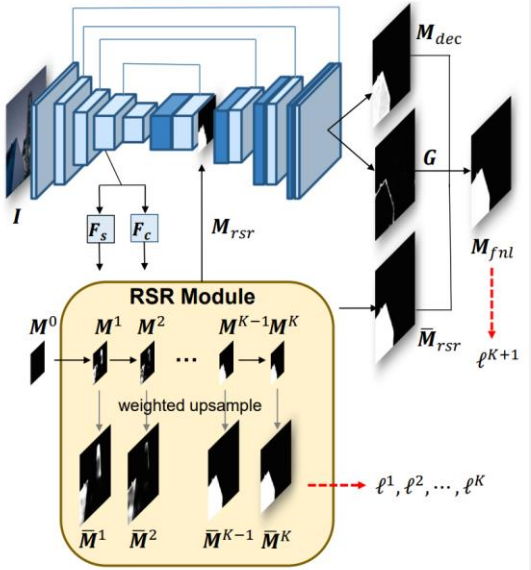
Inharmonious Region Localization via Recurrent Self-Reasoning

Penghao Wu, Li Niu, Jing Liang, Liqing Zhang

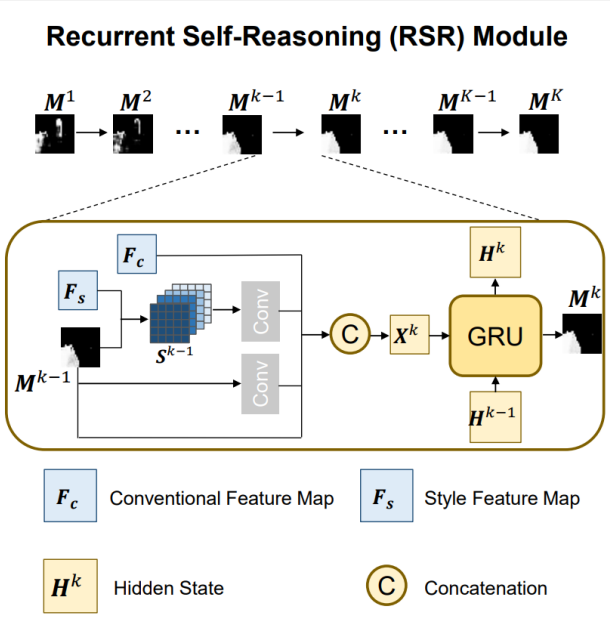
Inharmonious region localization

inconsistent color and illumination characteristics severely degrade the quality of synthetic images

In this work, we take inspiration from K-Means algorithm to propose our RSRNet



Recurrent Self-Reasoning (RSR) module



Assignment step: assign each pixel to a cluster based on similarity

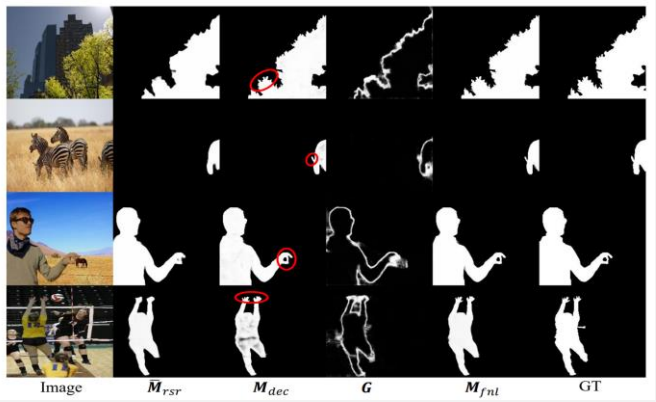
- Background style feature
- Multi-scale similarity map

Update step: update clusters based on associated pixels

- GRU Based Recurrent Unit

$$
\begin{aligned}
\boldsymbol{Z}^k &= \sigma(\mathrm{Conv}([\boldsymbol{H}^{k-1}, \boldsymbol{X}^k], \boldsymbol{W}_z)),\\
\boldsymbol{R}^k &= \sigma(\mathrm{Conv}([\boldsymbol{H}^{k-1}, \boldsymbol{X}^k], \boldsymbol{W}_r)),\\
\tilde{\boldsymbol{H}}^k &= \tanh(\mathrm{Conv}([\boldsymbol{R}^k \odot \boldsymbol{H}^{k-1}, \boldsymbol{X}^k], \boldsymbol{W}_h)),\\
\boldsymbol{H}^k &= (1-\boldsymbol{Z}^k) \odot \boldsymbol{H}^{k-1} + \boldsymbol{Z}^k \odot \tilde{\boldsymbol{H}}^k,
\end{aligned}
$$

Adaptive Combination



$$M_{fnl} = \boldsymbol{G} \odot \boldsymbol{M}_{dec} + (1-\boldsymbol{G}) \odot \bar{\boldsymbol{M}}_{rsr}.$$

Fine-grained details
&
Compact general shape

Experiments

Method	HCOCO			HAdobe5k			HFlickr			Hday2night			All		
	AP↑	F_1↑	IoU↑	AP↑	F_1↑	IoU↑	AP↑	F_1↑	IoU↑	AP↑	F_1↑	IoU↑	AP↑	F_1↑	IoU↑
UNet	68.11	0.5869	56.57	89.26	0.8380	80.85	80.72	0.7683	74.58	35.74	0.2362	19.60	74.90	0.6717	64.74
DeepLabv3	69.09	0.6070	58.21	90.20	0.8591	81.56	80.01	0.7698	74.91	35.87	0.2550	21.38	75.69	0.6902	66.01
HRNet-OCR	68.89	0.5981	57.69	89.63	0.8387	80.98	79.62	0.7489	74.55	34.98	0.2477	21.34	75.33	0.6765	65.49
SegFormer	72.46	0.6578	58.78	89.43	0.8531	80.44	85.19	0.7986	75.02	45.16	0.3856	32.75	78.05	0.7249	66.55
MFCN	37.36	0.3030	25.18	62.75	0.5365	36.63	49.89	0.4209	28.34	19.71	0.1426	11.88	45.63	0.3794	28.54
MantraNet	56.55	0.4811	41.04	81.07	0.7510	68.50	67.52	0.6302	58.51	28.88	0.2019	16.71	64.22	0.5691	50.31
MAGritte	64.75	0.6058	51.77	85.50	0.8630	76.36	75.02	0.7725	70.25	31.20	0.2549	17.05	71.16	0.6907	60.14
H-LSTM	52.29	0.4336	37.81	77.62	0.7056	65.19	63.12	0.5918	54.93	26.59	0.1977	15.91	60.21	0.5239	47.07
SPAN	58.41	0.4906	45.07	82.57	0.7786	72.49	69.22	0.6510	62.20	29.58	0.2171	19.41	65.94	0.5850	54.27
F3Net	54.17	0.4703	40.03	74.31	0.6944	60.08	72.53	0.6582	59.31	30.08	0.2563	20.83	61.46	0.5506	47.48
GATENet	55.07	0.4568	38.89	75.19	0.6634	59.18	74.13	0.6256	57.51	30.98	0.2174	19.38	62.43	0.5296	46.33
MINet	71.74	0.6022	55.79	89.58	0.8379	77.23	83.86	0.7761	72.51	37.82	0.2710	19.38	77.51	0.6822	63.04
DIRL	74.25	0.6701	60.85	92.16	0.8801	84.02	84.21	0.7786	73.21	38.74	0.2396	20.11	80.02	0.7317	67.85
RSRNet	**78.42**	**0.7131**	**65.85**	**93.10**	**0.8901**	**85.96**	**87.11**	**0.8048**	**76.84**	**47.34**	**0.3028**	**26.34**	**82.57**	**0.7607**	**71.63**

Qualitative Comparison