UMat: Uncertainty-Aware Single Image High Resolution Material Capture

Carlos Rodriguez-Pardo[1,2], Henar Dominguez-Elvira[1,2], David Pascual-Hernandez[1], Elena Garces[1,2]
SEDDI[1], Universidad Rey Juan Carlos[2]

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

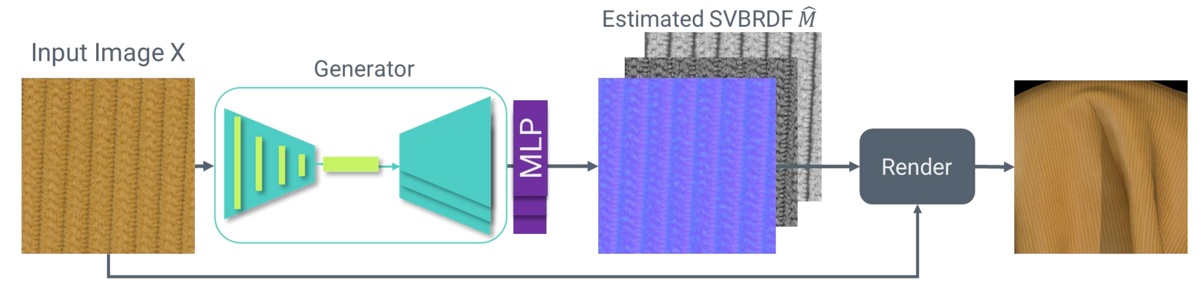
Summary

- Single Image High Resolution SVBRDF Estimation
- Flatbed Scanners as Capture Device
- Reflectance Estimation Using Microgeometry
- Uncertainty Quantification in Material Capture

SVBRDF

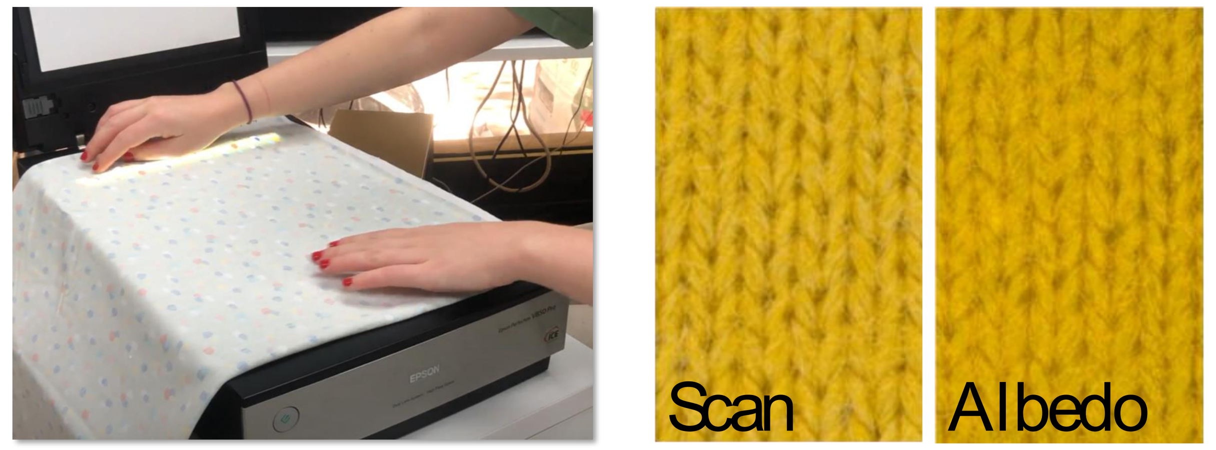
Burley. Physically-based shading at Disney (2012)

$$f_{l,v}(\mathbf{M},\mathbf{X}) = \frac{\mathbf{X}}{\pi} + s_{l,v}(\mathbf{M}) \in \mathbb{R}^{x \times y}$$



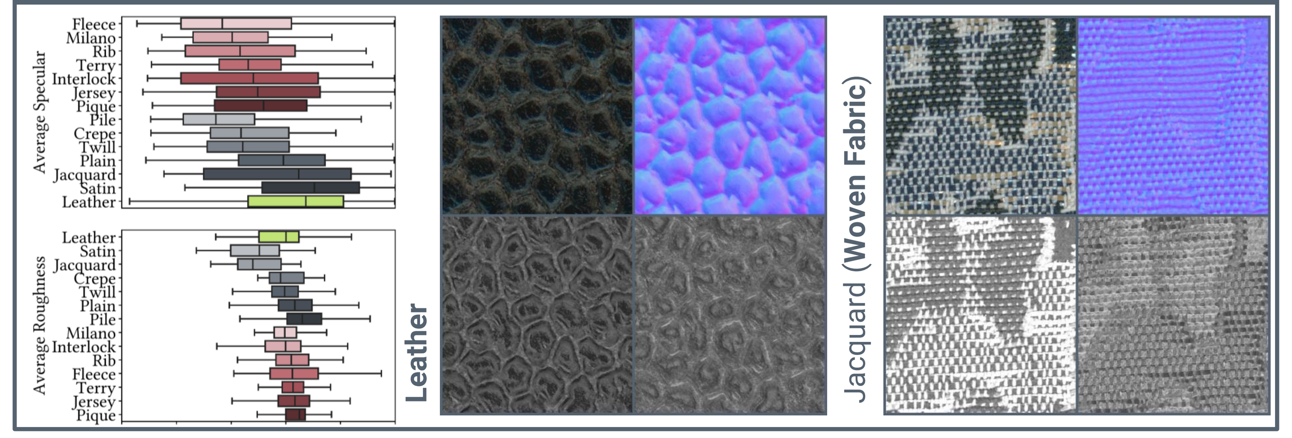
X: Albedo (Input) **M (G(X)):** Normals, Roughness, Specular

Capture with Flatbed Scanners



- Scalable digitization at **very high resolutions**.
- Some scanners provide **diffuse-like illumination** which resembles albedos.

Key Idea: Microgeometry as Reflectance Cue



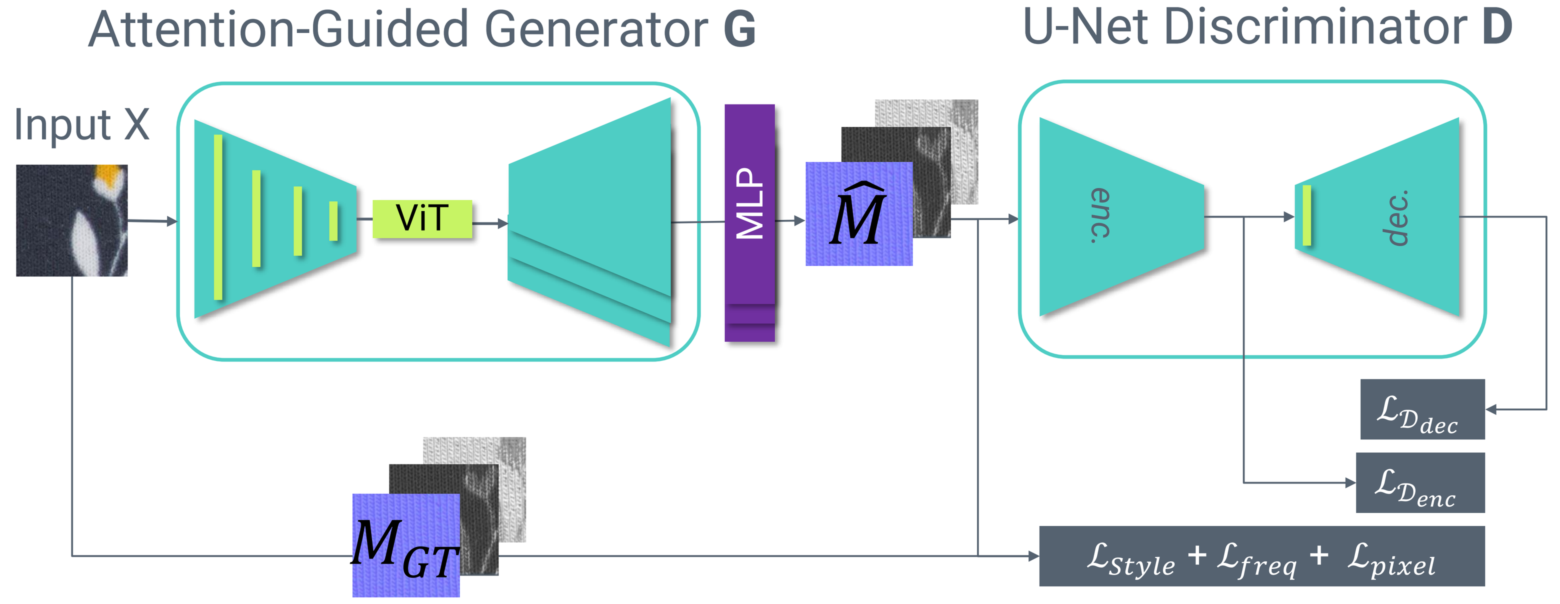
Model Overview

GAN tailored for material digitization.

Attention: ↑Accuracy and no artifacts.

One specialized decoder for each map: Improved quality and accuracy

U-Net Discriminator: Large improvements compared to baselines.



Multiple loss functions to maximize quality and accuracy.

$$\mathcal{L}_G = \sum_i \lambda_i \mathcal{L}_{pixel_i} + \lambda_{adv}\mathcal{L}_{adv} + \lambda_{style}\mathcal{L}_{style} + \lambda_{freq}\mathcal{L}_{freq}$$

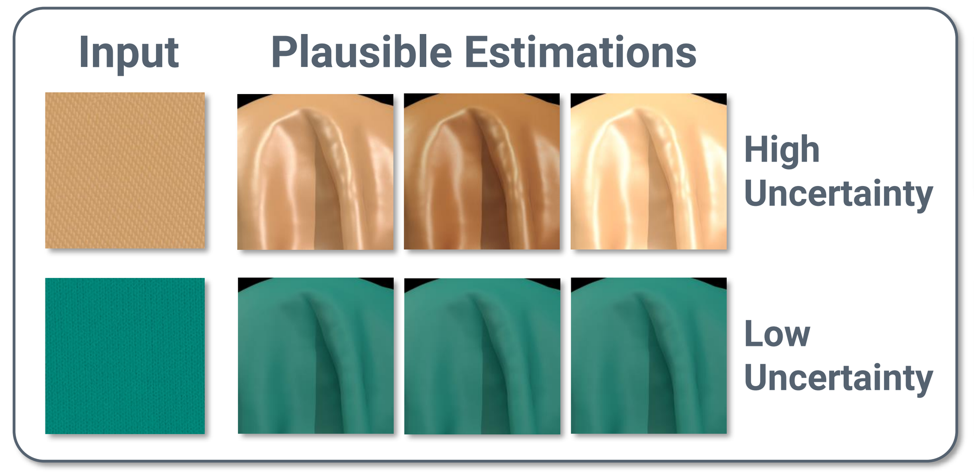
$$\mathcal{L}_{\mathcal{D}} = \mathcal{L}_{\mathcal{D}_{enc}} + \mathcal{L}_{\mathcal{D}_{dec}} + \lambda_{cons}\mathcal{L}^{cons}_{\mathcal{D}_{dec}}$$

$$\mathcal{L}_{adv} = \log(\mathcal{D}_{enc}(\mathrm{G}(\mathrm{X}))) + \log(\mathcal{D}_{dec}(\mathrm{G}(\mathrm{X})))$$

Frequency loss [5] compares the Fourier expectrum of $\hat{M}$ and M_{GT}

Style loss [6] compares $\hat{M}$ and M_{GT} perceptually, using LPIPS.

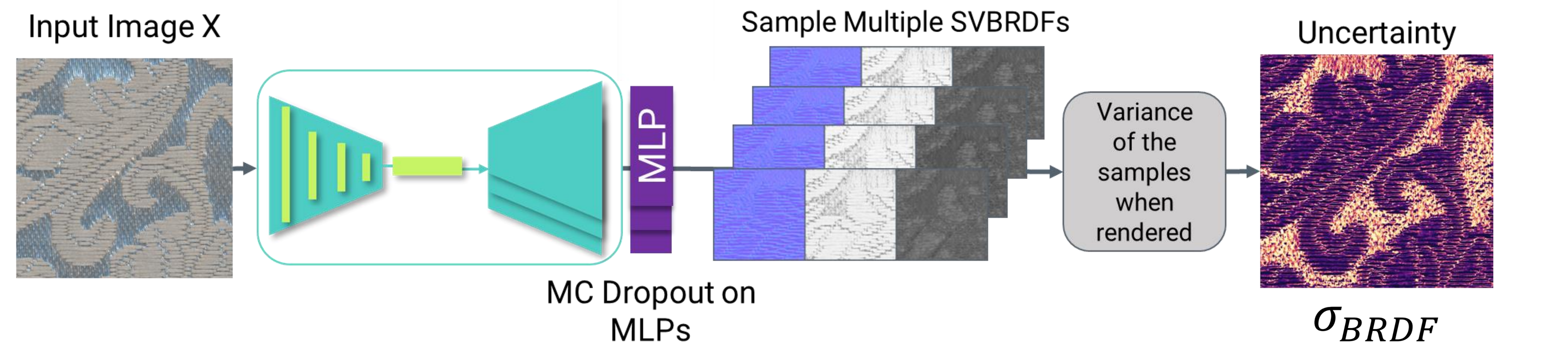
BRDF Evaluation and Uncertainty Quantification



Single image material estimation is an **ill-posed problem**

We introduce the **first uncertainty quantification method for material capture**

BRDF Estimation Uncertainty σ_{BRDF}



$$\sigma_{\text{BRDF}} = \frac{1}{|xy|}\sum_{xy}\log\left(\frac{1}{|S|}\sqrt{\sum_{(l,v)\in S}\sqrt[3]{\sigma_{l,v}(\{f_{l,v}(\mathrm{U}_j, K)\cos(\theta_l)\}_{j=1}^{N})}}\right)$$

We sample a set (**U**) of estimations using **Monte Carlo Dropout** on the generator MLPs.

Variance across renders generated at a set of light (**l**) and cameras (**v**) positions

We can efficiently measure uncertainty without dedicated training or model design.

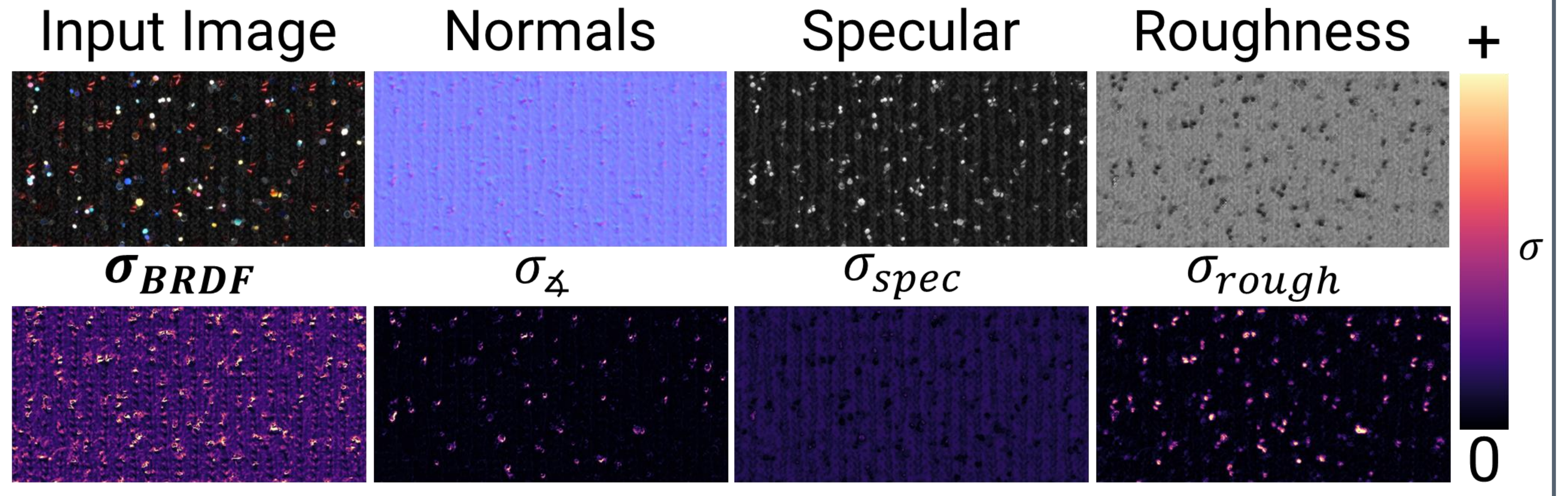
BRDF Estimation Error $\mathcal{L}_{BRDF}$

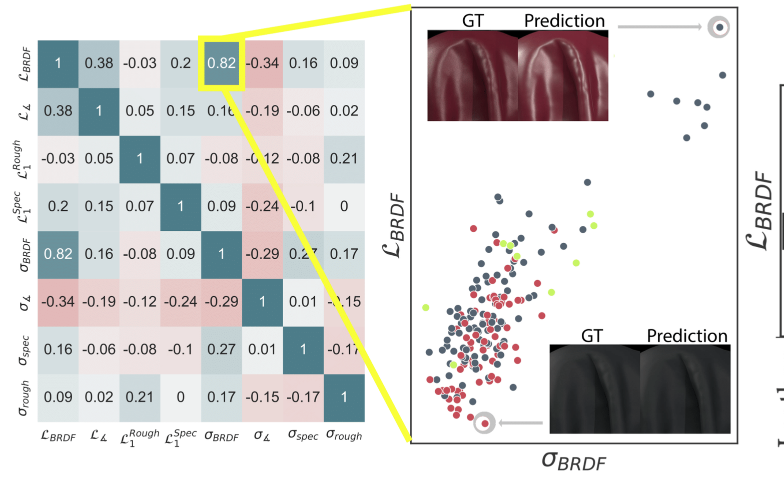
$$\mathcal{L}_{\text{BRDF}} = \frac{1}{|xy|}\sum_{xy}\sqrt{\frac{1}{|S|}\sum_{(l,v)\in S}\sqrt[3]{\cos^2(\theta_l)\left(f_{l,v}(\mathbf{M}_{GT},K) - f_{l,v}(\hat{\mathbf{M}},K)\right)^2}}$$

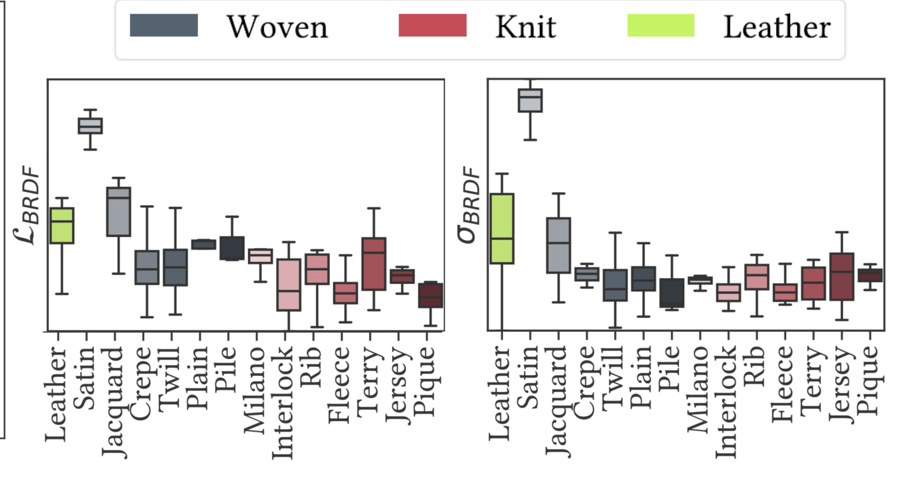
Average **distance** between renders generated at a set of light (**l**) and cameras (**v**) positions

Constant grayscale albedo K to isolate the impact of the accuracy of $\hat{M}$

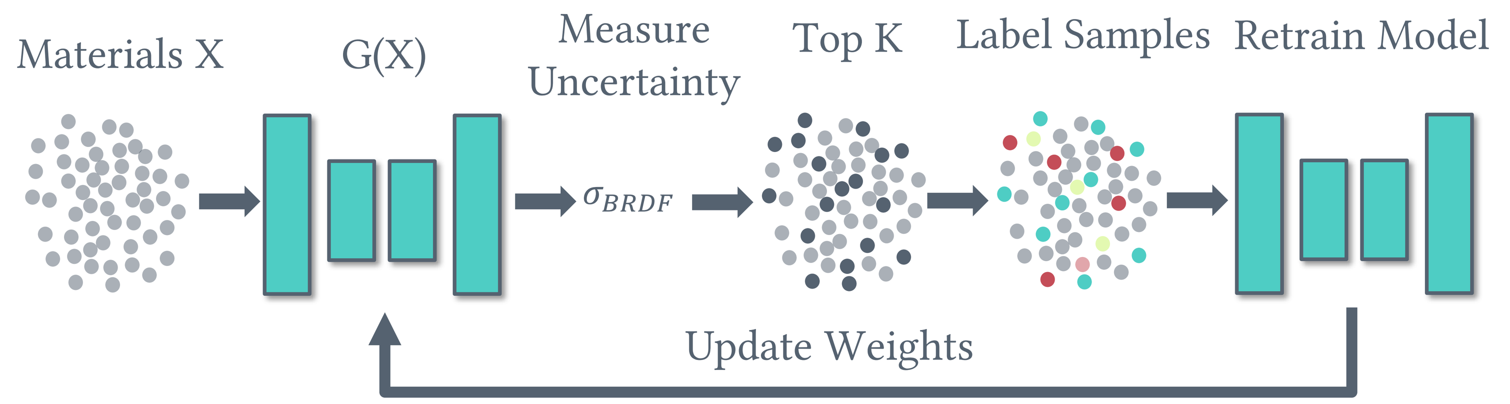
Perceptually Motivated: Specular Peak attenuation, Cosine Weighting

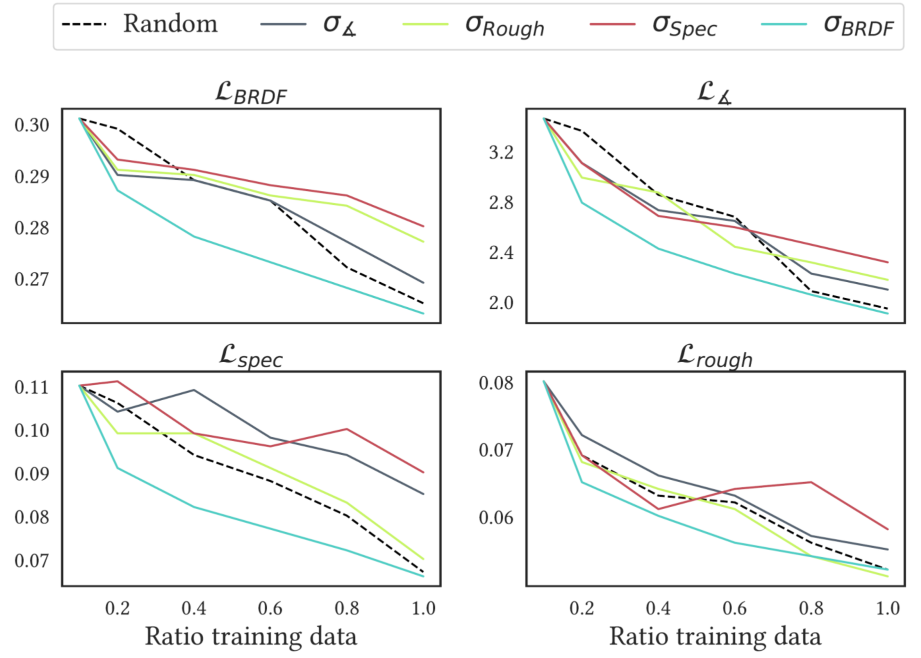






Active Learning

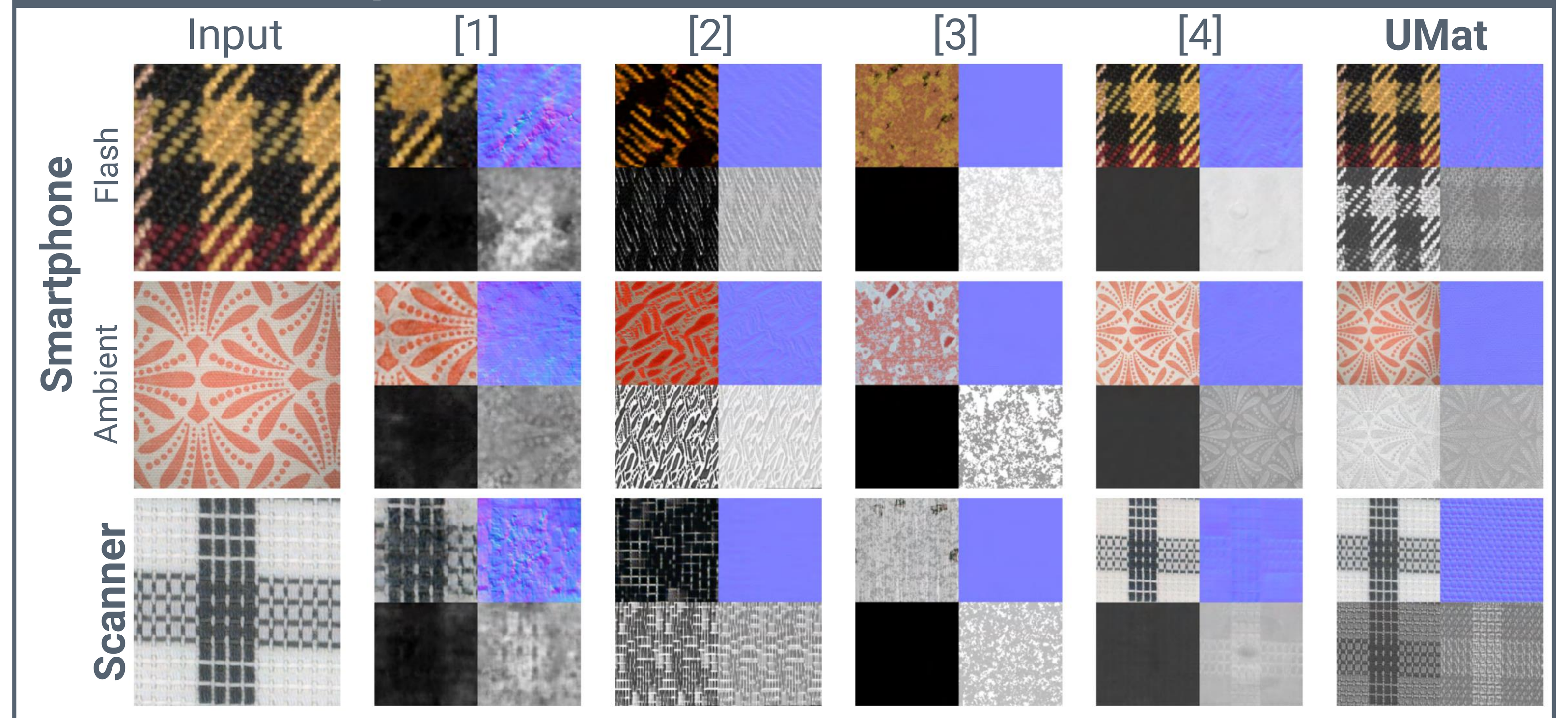




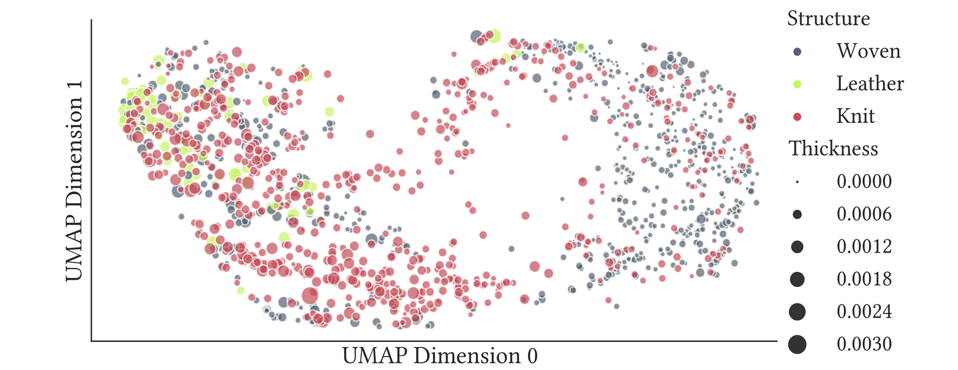
Uncertainty sampling to actively selecting which samples to label.

With active learning, we reduce the need of training data by factors of >2.

Comparisons with Previous Work



Latent Embeddings



References

[1] Duan Gao, Xiao Li, Yue Dong, Pieter Peers, Kun Xu, and Xin Tong. **Deep inverse rendering for high-resolution svbrdf estimation from an arbitrary number of images**. ACM Transactions on Graphics 2019
[2] Philipp Henzler, Valentin Deschaintre, Niloy J Mitra, and Tobias Ritschel. **Generative modelling of brdf textures from flash images**. ACM Transactions on Graphics (SIGGRAPH Asia) 2021
[3] Liang Shi, Beichen Li, Milos Ha ˇ san, Kalyan Sunkavalli, Tamy Boubekeur, Radomir Mech, and Wojciech Matusik. **Match: differentiable material graphs for procedural material capture**. ACM Transactions on Graphics, 2020.
[4] Xilong Zhou and Nima Khademi Kalantari. **Adversarial single-image svbrdf estimation with hybrid training**. In Computer Graphics Forum, 2021.
[5]: Liming Jiang, Bo Dai, Wayne Wu, and Chen Change Loy. **Focal frequency loss for image reconstruction and synthesis.** ICCV 2021
[6]: Thomas Chambon, Eric Heitz, and Laurent Belcour. Passing multi-channel material textures to a 3-channel loss. SIGGRAPH 2021 Talks. 2021.