Supplementary Materials: Reconstructing, Understanding, and Analyzing Relief-Type Cultural Heritage from a Single Old Photo

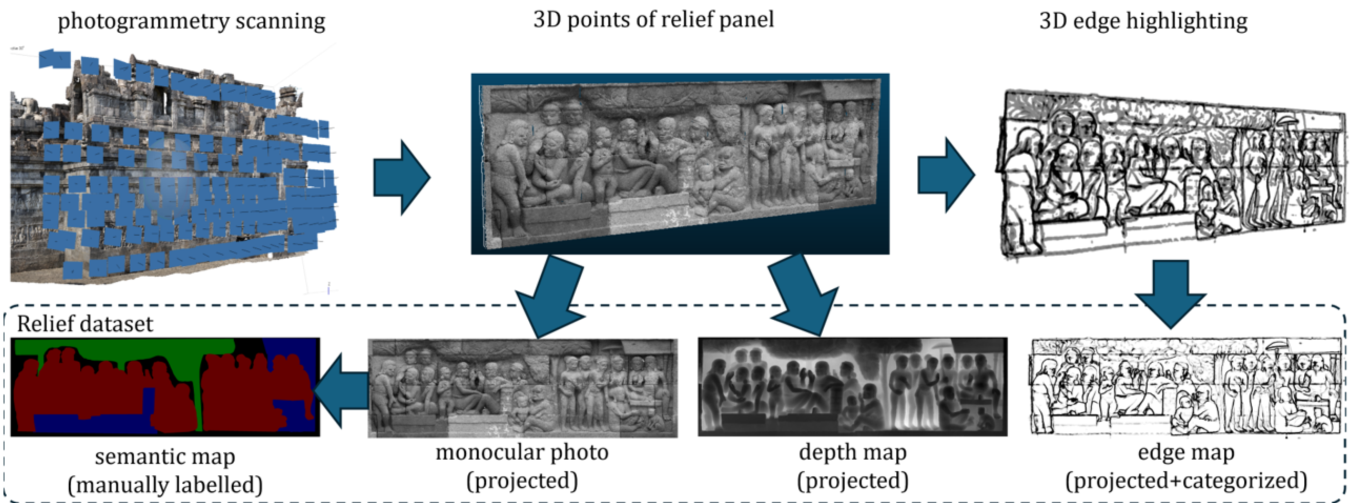

Figure 1: Creation process of the relief dataset.

1 RELIEF DATASET

As we mentioned in our paper, we primarily referenced Pan et al. [3] and made minor modifications to adapt it to our newly proposed task of soft edge detection. In this section, we review the relief dataset utilized in our experiments to provide a more detailed introduction to our readers.

Our dataset originates from the reliefs at Borobudur Temple, a UNESCO World Heritage Site in Indonesia. The temple is adorned with 2,672 relief panels, 156 of which are buried, leaving only monocular old photos as records. We constructed our dataset from reliefs that were not buried and utilized it to train the multi-task network proposed in this work. This training enables the network to learn the data characteristics specific to relief-type cultural heritage. Subsequently, this facilitates the objectives of 3D reconstruction, understanding, and analysis from old photos.

As shown in Figure 1, we initially collect the 3D points of relief parts using photogrammetry scanning. Photos were taken with a resolution of 6000 × 4000 pixels using a digital camera (RICOH GR III). The distance from the camera to the temple building is approximately 2 meters, and each photo overlaps by about 60%. The Structure from Motion-Multi-View Stereo (SfM-MVS) method is adopted for the photogrammetry. In our work, we use the commercial software Agisoft Metashape to generate dense 3D points from the photogrammetry data. These scanned points contain information of 3D coordinates and RGB values, so that we can project from the front view to obtain monocular images and their corresponding depth maps.

Moreover, to obtain the soft edge maps, we apply a 3D edge highlighting algorithm [2] to the scanned points. In this algorithm, the local point distribution is assessed by measuring changes in curvature, which typically disappear in a planar distribution. A critical aspect of this technique is that the local point distribution correlates with the opacity of the point [4, 5]. Thus, narrow areas around the central lines of the soft edge regions exhibit comparatively higher opacity and are highlighted with brighter colors. In this way, information about depth variations around edge regions is preserved to a certain extent. By projecting in the same manner as the monocular images, we can obtain the corresponding edge maps with continuous pixel values from 0 to 255. Then, we convert them into multi-class edge maps as edge labels for our proposed soft edge detection task according to Section 3.2 in our paper.

Furthermore, we annotated the projected monocular images with semantic segmentation category labels, assigning the data into four classes: background (black), characters (red), plants (green), and others (blue).

At this point, our training and validation dataset, which includes monocular images along with their corresponding depth, semantic, and edge labels, is complete. The Borobudur relief is 2.7 m wide, 0.92 m high, and 0.15 m (150 mm) deep. After performing the photogrammetry scanning, we project 3D points of each relief into a monocular image with 3072 × 1024 and prepare the corresponding labels following the process mentioned above.

In this paper, we used 11 relief panels as training data and one panel as validation data. Due to the structure of the proposed multi-task network, no labels are required for the test set. Inference can be conducted with just a single front-facing monocular old photo. Our approach requires a small training set to achieve high accuracy in predictive maps, especially in tasks of depth estimation, showing outstanding results. Thus, we demonstrate that our approach is broadly applicable to relief-type cultural heritage objects in various large-scale ancient sites other than the Borobudur temple.

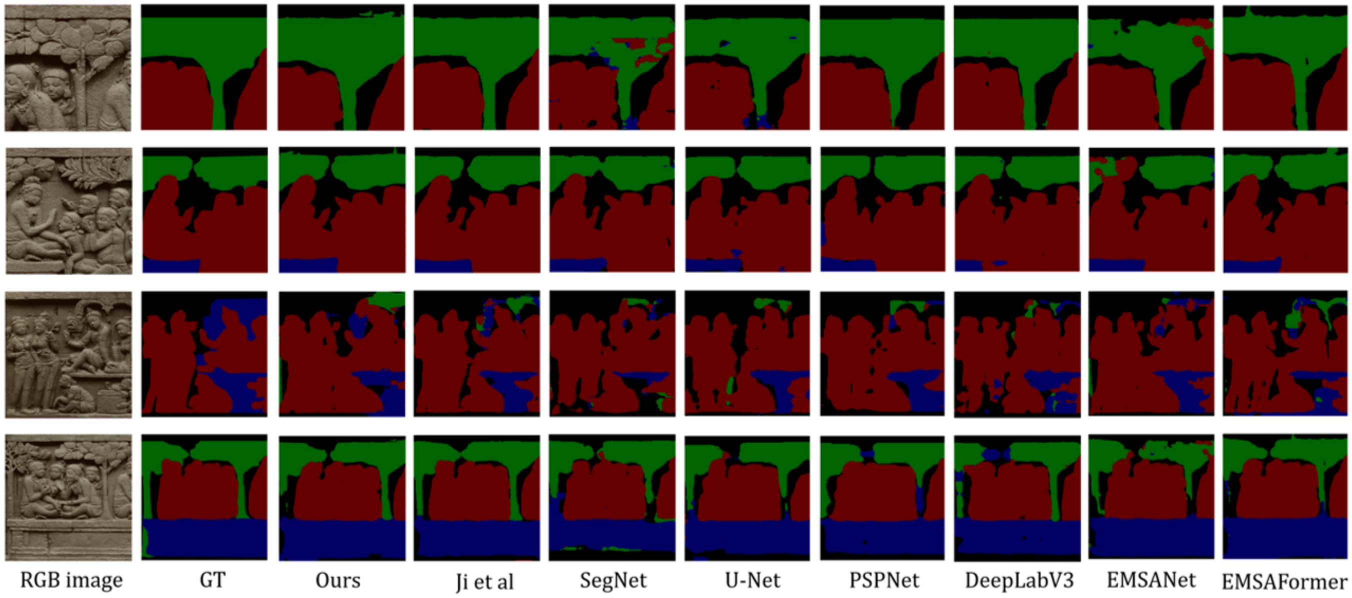

Figure 2: Comparison results of semantic segmentation task.

2 COMPARISON RESULT

We present additional visualization results from our comparison experiments to further demonstrate the advanced performance of our proposed method.

First, in Section 2.1, we present the visualization results of semantic segmentation task, corresponding to Table 4 in our paper. Second, in Section 2.2, we present the 3D reconstruction models with their corresponding depth prediction maps in addition to Figure 6 in our paper.

2.1 Semantic segmentation result

We present the visualization results of semantic segmentation task in Figure 2. As mentioned in our paper, the comparison results demonstrate that our model performs optimally in semantic segmentation tasks under relief scenes. We demonstrate again that JI et al [1]. used a dataset twice as large as ours and introduced two additional inputs: depth and edge. Yet, their numerical results are only marginally higher than ours according to Figure 2 in this material and Table 4 in our paper.

2.2 3D reconstruction result

We present more 3D reconstruction models of the buried reliefs in Borobudur temple. Due to the absence of ground truth, we only present the comparison between our method and Pan et al. [3]. Pan et al. provides currently the only method, aside from ours, capable of predicting detailed features. However, the details it provides largely depend on the original edge priors used as additional inputs, whereas our model requires no such inputs as we mentioned in our paper.

As shown in Figure 3 and Figure 4, we present the old photos taken in 1890s with their corresponding 3D reconstruction models and the predicted depth maps. In our predicted depth maps, black lines between patches are eliminated by voting algorithm. Please note that our method provides more detailed features than Pan et al. [3]. In 3D reconstruction models, our method offers a smoother and more harmonious sense of 3D structure due to the accurate depth values provided. Please focus on comparing the depth variations in the soft edges of the character's face and the floral patterns of the carvings in the image.

Additionally, we have included rotating video of the 3D reconstruction models corresponding to the old photo displayed in the first row, available in the supplementary materials.

REFERENCES

[1] Shenyu Ji, Jiao Pan, Liang Li, Kyoko Hasegawa, Hiroshi Yamaguchi, Fadjar I. Thufail, Brahmantara, Upik Sarjiati, and Satoshi Tanaka. 2023. Semantic Segmentation for Digital Archives of Borobudur Reliefs Based on Soft-Edge Enhanced Deep Learning. *Remote Sensing* 15, 4 (Feb. 2023), 956. https://doi.org/10/gs6hnx

[2] K. Kawakami, K. Hasegawa, L. Li, H. Nagata, M. Adachi, H. Yamaguchi, F. Thufail, Setyo Riyanto, and S. Tanaka. 2020. OPACITY-BASED EDGE HIGHLIGHTING FOR TRANSPARENT VISUALIZATION OF 3D SCANNED POINT CLOUDS. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* V-2-2020 (Aug. 2020), 373–380. https://doi.org/10/gs6hns

[3] Jiao Pan, Liang Li, Hiroshi Yamaguchi, Kyoko Hasegawa, Fadjar I. Thufail, Brahmantara, and Satoshi Tanaka. 2022. 3D reconstruction of Borobudur reliefs from 2D monocular photographs based on soft-edge enhanced deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing* 183 (Jan. 2022), 439–450. https://doi.org/10/gp59q7

[4] S. Tanaka, K. Hasegawa, N. Okamoto, R. Umegaki, S. Wang, M. Uemura, A. Okamoto, and K. Koyamada. 2016. SEE-THROUGH IMAGING OF LASER-SCANNED 3D CULTURAL HERITAGE OBJECTS BASED ON STOCHASTIC RENDERING OF LARGE-SCALE POINT CLOUDS. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* III-5 (June 2016), 73–80. https://doi.org/10/gcc6pg EI: .

[5] Tomomasa Uchida, Kyoko Hasegawa, Liang Li, Motoaki Adachi, Hiroshi Yamaguchi, Fadjar I. Thufail, Sugeng Riyanto, Atsushi Okamoto, and Satoshi Tanaka. 2020. Noise-robust transparent visualization of large-scale point clouds acquired by laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing* 161 (March 2020), 124–134. https://doi.org/10/gp5t5z JCR: Q1 : 1 : 12.7 5: 12.4 EI: : A.

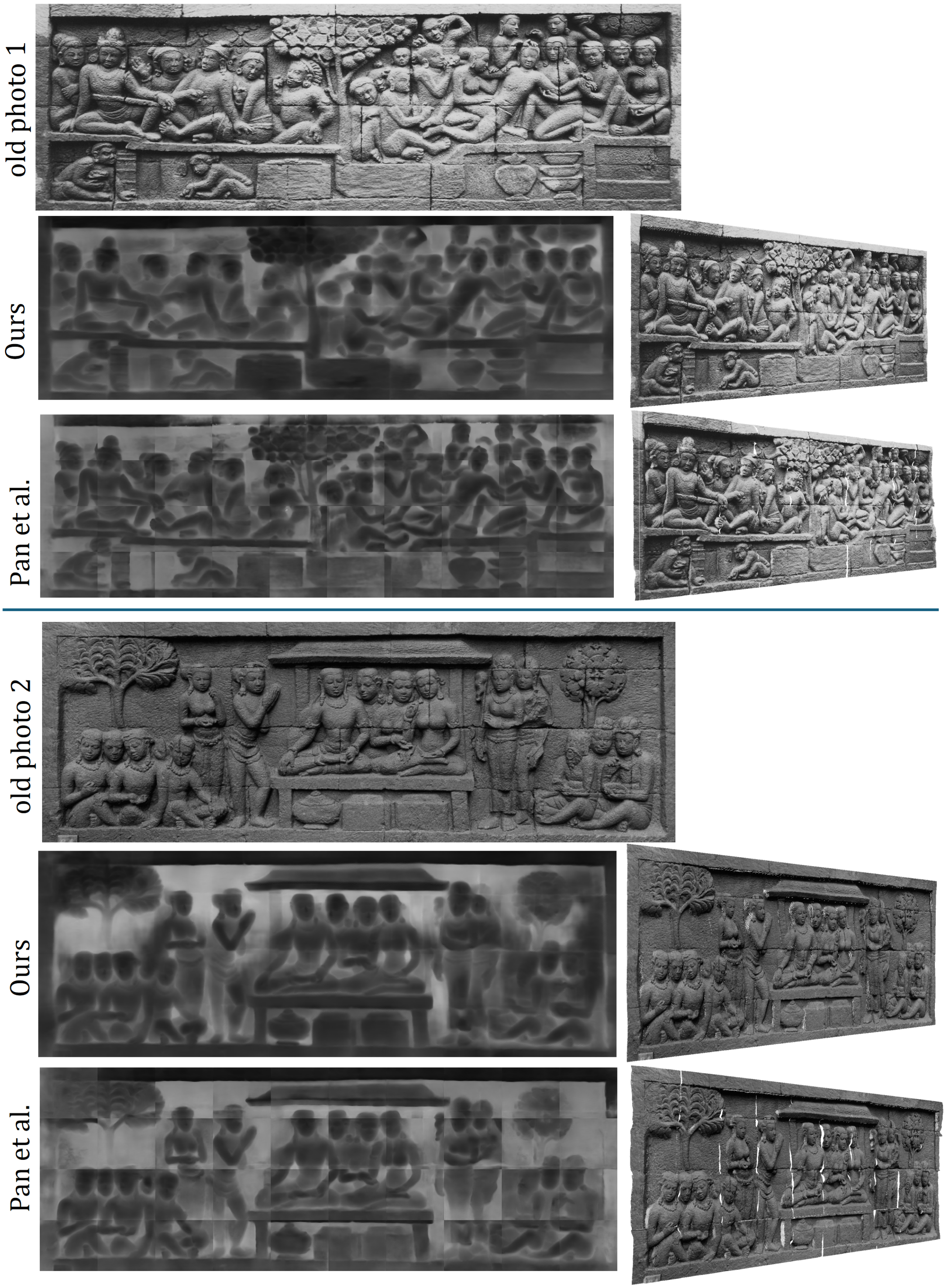

Figure 3: Additional 3D reconstruction results of the buried reliefs in the Borobudur temple (1/2).

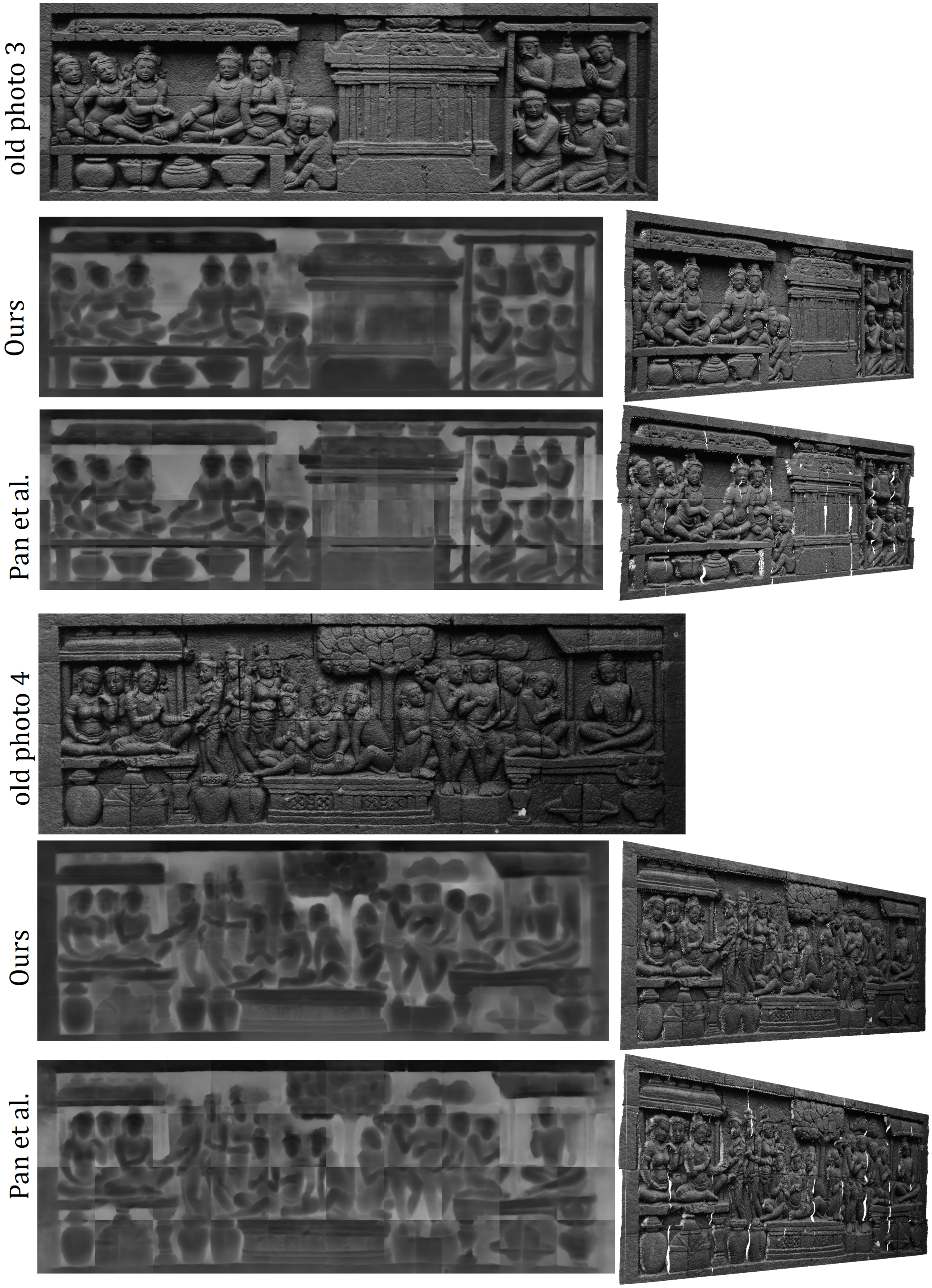

Figure 4: Additional 3D reconstruction results of the buried reliefs in the Borobudur temple (2/2).