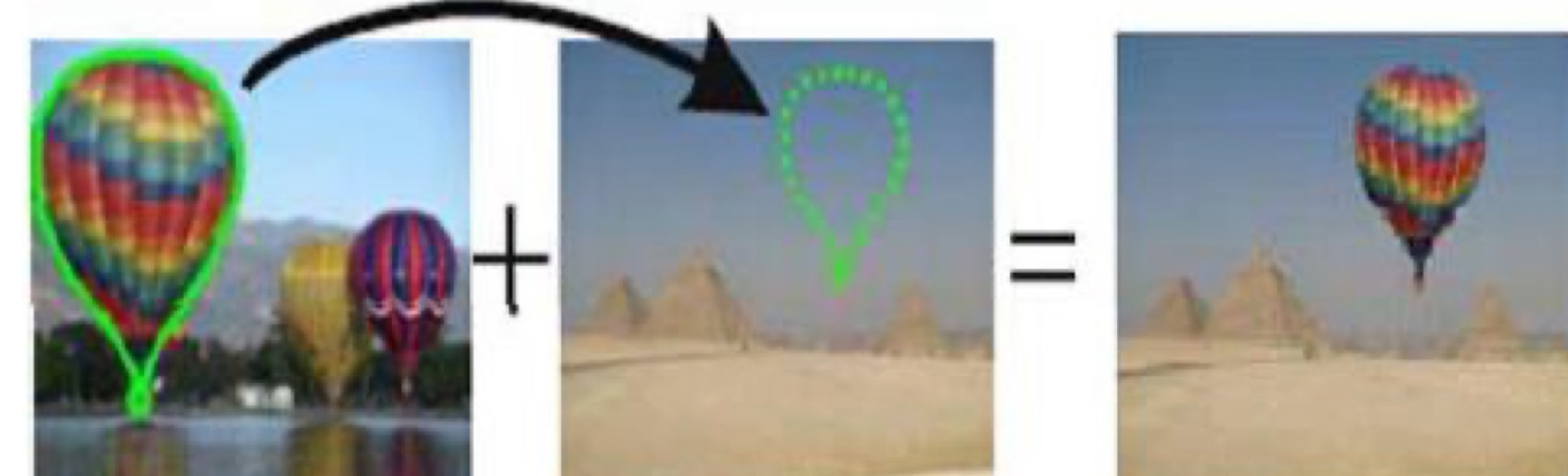


## Weakly-supervised image manipulation detection

- Typical image manipulations
  - They typically have pixel-level masks



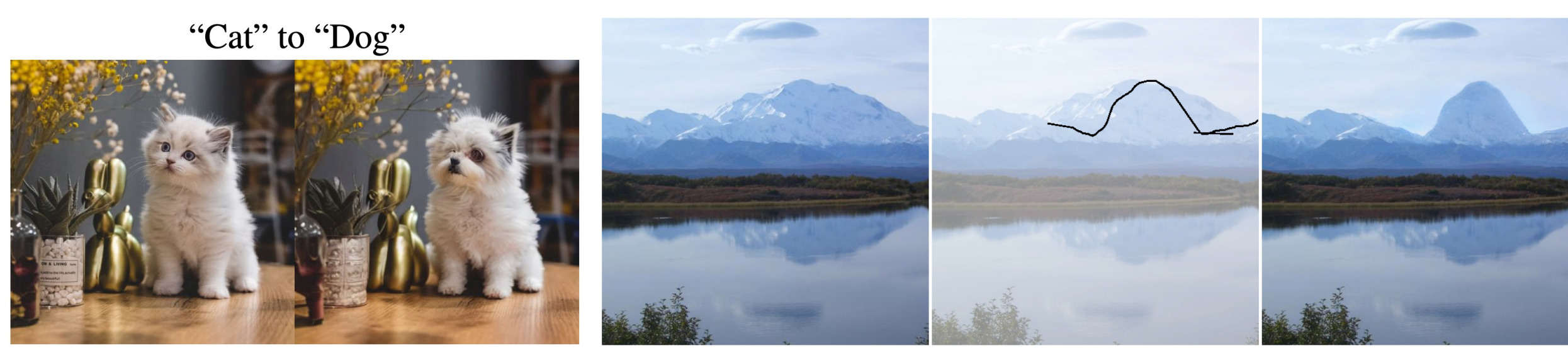
Splicing [Sharma et al.]



Inpainting [Trung et al.]

Copy-move [Mahdi et al.]

- Emerging editing methods do not necessarily generate such masks

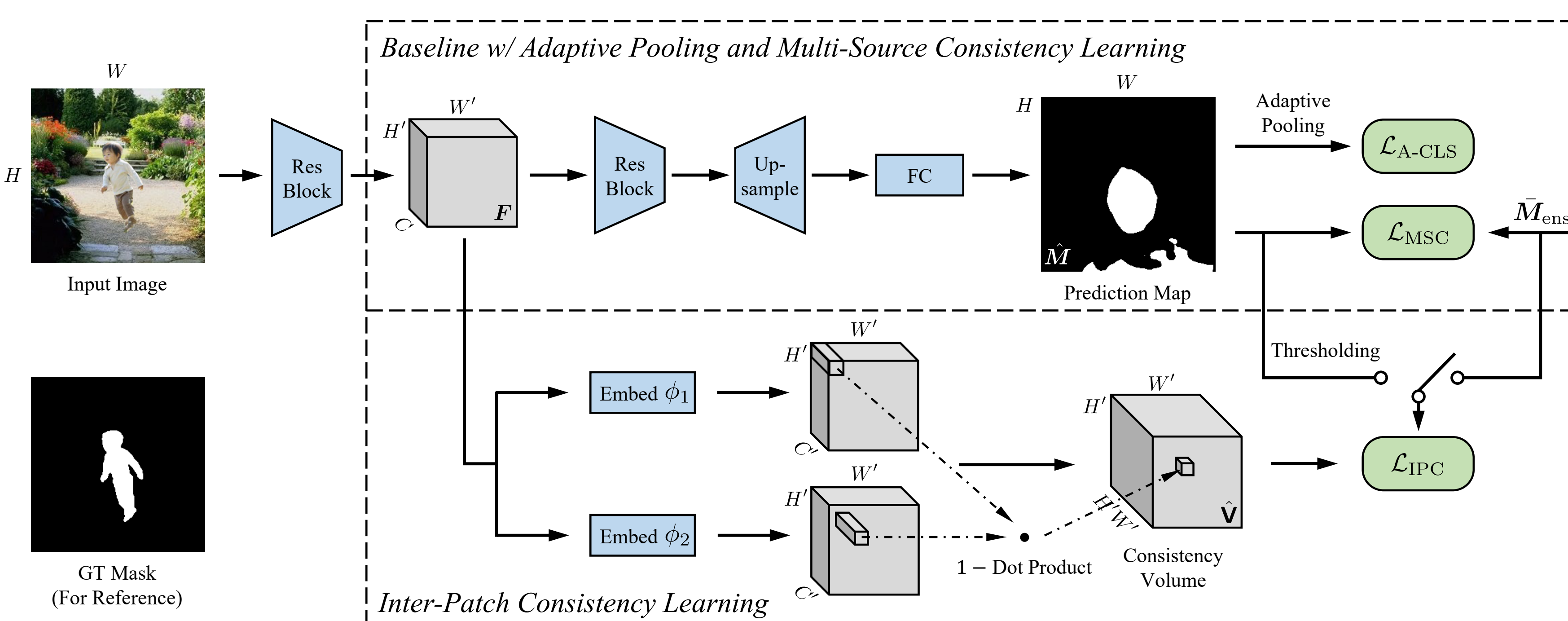


Language-guided editing [Wang et al.]

Sketch-based editing [Zeng et al.]

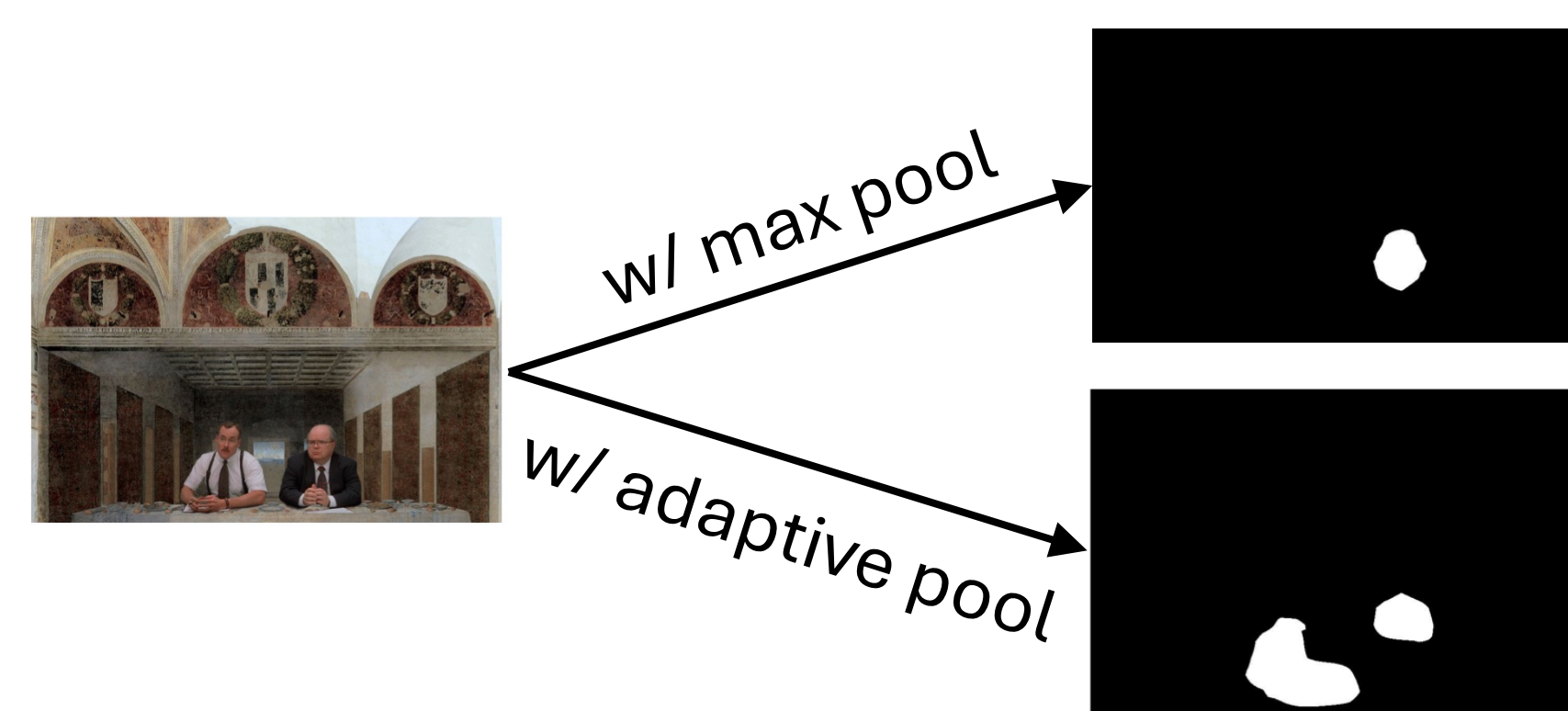
- New task: weakly-supervised image manipulation detection (W-IMD)
  - Given only **binary image-level labels** (real or fake), predict whether an image is manipulated, and localize the manipulation at the pixel level.
  - Such a paradigm eliminates the need for pixel-level masks, and can quickly adapt to novel manipulations without such annotations

## Weakly-supervised self-consistency learning



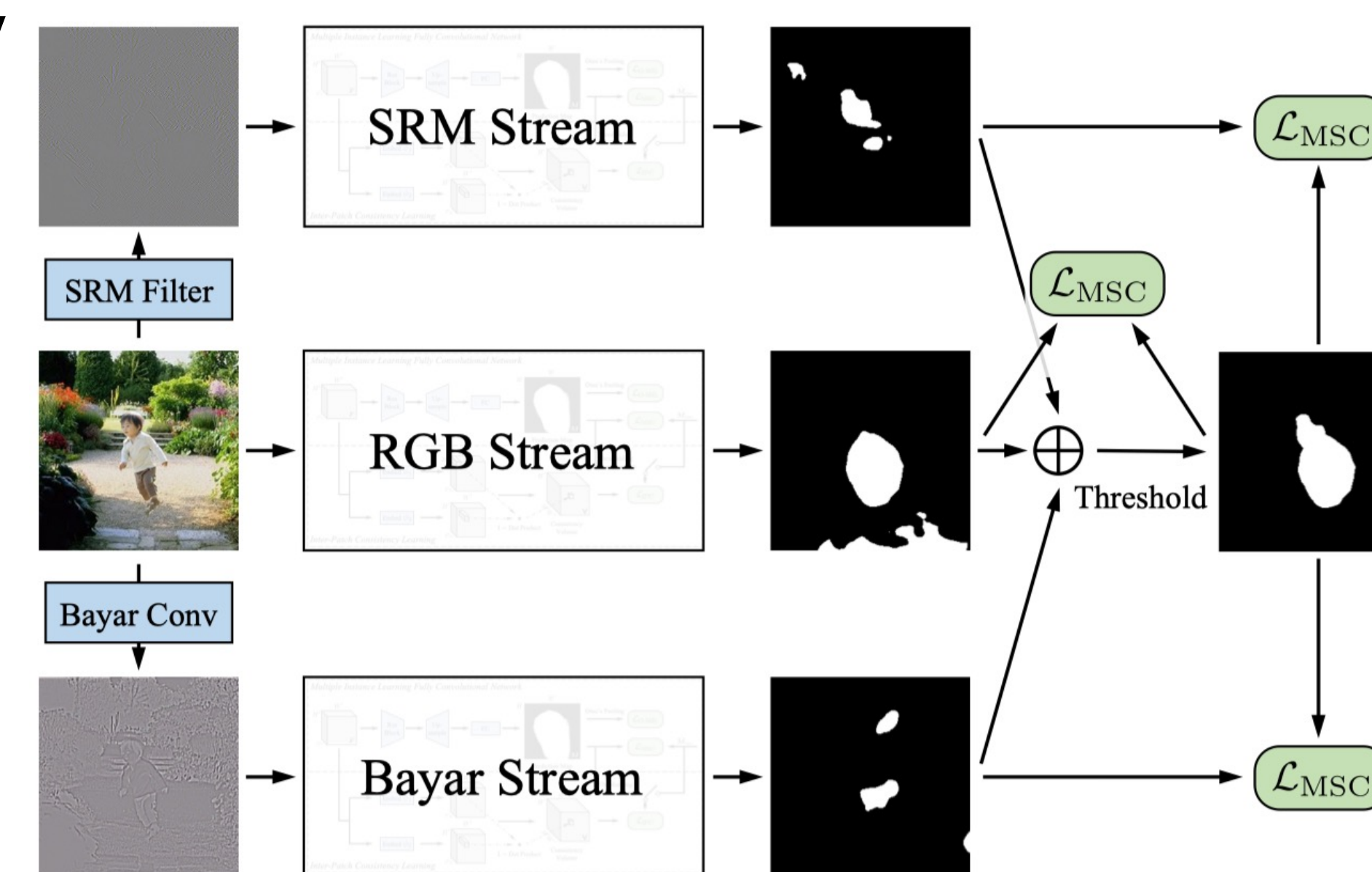
- Adaptive pooling

- Based on Otsu's binarization method, dynamically select the pixel-level responses from the prediction map



- Multi-source consistency (MSC) learning

- Fuse information from different noise sources (RGB, SRM, and Bayar)
- The ensemble prediction is in turn used as pseudo ground truth, and supervises each individual stream



- Inter-patch consistency learning

- The global patch-patch similarity is computed, and forms a 4D consistency volume
- The pseudo ground truth from MSC is used to supervised the consistency volume, thus to enhance low-level features

## Results

- Strong image-level detection results

	Method	CASIv1				Columbia				Coverage				IMD2020				Avg	
		AUC	Spe.	Sen.	F1	AUC	Spe.	Sen.	F1	AUC	Spe.	Sen.	F1	AUC	Spe.	Sen.	F1	AUC	F1
Full	NOH [29]	0.500	0.000	1.000	0.000	0.500	0.000	1.000	0.000	0.500	0.000	1.000	0.000	0.500	0.000	1.000	0.000	0.500	0.000
	CPAI [13]	0.482	0.000	1.000	0.000	0.344	0.000	1.000	0.000	0.525	0.000	1.000	0.000	0.500	0.000	1.000	0.000	0.500	0.000
	Mantra-Net [55]	0.341	0.000	1.000	0.000	0.701	0.000	1.000	0.000	0.491	0.000	1.000	0.000	0.719	0.000	1.000	0.000	0.513	0.000
	CR-CNN [57]	0.766	0.224	0.930	0.361	0.783	0.246	0.961	0.392	0.566	0.070	0.967	0.131	0.617	0.112	0.936	0.200	0.683	0.271
	GSR-Net [65]	0.502	0.011	0.994	0.022	0.502	0.011	1.000	0.022	0.515	0.000	1.000	0.000	0.505	0.008	0.998	0.014	0.506	0.019
	CAT-Net [22]	0.690	0.328	0.762	0.459	0.849	0.373	0.782	0.505	0.572	0.093	0.902	0.169	0.721	0.132	0.872	0.229	0.693	0.157
Weak	FCN+DA [6]	0.795	0.844	0.717	0.775	0.762	0.322	0.950	0.481	0.541	0.100	0.900	0.180	0.746	0.100	0.981	0.182	0.711	0.404
	MIL-FCN [37]	0.647	0.538	0.569	0.553	0.807	0.220	0.732	0.338	0.542	0.062	0.793	0.115	0.578	0.116	0.886	0.205	0.644	0.303
	MIL-FCN [37] + WSCL	<b>0.829</b>	0.795	0.690	0.738	<b>0.920</b>	0.519	0.983	<b>0.680</b>	<b>0.584</b>	0.440	0.714	<b>0.544</b>	<b>0.733</b>	0.221	0.966	<b>0.360</b>	<b>0.766</b>	<b>0.580</b>
	Araslanov and Roth [1]	0.642	0.458	0.542	0.496	0.773	0.127	0.902	0.223	0.560	0.077	0.746	0.140	0.665	0.126	0.832	0.219	0.660	0.270
Araslanov and Roth [1] + WSCL	0.795	0.638	0.726	0.679	0.917	0.324	0.948	0.483	<b>0.591</b>	0.220	0.838	0.348	0.701	0.193	0.872	0.316	0.751	0.456	

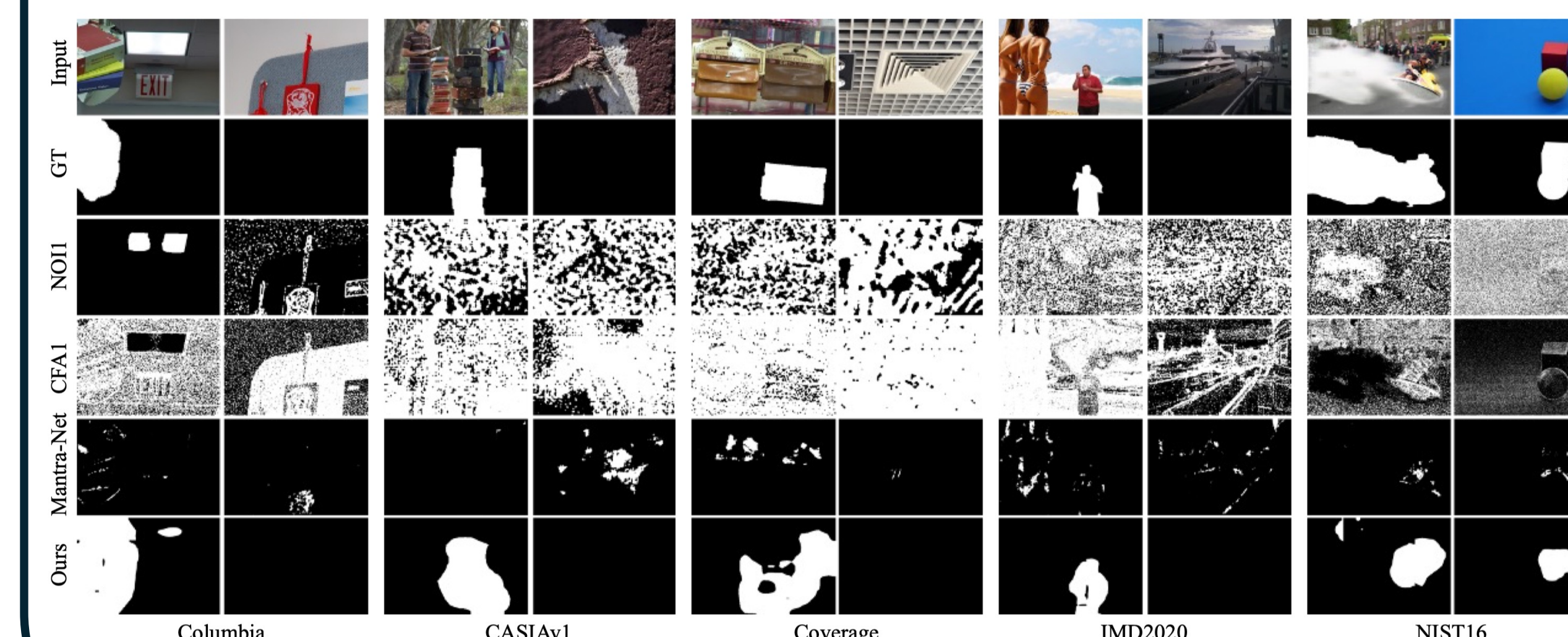
- Adaptation to novel manipulations

	Method	GIER [43]		IEdit [45]		Avg	
		AUC	F1	AUC	F1	AUC	F1
Full	CAT-Net [22]	0.508	0.336	0.532	0.476	0.502	0.406
	FCN+DA [6]	0.507	0.428	0.539	0.489	0.523	0.458
	MVSS-Net [6]	0.510	0.325	0.537	0.522	0.523	0.423
Weak	MIL-FCN [37] + WSCL	0.574	0.320	0.563	0.556	0.568	0.438
	MIL-FCN [37] + WSCL w/ fine-tune	<b>0.621</b>	<b>0.533</b>	<b>0.617</b>	<b>0.602</b>	<b>0.619</b>	<b>0.568</b>

- Reasonable pixel-level localization results

	Method	Pixel-Level F1						Combined F1					
		CASIv1	Columbia	Coverage	IMD2020	NIST16	Avg	CASIv1	Columbia	Coverage	IMD2020	Avg	
Full	NOH [29]	0.157	0.311	0.205	0.124	0.089	0.190	0.000	0.000	0.000	0.000	0.000	
	CPAI [13]	0.140	0.320	0.188	0.111	0.106	0.188	0.000	0.000	0.000	0.000	0.000	
	Mantra-Net [55]	0.155	0.364	0.286	0.122	0.090	0.185	0.000	0.000	0.000	0.000	0.000	
	CR-CNN [57]	0.405	0.436	0.291	0.122	0.238	0.382	0.413	0.181	0.026	0.026	0.028	
	GSR-Net [65]	0.387	0.613	0.285	0.175	0.283	0.349	0.042	0.042	0.000	0.026	0.028	
	CAT-Net [22]	0.276	0.352	0.134	0.102	0.138	0.200	0.345	0.406	0.149	0.144	0.261	
Weak	FCN+DA [6]	0.441	0.223	0.199	0.270	0.167	0.260	0.562	0.305	0.189	0.217	0.318	
	MIL-FCN [37]	0.117	0.089	0.121	0.097	0.024	0.090	0.193	0.141	0.118	0.131	0.146	
	MIL-FCN [37] + WSCL	0.172	0.270	0.178	0.193	0.110	0.185	0.280	0.386	0.268	0.252	0.296	
	Araslanov and Roth [1]	0.112	0.102	0.127	0.094	0.026	0.092	0.182	0.140	0.133	0.046	0.125	
Araslanov and Roth [1] + WSCL	0.153	0.362	0.201	0.173	0.099	0.198	0.250	0.414	0.255	0.159	0.270		

- Qualitative results



## Open source



Code



Paper



Project