**0.9580**	**0.9486**	0.6130	0.9647	**0.9170**	**0.9403**	**0.9598**	**0.9512**	0.6164
2	YOLOv3	0.9214	**0.8855**	**0.9031**	0.9488	0.9103	**0.6034**	**0.9298**	**0.9056**	**0.9176**	**0.9593**	**0.9227**	**0.6134**	**0.9298**	**0.8954**	**0.9123**	**0.9586**	**0.9235**	**0.6130**
	YOLOv3-TST	**0.9262**	0.8768	0.9008	**0.9496**	**0.9114**	0.5891	0.9259	0.8807	0.9027	0.9446	0.9057	0.5915	0.9219	0.8896	0.9055	0.9504	0.9131	0.5971
	YOLOv8m	0.9189	0.8550	0.8858	0.9348	**0.8991**	**0.6025**	**0.9308**	0.8494	0.8883	0.9305	0.8955	**0.6053**	**0.9362**	0.8672	0.9004	0.9420	**0.9117**	**0.6090**
	YOLOv8m-TST	**0.9244**	**0.8748**	**0.8989**	**0.9422**	0.8972	0.5752	0.9219	**0.8868**	**0.9040**	**0.9481**	**0.9096**	0.5915	0.9230	**0.8888**	**0.9056**	**0.9524**	0.9116	0.5907
	YOLOv5m	**0.9316**	0.8309	0.8784	0.9069	0.8672	**0.5334**	**0.9334**	0.8222	0.8743	0.9064	0.8670	**0.5415**	**0.9341**	0.8590	**0.8950**	**0.9301**	**0.8897**	**0.5569**
	YOLOv5m-TST	0.9244	**0.8607**	**0.8914**	**0.9256**	**0.8777**	0.5262	0.9130	**0.8686**	**0.8902**	**0.9285**	**0.8897**	0.5406	0.9246	**0.8622**	0.8923	0.9280	0.8877	0.5482
3	YOLOv3	**0.9136**	**0.9097**	**0.9117**	0.9531	0.9144	**0.6078**	**0.9208**	**0.9129**	**0.9168**	0.9549	0.9128	0.6063	**0.9351**	0.8819	0.9077	0.9513	0.9228	**0.6282**
	YOLOv3-TST	0.9117	0.8976	0.9046	**0.9541**	**0.9185**	0.5970	0.9108	0.9075	0.9092	**0.9558**	**0.9203**	**0.6127**	0.9225	**0.9066**	**0.9145**	**0.9577**	**0.9230**	0.6098
	YOLOv8m	**0.9134**	0.8778	0.8952	0.9382	0.8940	**0.5931**	0.9155	**0.8903**	0.9028	0.9483	0.9057	**0.6047**	0.9148	0.8827	0.8985	**0.9494**	**0.9109**	**0.6052**
	YOLOv8m-TST	0.9112	**0.8890**	**0.9000**	**0.9492**	**0.9107**	0.5904	**0.9214**	0.8860	**0.9034**	**0.9511**	**0.9100**	0.5952	**0.9148**	**0.8896**	**0.9020**	0.9457	0.9062	0.5976
	YOLOv5m	0.9158	0.8538	0.8837	0.9226	0.8757	**0.5386**	**0.9229**	0.8542	**0.8872**	0.9173	0.8688	**0.5316**	**0.9326**	0.8491	**0.8889**	0.9261	0.8715	0.5407
	YOLOv5m-TST	**0.9257**	**0.8736**	**0.8989**	**0.9330**	**0.8866**	0.5288	0.8993	**0.8657**	0.8822	**0.9241**	**0.8768**	0.5310	0.9137	**0.8634**	0.8878	**0.9297**	**0.8918**	**0.5477**
4	YOLOv3	0.9203	0.8336	0.8748	0.9205	0.8987	**0.5944**	**0.9179**	0.8315	0.8726	0.9235	0.8990	**0.5874**	**0.9239**	0.8451	0.8827	0.9334	0.9105	**0.5988**
	YOLOv3-TST	**0.9204**	**0.8711**	**0.8951**	**0.9416**	**0.9120**	0.5856	0.9020	**0.8617**	**0.8814**	**0.9339**	**0.9080**	0.5857	0.9151	**0.8702**	**0.8921**	**0.9392**	**0.9150**	0.5808
	YOLOv8m	**0.9406**	0.8419	**0.8885**	**0.9320**	**0.9087**	**0.5932**	0.9211	0.8285	0.8723	0.9188	0.8944	**0.5864**	**0.9256**	0.8442	0.8830	0.9311	0.9029	**0.5903**
	YOLOv8m-TST	0.9136	**0.8420**	0.8763	0.9317	0.8992	0.5701	**0.9358**	**0.8391**	**0.8848**	**0.9356**	**0.9060**	0.5715	0.9124	**0.8625**	**0.8868**	**0.9380**	**0.9098**	0.5761
	YOLOv5m	0.9170	0.7878	0.8475	0.8811	0.8372	0.5166	**0.9240**	0.8017	0.8586	0.8928	0.8554	0.5197	0.9136	0.8048	0.8558	0.8952	0.8605	0.5328
	YOLOv5m-TST	**0.9193**	**0.8294**	**0.8720**	**0.9108**	**0.8717**	**0.5206**	0.9238	**0.8529**	**0.8869**	**0.9200**	**0.8822**	**0.5341**	**0.9186**	**0.8420**	**0.8786**	**0.9192**	**0.8810**	**0.5349**
5	YOLOv3	**0.9505**	0.8373	0.8903	0.9224	0.8885	**0.6011**	**0.9548**	**0.8595**	**0.9046**	0.9311	0.8939	**0.6052**	**0.9576**	0.8426	0.8964	0.9191	0.8875	**0.6137**
	YOLOv3-TST	0.9484	**0.8526**	**0.8980**	**0.9286**	**0.8945**	0.6001	0.9526	0.8570	0.9023	**0.9372**	**0.8990**	0.5967	0.9450	**0.8735**	**0.9078**	**0.9435**	**0.9109**	0.6035
	YOLOv8m	**0.9477**	0.8162	0.8771	0.9037	0.8705	**0.5934**	0.9431	0.8221	0.8784	0.9122	0.8747	0.5993	**0.9537**	0.8308	0.8881	0.9177	0.8848	**0.6079**
	YOLOv8m-TST	0.9395	**0.8582**	**0.8970**	**0.9363**	**0.8987**	0.5933	**0.9445**	**0.8745**	**0.9082**	**0.9429**	**0.9060**	**0.6020**	0.9468	**0.8464**	**0.8938**	**0.9286**	**0.8942**	0.6020
	YOLOv5m	0.9231	0.7978	0.8559	0.8846	0.8443	0.5303	**0.9533**	0.8065	0.8738	0.8929	0.8572	**0.5424**	0.9388	0.8135	0.8717	**0.9050**	**0.8625**	0.5471
	YOLOv5m-TST	**0.9364**	**0.8365**	**0.8837**	**0.9076**	**0.8661**	**0.5348**	0.9366	**0.8285**	**0.8792**	**0.9062**	**0.8610**	0.5354	**0.9544**	**0.8168**	**0.8802**	0.8957	0.8567	**0.5538**
Avg.	YOLOv3	0.9342	0.8795	0.9056	0.9438	0.9157	**0.6170**	**0.9398**	0.8890	**0.9133**	0.9491	0.9197	**0.6210**	**0.9434**	0.8803	0.9105	0.9480	0.9229	**0.6296**
	YOLOv3-TST	**0.9347**	**0.8873**	**0.9102**	**0.9500**	**0.9210**	0.6093	0.9316	**0.8902**	0.9101	**0.9499**	**0.9204**	0.6127	0.9342	**0.8954**	**0.9143**	**0.9528**	**0.9257**	0.6123
	YOLOv8m	**0.9365**	0.8651	0.8989	0.9363	0.9077	**0.6119**	0.9351	0.8627	0.8971	0.9363	0.9069	**0.6161**	**0.9398**	0.8683	0.9024	0.9413	0.9142	**0.6189**
	YOLOv8m-TST	0.9316	**0.8775**	**0.9036**	**0.9460**	**0.9139**	0.5991	**0.9381**	**0.8848**	**0.9104**	**0.9507**	**0.9197**	0.6069	0.9327	**0.8847**	**0.9078**	**0.9481**	**0.9180**	0.6097
	YOLOv5m	0.9302	0.8329	0.8786	0.9082	0.8712	0.5440	**0.9404**	0.8375	0.8857	0.9118	0.8779	**0.5520**	**0.9381**	0.8447	0.8889	0.9215	0.8849	0.5590
	YOLOv5m-TST	**0.9336**	**0.8627**	**0.8967**	**0.9272**	**0.8902**	**0.5446**	0.9256	**0.8662**	**0.8947**	**0.9274**	**0.8917**	0.5508	0.9352	**0.8603**	**0.8958**	**0.9265**	**0.8937**	**0.5602**

Table 3. Comparison of the performance between our proposed method applied to YOLOv3 [8] and YOLOv8m [3] for examining the compatibility of the model. Each model is evaluated on the combined datasets of SUN [6], KUMC [5], and Kvasir [1]. The best scores are bold-faced.

# Fold	Key Dim.	320					
		Precision	Recall	F1	mAP25	mAP50	mAP50:95
1	4	0.9623	*0.9135*	0.9373	0.9590	0.9488	0.6124
	64	*0.9688*	**0.9216**	**0.9446**	**0.9646**	**0.9561**	**0.6247**
	128	0.9617	0.9133	0.9369	0.9616	0.9513	0.6183
	256	**0.9712**	0.9083	*0.9387*	*0.9637*	*0.9554*	*0.6232*
	YOLOv5m	0.9636	0.8941	0.9275	0.9457	0.9318	0.6010
2	4	0.9244	0.8607	0.8914	0.9256	0.8777	0.5262
	64	0.9279	**0.8666**	**0.8962**	*0.9359*	*0.8948*	**0.5511**
	128	*0.9283*	0.8626	0.8942	**0.9378**	**0.8958**	0.5452
	256	0.9274	*0.8665*	*0.8959*	0.9321	0.8906	*0.5466*
	YOLOv5m	**0.9316**	0.8309	0.8784	0.9069	0.8672	0.5334
3	4	*0.9257*	0.8736	0.8989	0.9330	0.8866	0.5288
	64	0.9221	0.8778	0.8994	0.9412	0.8955	0.5516
	128	**0.9298**	*0.8782*	**0.9033**	**0.9469**	**0.9029**	**0.5546**
	256	0.9220	**0.8793**	*0.9002*	*0.9435*	*0.9028*	*0.5531*
	YOLOv5m	0.9158	0.8538	0.8837	0.9226	0.8757	0.5386
4	4	*0.9193*	0.8294	0.8720	0.9108	0.8717	0.5206
	64	0.9182	*0.8389*	*0.8768*	*0.9137*	**0.8795**	**0.5320**
	128	0.9147	0.8265	0.8684	0.9070	0.8663	0.5230
	256	**0.9243**	**0.8398**	**0.8800**	**0.9153**	*0.8775*	*0.5286*
	YOLOv5m	0.9170	0.7878	0.8475	0.8811	0.8372	0.5166
5	4	0.9364	0.8365	0.8837	0.9076	0.8661	0.5348
	64	0.9451	*0.8378*	*0.8882*	0.9125	0.8770	0.5418
	128	**0.9506**	**0.8419**	**0.8929**	*0.9172*	*0.8787*	*0.5462*
	256	*0.9457*	0.8357	0.8873	**0.9228**	**0.8809**	**0.5490**
	YOLOv5m	0.9231	0.7978	0.8559	0.8846	0.8443	0.5303
Avg.	4	0.9336	0.8627	0.8967	0.9272	0.8902	0.5446
	64	0.9364	**0.8686**	**0.9011**	0.9336	*0.9006*	**0.5602**
	128	*0.9370*	0.8645	0.8991	*0.9341*	0.8990	0.5575
	256	**0.9381**	*0.8659*	*0.9004*	**0.9355**	**0.9014**	*0.5601*
	YOLOv5m	0.9302	0.8329	0.8786	0.9082	0.8712	0.5440

Table 4. Comparison of the performance between our proposed method with varying key dimension sizes and the original YOLOv5m [2] model. Result of models evaluated on the combined SUN [6], KUMC [5], and Kvasir [1] datasets. The best, second-best, and third-best scores are highlighted in bold, italics, and underlined, respectively.

original YOLOv5m model for all experiments of 5-fold cross-validation on combined datasets.

References

[1] Debesh Jha, Pia H Smedsrud, Michael A Riegler, Pål Halvorsen, Thomas de Lange, Dag Johansen, and Håvard D Johansen. Kvasir-seg: A segmented polyp dataset. In *MultiMedia Modeling: 26th International Conference, MMM 2020, Daejeon, South Korea, January 5–8, 2020, Proceedings, Part II 26*, pages 451–462. Springer, 2020. 1, 2, 3, 4

[2] Glenn Jocher. Ultralytics yolov5, 2020. 1, 2, 3, 4

[3] Glenn Jocher, Ayush Chaurasia, and Jing Qiu. Ultralytics yolov8, 2023. 3, 4

[4] Adrian Krenzer, Michael Banck, Kevin Makowski, Amar Hekalo, Daniel Fitting, Joel Troya, Boban Sudarevic, Wolfgang G Zoller, Alexander Hann, and Frank Puppe. A real-time polyp-detection system with clinical application in colonoscopy using deep convolutional neural networks. *Journal of Imaging*, 9(2):26, 2023. 3

[5] Kaidong Li, Mohammad I Fathan, Krushi Patel, Tianxiao Zhang, Cuncong Zhong, Ajay Bansal, Amit Rastogi, Jean S Wang, and Guanghui Wang. Colonoscopy polyp detection and classification: Dataset creation and comparative evaluations. *Plos one*, 16(8):e0255809, 2021. 1, 2, 3, 4

[6] Masashi Misawa, Shin-ei Kudo, Yuichi Mori, Kinichi Hotta, Kazuo Ohtsuka, Takahisa Matsuda, Shoichi Saito, Toyoki Kudo, Toshiyuki Baba, Fumio Ishida, et al. Development of a computer-aided detection system for colonoscopy and a publicly accessible large colonoscopy video database (with video). *Gastrointestinal endoscopy*, 93(4):960–967, 2021. 1, 3, 4

[7] Shimin Ou, Yixing Gao, Zebin Zhang, and Chenjian Shi. Polyp-yolov5-tiny: A lightweight model for real-time polyp detection. In *2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)*, volume 2, pages 1106–1111. IEEE, 2021. 3

[8] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018. 3, 4