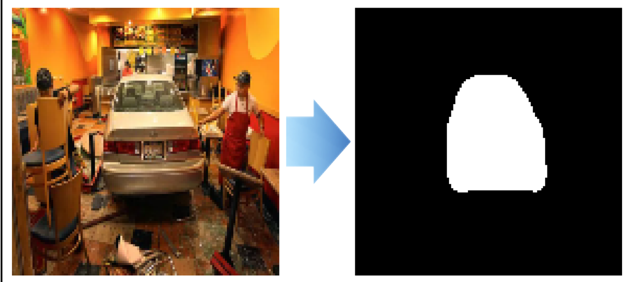
Unsupervised Segmentation incorporating Shape Prior via Generative Adversarial Networks

Dahye Kim, Byung-Woo Hong
Chung-Ang University, Seoul, Korea

2021 ICCV OCTOBER 11-17 VIRTUAL

I. Motivation



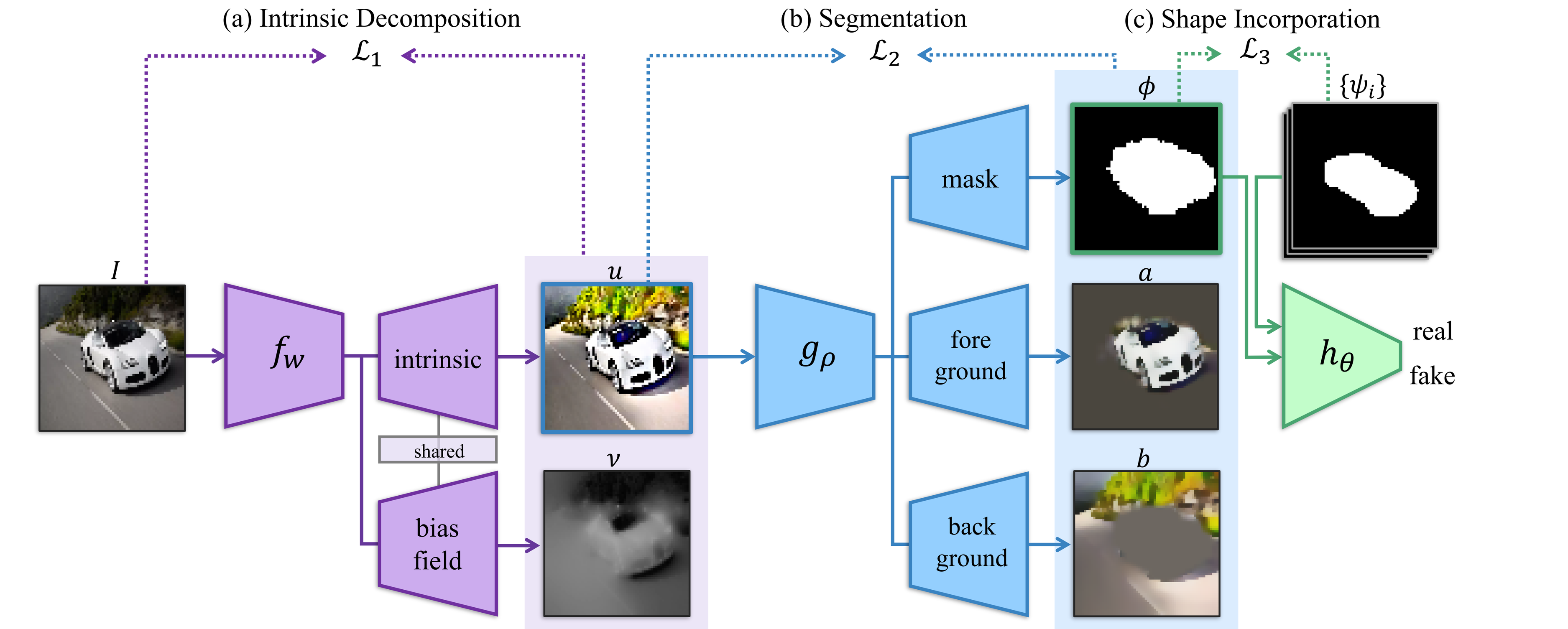
Bi-partitioning

Challenges in object segmentation:

- Interruption by bias fields and occlusions.
- Poor generalization and expensive annotations problems of supervised learning.
- GAN-based unsupervised approaches struggle to learn distributions of enormous dimensions.

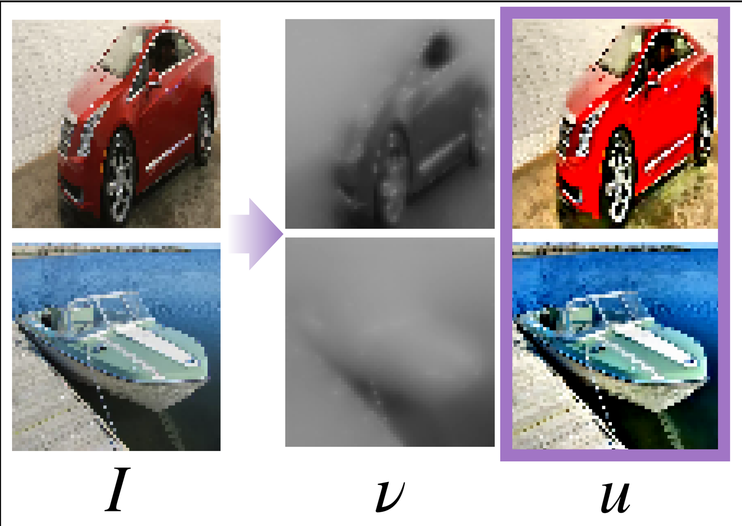
Our solution: retrieve intrinsic representations to remove bias fields, train segmentation network based on an unsupervised energy and impose shape constraints by applying GAN framework that learns only geometric distributions.

II. Method Overview



(a) f obtains an intrinsic representation u that is robust to a bias field ν.
(b) g derives a segmenting function ϕ from u in an unsupervised way.
(c) h imposes a geometric constraint to ϕ using a set of prior shapes $\{\psi_i\}$.

III. Intrinsic Decomposition



- An image I is decomposed into a bias field ν and an intrinsic representation u using the objective function $\mathcal{L}_1$ derived from an image formation model using an additive noise with a multiplicative bias field.
- The model parameter w of f is optimized based on $\mathcal{L}_1$.

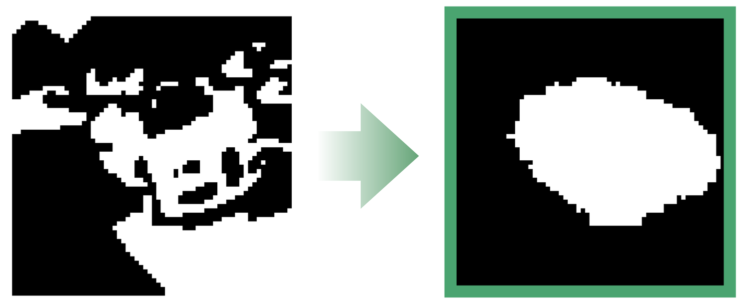
$$I = \nu(u + \eta) \Rightarrow \mathcal{L}_1(u, \nu; I) = \|\frac{I}{\nu} - u\|_2^2 + \lambda\|\nabla u\| + \alpha\|\nabla \nu\|_2^2 + \beta\|\nu - 1\|_2^2$$

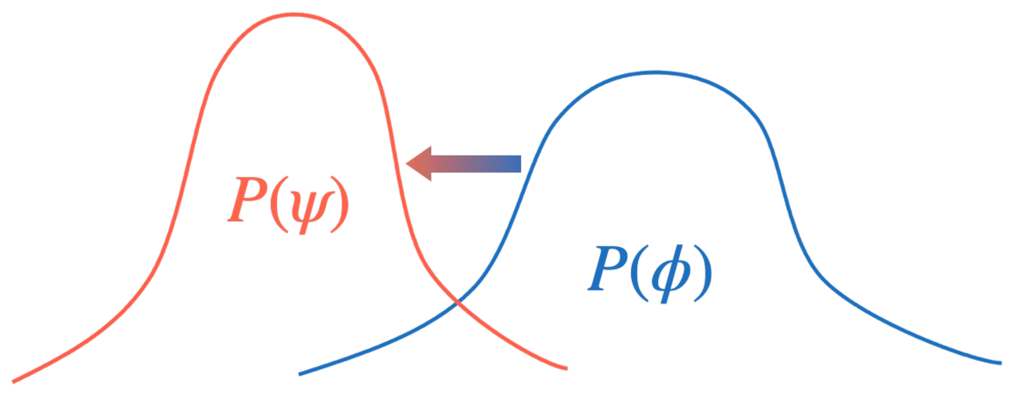
IV. Unsupervised Segmentation

- A piecewise smooth Mumford-Shah[1] model is applied for the objective function $\mathcal{L}_2$ for unsupervised segmentation.
- The segmenting function ϕ and its associated estimates a and b are directly achieved by the deep neural network g.
- The obtained u is used as input for g.

$$\mathcal{L}_2(\phi, a, b; u) = \int_\Omega |u(x) - a(x)|^2 \phi(x) \mathrm{d}x + \int_\Omega |u(x) - b(x)|^2 (1 - \phi(x)) \mathrm{d}x + \gamma_1\|\nabla \phi(x)\| + \gamma_2\|\nabla a(x)\| + \gamma_2\|\nabla b(x)\|$$

V. Shape Incorporation

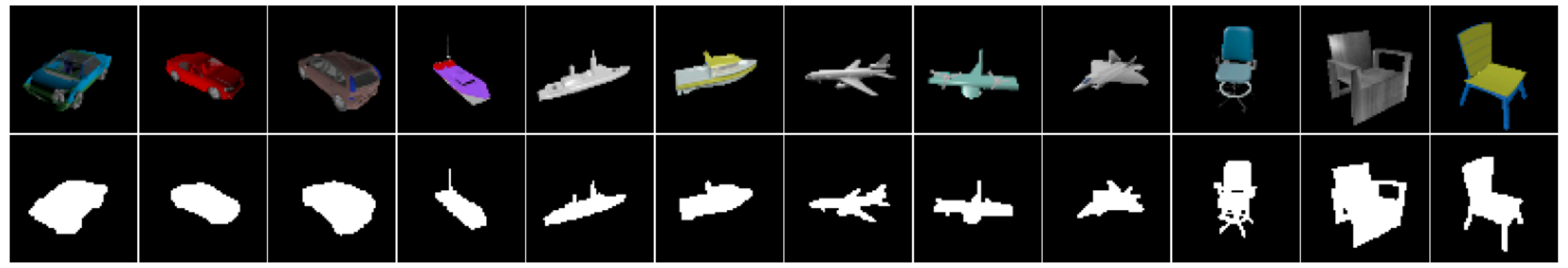




$$\mathcal{L}_3(\rho, \theta; \mathcal{S}, \mathcal{R}) = \mathbb{E}_{\psi \sim P(\psi)}[\log(h(\psi))] - \mathbb{E}_{\phi \sim Q(\phi)}[\log(h(\phi))] - \frac{\kappa}{2}\mathbb{E}_{\psi \sim P(\psi)}[\|\nabla h(\psi)\|_2^2]$$

- Shape constraints are imposed by minimizing the discrepancy between probability distributions of partitioning function $Q(\phi)$ and prior shape $P(\psi)$ using the GAN loss function $\mathcal{L}_3$.
- $\mathcal{L}_3$ is used to optimize the model parameters of both the segmentation network g and the discriminator h.

Shape prior construction with a set of deformed shapes $\psi_i(x) = \chi_S \circ t_i(x)$:

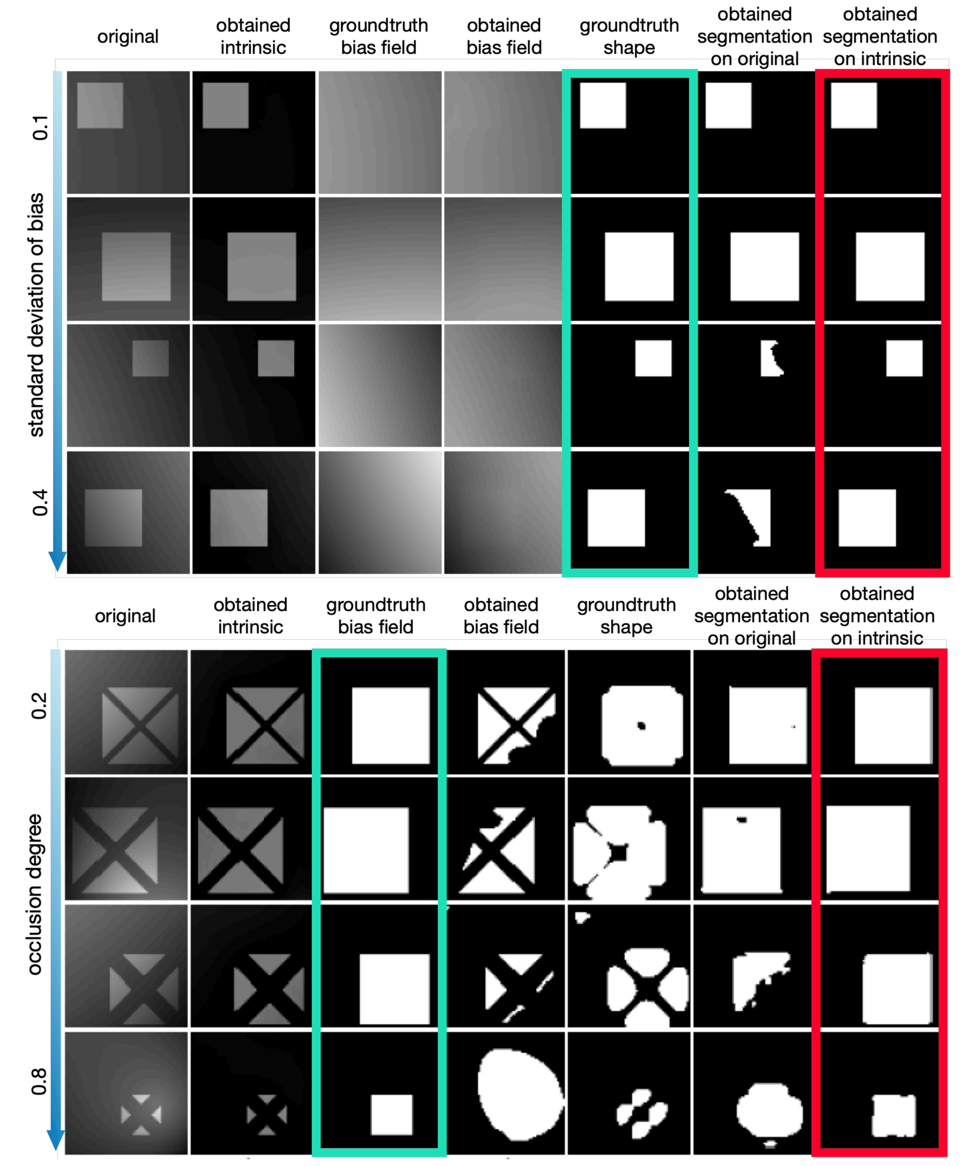


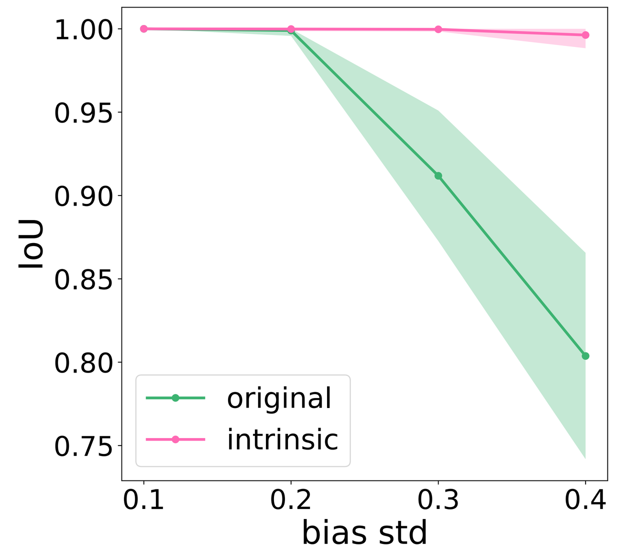
Examples of shape prior data

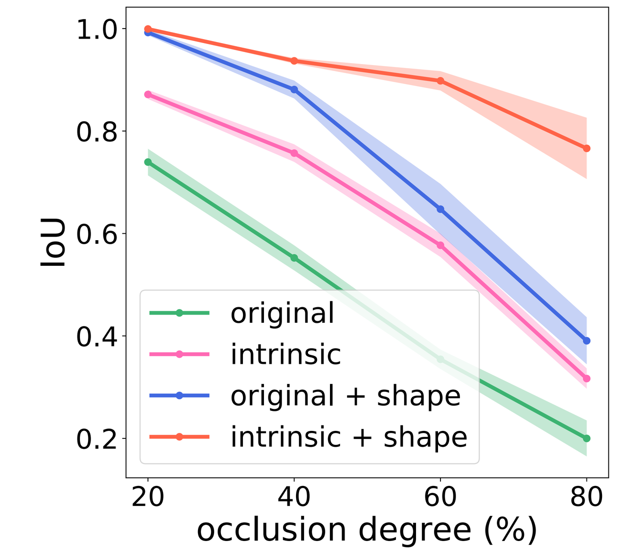
- $P(\psi)$ is constructed by the equivalence class $\mathcal{S} = \{\psi_i = \chi_S \circ t_i \mid t_i \in \mathcal{T}\}$ of shape $\mathcal{S}$ under the action of the transformation group $\mathcal{T}$.

VI. Results

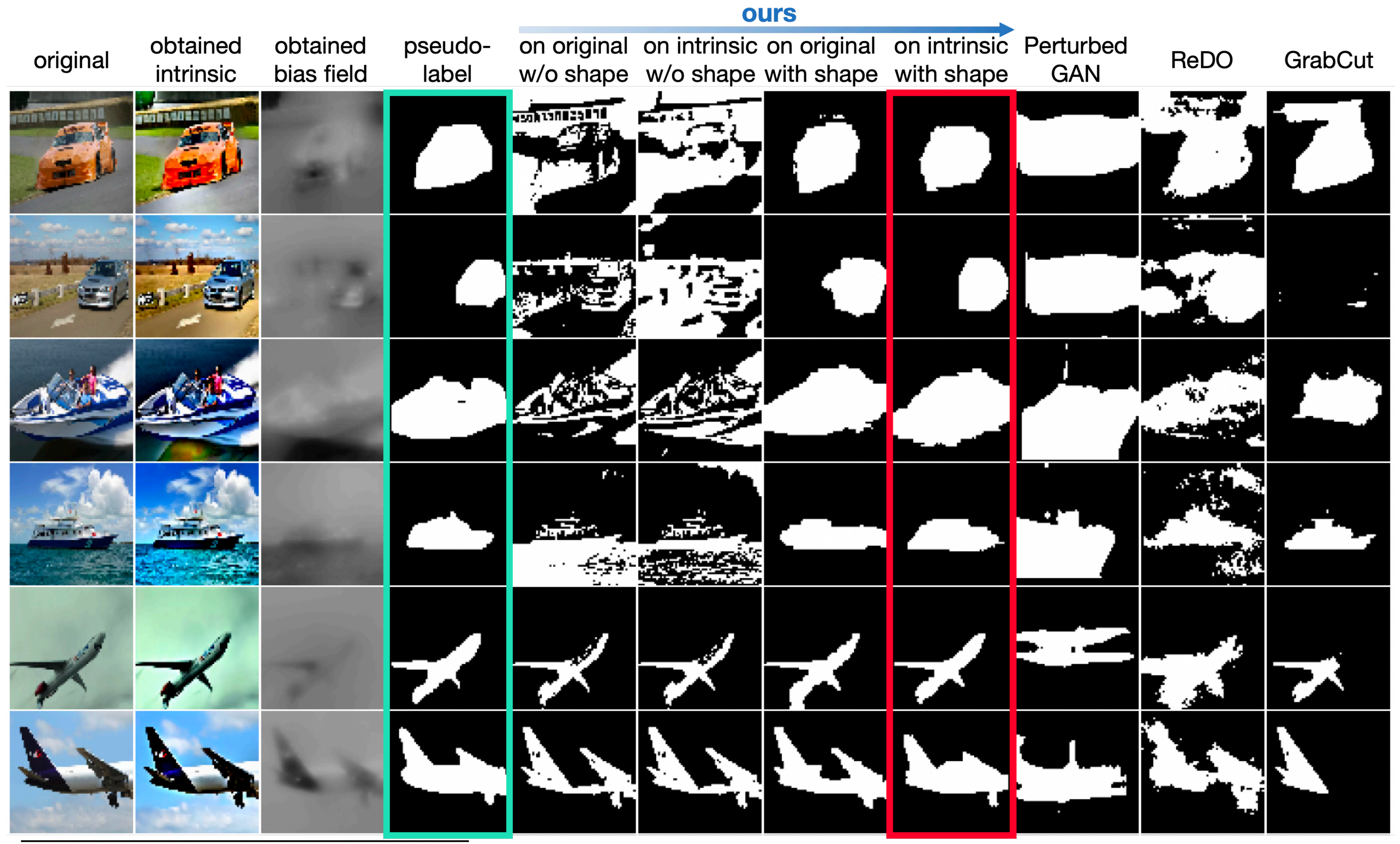
Synthetic







LSUN



method	car	boat	airplane	chair
original	0.3126	0.2002	0.2131	0.3543
intrinsic	0.3250	0.2348	0.2307	0.3683
original + shape	0.6303	0.4756	0.4544	0.4824
intrinsic + shape	**0.6340**	**0.4901**	**0.4714**	0.4776
PerturbedGAN	0.5026	0.3122	0.3049	0.3902
ReDO	0.4637	0.3618	0.4110	0.4181
GrabCut	0.5122	0.3325	0.4026	**0.5127**

Reference

1. Mumford, D.B. and Shah, J., 1989. Optimal approximations by piecewise smooth functions and associated variational problems. Communications on pure and applied mathematics.