LumiNet: Latent Intrinsics Meets Diffusion Models for Indoor Scene Relighting

Supplementary Material

Outline. We begin by detailing the training setup and implementation, followed by additional relit results and an illustration of the nearest-neighbor search.

A. Training details

We use Stable Diffusion 2.1 [51] as our base model to balance performance and training costs. To better preserve the details of the input images, we jointly estimate the denoised image and noise map at each denoising step (known as the v-prediction). Our method also applies to other objective functions, such as ϵ (only predicts the noise map). All training and testing are conducted on an 8-GPU NVIDIA A6000 Ada 48GB node. For the SD2.1 base model, we train on images with a resolution of 512×512. An AdamW [41] optimizer with a learning rate of 4×10^{-5} and a decay rate of 0.9 is used. Training requires approximately 120 hours on a single GPU. At inference time, LumiNet outputs a relighted image (resolution: 512×512) in 5 seconds with 50 DDIM steps.

B. More results

Nearest neighbor search In complex real-world scene relighting, particularly for fine-grained light control, we observe that light transfer is highly sensitive to the choice of seed. To address this issue, we propose a nearest neighbor search over multiple seed candidates to identify the relit result that best approximates the target lighting.

Fig. S.1 illustrates the nearest neighbor search over multiple seed candidates to identify the relit result that best approximates the target lighting.

Additional relit images We present additional relit results in the following pages to demonstrate the robustness of our method under varying lighting conditions and across diverse scenes. These include an extension of our teaser figures (Fig. S.2), a scenario showcasing the effect of turning on ceiling lights (Fig. S.3), an example illustrating reduced ambient lighting in the room (Fig. S.4), and results depicting the effect of turning on lamps (Fig. S.5). We provide a detailed analysis of these phenomena in the figure captions and highlight the lighting effects using red bounding boxes within the figures.

Figure S.1. **Nearest Neighbor Search.** Diffusion models are sensitive to seed choice [61]. We observed that the choice of random seeds significantly impacts relighting quality. Here, we present sampled relights generated from 30 random seeds, sorted by their match to the target lighting image. Sorting is based on nearest-neighbor matching of the latent extrinsic (a low-dimensional lighting vector) to the target.

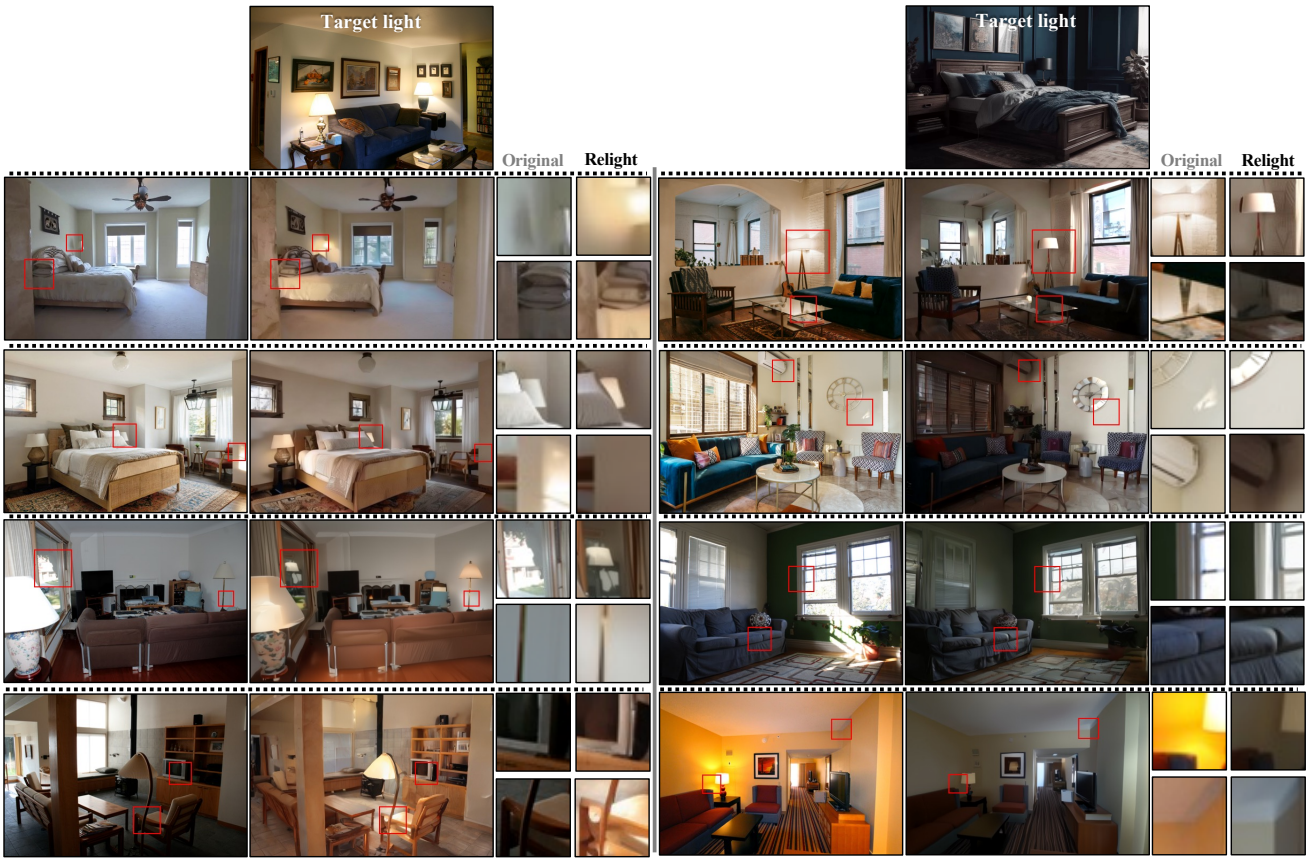

Figure S.2. Our LumiNet architecture transfers complex lighting conditions between indoor scenes using latent intrinsic representations while preserving scene layout, geometry, and albedo. Each scene shows an *original image* (left) paired with its *relighted* version (right) matching the target lighting shown at the top. Our method preserves scene structure and materials while accurately transferring lighting characteristics. **Left panel** demonstrates our method can adjust luminaires to match lighting conditions: it "knows" that to get more light in the right place in the room, it must switch on bedside lights (first row and second row) or table lamps (third row and fourth row), showing our model's ability to handle direct illumination. Zoomed-in crops highlight the changes in images caused by relighting. In the first row, observe the added gloss on the wall behind the lamp in the top crop, as well as the effects on the side of the bed in the bottom crop, influenced by the invisible luminaire. In the second row, note the gloss removal on the side wall, as shown in the bottom crop. In the third row, you can see the reflection of the lamp on the large stationary glass window on the left, highlighted in the top crop. Finally, in the bottom row, observe the strong gloss added to the chair and the faint inter-reflection on the TV screen. **Right panel** shows natural lighting scenarios where bedside lamps are off. Top row's crop shows suppressed specular reflections on the glass table and realistic lamp pole shadows added after relighting. Second row shows strong specular highlights on the wall clock and strong cast shadows from the AC unit. Third row captures soft ambient lighting with intricate specular details on window frames and appropriate surface sheen on furniture. Fourth row demonstrates the removal of bright light from the lamps and all indirect effects, including the recovery of sharp edges at the intersection of the ceiling and side walls.

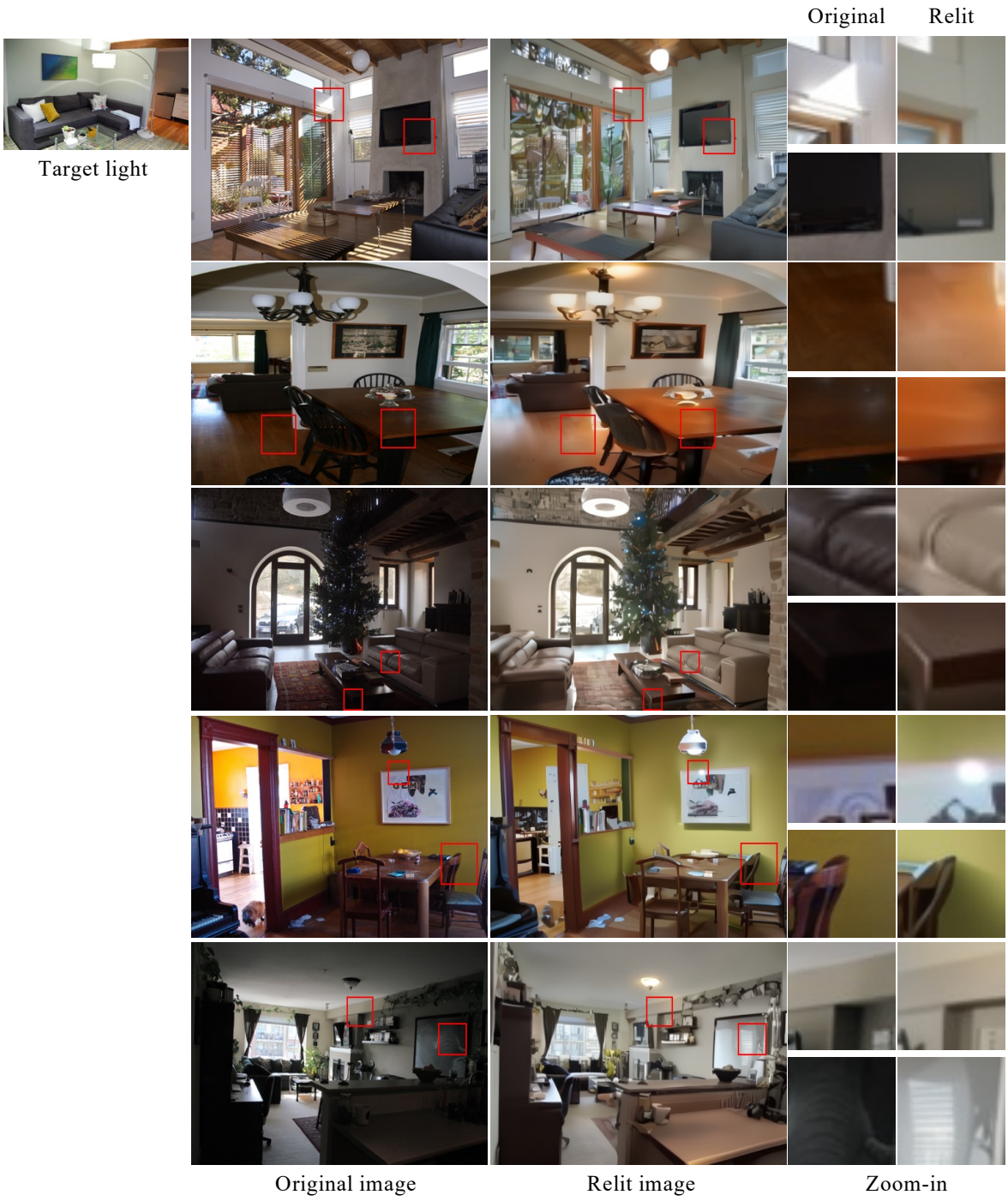

Figure S.3. **Additional Relit Images (switching on ceiling lamps)**. The target lighting is shown in the top-left image, where a ceiling lamp is switched on. Ceiling lamps are very rare in our training data; however, we find that LumiNet is still able to understand them and synthesize plausible relit images, as shown in the third column. In the first row, notice the suppression of gloss near the window at the top (see crop) and the added gloss due to inter-reflection on the TV screen. Also, note how the shaft lighting effect from the source image is suppressed. In the second row, observe how three ceiling lamps significantly brighten the room, with strong gloss visible on both the wooden floor and the dining table. In the third row, notice the sheen on the sofa and the edge of the coffee table, which become clearly visible after relighting. In the fourth row, see how the reflection of the lamp appears on the painting on the side wall. Also, note the shadow cast by the chair on the side wall below the painting. Finally, in the last row, observe how soft shadows along the edges of the ceiling and side wall are suppressed, while soft-light gloss becomes visible. Further, note the reflection on a mirror-like object in the bottom crop.

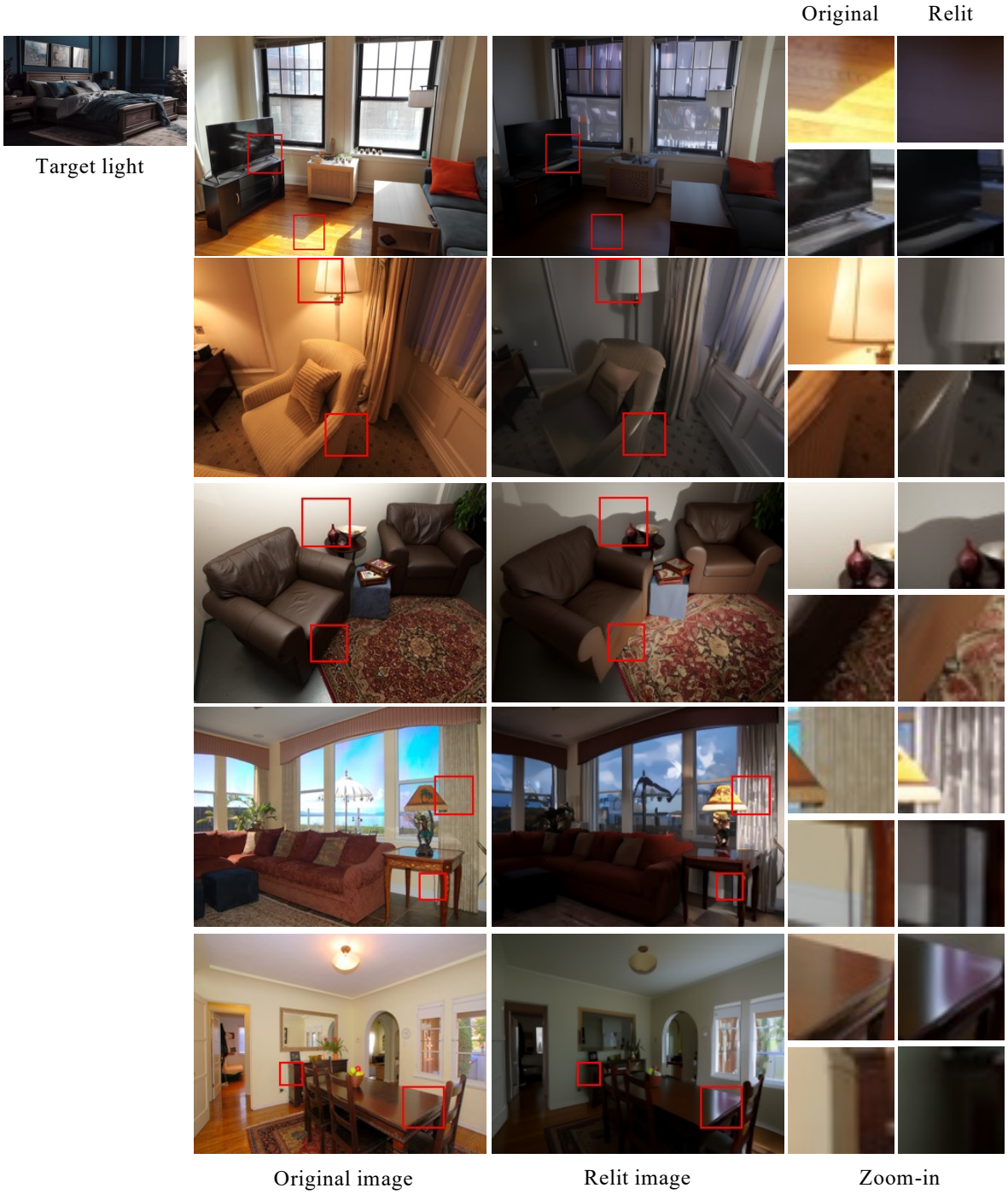

Figure S.4. **Additional Relit Images**. The target light is shown in the top-left image, where all lamps are switched off, and the only illumination comes from diffused natural light entering through a window on the right. The second column displays the source images to be relit to match the target light, while the third column presents the relit images. The final column highlights cropped regions before and after relighting, emphasizing the second-order lighting effects captured by LumiNet. In the top row (first relit image), note the table's reflection in the TV and the strong gloss on the table from the directional window light. In the fourth row, observe how the sky changes to reflect the ambiance of the target light. In the last row, notice specular highlights on the table because of the direction light from the window. Also, notice the shadow cast by the cabinet in the bottom crop.

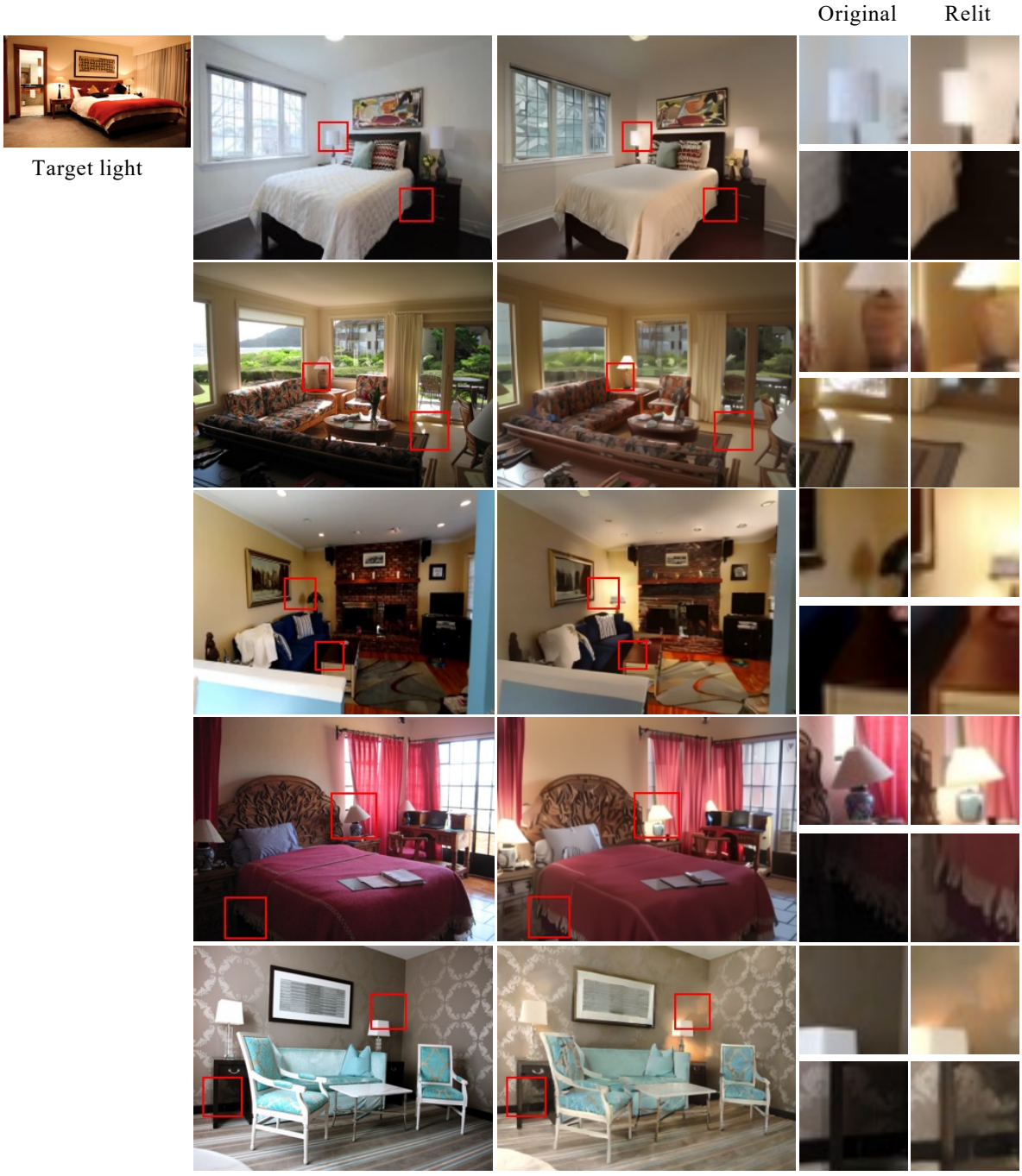

Figure S.5. **Additional Relit Images**. The target lighting is shown in the top-left image, where all lamps are switched on. The second column displays the source images to be relit to match the target lighting, where all lamps are switched off, and the third column presents the relit images. The final column highlights cropped regions before and after relighting. In the top row (first relit image), note the overall change in the room's color and the colored gloss added to the side of the bedsheet. In the second row, notice that the strong gloss on the carpet is removed. In the third row, switching on the side lamps removes the lamp shadow; also, observe the effect of the lamp on the ceiling and the gloss added to the edge of the table, as shown in the crop. In the fourth row, notice that the left side of the bed is now well-lit due to the lamp. Finally, in the last row, observe the gloss added to the wallpaper because of switching on the lamp