SIGGRAPH 2023
LOS ANGELES+ 6-10 AUG

THE PREMIER CONFERENCE & EXHIBITION ON
COMPUTER GRAPHICS & INTERACTIVE TECHNIQUES

ULTRA-HIGH RESOLUTION SVBRDF RECOVERY FROM A SINGLE IMAGE

Jie Guo[1] Shuichang Lai[1] Qinghao Tu[1] Chengzhi Tao[1]
Changqing zou[2] Yanwen Guo[1]

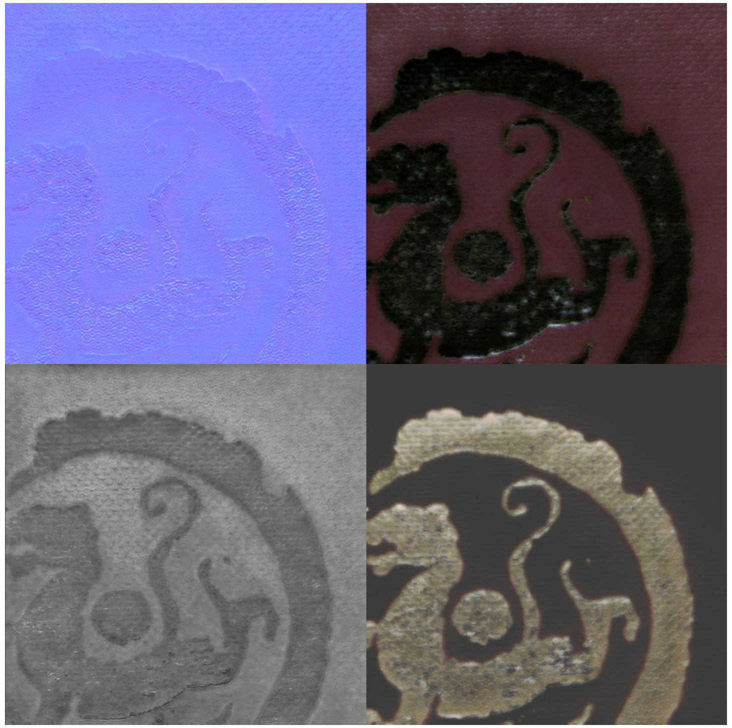
[1] STATE KEY LAB FOR NOVEL SOFTWARE TECHNOLOGY, NANJING UNIVERSITY
[2] STATE KEY LAB OF CAD&CG, ZHEJIANG UNIVERSITY

SVBRDF RECOVERY

Input image

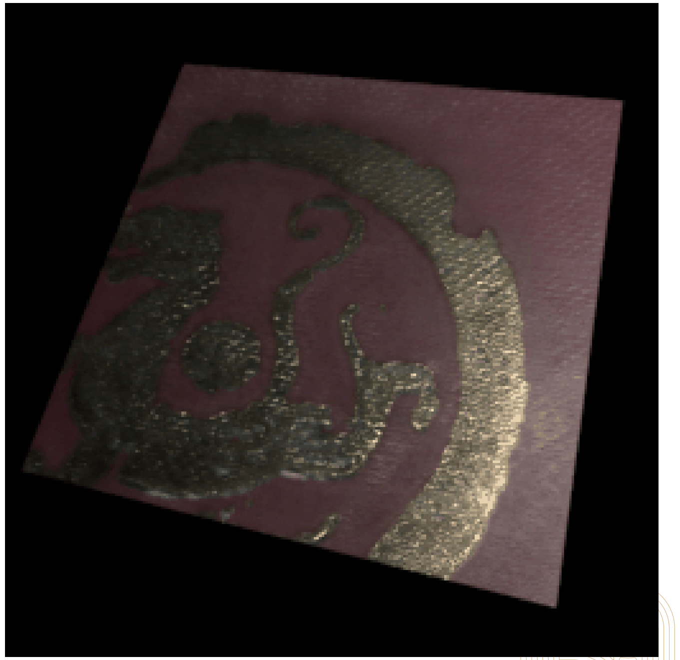
SVBRDF maps

normal | diffuse



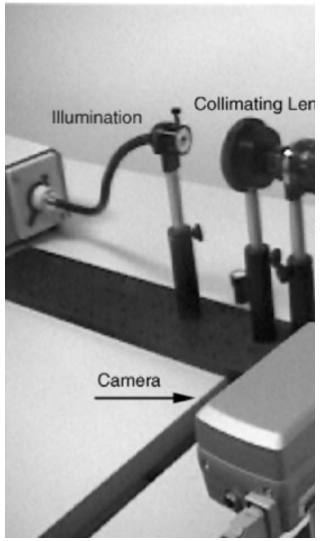
roughness | specular

Rendering



Specialized hardware

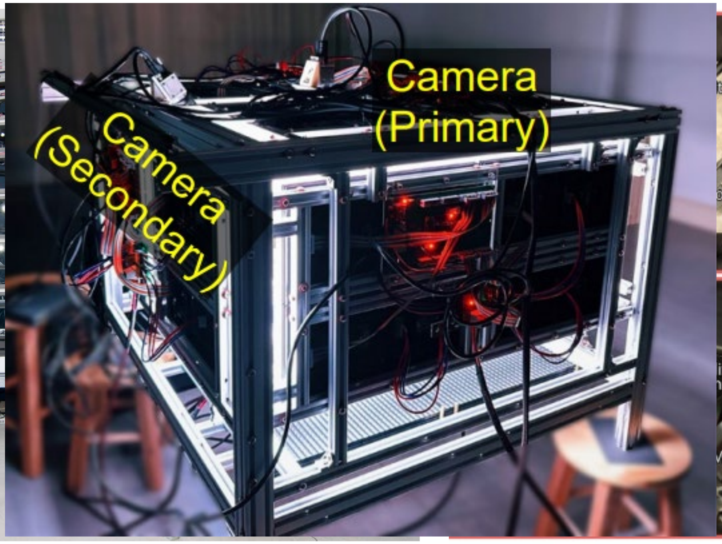
- High quality
- Dense sampling of light/view angle pairs
- Requiring a prohibitive amount of time



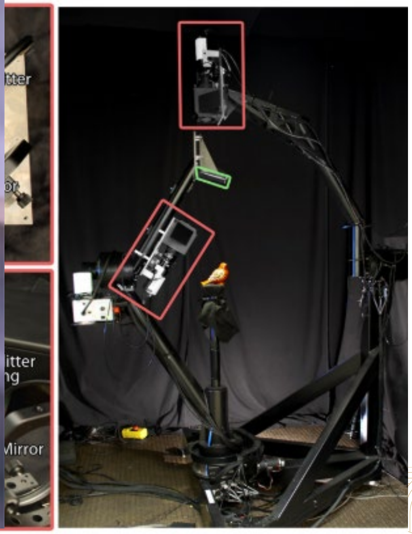
Dana and Wang 2004

Ashok et al. 2020

C. Schwartz et al. University of Bonn's Dome II



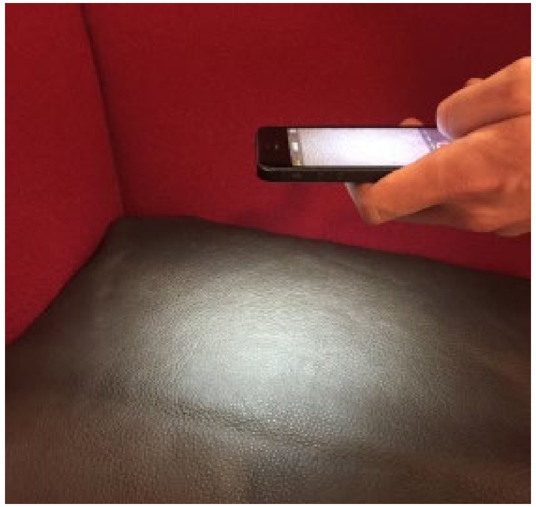
Masia et al. 2023



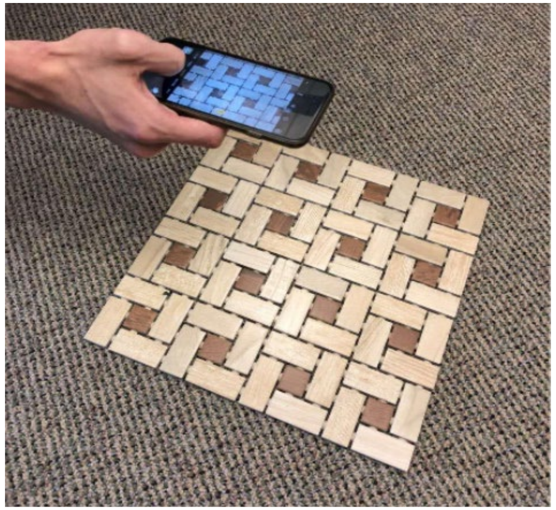
Holroyd et al. 2010

Light-weight setups

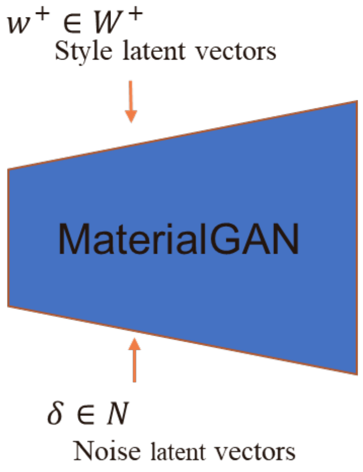
- A great trend in SVBRDF acquisition
- Fast, convenient, low-cost
- Generally only using a cell phone with flash
- Relying on **deep neural networks** (in particular **CNNs)**



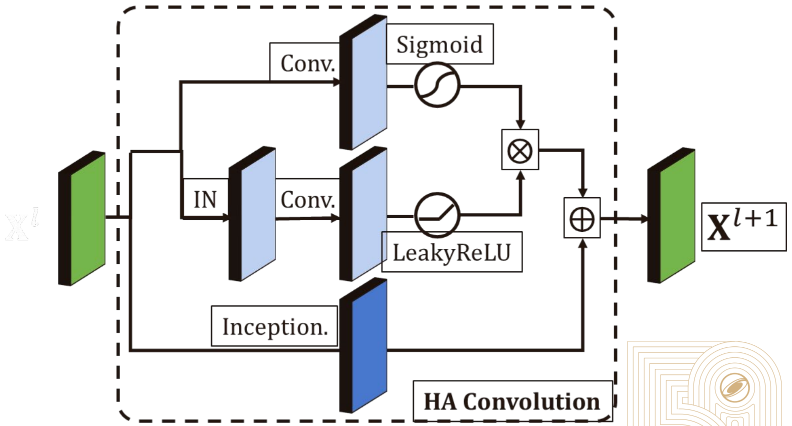
Aittala et al. 2015



Li et al. 2018



Guo et al. 2020



Guo et al. 2021

RELATED WORK

- Potential limitations of CNN-based SVBRDF acquisition methods when facing **Ultra-High Resolution** inputs:
  - Memory constraints of GPUs
  - Difficulties in capturing global features

4K input

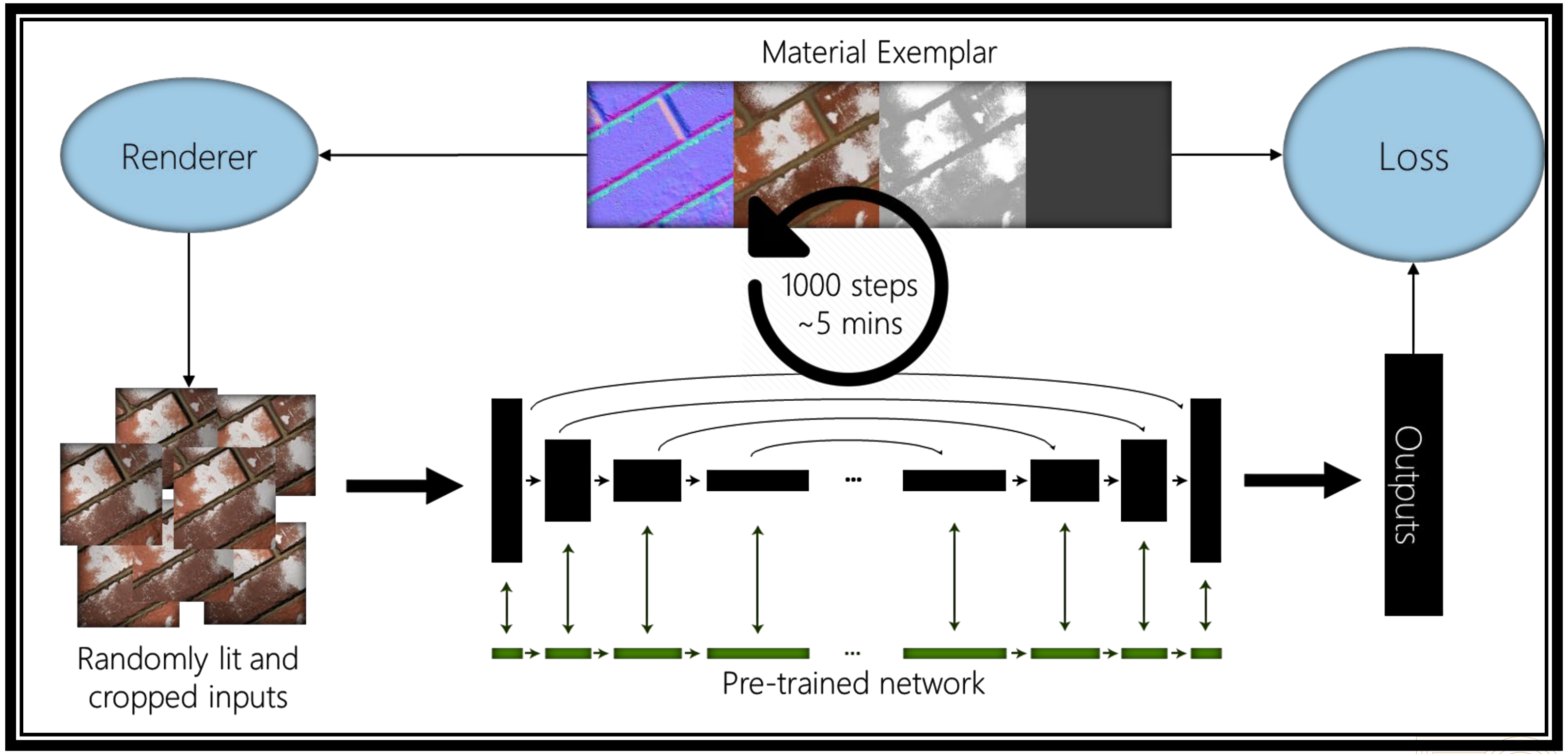
Out of Memory

RELATED WORK

SIGGRAPH 2023
LOS ANGELES+ 6-10 AUG

- HR Image cropping
- Material Exemplar prediction
- Exemplar-specific fine-tuning
- Post-processing: image stitching

Runs extremely slow even with 1K inputs (> 200 seconds per-image)

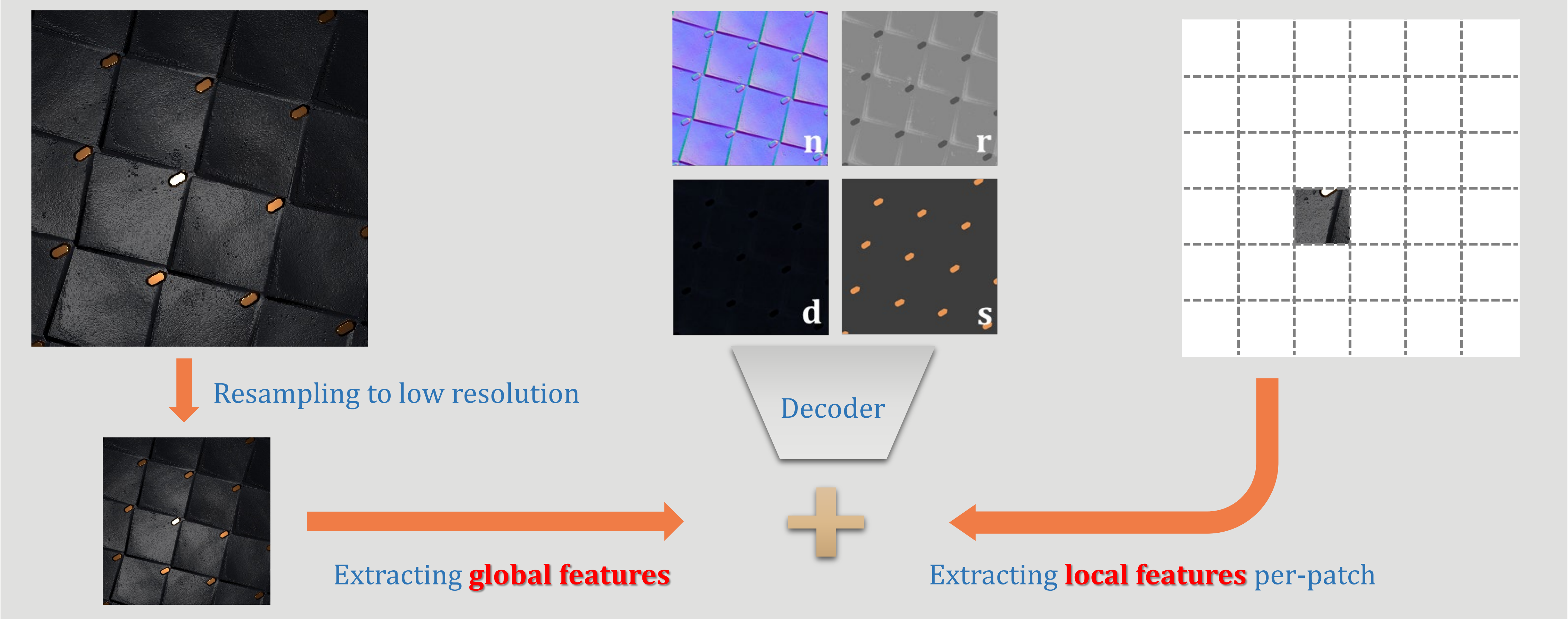


Deschaintre et al. 2020

OUR GOAL

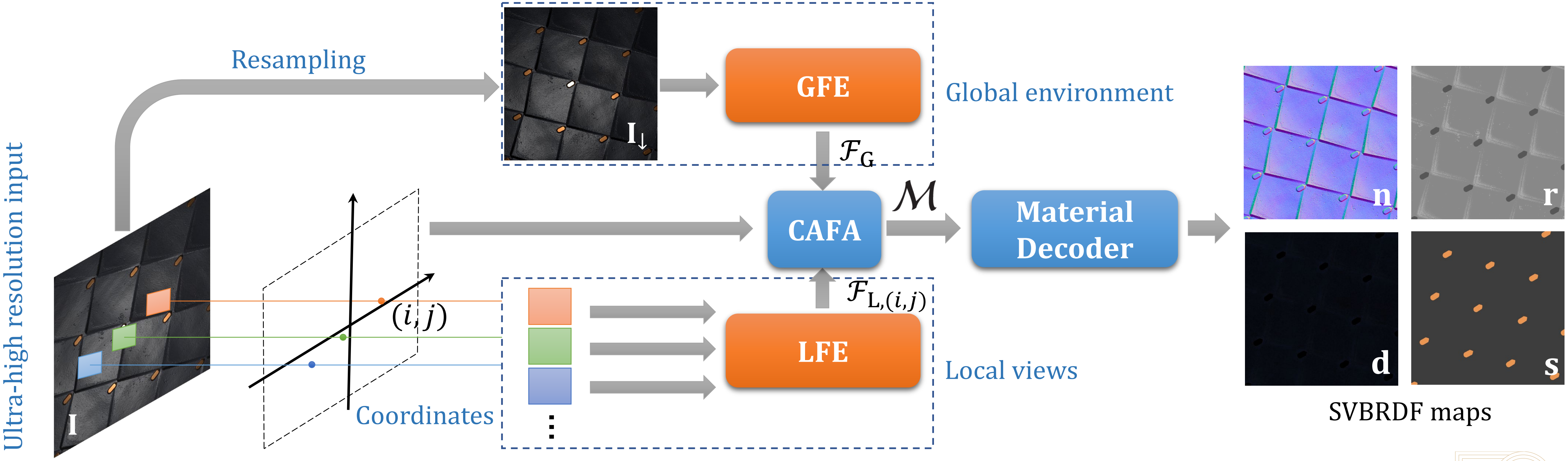
- Single, casually captured image as input (without camera calibration)
- Supporting Ultra-high resolution inputs (2K or 4K)
- Light-weight setup
- Low computational cost
- Fast evaluation with end-to-end prediction

KEY IDEA



A divide-and-conquer solution

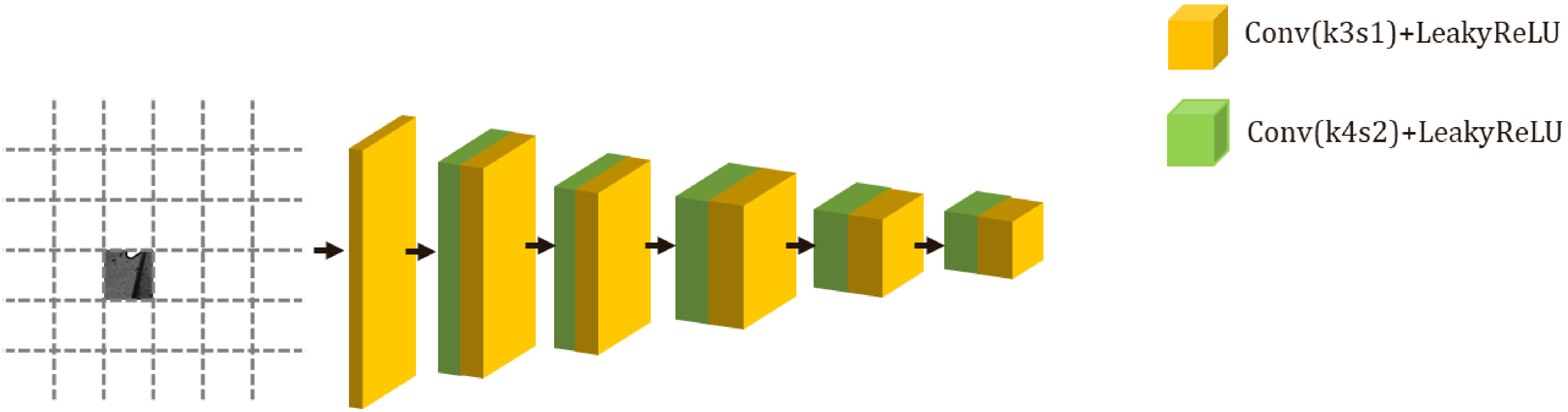
OVERVIEW



$$\{n, d, r, s\} = \mathrm{De}(\mathcal{M}, C)$$

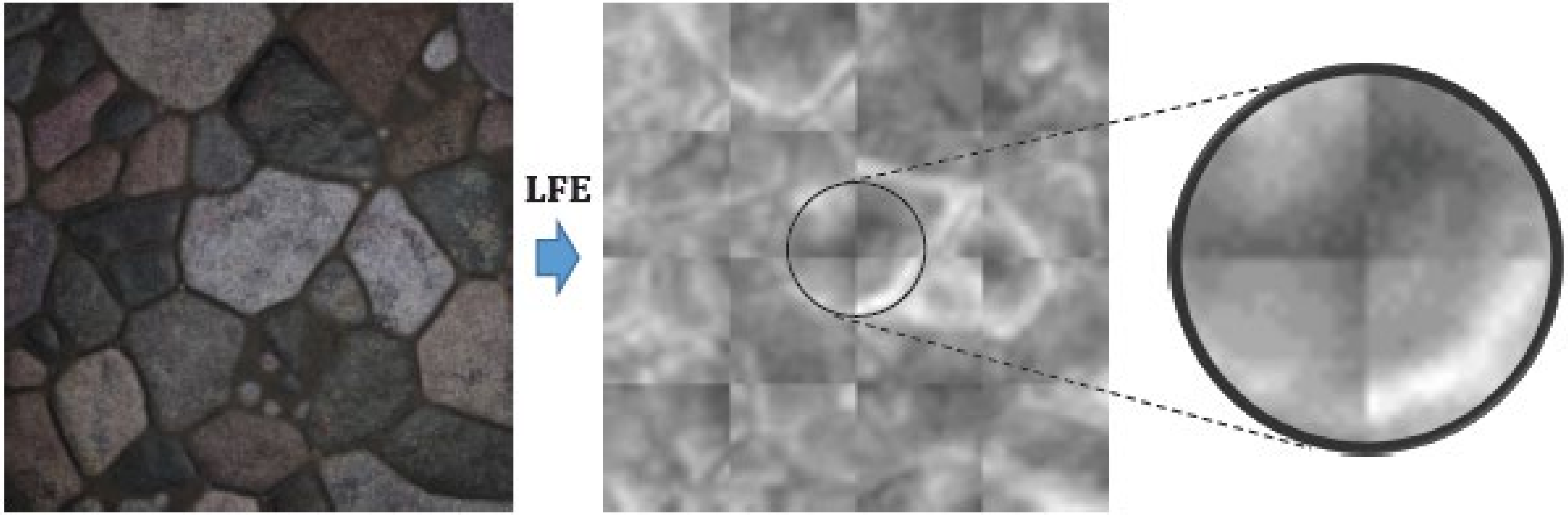
LOCAL FEATURE EXTRACTOR

Providing a "**local view**" of the underlying material



stacked convolutional blocks

LOCAL FEATURE EXTRACTOR

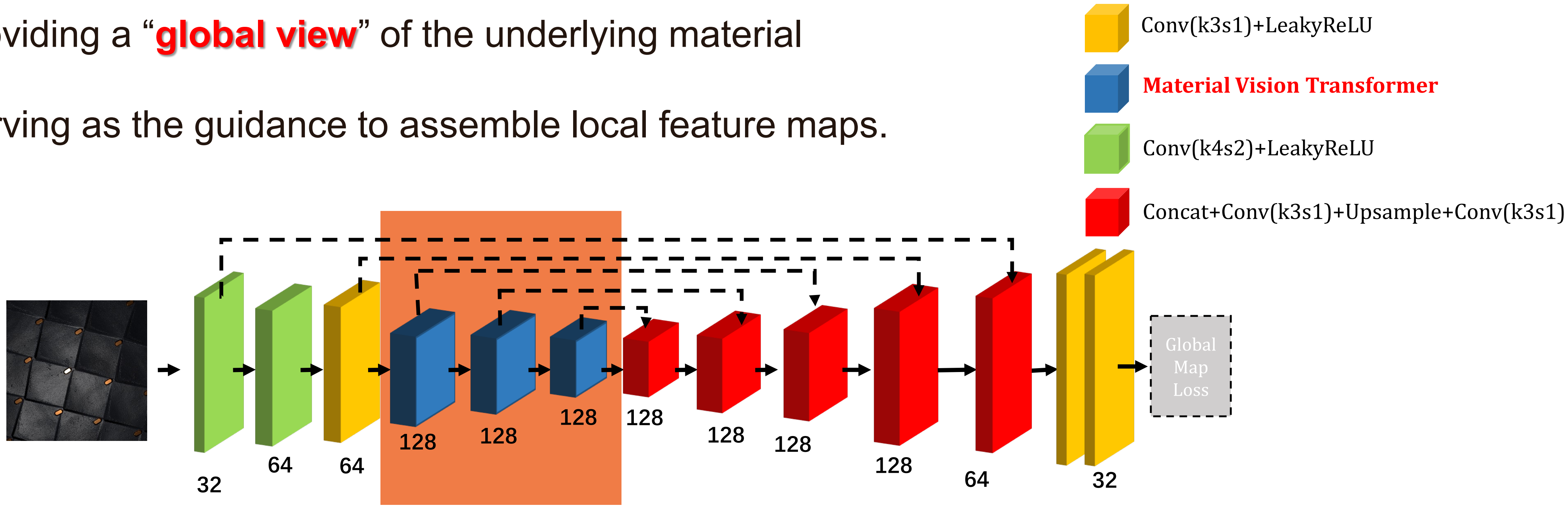


Simply stitching local feature maps generated by **LFE** will cause noticeable inconsistency among different patches.

GLOBAL FEATURE EXTRACTOR

Providing a "**global view**" of the underlying material

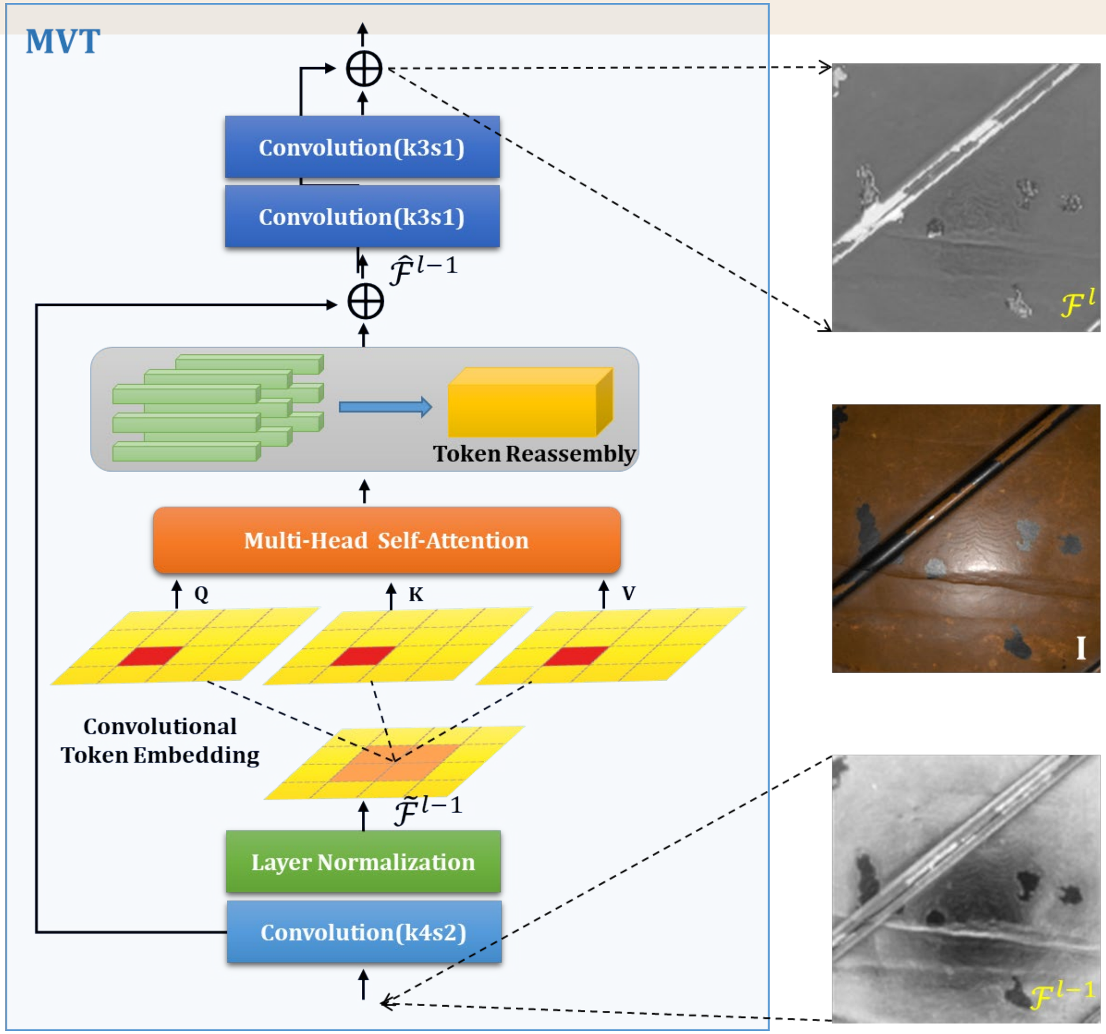
Serving as the guidance to assemble local feature maps.



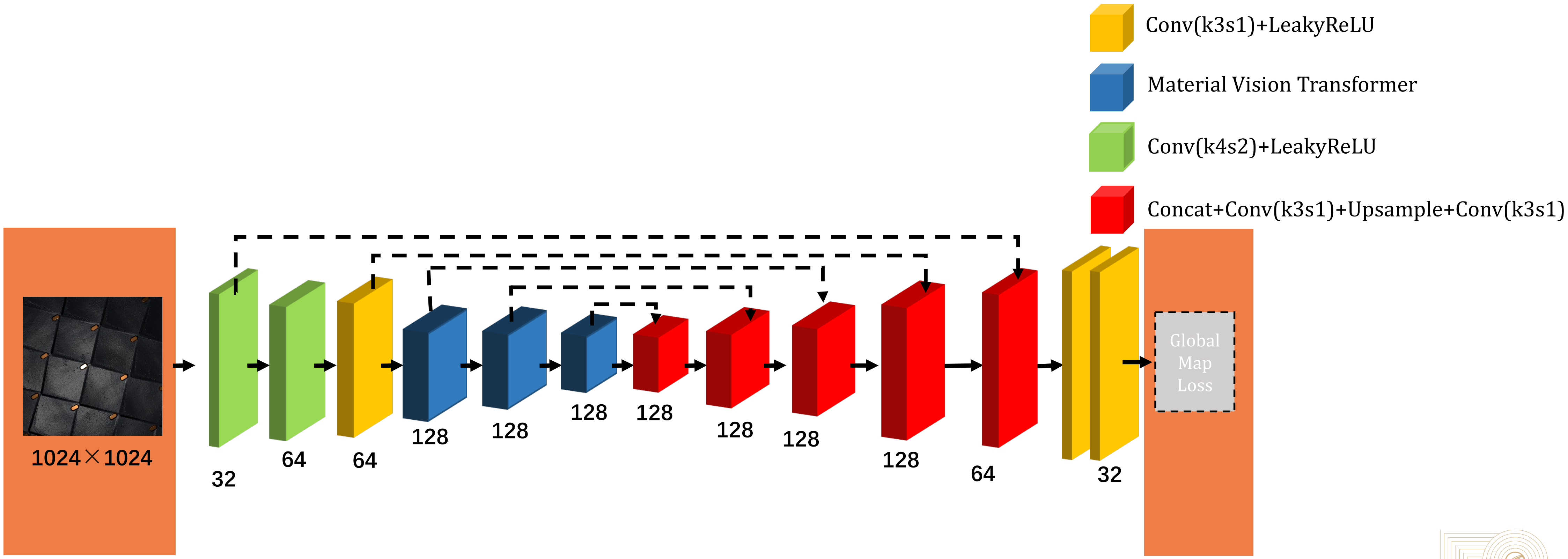
Material Vision Transformers

MATERIAL VISION TRANSFORMER

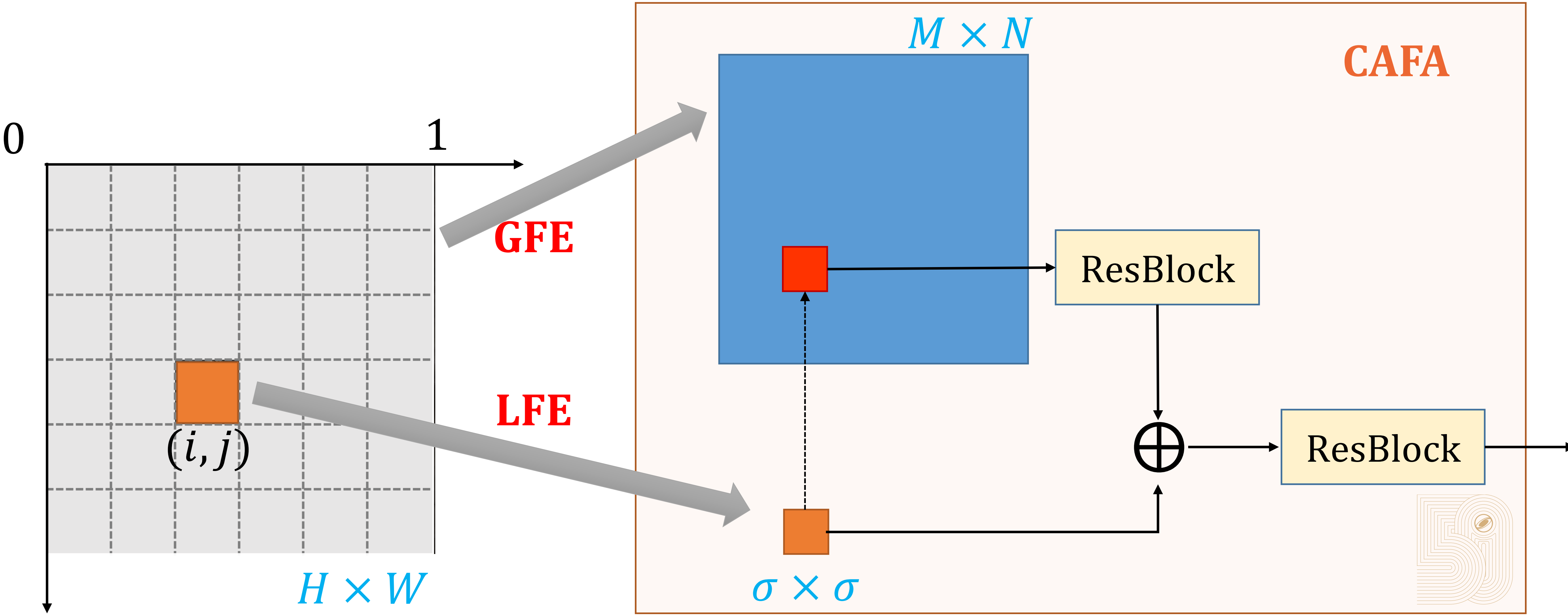
Each MVT has a **global receptive field** that is beneficial for preserving globally coherent material properties



GLOBAL FEATURE EXTRACTOR



COORDINATE-AWARE FEATURE ASSEMBLY MODULE

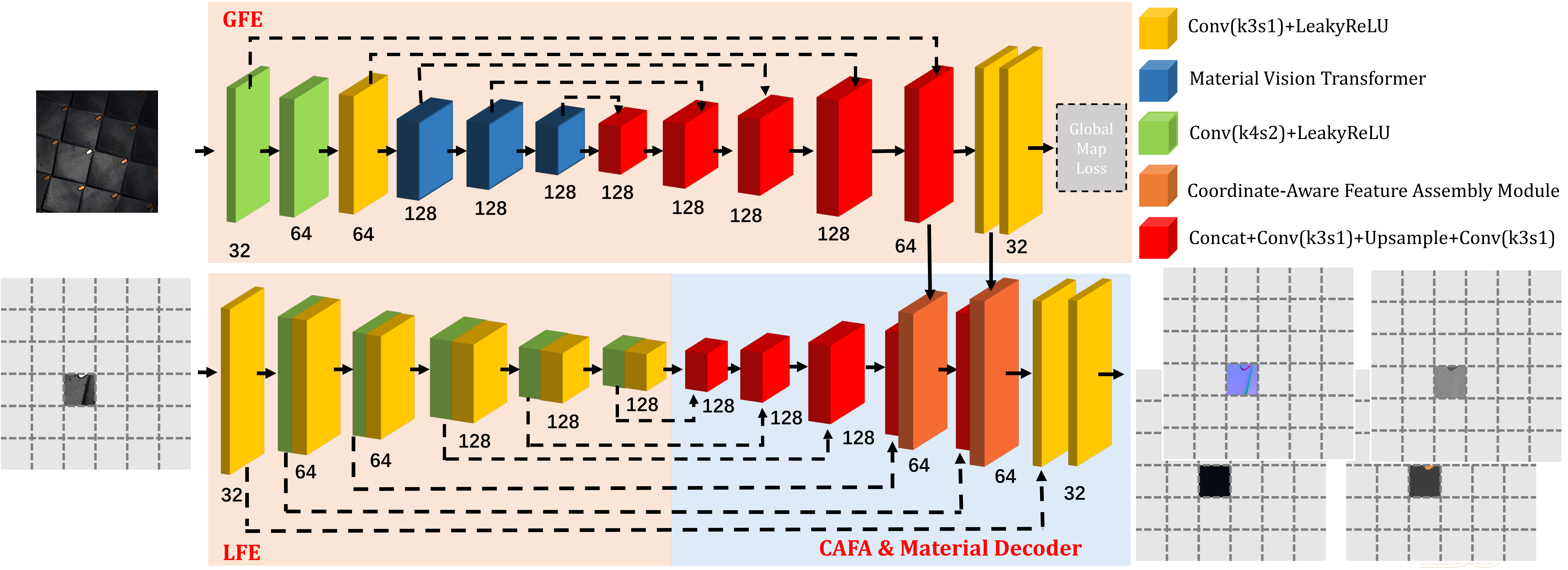


COORDINATE-AWARE FEATURE ASSEMBLY MODULE

(a) Simple stitching (b) Ours without overlaps (c) Ours with overlaps

Removing seams in a recovered diffuse map by introducing some overlaps between adjacent patches.

COMPLETE NETWORK ARCHITECTURE



DATASET AND LOSS FUNCTION

SIGGRAPH 2023
LOS ANGELES+ 6-10 AUG

- UHR dataset

	4K		2K		1K	
	Part 1	Part 2	Part 1	Part 2	Part 1	Part 2
Training	-	160	2,105	-	6,315	2,400
Testing	-	25	24	26	-	-

- Part1 Collected by Deschaintre et al. [2020]

- Part2
  - https://ambientcg.com/
  - https://polyhaven.com/
  - https://www.sharetextures.com
  - https://www.cgbookcase.com/textures/

- Loss Function

$$\mathcal{L} = \mathcal{L}_{\text{map}} + \mathcal{L}_{\text{render}} + \lambda_{\text{percept}}\mathcal{L}_{\text{percept}}$$

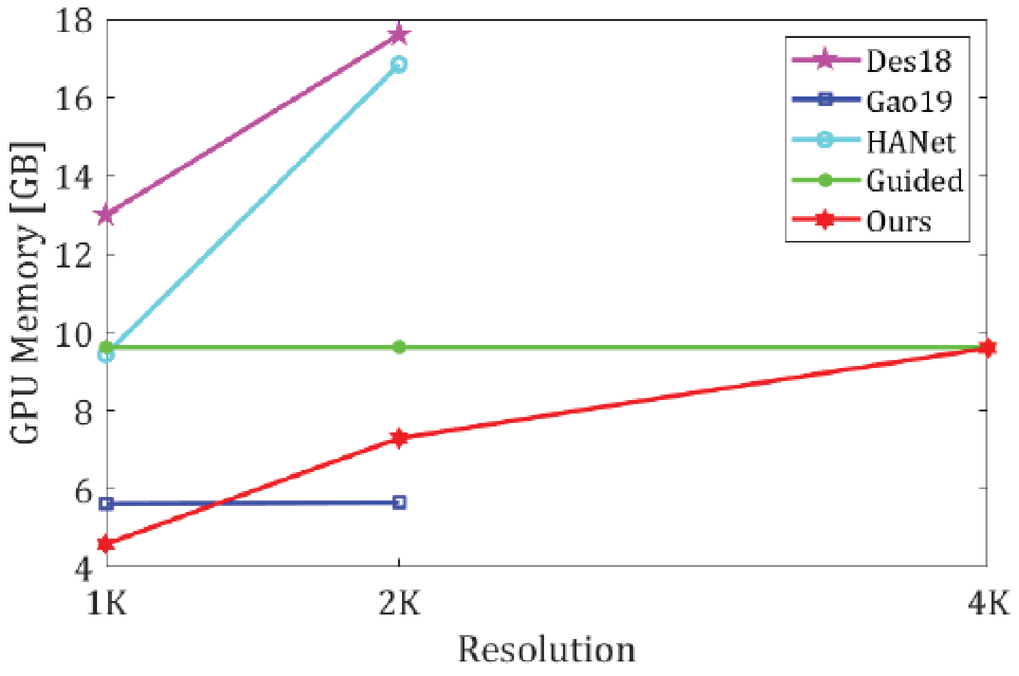
$$\mathcal{L}_{\text{all}} = \lambda_1\mathcal{L}_{\text{global}} + \lambda_2\mathcal{L}_{\text{patch}}$$

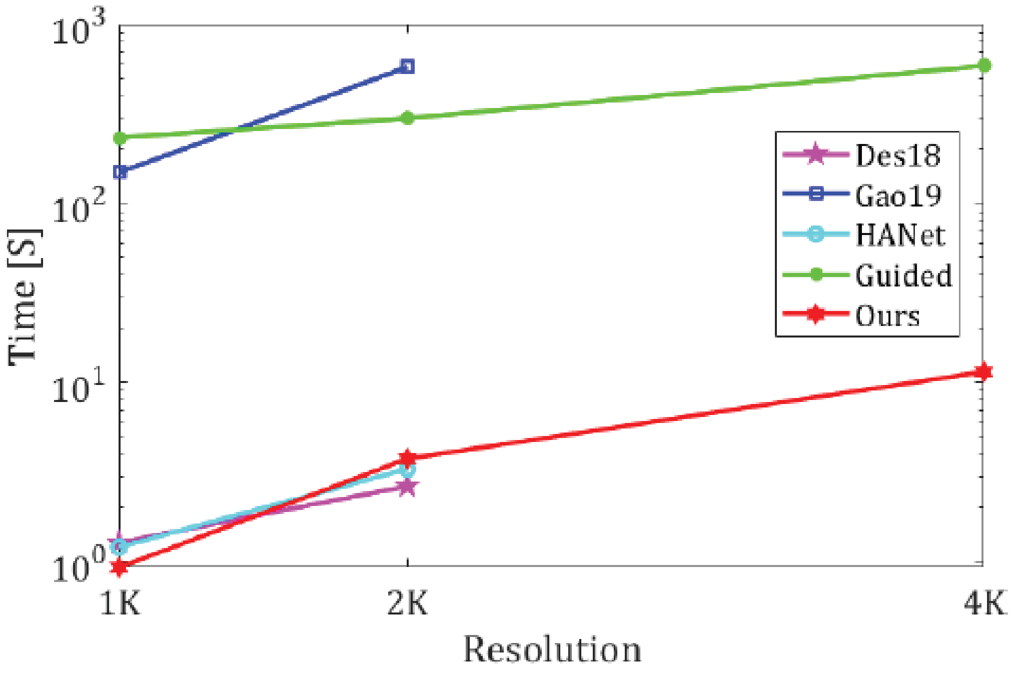
RESULTS

We compare our method to

- **Des18**: a classic U-Net based end-to-end method that can recover SVBRDFs from a single image, proposed by Deschaintre et al. [2018],
- **Gao19**: a deep inverse rendering method that can recover high-resolution SVBRDFs from multiple images, proposedby Gao et al. [2019],
- **HANet**: a state-of-the-art method for singe-image (usually low-resolution) SVBRDF recovery, proposed by Guo et al.[2021],
- **Guided**: a recent method that can support 4K SVBRDF recovery by transferring low-resolution exemplar SVBRDFs to a target UHR image, proposed by Deschaintre et al. [2020].

PERFORMANCE ANALYSIS

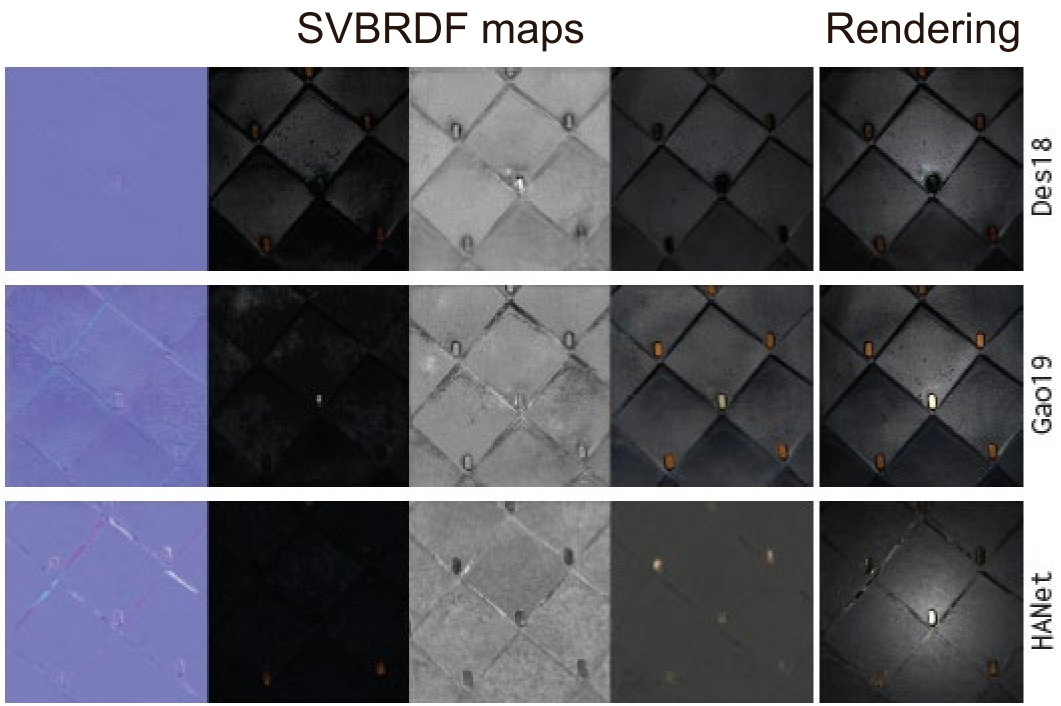


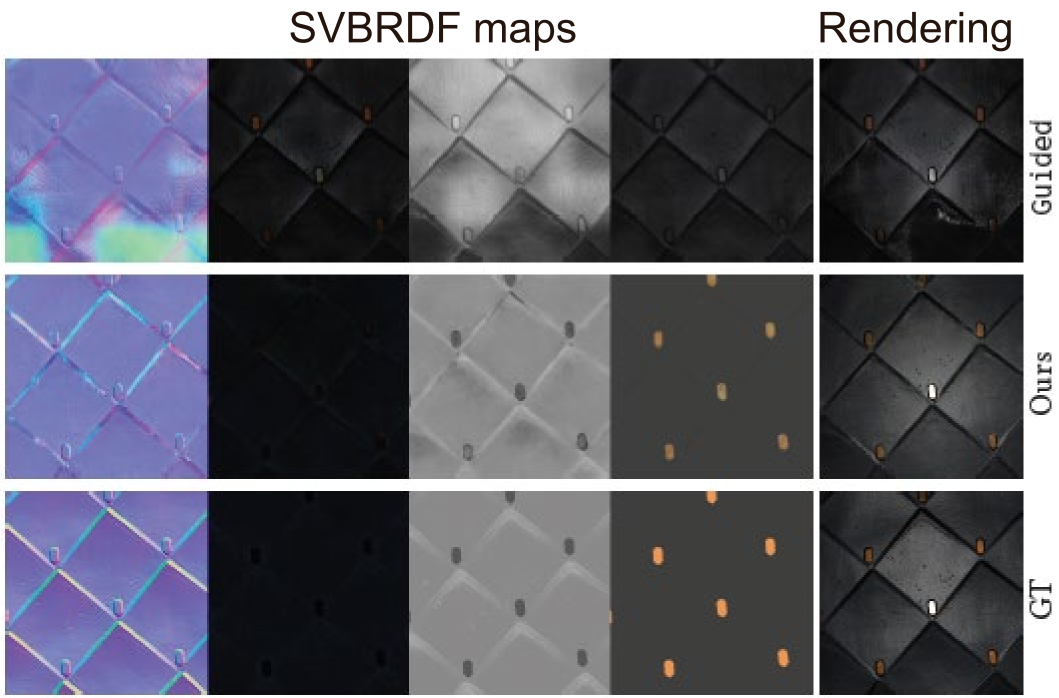


GPU memory consumption

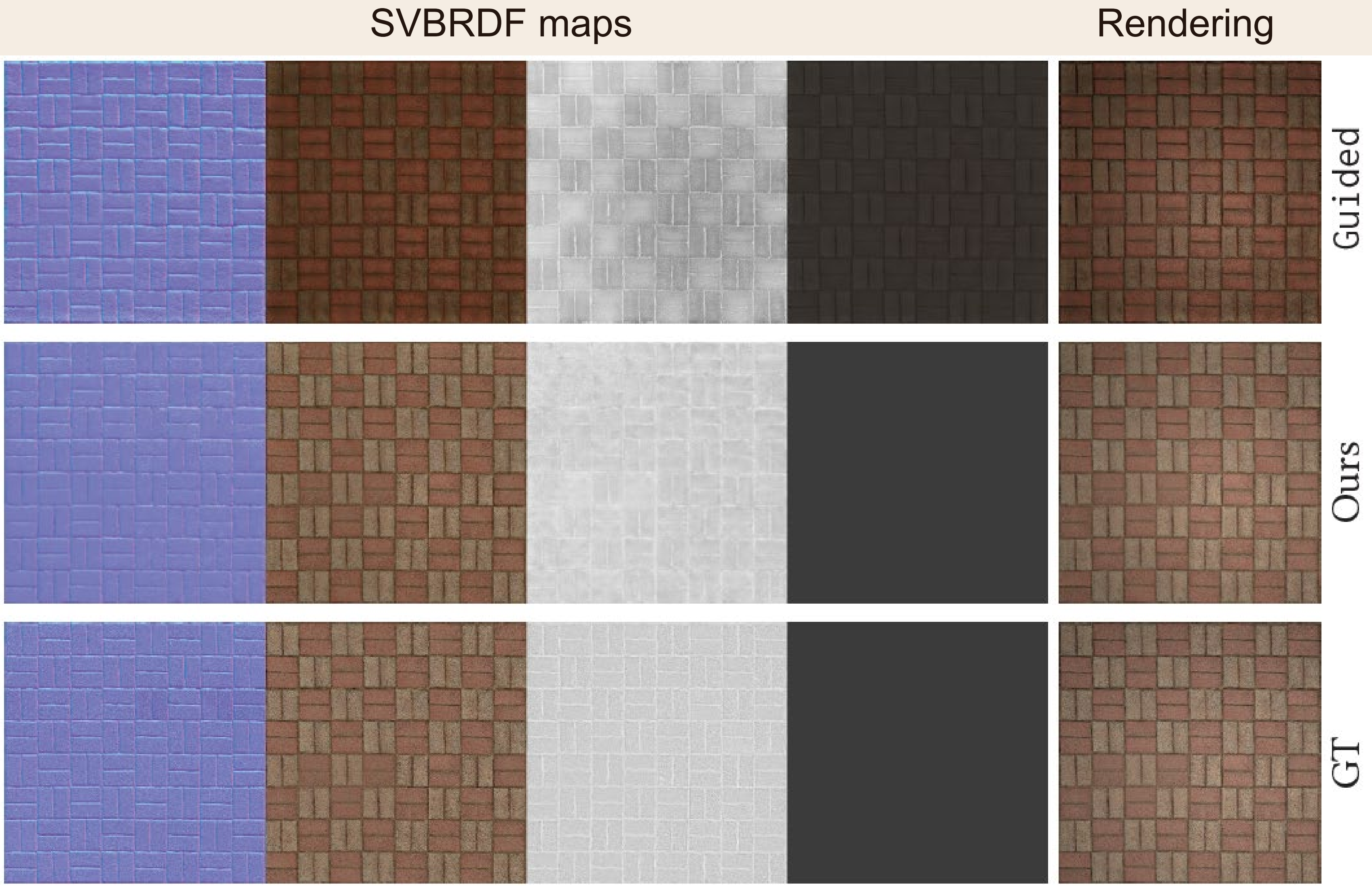
Runtime performance

COMPARISONS ON SYNTHETIC DATA 2K

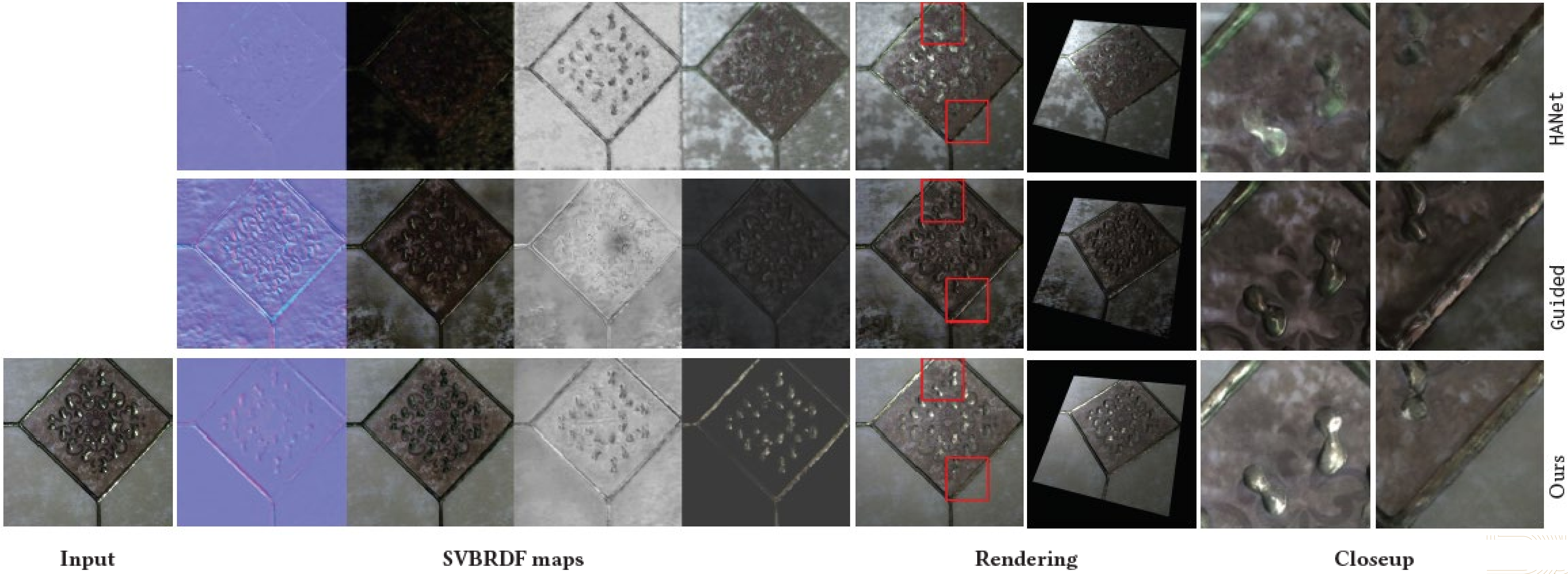




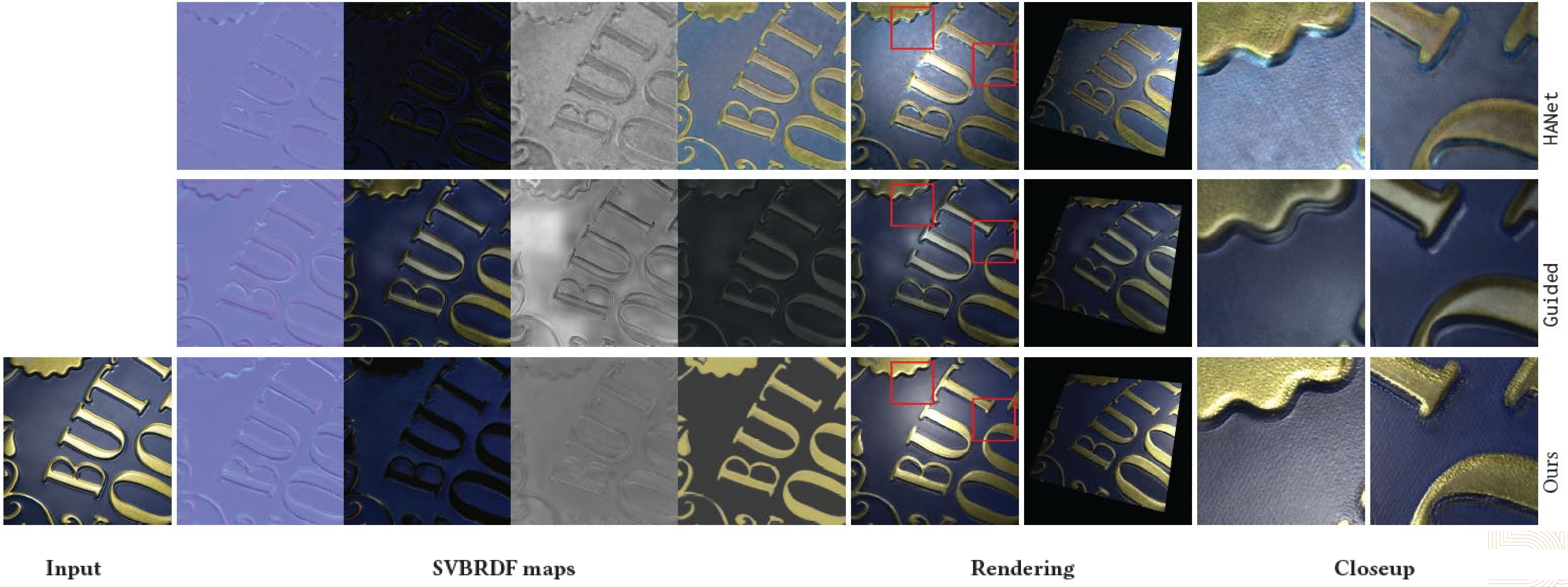
COMPARISONS ON SYNTHETIC DATA 4K



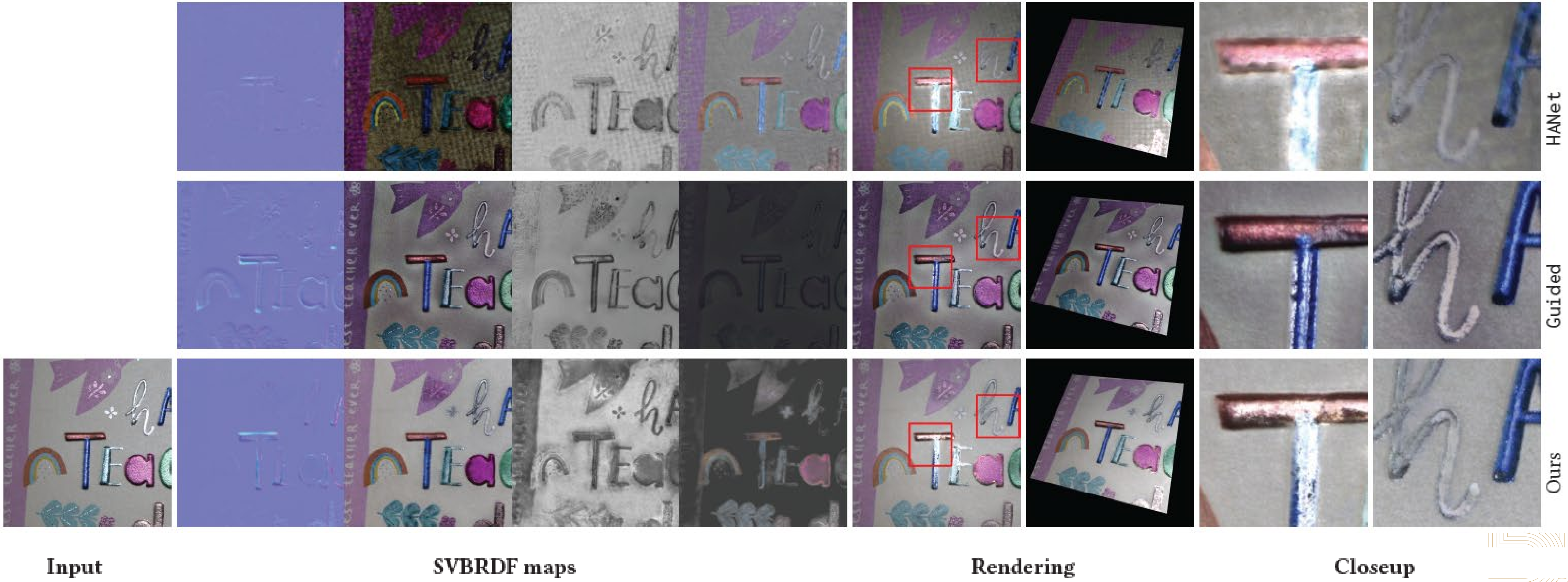
COMPARISONS ON REAL DATA 2K



COMPARISONS ON REAL DATA 2K



COMPARISONS ON REAL DATA 2K



COMPARISONS ON REAL DATA 2K

Guided

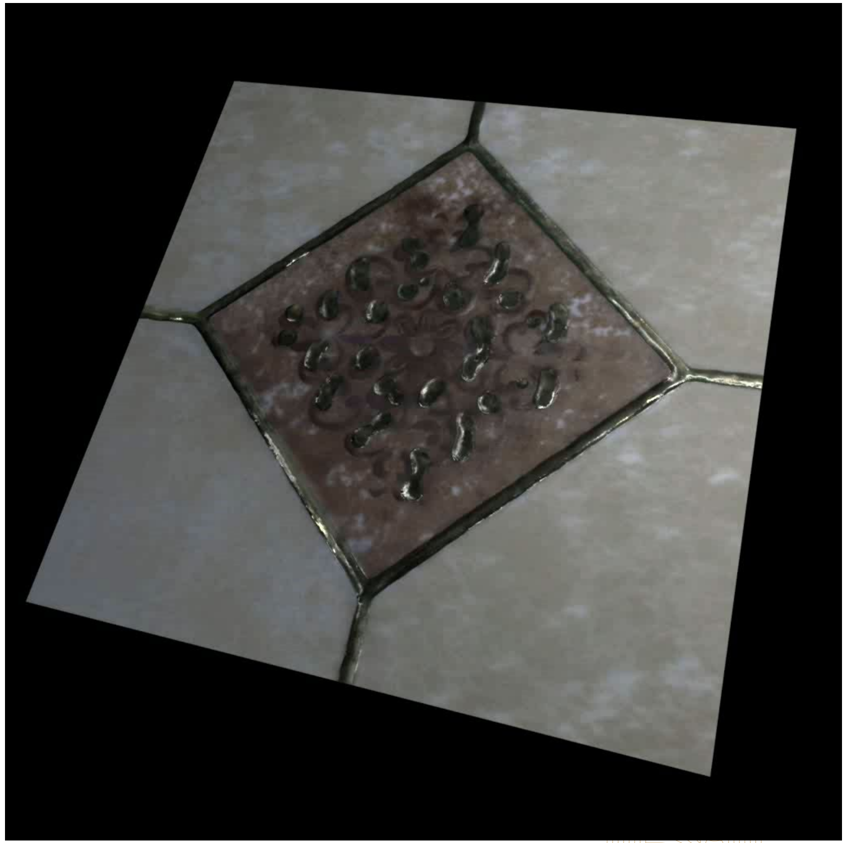
HANet

Ours

COMPARISONS ON REAL DATA 2K

Guided

HANet

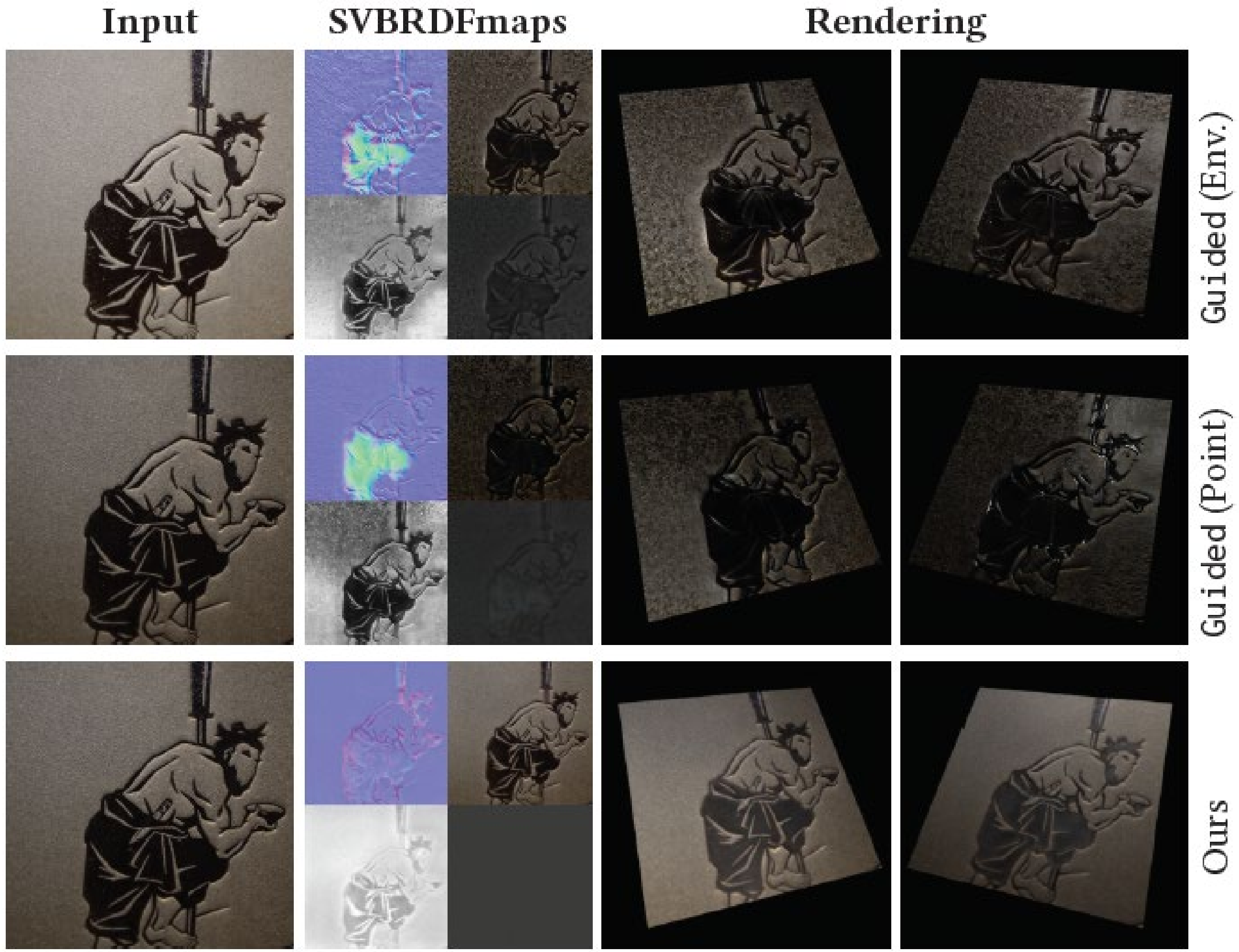


Ours

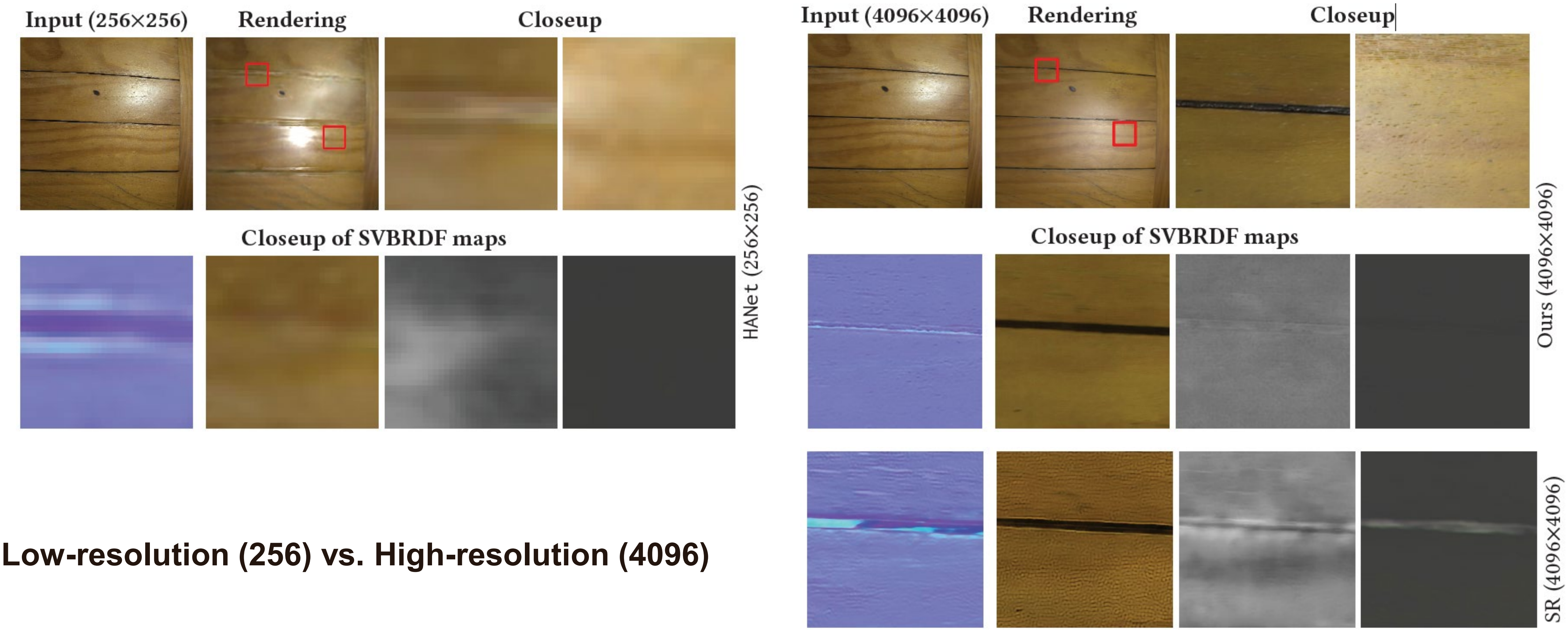
COMPARISONS ON REAL DATA 4K

The choice of exemplars to Guided

For **Guided**, we show the prediction results from two input images of the same material taken under environment lighting (Env.) and a flash point light (Point), respectively.



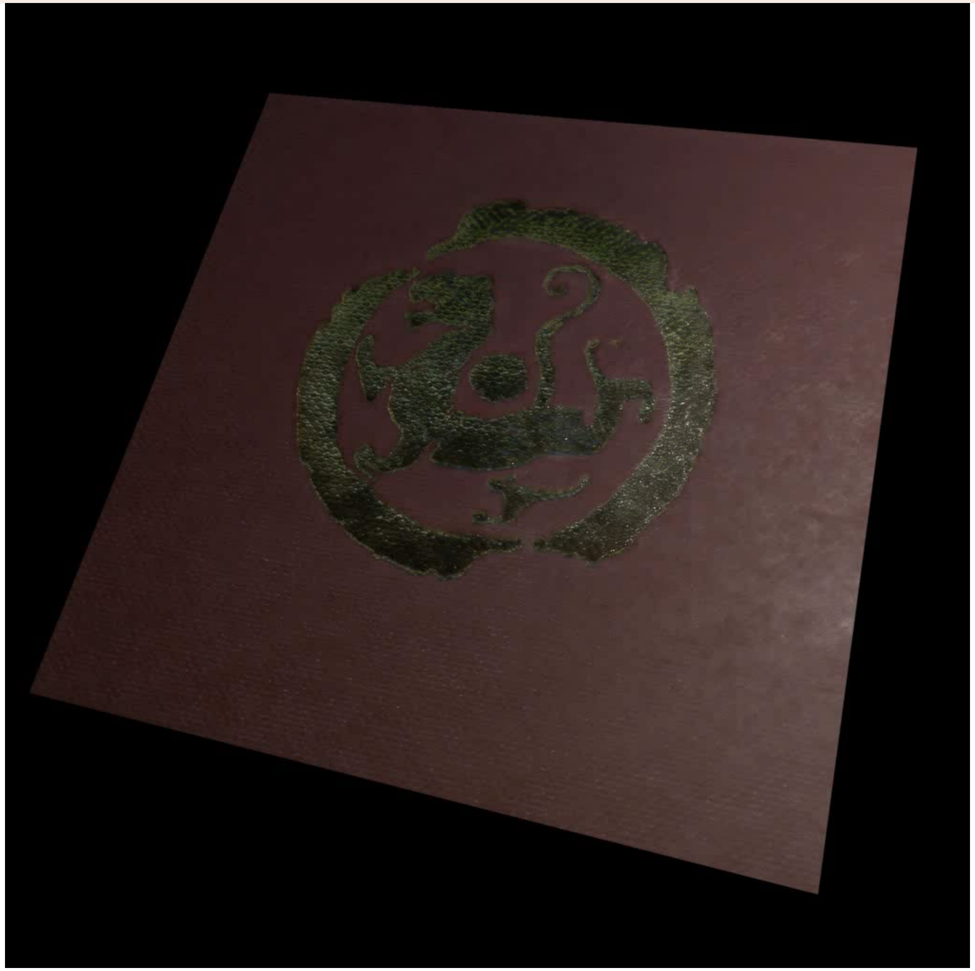
COMPARISONS ON REAL DATA 4K



Low-resolution (256) vs. High-resolution (4096)

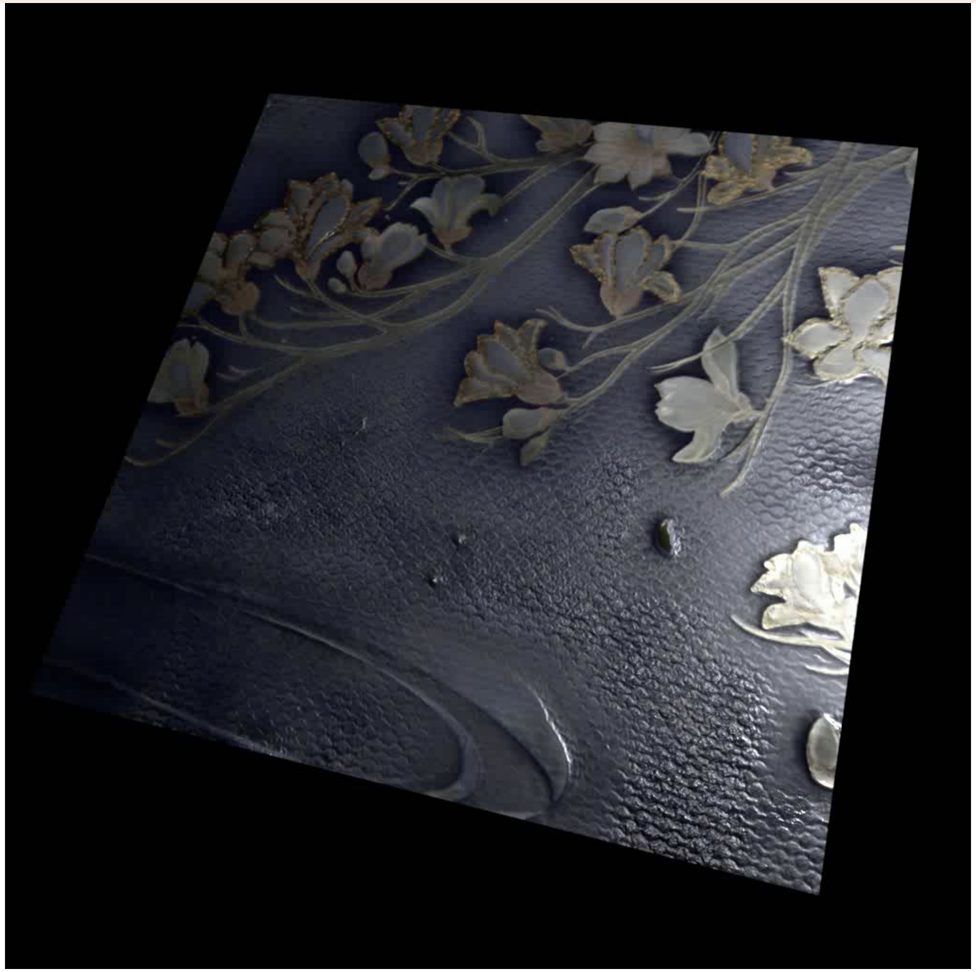
COMPARISONS ON REAL DATA 4K

Guided

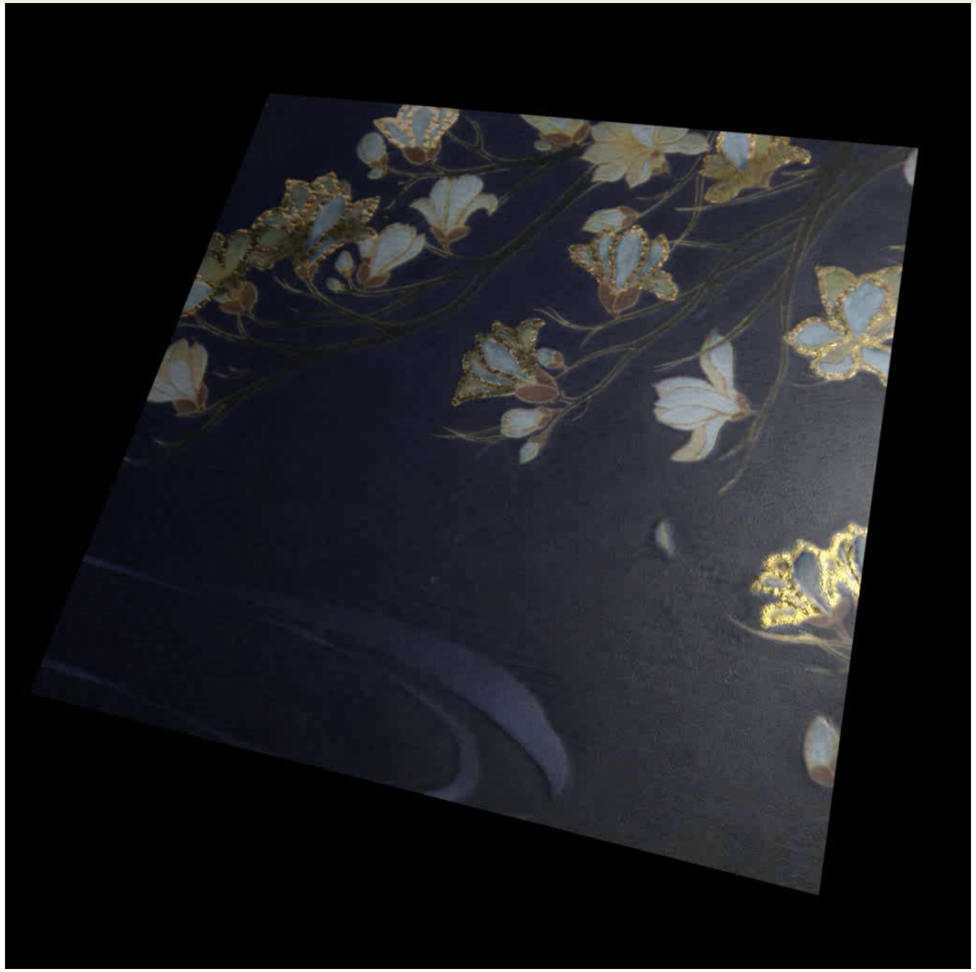


Ours

COMPARISONS ON REAL DATA 4K



Guided



Ours

COMPARISONS ON REAL DATA 4K

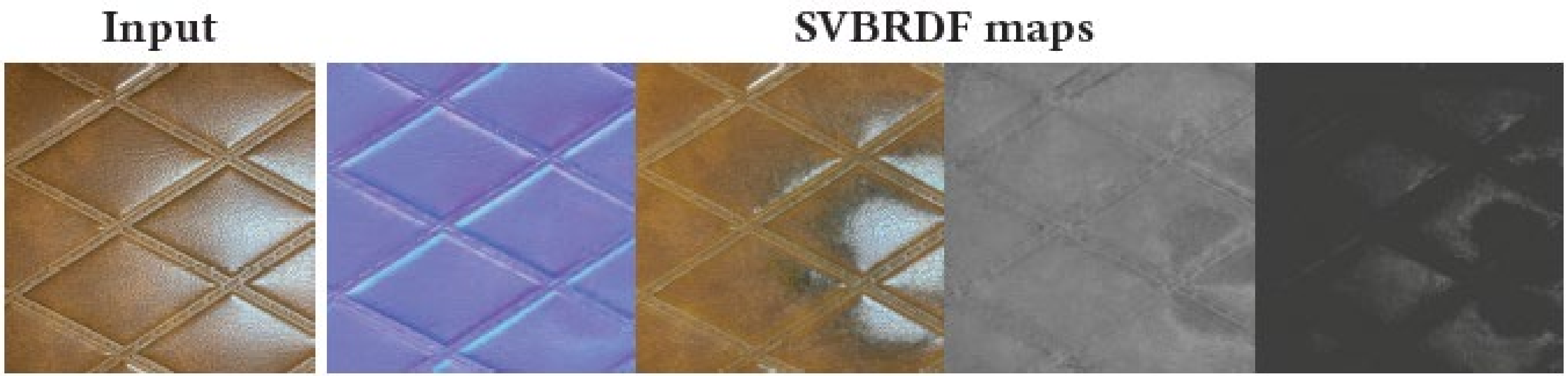
COMPARISONS ON REAL DATA 4K

CONCLUSION

- A consistent and implicit neural representation for SVBRDF recovery that can preserve both global structures.
- Material vision Transformer, a convolution-augmented vision transformer, to extract rich global features from the UHR input, providing the “global environment” of the material.
- A coordinate-aware feature assembly module to assemble “local views” of the underlying material in the feature space, guaranteeing spatial coherency.
- Our method is able to recover material maps as large as 4K, which is a challenge for previous learning-based methods.

FUTURE WORK

- Increasing the normal map diversity in the dataset would further improve the quality of reconstructed normal maps.
- Incorporating some special designs, e.g. highlight-aware convolution.
- Extending our method to curved surfaces.



SIGGRAPH 2023
LOS ANGELES+ 6-10 AUG

THE PREMIER CONFERENCE & EXHIBITION ON COMPUTER GRAPHICS & INTERACTIVE TECHNIQUES

THANK YOU

Email: guojie@nju.edu.cn