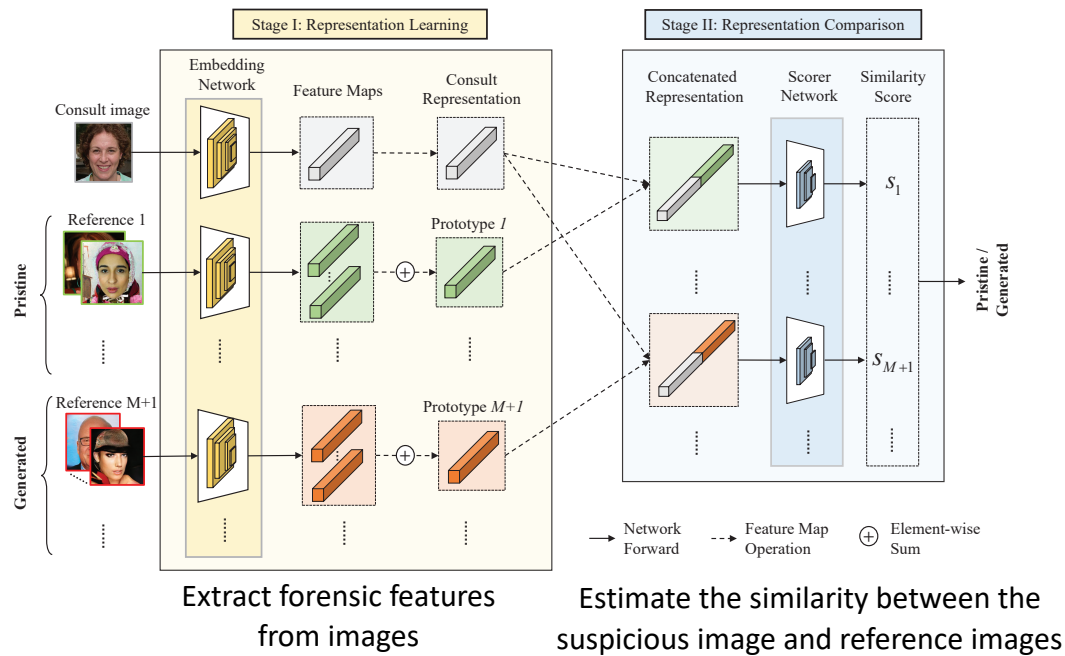


How to distinguish GAN-generated images?

Previous work: extract artifacts and classify

Our method: extract artifacts and compare with reference images



CDALoss

$$\hat{\theta}, \hat{\phi} = \arg \min_{g(\cdot; \theta), f(\cdot; \phi)} \sum_{y^* \neq y^n} f(g(x^*; \theta), p^n; \phi) - \sum_{y^* = y^n} f(g(x^*; \theta), p^n; \phi)$$

Utilize the **category** and **domain** information simultaneously to enlarge inter-class discrepancy and improve intra-class compactness.

ResNet-INS

Equip ResNet-50 with instance normalization in shallow layers.

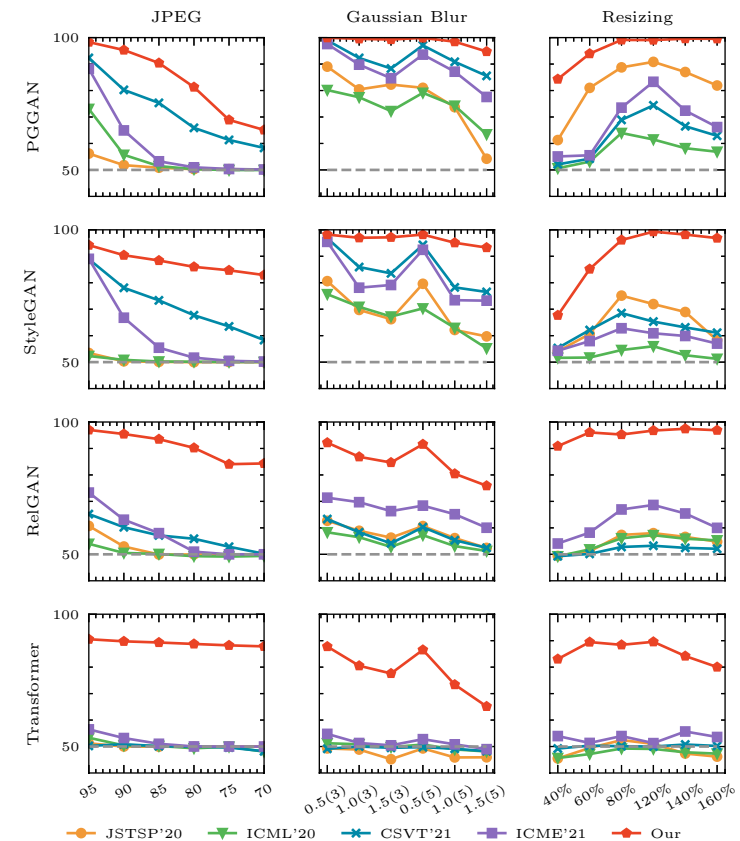
1. Enhance the ability to extract forensic traces
2. Make the trace invariant to image contents and source attributes

Experiments

Our method have better **robustness** and **generalization** over SOTAs

Detection performance over images generated by different models(%)

Model	Intra-domain		Cross-domain			
	PGGAN	StyleGAN	StyleGAN2	StarGAN	RelGAN	Transformer
JSTSP'20	99.55	99.02	49.25	94.01	65.60	49.88
ICML'20	99.30	99.53	52.28	99.07	63.72	50.66
CSV'T'21	99.25	99.26	61.23	73.32	65.67	51.09
ICME'21	99.65	99.13	75.53	97.32	78.03	59.87
Our	100.00	99.02	93.45	99.35	96.82	90.10



Robustness against different post-processing operations(%)