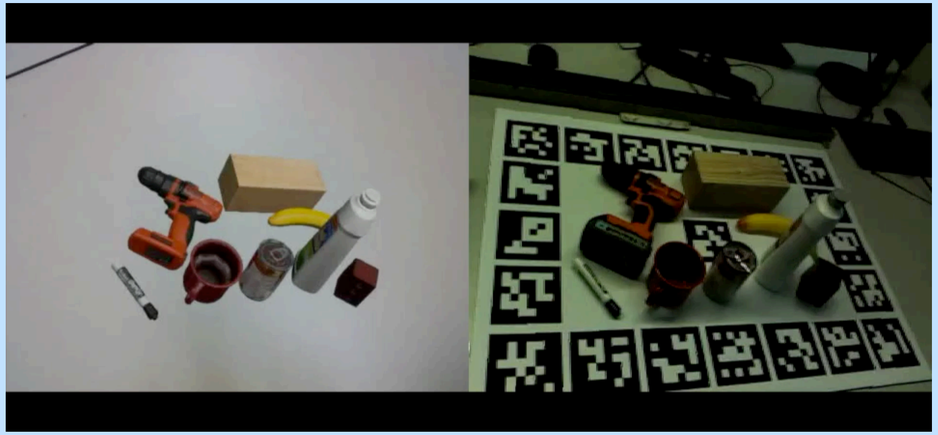
Quantifying the use of Domain Randomization

Mohammad Ani, Hector Basevi, Ales Leonardis

UNIVERSITY OF BIRMINGHAM

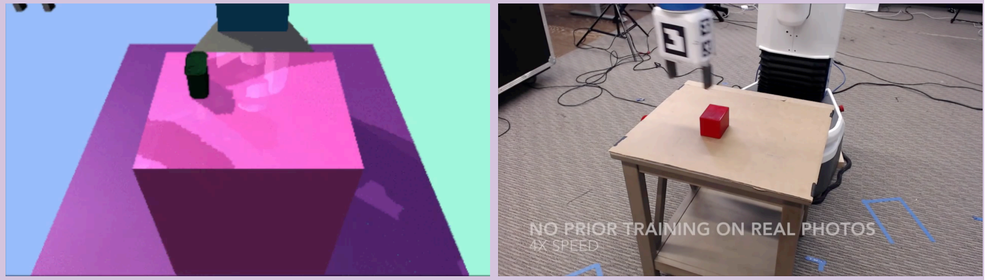
ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Motivation



Synthetic world **Real world**
(Grenzdörffer et al.)

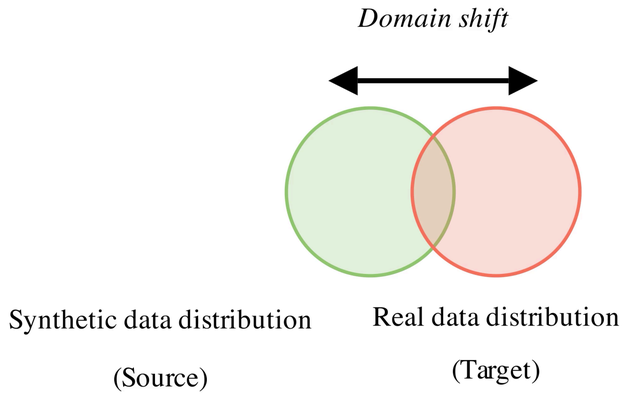
Similar target textures

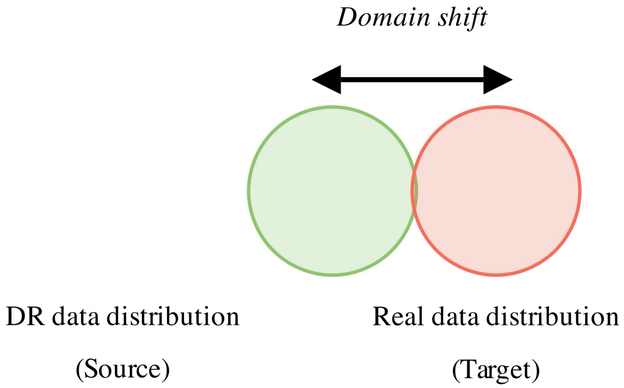


Domain Randomized (DR) Synthetic Images
(Tobin et al.)

Real world
(Tobin et al.)

Unknown target textures

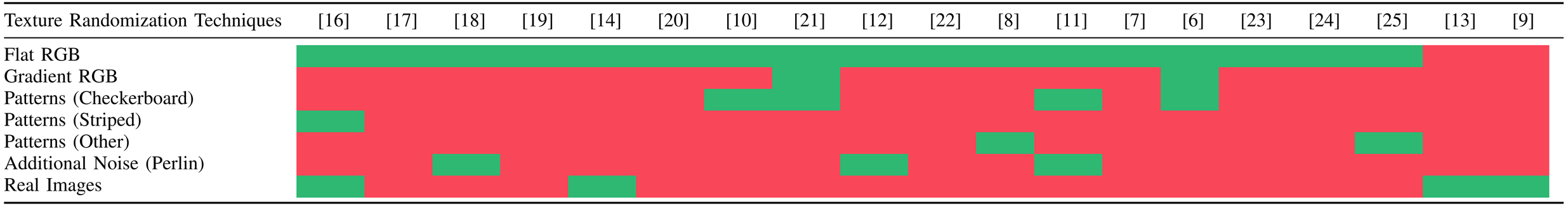




T. Grenzdörffer, M. Günther, and J. Hertzberg, “YCB-M: A Multi-Camera RGB-D Dataset for Object Recognition and 6DoF Pose Estimation,” in *2020 IEEE International Conference on Robotics and Automation, ICRA 2020, Paris, France, May 31-June 4, 2020*. IEEE, 2020.
Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel,“Domain randomization for transferring deep neural networks from simulation to the real world,” in IEEE/RSJ International Conference on Intelligent Robots and Systems IROS, 2017, pp. 23–30.

Motivation

Domain Randomization (DR) can broaden a synthetic data distribution to encompass a realistic data distribution

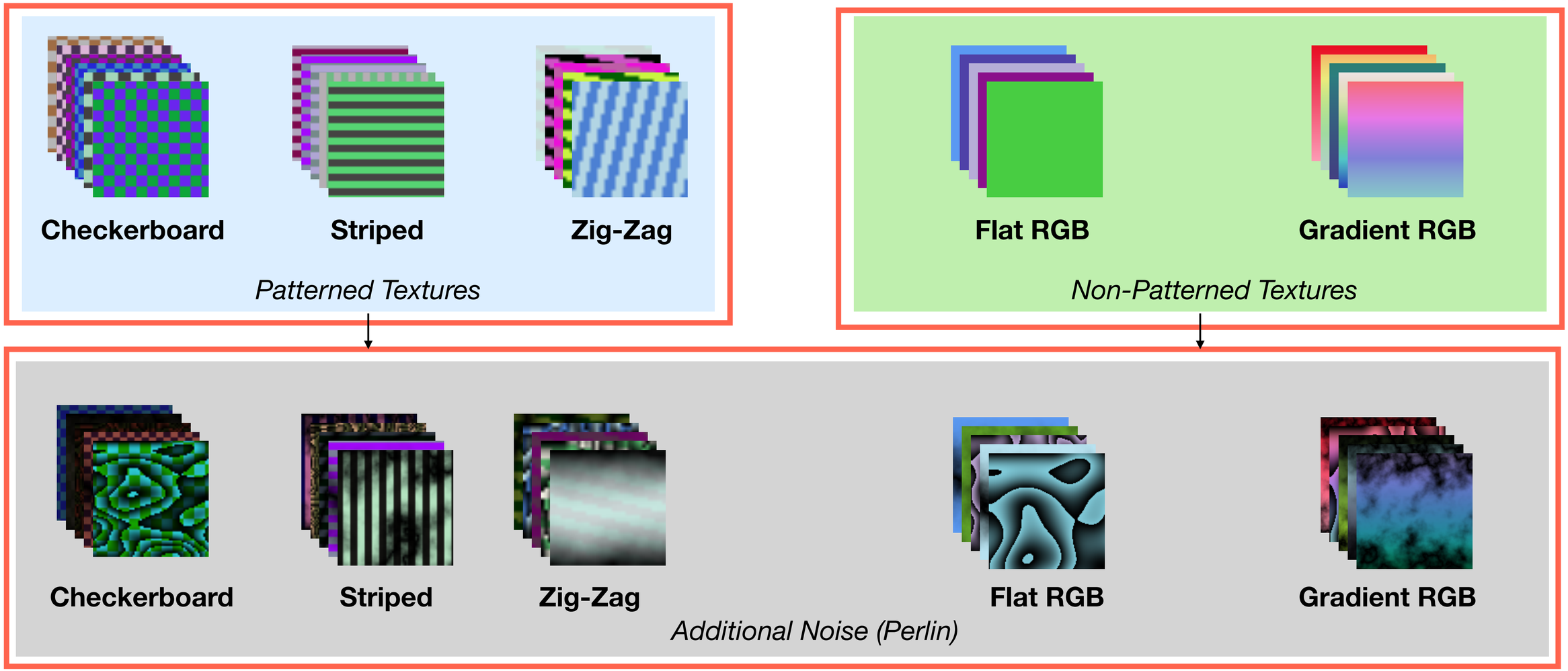


Detailed citations available in the paper

No consensus on the optimal approach to applying DR — leading to inefficient sampling and unguided application

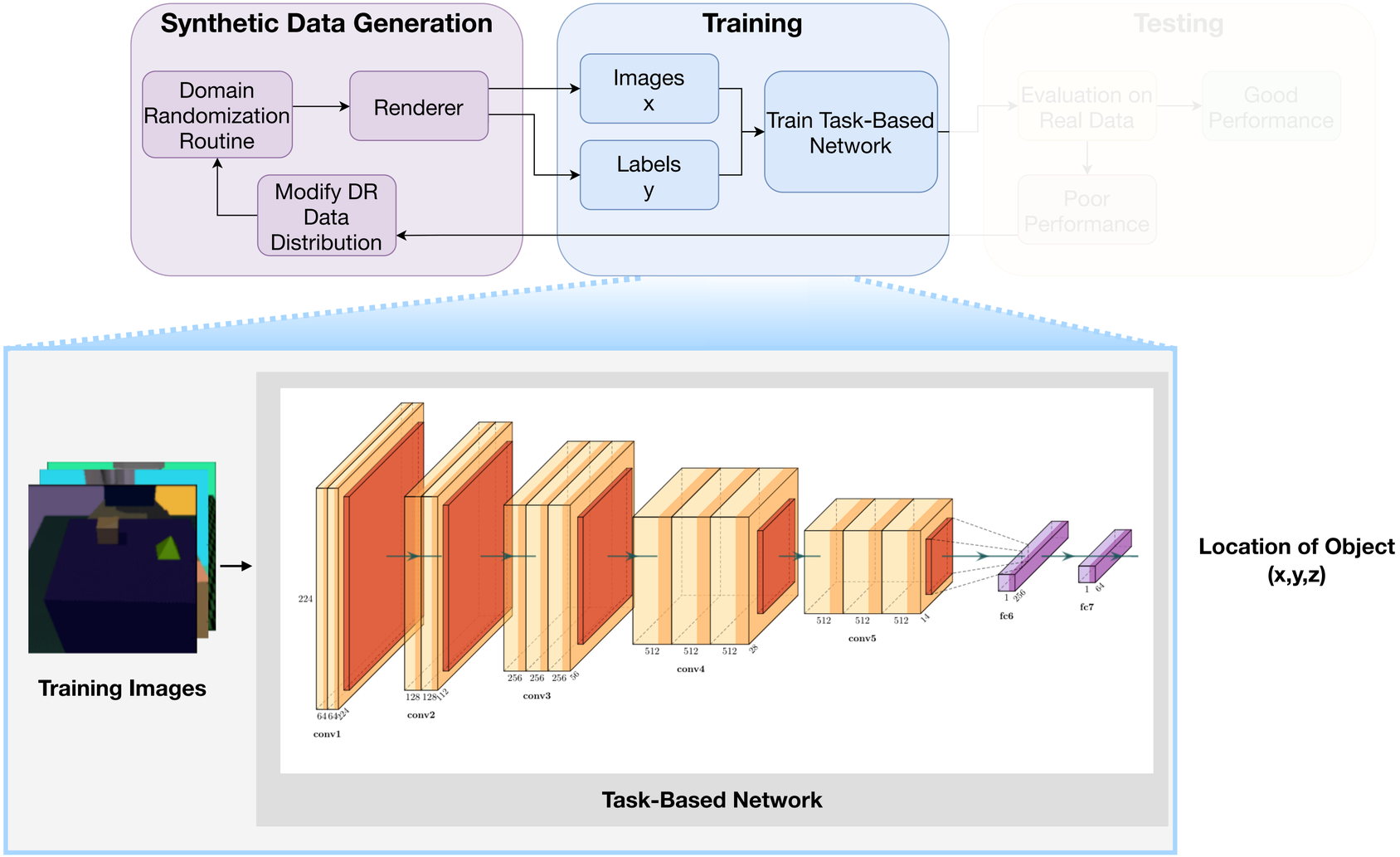
Texture Randomization

Large amount of possible textures — which textures would result in the highest performance to solve a given task?



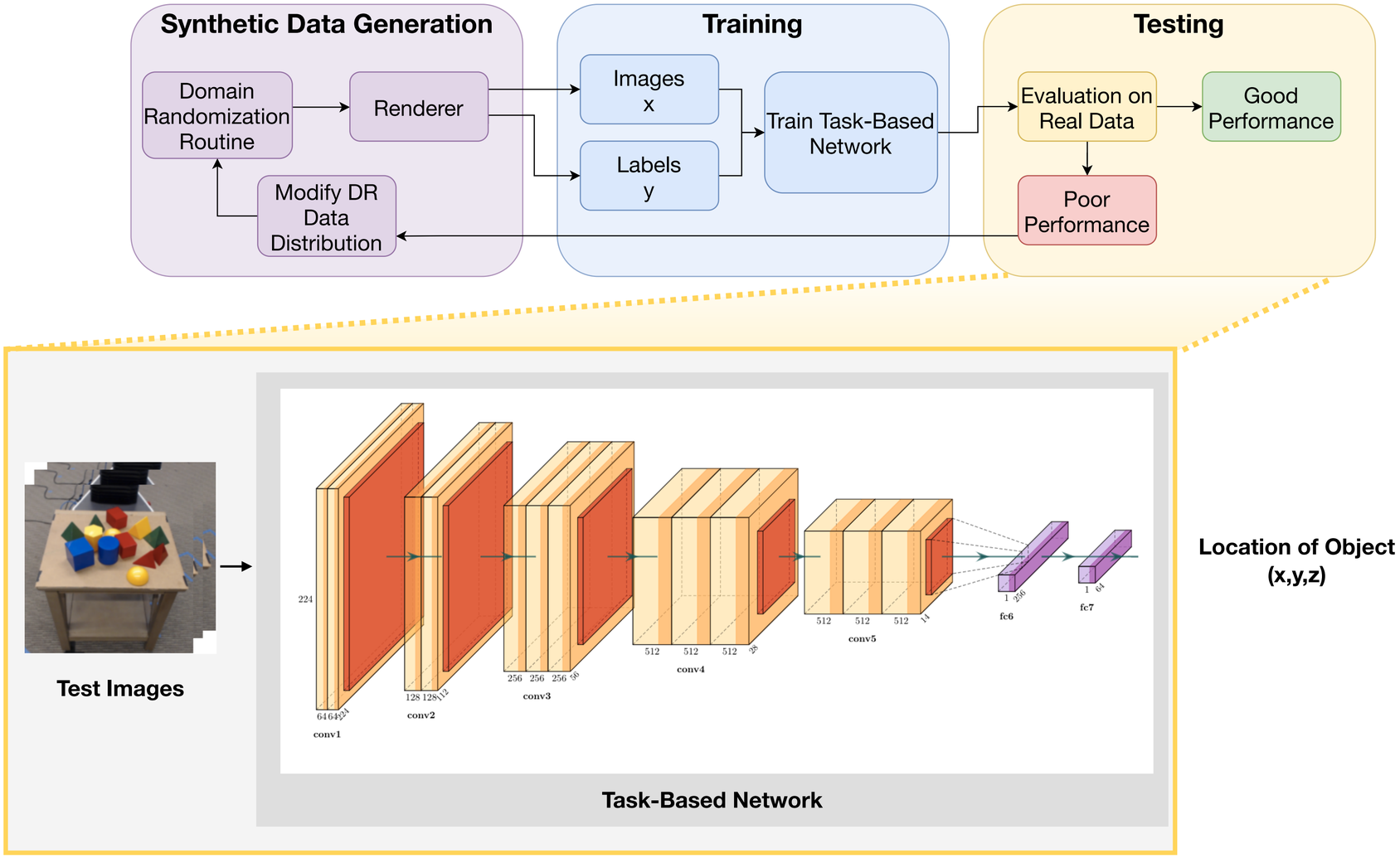
Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel,"Domain randomization for transferring deep neural networks from simulation to the real world," in IEEE/RSJ International Conference on Intelligent Robots and Systems IROS, 2017, pp. 23–30.
S. James, A. J. Davison, and E. Johns,"Transferring End-to-EndVisuomotor Control from Simulation to Real World for a Multi-StageTask," 1st Conference on Robot Learning (CoRL), 2017
J. Tremblay, T. To, B. Sundaralingam, Y. Xiang, D. Fox, and S. Birch-field, "Deep Object Pose Estimation for Semantic Robotic Grasping ofHousehold Objects," inConference on Robot Learning (CoRL), 2018

Typical Approach



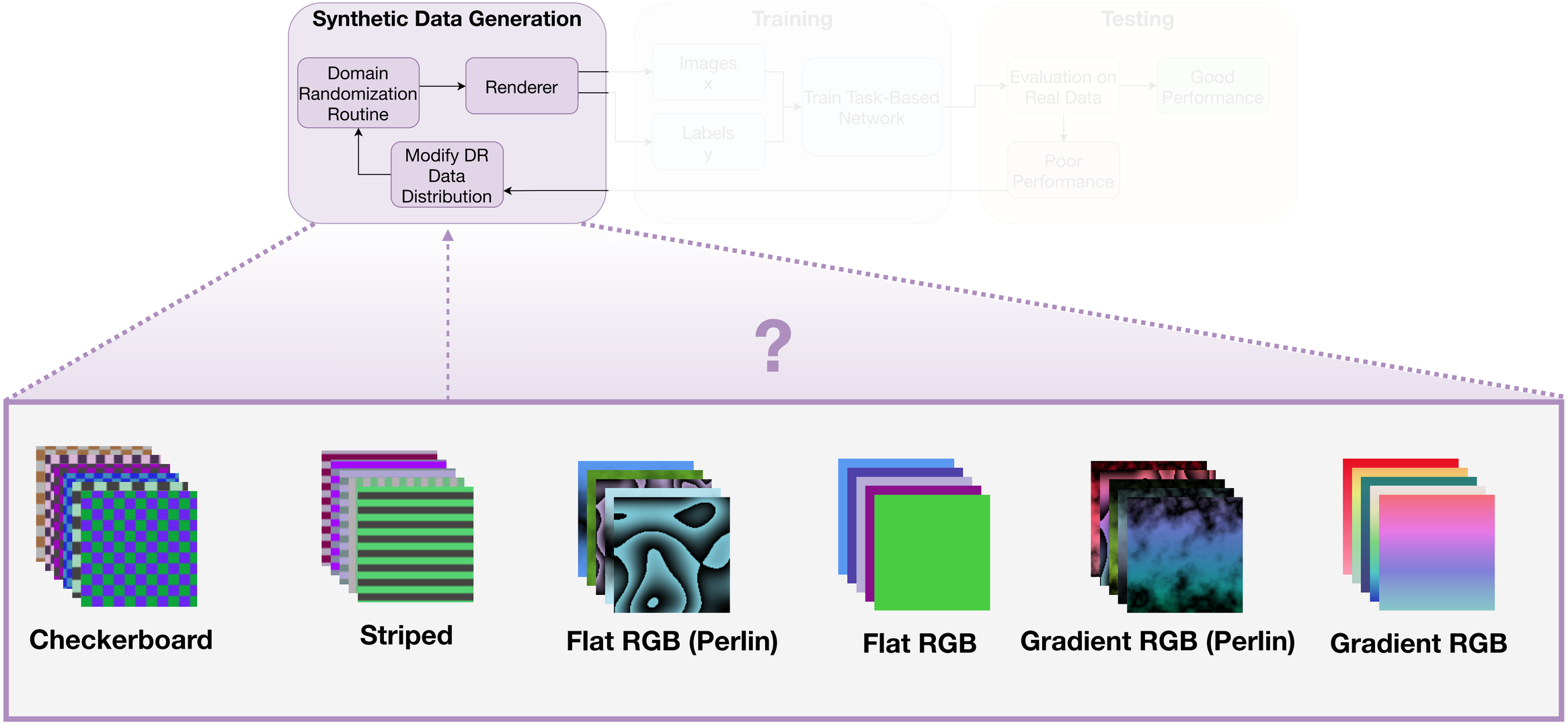
Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel,"Domain randomization for transferring deep neural networks from simulation to the real world," in IEEE/RSJ International Conference on Intelligent Robots and Systems IROS, 2017, pp. 23–30.

Typical Approach



Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel,"Domain randomization for transferring deep neural networks from simulation to the real world," in IEEE/RSJ International Conference on Intelligent Robots and Systems IROS, 2017, pp. 23–30.

Typical Approach

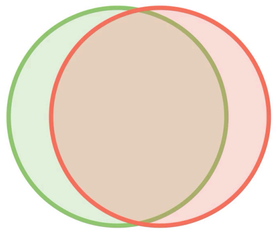


Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel,"Domain randomization for transferring deep neural networks from simulation to the real world," in IEEE/RSJ International Conference on Intelligent Robots and Systems IROS, 2017, pp. 23–30.

Goal

Quantifying the difference between **realistic** and **DR** data distributions.

DR data distribution



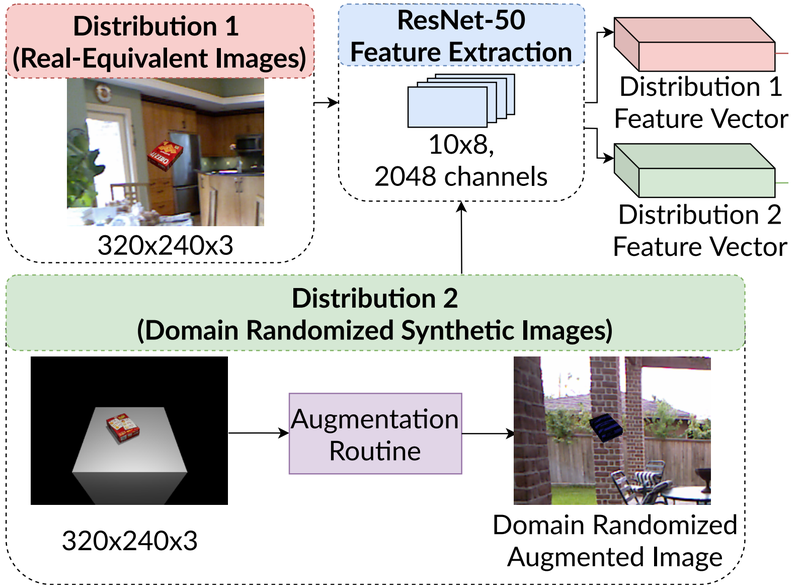
Real data distribution

	DR Technique	Distance Estimate
Lowest estimate yields	$\boxed{\theta_2}$	0.15
highest task-based performance	θ_3	0.39
	θ_1	0.70

Contributions

- **Novel** method of **quantifying** differences between **DR** data distributions and **real-equivalent** samples
- Our method **avoids** the **expense** of **task-specific** training and evaluation
- We show **lower distance estimates** between **DR** and **real-equivalent** distributions, generated **without** task-based training, corresponds to **higher task-based performance**

Proposed Solution



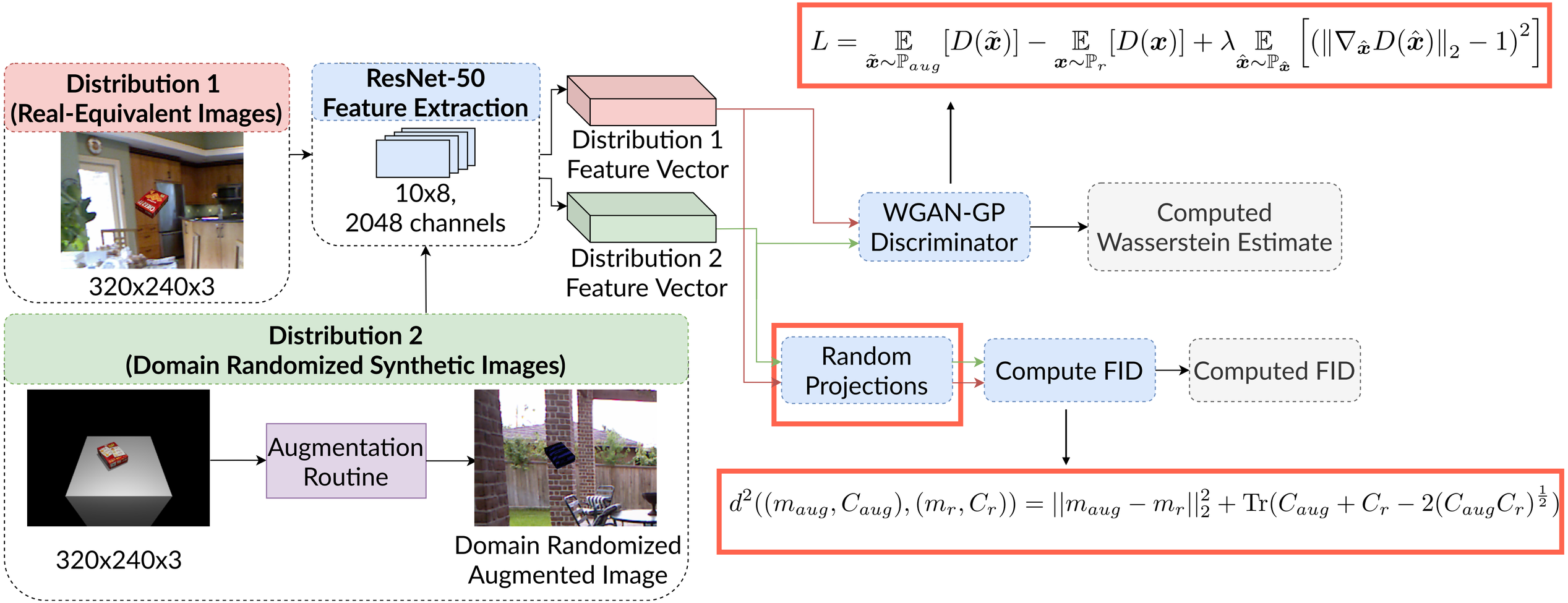
Fréchet Inception Distance (FID)

$$d^2((m_{aug}, C_{aug}), (m_r, C_r)) = ||m_{aug} - m_r||_2^2 + \mathrm{Tr}(C_{aug} + C_r - 2(C_{aug}C_r)^{\frac{1}{2}})$$

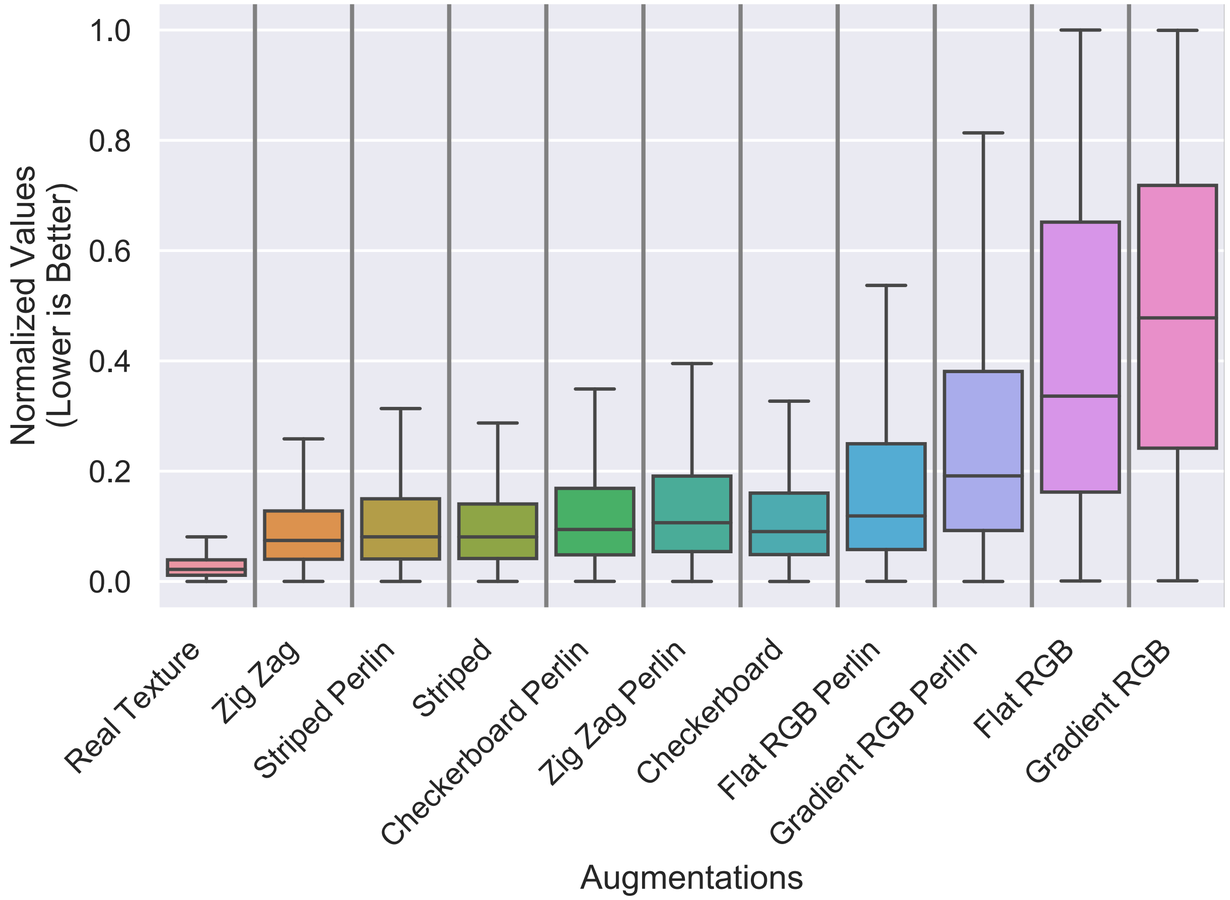
Wasserstein Distance

$$W(P_r, P_{aug}) = \inf_{\gamma \in \Pi(P_r, P_{aug})} \mathbb{E}_{(x,y)\sim\gamma}\left[||x - y||\right]$$

Method

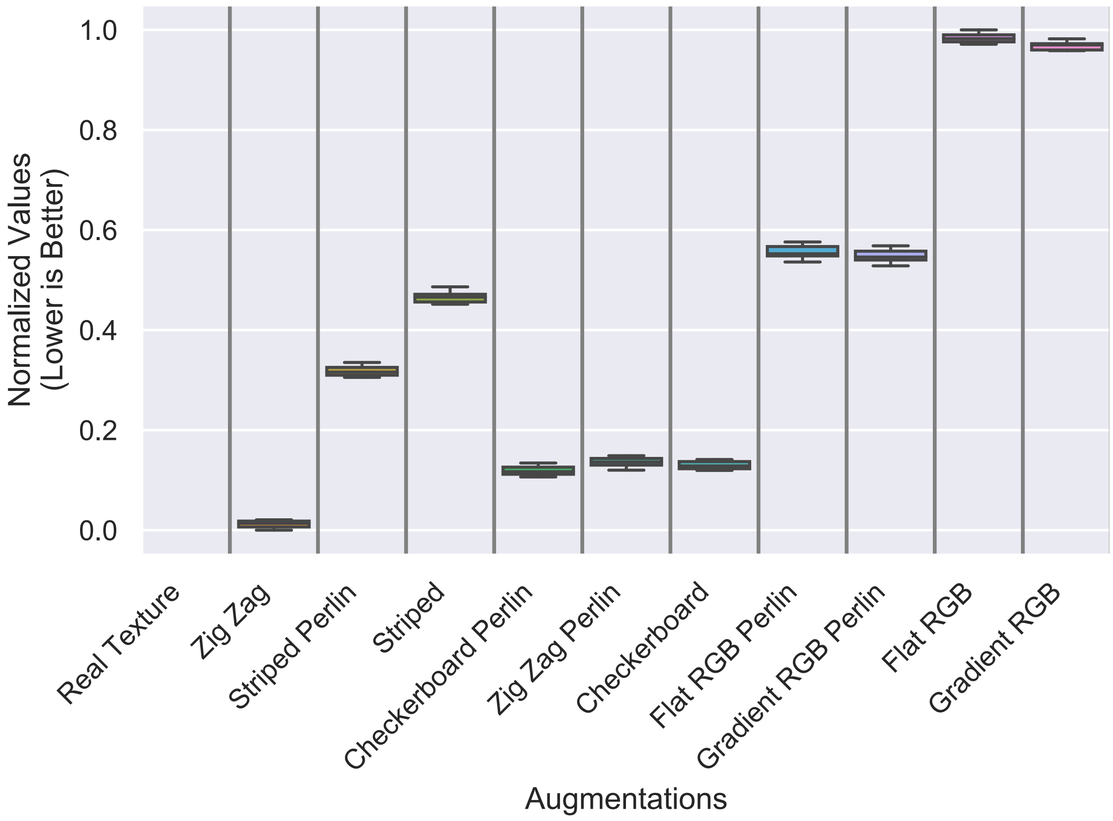

Distribution 1
(Real-Equivalent Images)
320x240x3
ResNet-50
Feature Extraction
10x8,
2048 channels
Distribution 1
Feature Vector
Distribution 2
Feature Vector
$L = \mathop{\mathbb{E}}_{\tilde{\boldsymbol{x}} \sim \mathbb{P}_{aug}} [D(\tilde{\boldsymbol{x}})] - \mathop{\mathbb{E}}_{\boldsymbol{x} \sim \mathbb{P}_r} [D(\boldsymbol{x})] + \lambda \mathop{\mathbb{E}}_{\hat{\boldsymbol{x}} \sim \mathbb{P}_{\hat{\boldsymbol{x}}}} \left[(\|\nabla_{\hat{\boldsymbol{x}}} D(\hat{\boldsymbol{x}})\|_2 - 1)^2 \right]$
WGAN-GP
Discriminator
Computed
Wasserstein Estimate
Distribution 2
(Domain Randomized Synthetic Images)
Augmentation
Routine
320x240x3
Domain Randomized
Augmented Image
Random
Projections
Compute FID
Computed FID
$d^2((m_{aug}, C_{aug}), (m_r, C_r)) = ||m_{aug} - m_r||_2^2 + \mathrm{Tr}(C_{aug} + C_r - 2(C_{aug} C_r)^{\frac{1}{2}})$

Results

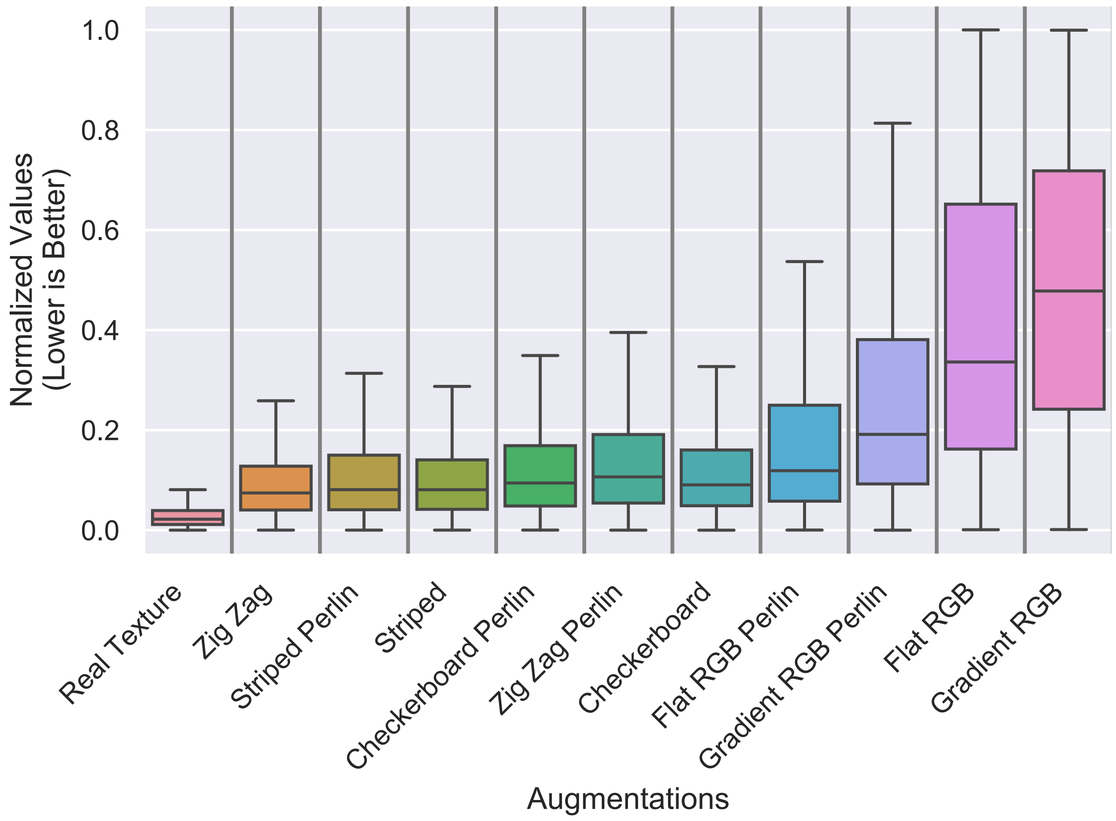


Localization Task Error

Results

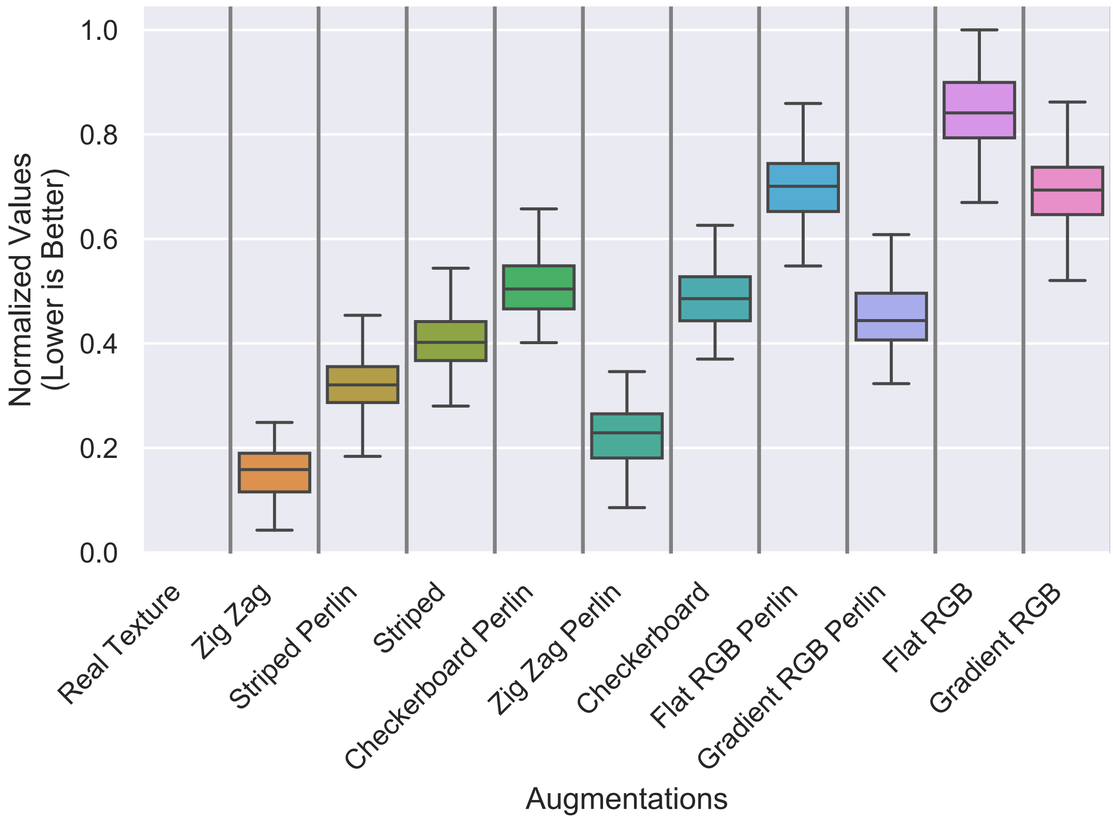


FID Estimate

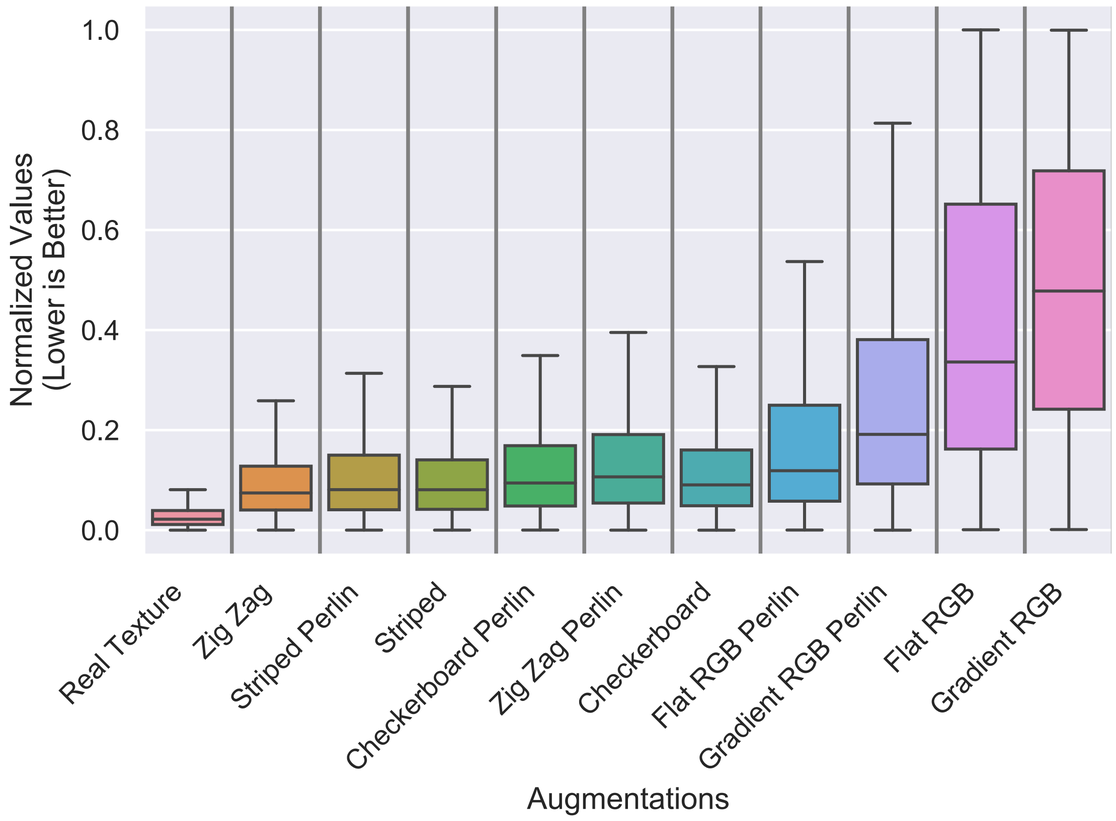


Localization Task Error

Results

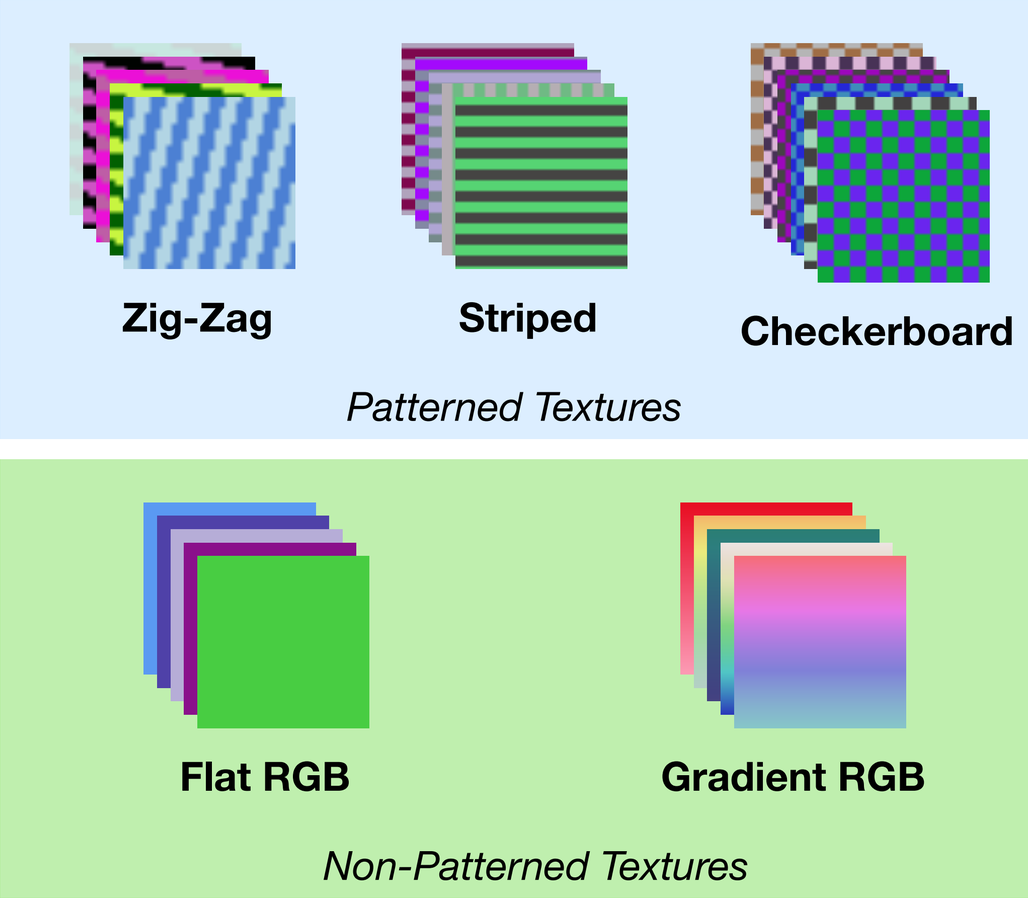


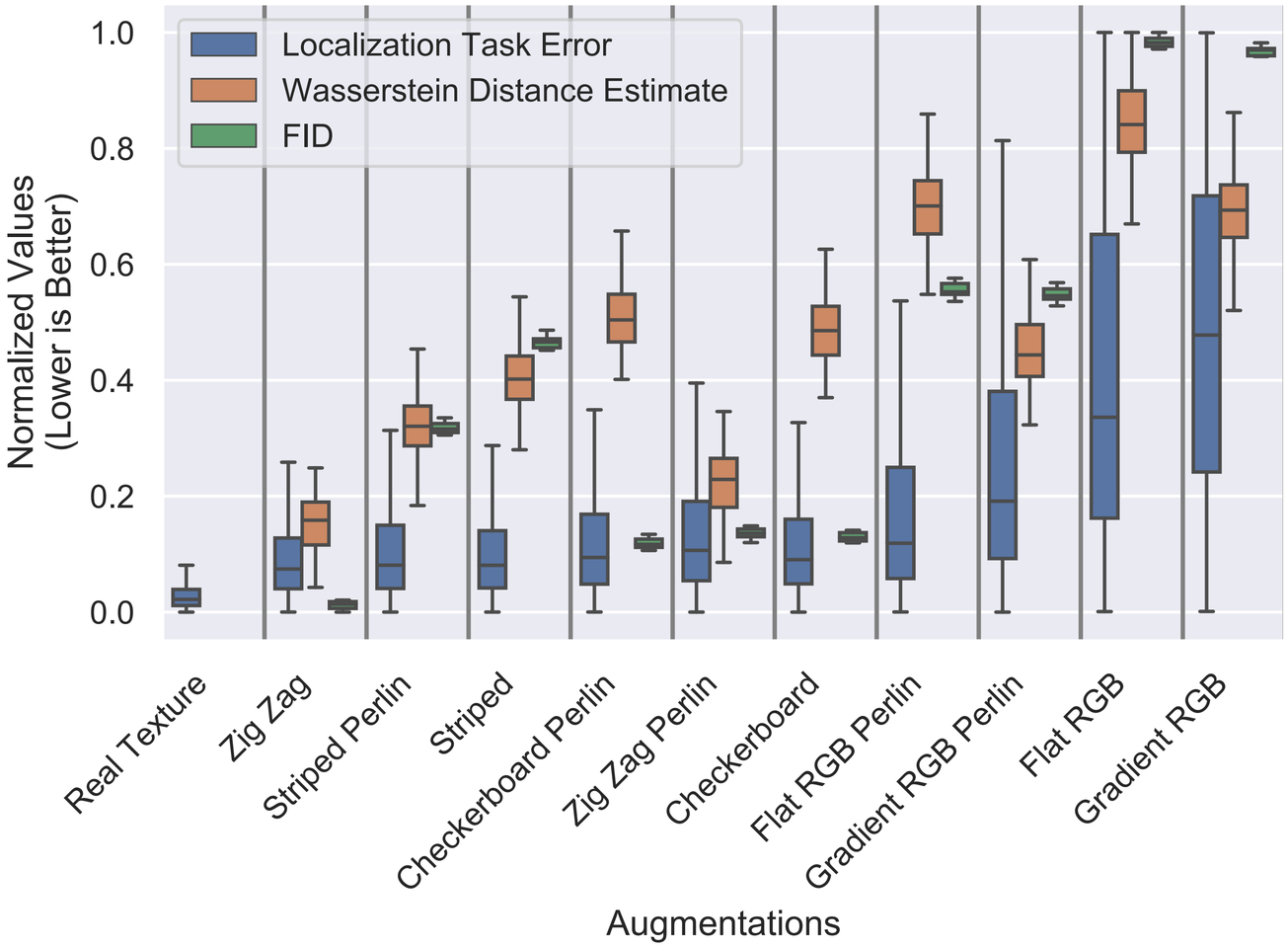
Wasserstein Distance Estimate



Localization Task Error

Results


Zig-Zag
Striped
Checkerboard
Patterned Textures
Flat RGB
Gradient RGB
Non-Patterned Textures

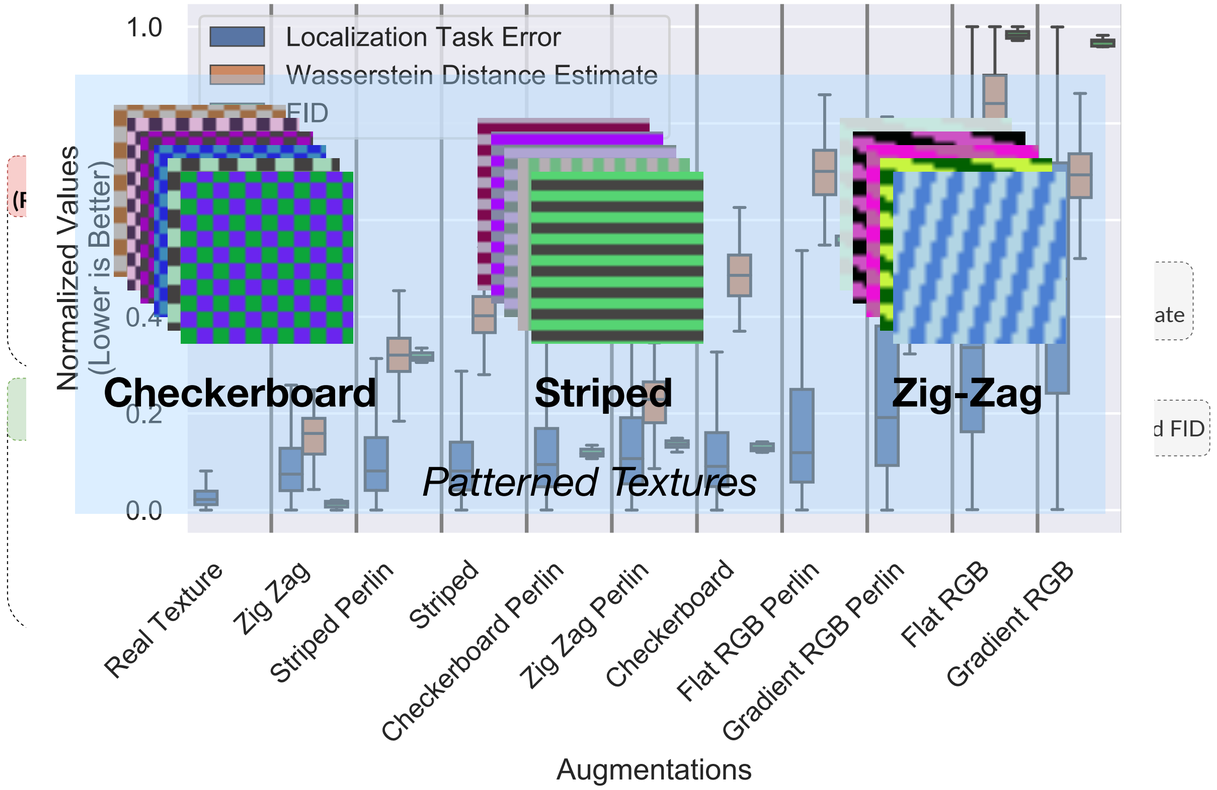

Localization Task Error
Wasserstein Distance Estimate
FID
Normalized Values
(Lower is Better)
1.0
0.8
0.6
0.4
0.2
0.0
Real Texture
Zig Zag
Striped Perlin
Striped
Checkerboard Perlin
Zig Zag Perlin
Checkerboard
Flat RGB Perlin
Gradient RGB Perlin
Flat RGB
Gradient RGB
Augmentations

Conclusion

- We propose a novel method of **quantifying** differences between **DR** data distributions and real-equivalent samples

- We demonstrate that the method is capable of **ranking** the different augmentations which is **reflected** in the performance of an object localization task

- Based on the produced ranking, generated **without** task-based training, we recommend using more **complex patterned** textures when generating DR synthetic data



Acknowledgement

We acknowledge MoD/Dstl and EPSRC for providing the grant to support the UK academics involvement in a Department of Defense funded MURI project through EPSRC grant EP/N019415/1.

UNIVERSITY OF BIRMINGHAM

UKRI Engineering and Physical Sciences Research Council

ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021