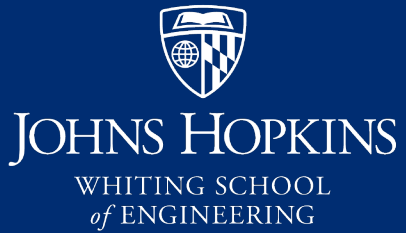
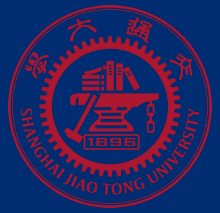




Shape-Texture Debiased Neural Network Training

Yingwei Li, Qihang Yu, Mingxing Tan, Jieru Mei, Peng Tang
Wei Shen, Alan Yuille, Cihang Xie

Google

Background: ICLR'19 Geirhos *et al.*

- ImageNet-trained CNNs are biased towards texture

(a) Texture image

81.4%	**Indian elephant**
10.3%	indri
8.2%	black swan

(b) Content image

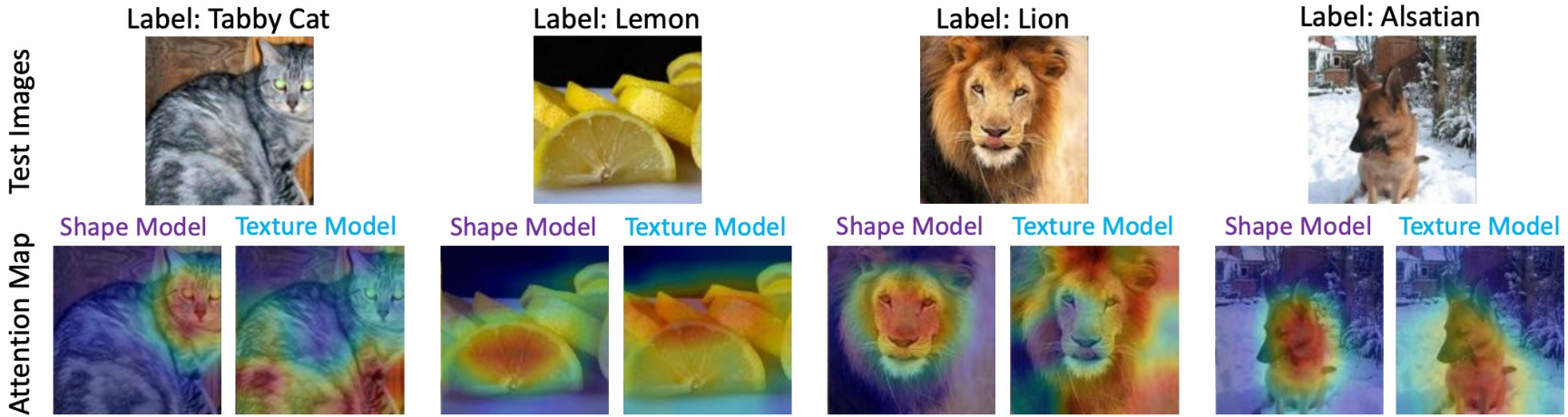
71.1%	**tabby cat**
17.3%	grey fox
3.3%	Siamese cat

(c) Texture-shape cue conflict

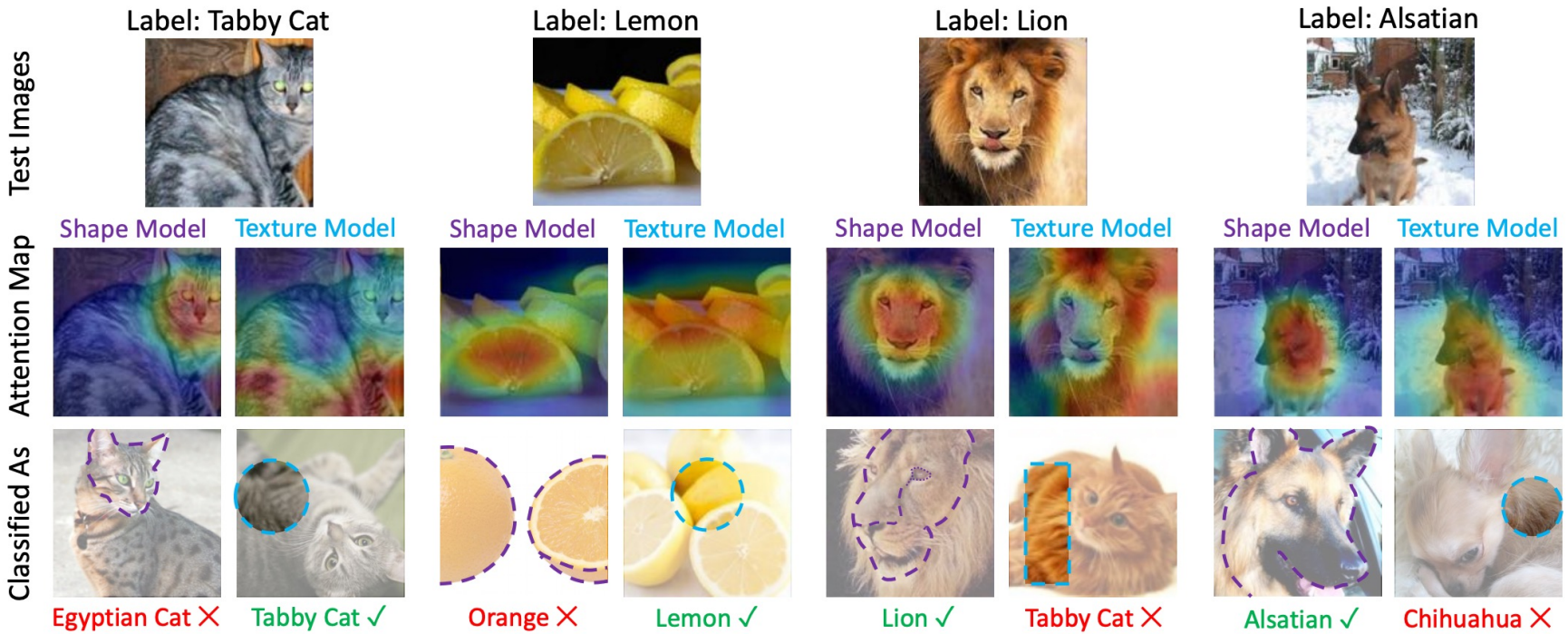
63.9%	**Indian elephant**
26.4%	indri
9.6%	black swan

Robert Geirhos et al. (2019). "Imagenet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness." In: ICLR.

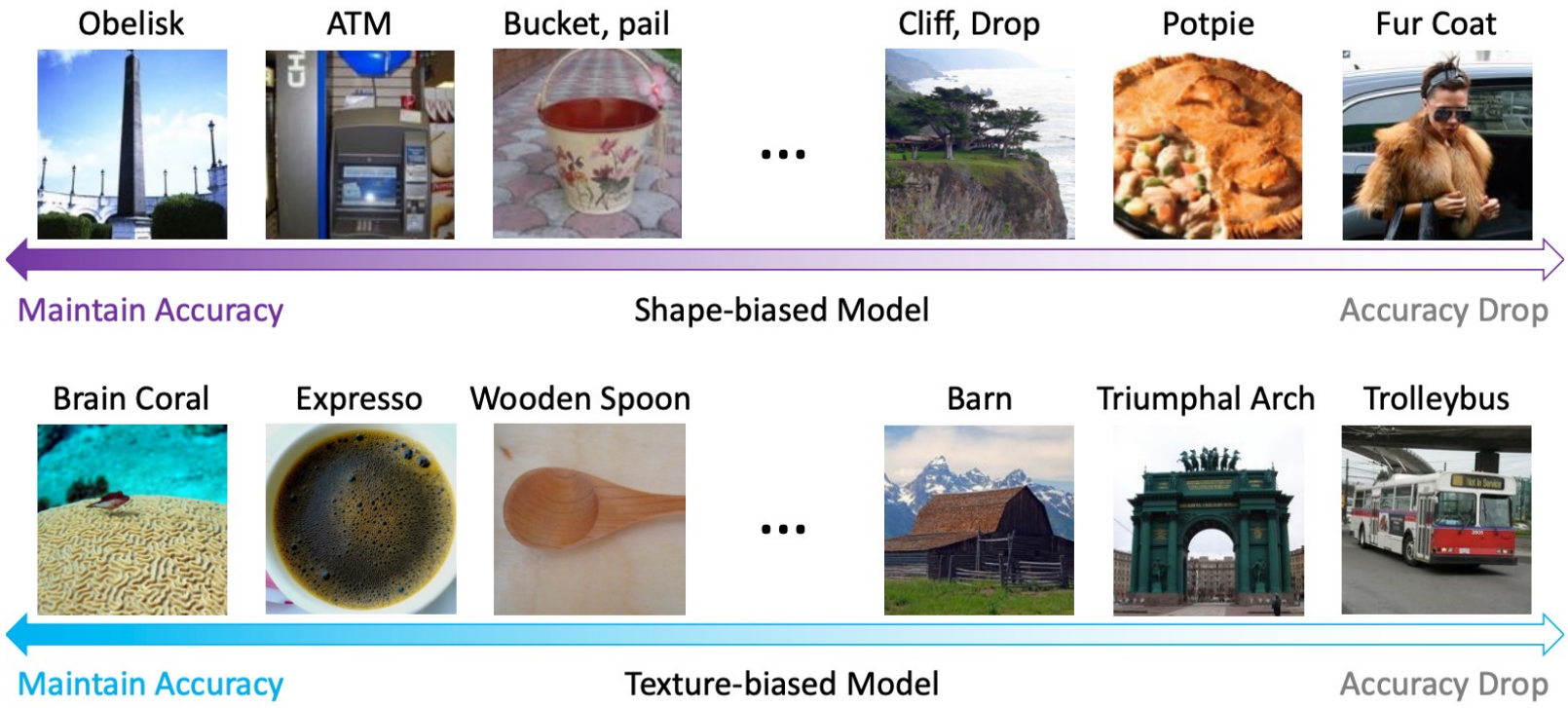
Shape and Texture: A Pair of Complementary Cues



Shape and Texture: A Pair of Complementary Cues



Shape and Texture: A Pair of Complementary Cues



Goal: Use Both Cues to Improve the Recognition Ability

Test Image Label: Fur Coat

Shape-biased Model
Prediction: Poncho ✗

✓Shape ✗Texture

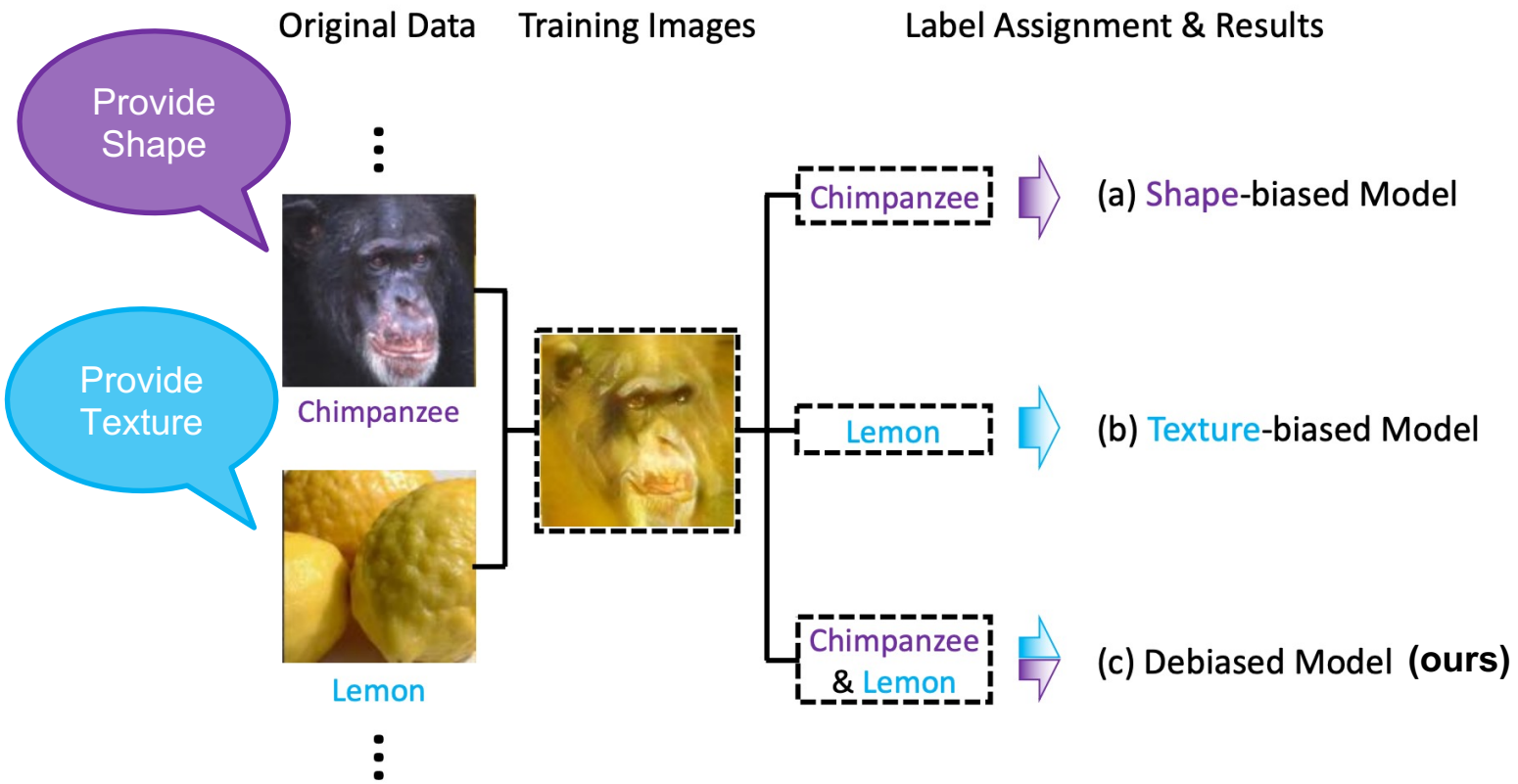
Texture-biased Model
Prediction: Egyptian cat ✗

✗Shape ✓Texture

Debiased Model
Prediction: Fur Coat ✓

✓Shape ✓Texture

Shape-Texture Debiased Neural Network Training



For Segmentation

Texture Source Image & Label

Texture Source Object

Shape Source Image & Label

Generated Image & Label

Improves Accuracy and Robustness

	CLEAN Top-1 Acc.↑	IMAGENET-A Top-1 Acc.↑	IMAGENET-C mCE↓	S-IMAGENET Top-1 Acc.↑	FGSM Top-1 Acc.↑
ResNet-50	76.4	2.0	75.0	7.4	17.1
Debiased	76.9(+0.5)	3.5(+1.5)	67.5(-7.5)	17.4(+10.0)	27.4(+10.3)
ResNet-101	77.9	5.6	69.8	9.9	23.1
Debiased	78.9(+1.0)	9.1(+3.5)	62.2(-7.6)	22.0(+12.1)	34.4(+11.3)
ResNet-152	78.6	7.4	67.2	11.3	25.2
Debiased	79.8(+1.2)	12.6(+5.2)	58.9(-8.3)	22.4(+11.1)	39.6(+14.4)

Table 2: Model robustness comparison evaluated on ImageNet-A (%), on ImageNet-C (mCE), on Stylized-ImageNet (S-ImageNet, %), and under FGSM attack (%). The debiased neural network training significantly improves the model robustness.

Compare with State-of-the-art Methods

	IN Acc. ↑	IN-A Acc. ↑	IN-C mCE ↓	S-IN Acc. ↑	FGSM Acc. ↑
ResNet-50	76.4	2.0	75.0	7.4	17.1
CutMix + MoEx (Li et al., 2021)	79.0	8.0	74.8	**5.0**	41.0
DeepAugment + AugMix (Hendrycks et al., 2020)	**75.8**	3.9	53.6	21.2	18.8
SIN (Geirhos et al., 2019)	**60.2**	2.4	**77.3**	56.2	**5.6**
Shape-Texture Debiased Training (ours)	76.9	3.5	67.5	17.4	27.4

Results on Shape / Texture Datasets

Datasets	VANILLA	S-BIASED	T-BIASED	DEBIASED
ImageNet-Sketch	23.8	27.9	24.3	**28.4**
ImageNet-R	36.2	40.6	36.7	**40.8**
Kylberg Texture	99.5	99.1	**99.6**	99.5
Flicker Material	74.6	73.3	**79.2**	75.8

Thank you.

Any questions?

Email: yingwei.li@jhu.com
Website: https://yingwei.li/
Arxiv: https://arxiv.org/pdf/2010.05981.pdf
Code: https://github.com/LiYingwei/ShapeTextureDebiasedTraining