

# Supplementary Materials

## I. Overview

In the supplementary materials, we provide more comparison results and detailed results of the user study. We compared 10 existing methods, including BTF [8], SSBF [9], IG [6], WOTF [11], RTV [5], EAP [23], STDN [24], GFES [27], FIP [32] and SemiGANTF [34], except EPS [33] due to its lower generalizability. In applying the methods in comparison, we used their suggested parameters or carefully tuned them to optimize the filtering results. Here, the tested 20 images are from [26].

## II. Comparison with state-of-the-art methods

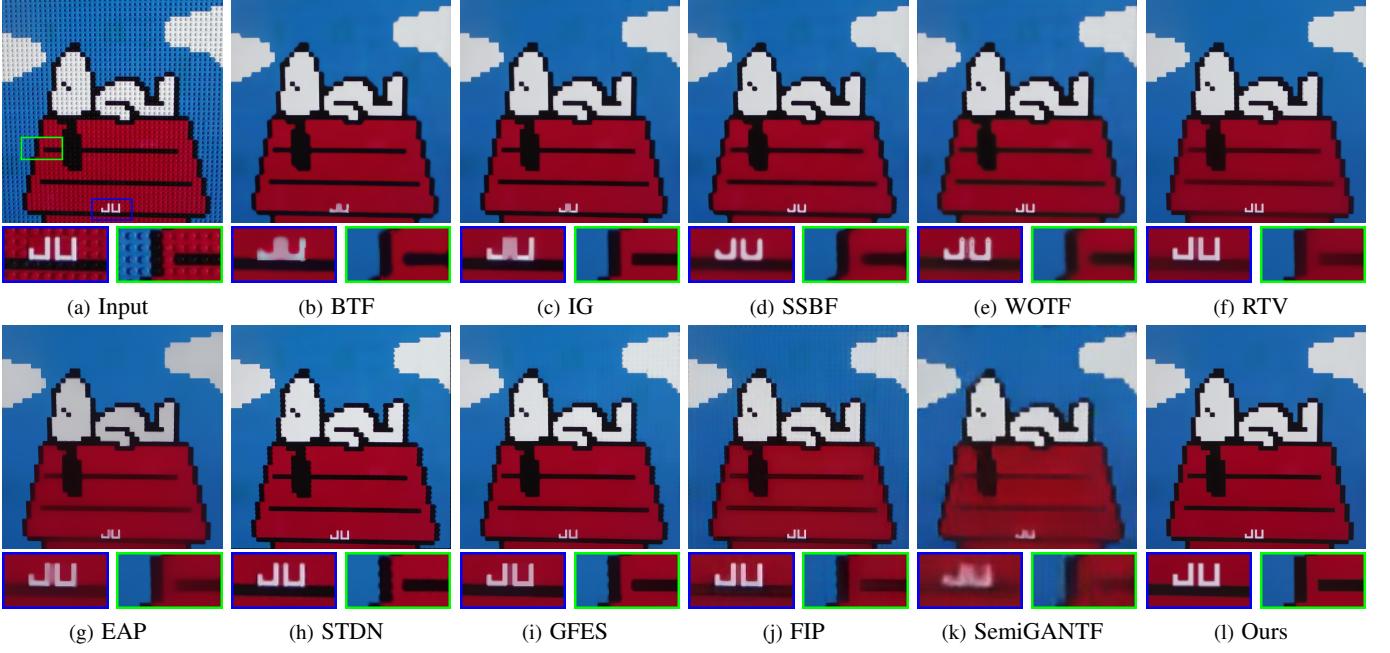


Fig. 1: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 7$ ,  $n_{iter} = 4$ ), (c) IG ( $\sigma = 4$ ), (d) SSBF ( $\sigma = 4$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 4$ ), (e) WOTF ( $k = 9$ ,  $n_{iter} = 4$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 0.75$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 15$ ,  $n_{iter} = 4$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

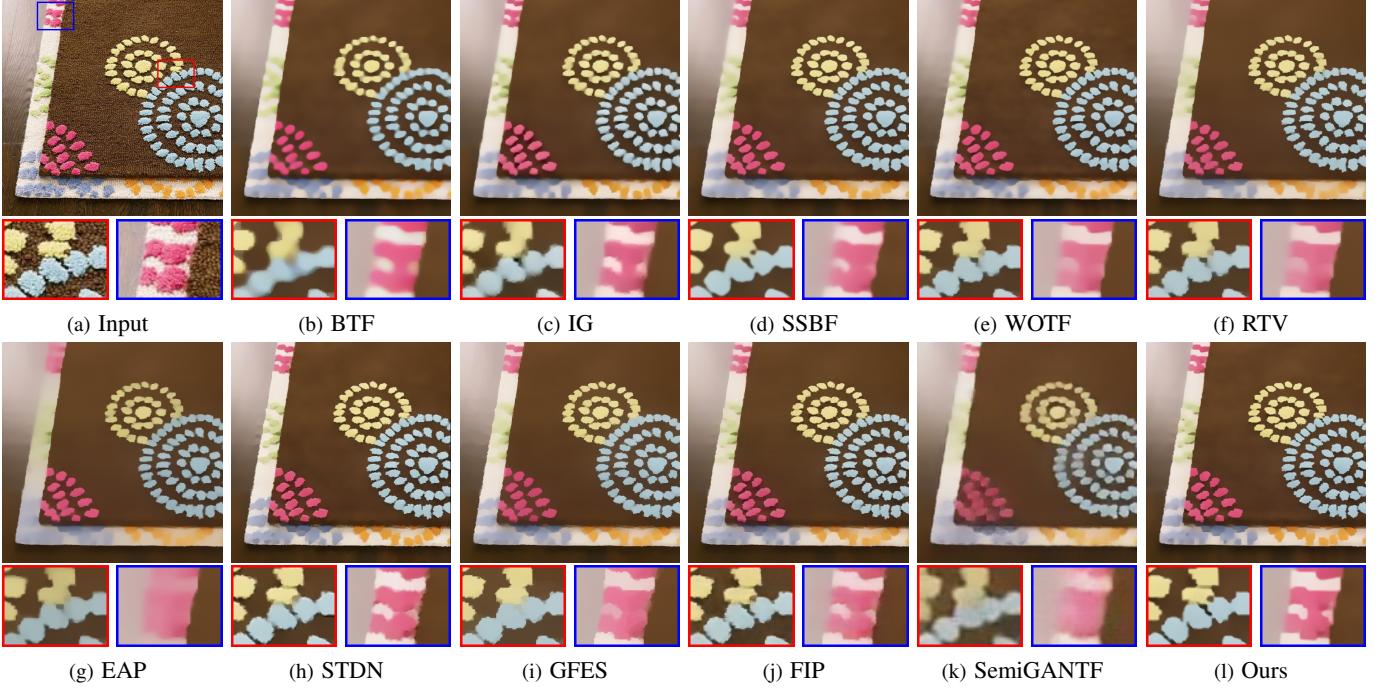


Fig. 2: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 9$ ,  $n_{iter} = 3$ ), (c) IG ( $\sigma = 4$ ), (d) SSBF ( $\sigma = 5$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 3$ ), (e) WOTF ( $k = 7$ ,  $n_{iter} = 6$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 6$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 6$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 1.25$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 13$ ,  $n_{iter} = 3$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.



Fig. 3: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 9$ ,  $n_{iter} = 3$ ), (c) IG ( $\sigma = 5$ ), (d) SSBF ( $\sigma = 4$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 4$ ), (e) WOTF ( $k = 9$ ,  $n_{iter} = 4$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 6$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 6$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 1.05$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 9$ ,  $n_{iter} = 4$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

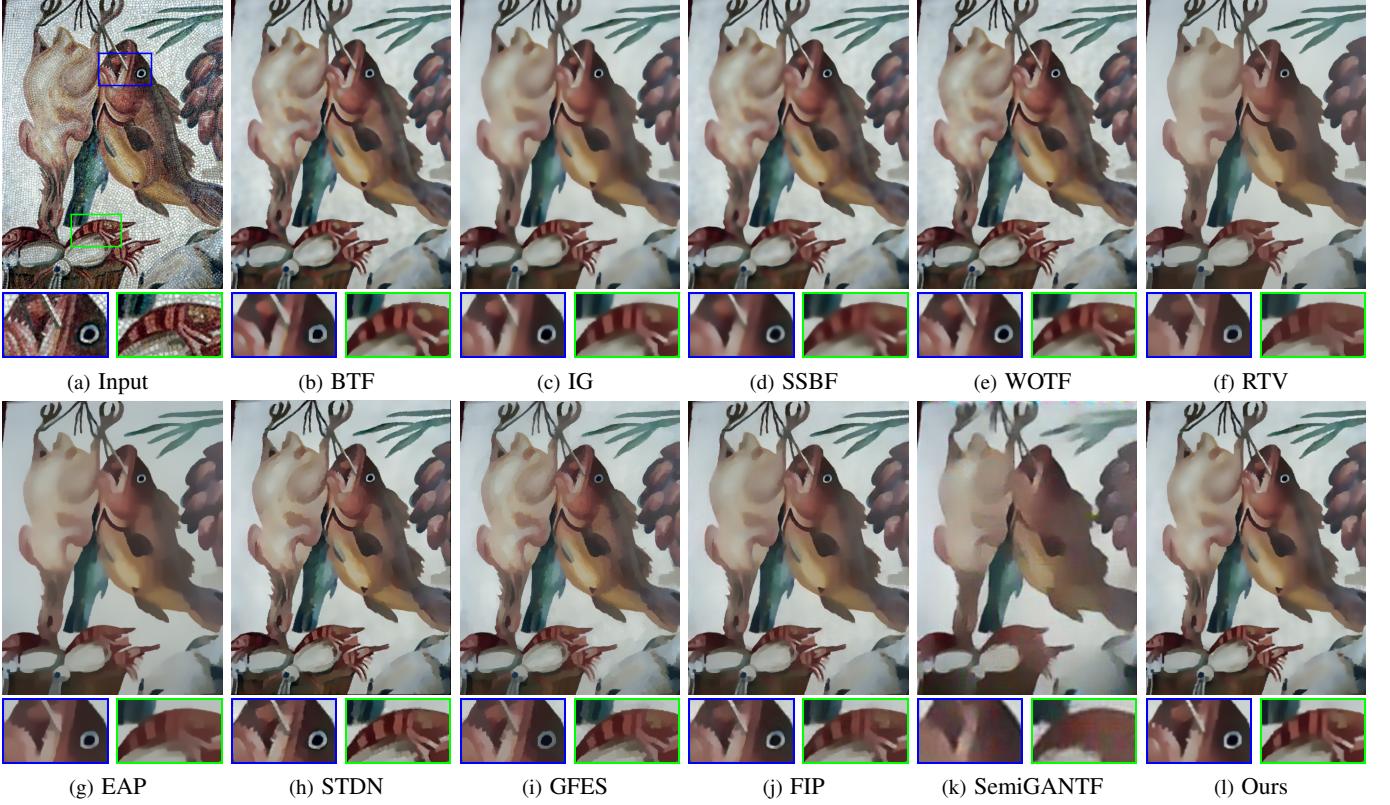


Fig. 4: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 5, n_{iter} = 4$ ), (c) IG ( $\sigma = 3$ ), (d) SSBF ( $\sigma = 3, \sigma_r = 0.1, n_{iter} = 4$ ), (e) WOTF ( $k = 7, n_{iter} = 4$ ), (f) RTV ( $\lambda = 0.015, \sigma = 4, n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015, \sigma = 4, n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 0.8, n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 13, n_{iter} = 3$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

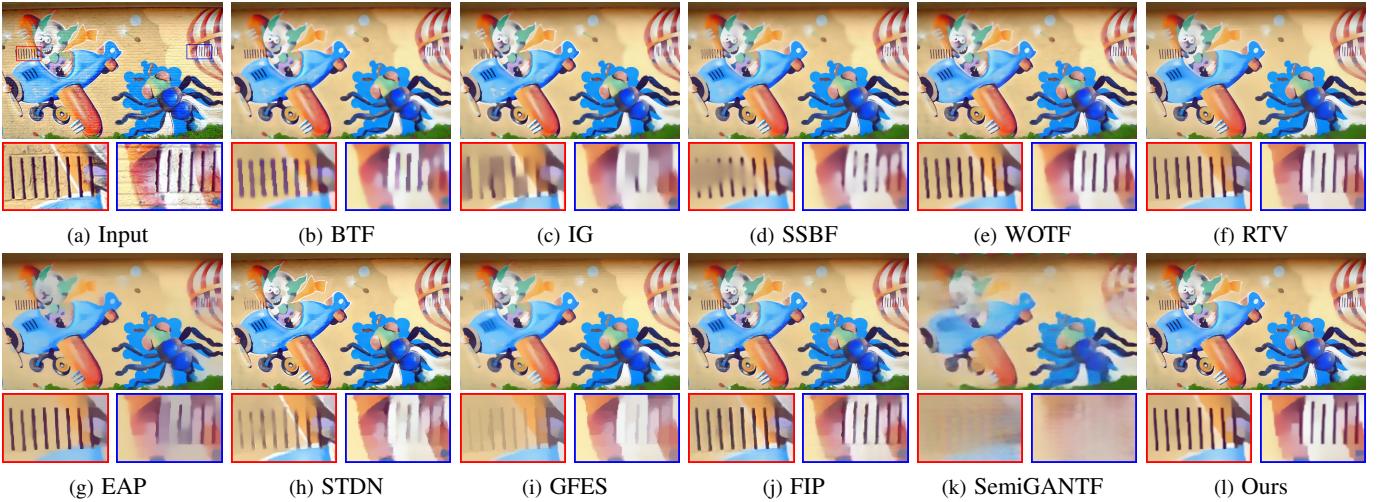


Fig. 5: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 7, n_{iter} = 4$ ), (c) IG ( $\sigma = 4$ ), (d) SSBF ( $\sigma = 5, \sigma_r = 0.1, n_{iter} = 3$ ), (e) WOTF ( $k = 7, n_{iter} = 5$ ), (f) RTV ( $\lambda = 0.015, \sigma = 4, n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015, \sigma = 4, n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 1.25, n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 13, \sigma_s = t, n_{iter} = 4$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

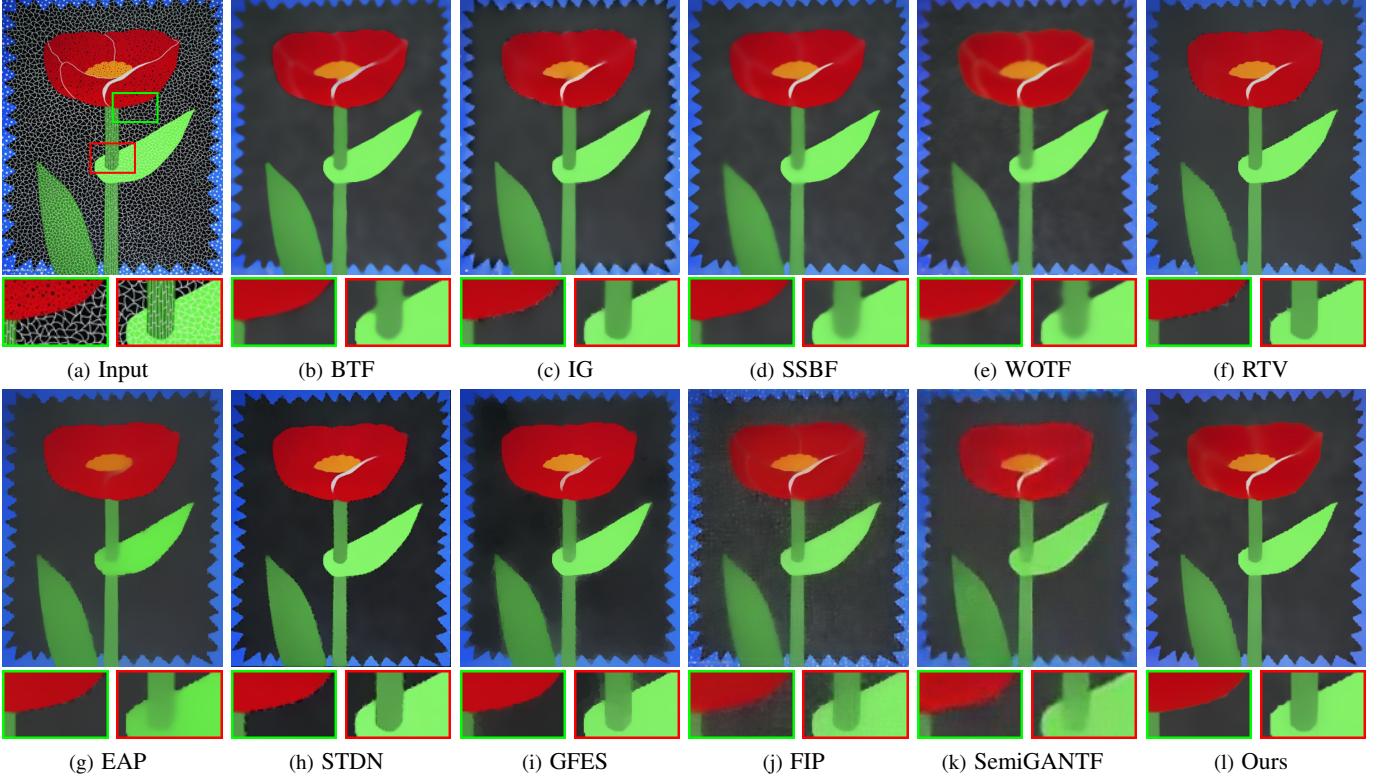


Fig. 6: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 7$ ,  $n_{iter} = 4$ ), (c) IG ( $\sigma = 4$ ), (d) SSBF ( $\sigma = 4$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 4$ ), (e) WOTF ( $k = 9$ ,  $n_{iter} = 4$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 1.25$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 13$ ,  $n_{iter} = 5$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

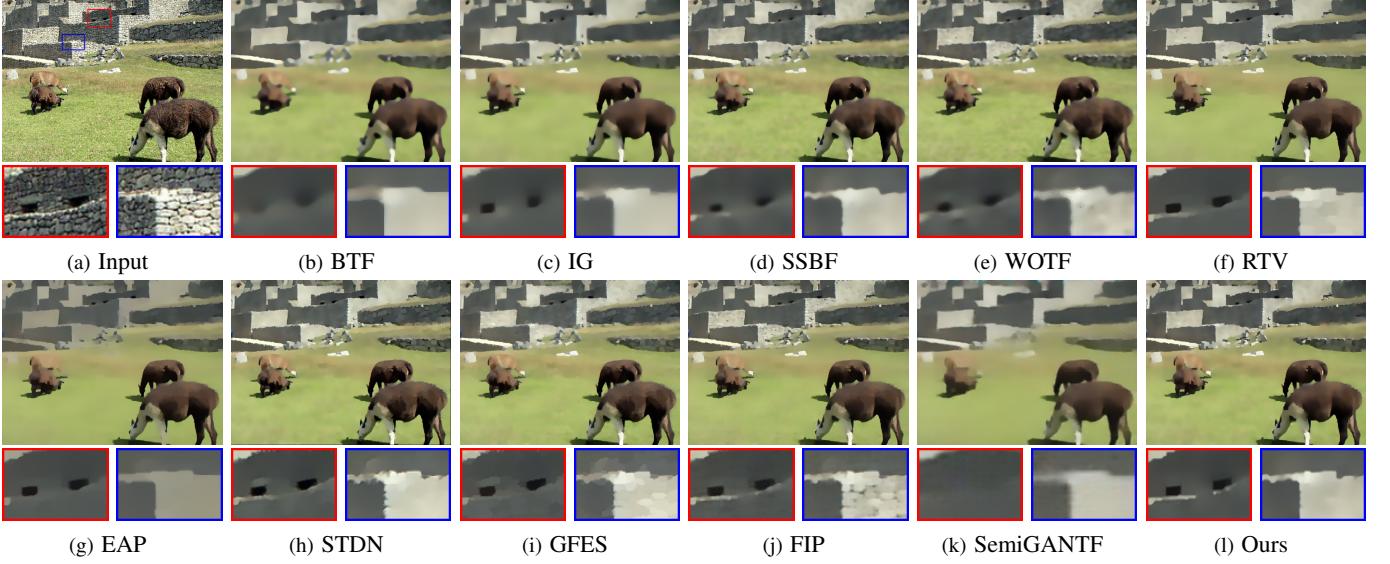


Fig. 7: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 11$ ,  $n_{iter} = 3$ ), (c) IG ( $\sigma = 5$ ), (d) SSBF ( $\sigma = 4$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 4$ ), (e) WOTF ( $k = 11$ ,  $n_{iter} = 3$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 0.75$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 13$ ,  $n_{iter} = 4$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

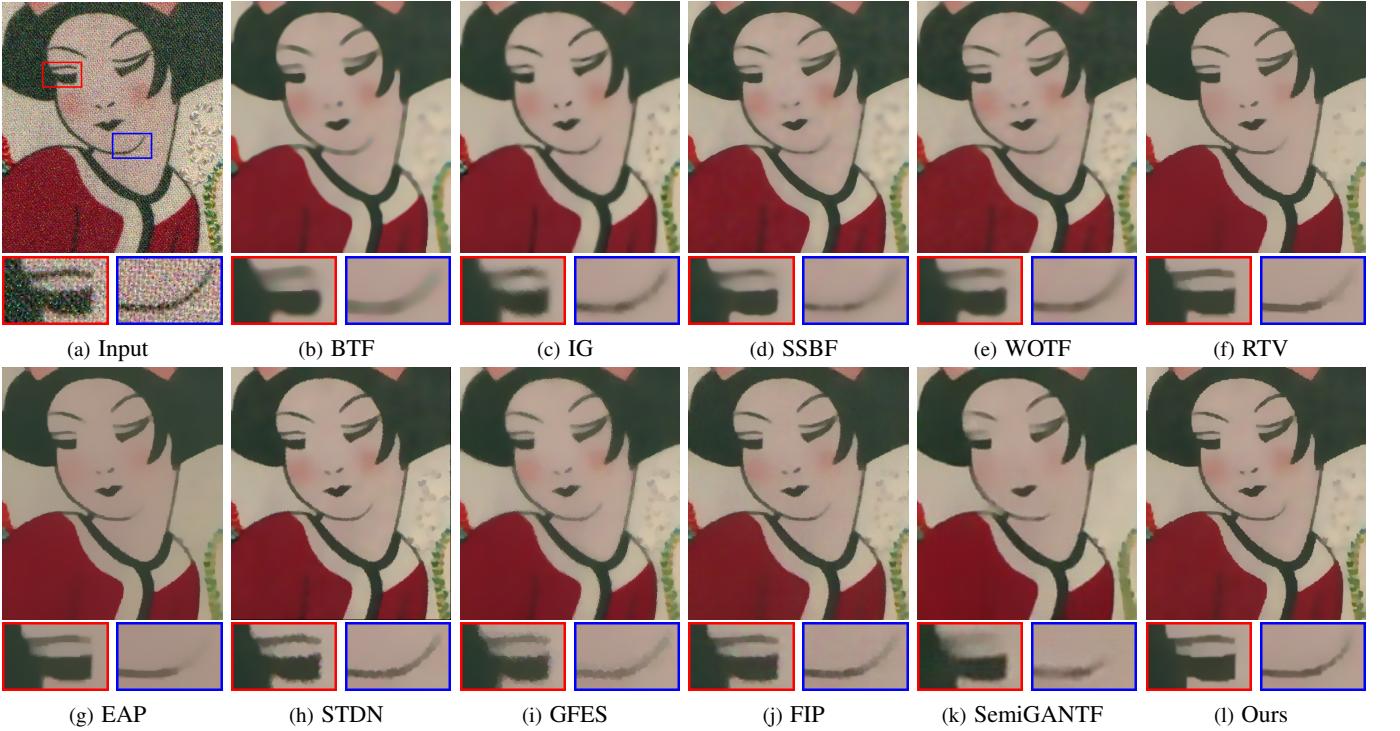


Fig. 8: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 9$ ,  $n_{iter} = 4$ ), (c) IG ( $\sigma = 4$ ), (d) SSBF ( $\sigma = 4$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 4$ ), (e) WOTF ( $k = 9$ ,  $n_{iter} = 4$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 0.75$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 13$ ,  $n_{iter} = 4$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

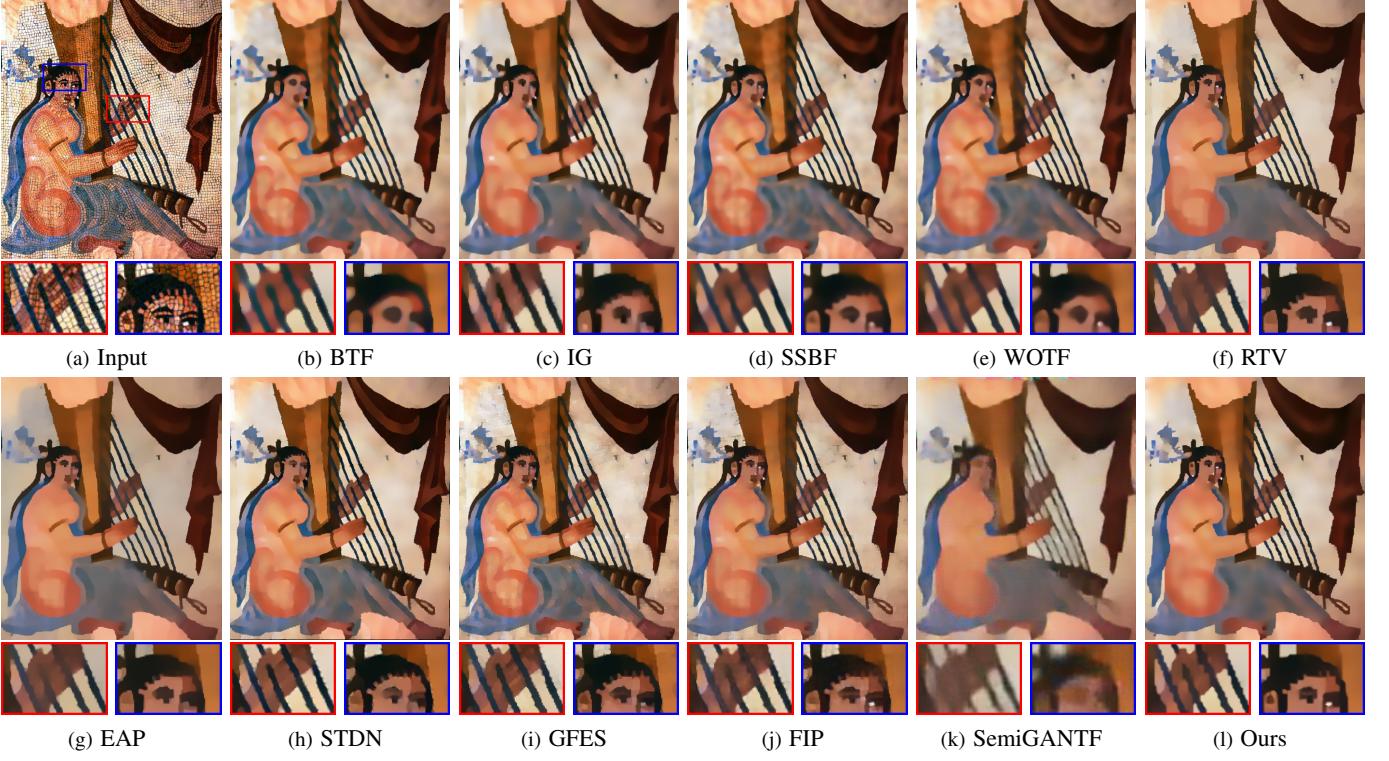


Fig. 9: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 7$ ,  $n_{iter} = 3$ ), (c) IG ( $\sigma = 4$ ), (d) SSBF ( $\sigma = 3$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 4$ ), (e) WOTF ( $k = 9$ ,  $n_{iter} = 4$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 0.5$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 13$ ,  $n_{iter} = 5$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

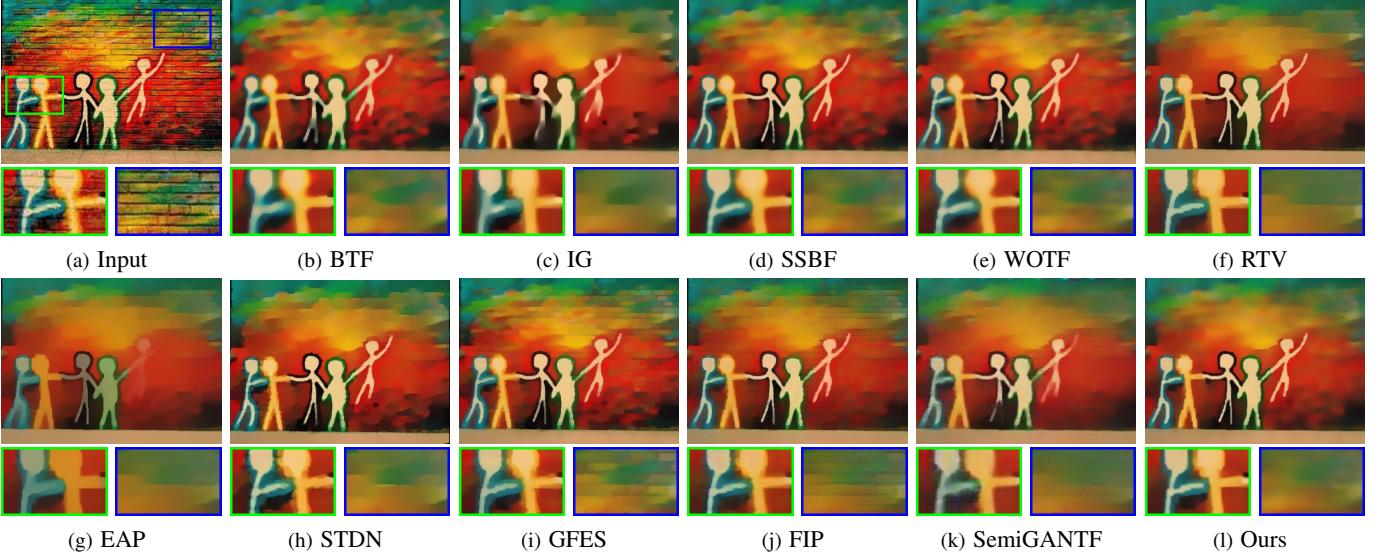


Fig. 10: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 7$ ,  $n_{iter} = 3$ ), (c) IG ( $\sigma = 5$ ), (d) SSBF ( $\sigma = 3$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 3$ ), (e) WOTF ( $k = 9$ ,  $n_{iter} = 3$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 3$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 3$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 0.5$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 11$ ,  $\sigma_s = t$ ,  $n_{iter} = 3$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

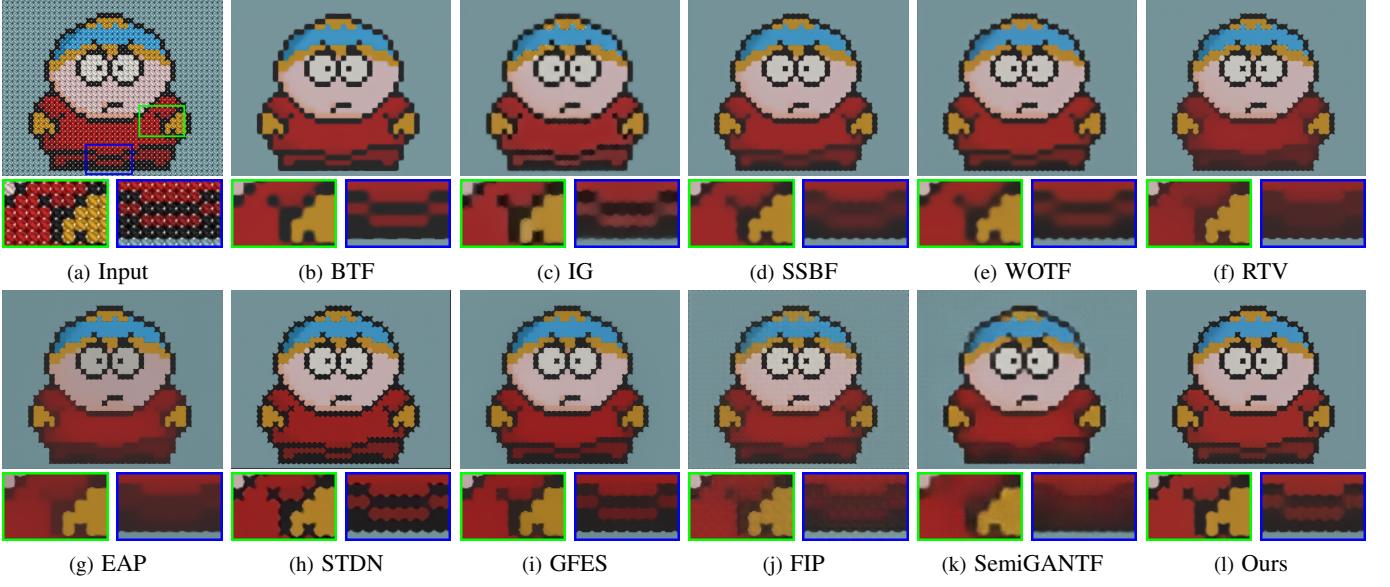


Fig. 11: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 7$ ,  $n_{iter} = 3$ ), (c) IG ( $\sigma = 4$ ), (d) SSBF ( $\sigma = 4$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 3$ ), (e) WOTF ( $k = 9$ ,  $n_{iter} = 4$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 0.75$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 13$ ,  $n_{iter} = 4$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

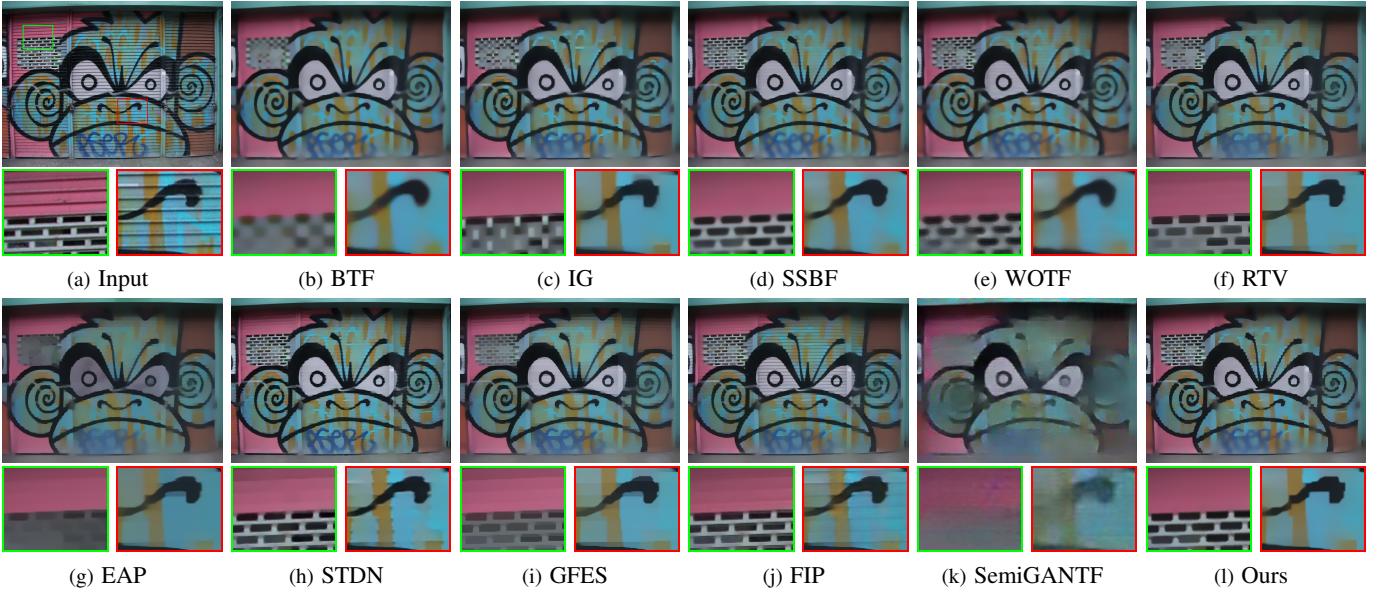


Fig. 12: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 9$ ,  $n_{iter} = 4$ ), (c) IG ( $\sigma = 5$ ), (d) SSBF ( $\sigma = 4$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 4$ ), (e) WOTF ( $k = 11$ ,  $n_{iter} = 4$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 5$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 5$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 1.25$ ,  $n_{iter} = 12$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 13$ ,  $n_{iter} = 5$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

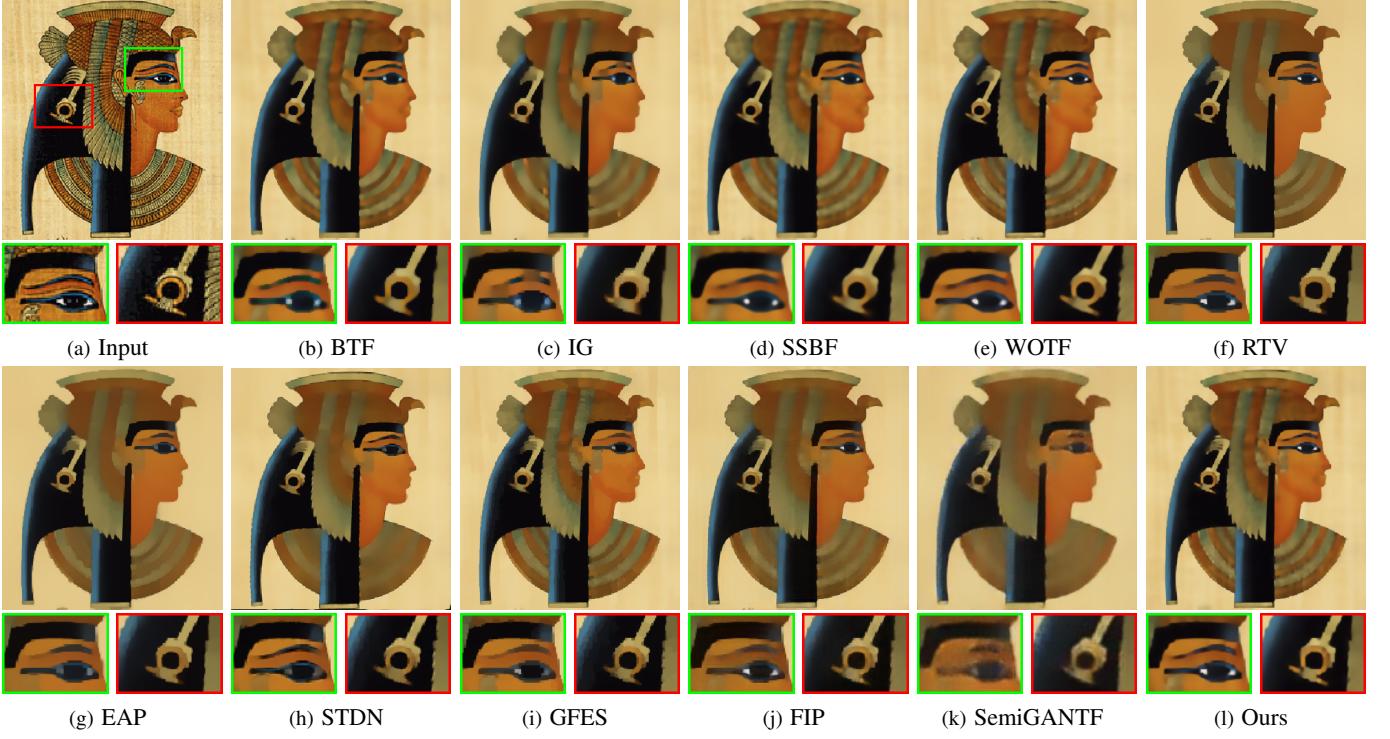


Fig. 13: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 5$ ,  $n_{iter} = 3$ ), (c) IG ( $\sigma = 3$ ), (d) SSBF ( $\sigma = 2$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 4$ ), (e) WOTF ( $k = 7$ ,  $n_{iter} = 3$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 3$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 3$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 0.5$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 7$ ,  $n_{iter} = 3$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.



Fig. 14: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 9$ ,  $n_{iter} = 3$ ), (c) IG ( $\sigma = 4$ ), (d) SSBF ( $\sigma = 5$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 5$ ), (e) WOTF ( $k = 9$ ,  $n_{iter} = 5$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 6$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 6$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 0.9$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 15$ ,  $\sigma_s = t$ ,  $n_{iter} = 6$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

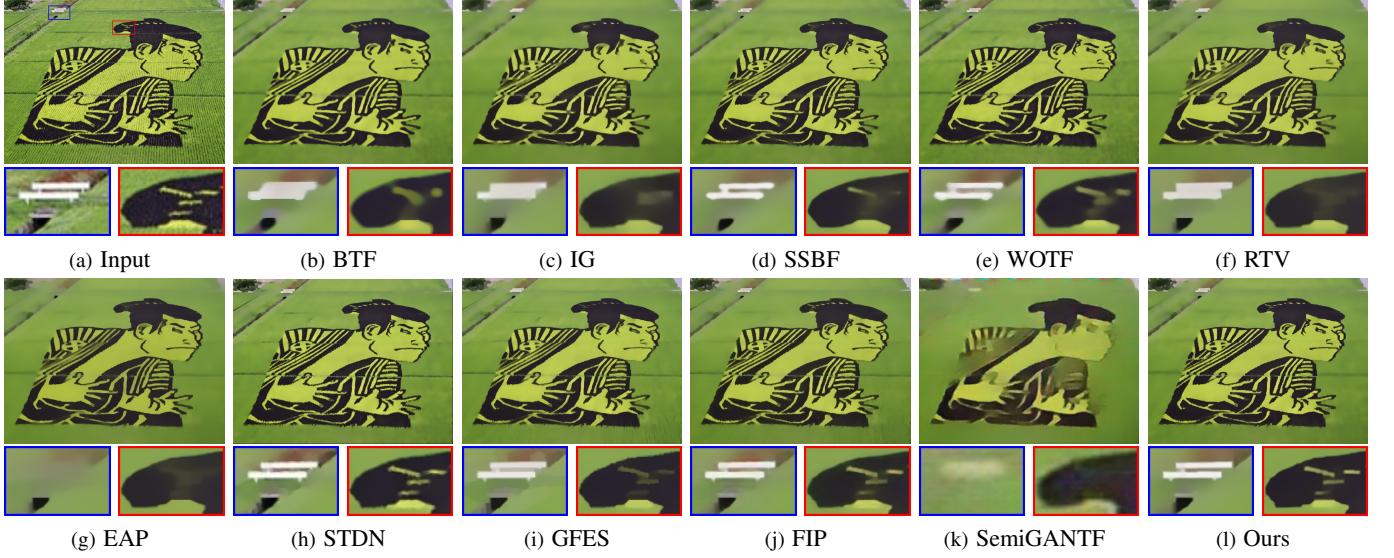


Fig. 15: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 7$ ,  $n_{iter} = 6$ ), (c) IG ( $\sigma = 6$ ), (d) SSBF ( $\sigma = 6$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 4$ ), (e) WOTF ( $k = 9$ ,  $n_{iter} = 4$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 6$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 5$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 0.8$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 15$ ,  $n_{iter} = 4$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

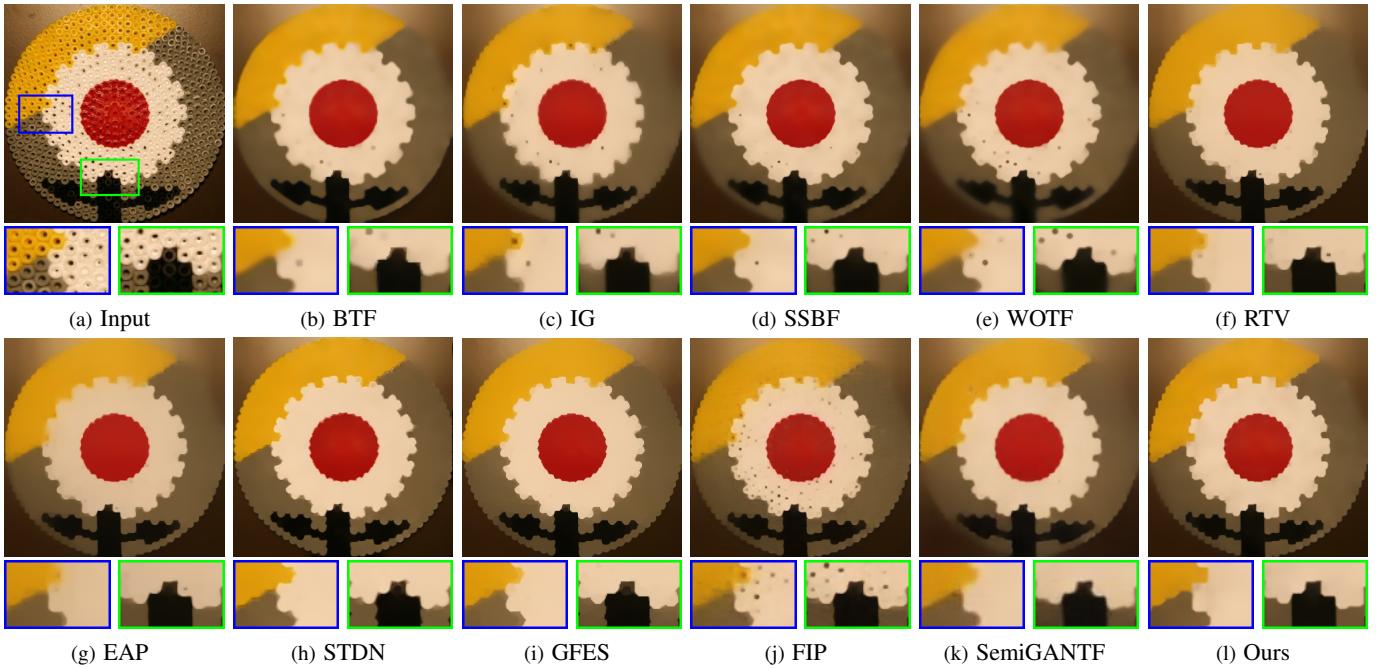


Fig. 16: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 7$ ,  $n_{iter} = 4$ ), (c) IG ( $\sigma = 5$ ), (d) SSBF ( $\sigma = 4$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 4$ ), (e) WOTF ( $k = 9$ ,  $n_{iter} = 6$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 5$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 5$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 1.25$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 15$ ,  $n_{iter} = 5$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

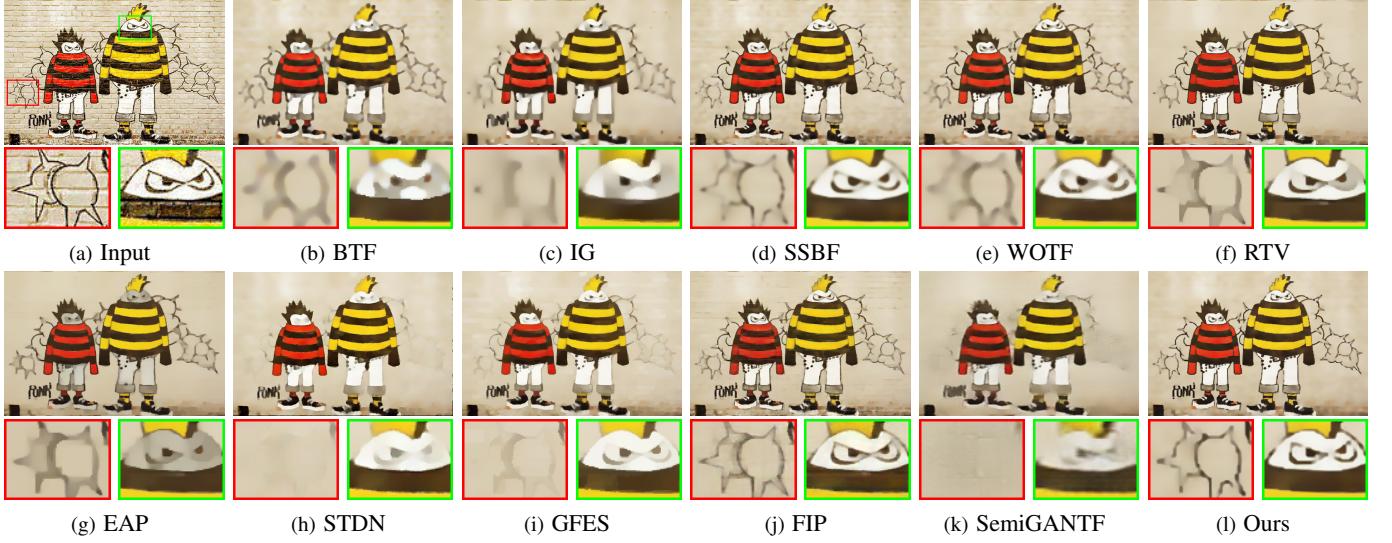


Fig. 17: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 7$ ,  $n_{iter} = 3$ ), (c) IG ( $\sigma = 4$ ), (d) SSBF ( $\sigma = 3$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 3$ ), (e) WOTF ( $k = 9$ ,  $n_{iter} = 4$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 0.75$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 9$ ,  $n_{iter} = 5$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.



Fig. 18: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 7$ ,  $n_{iter} = 4$ ), (c) IG ( $\sigma = 4$ ), (d) SSBF ( $\sigma = 4$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 4$ ), (e) WOTF ( $k = 7$ ,  $n_{iter} = 5$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 4$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 0.9$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 15$ ,  $n_{iter} = 4$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

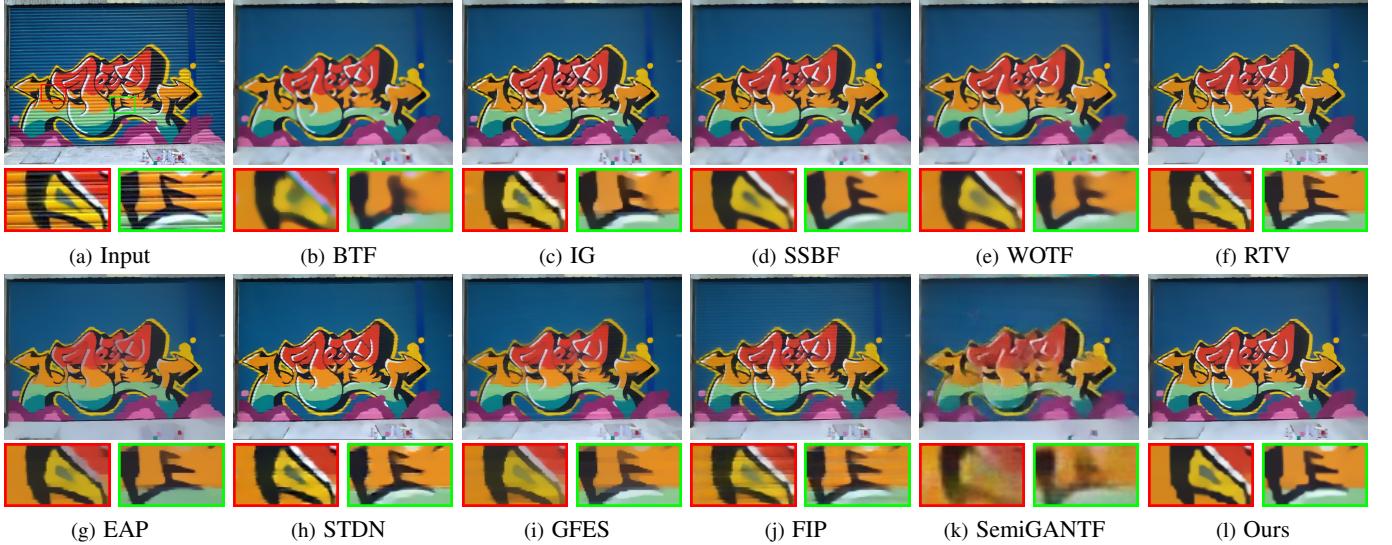


Fig. 19: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 7$ ,  $n_{iter} = 4$ ), (c) IG ( $\sigma = 3$ ), (d) SSBF ( $\sigma = 4$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 4$ ), (e) WOTF ( $k = 9$ ,  $n_{iter} = 5$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 3$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 3$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 0.75$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 13$ ,  $n_{iter} = 4$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.



Fig. 20: Comparison of texture filtering results.(a) Input image, (b) BTF ( $k = 9$ ,  $n_{iter} = 4$ ), (c) IG ( $\sigma = 5$ ), (d) SSBF ( $\sigma = 5$ ,  $\sigma_r = 0.1$ ,  $n_{iter} = 3$ ), (e) WOTF ( $k = 9$ ,  $n_{iter} = 5$ ), (f) RTV ( $\lambda = 0.015$ ,  $\sigma = 6$ ,  $n_{iter} = 4$ ), (g) EAP ( $\lambda = 0.015$ ,  $\sigma = 6$ ,  $n_{iter} = 4$ ), (h) STDN, (i) GFES ( $\lambda = 0.9$ ,  $n_{iter} = 10$ ), (j) FIP (RTV is used here.), (k) SemiGANTF, (l) Ours ( $t = 13$ ,  $n_{iter} = 4$ ). Here, we scale the filtered results of SemiGANTF to the original image size using bi-triple interpolation for comparison, since its predictions are all of  $256 \times 256$  pixels.

### III. The user study

We invited 30 graduate students (ten female and twenty male) to participate this user study. For this, we designed an interactive interface to facilitate the user to select the better filtering results, as shown in Fig. 21. In the study, we selected the 20 texture images and their filtering results listed above, and divided these participants into 10 groups for one-by-one comparison. Every time, we displayed a pair of filtered images, which were filtered by our method and a compared method, and inquired them to respectively select the better one. Here, we displayed the filtered images randomly in the window. Finally, the statistics about their percentages are shown in Fig. 22.

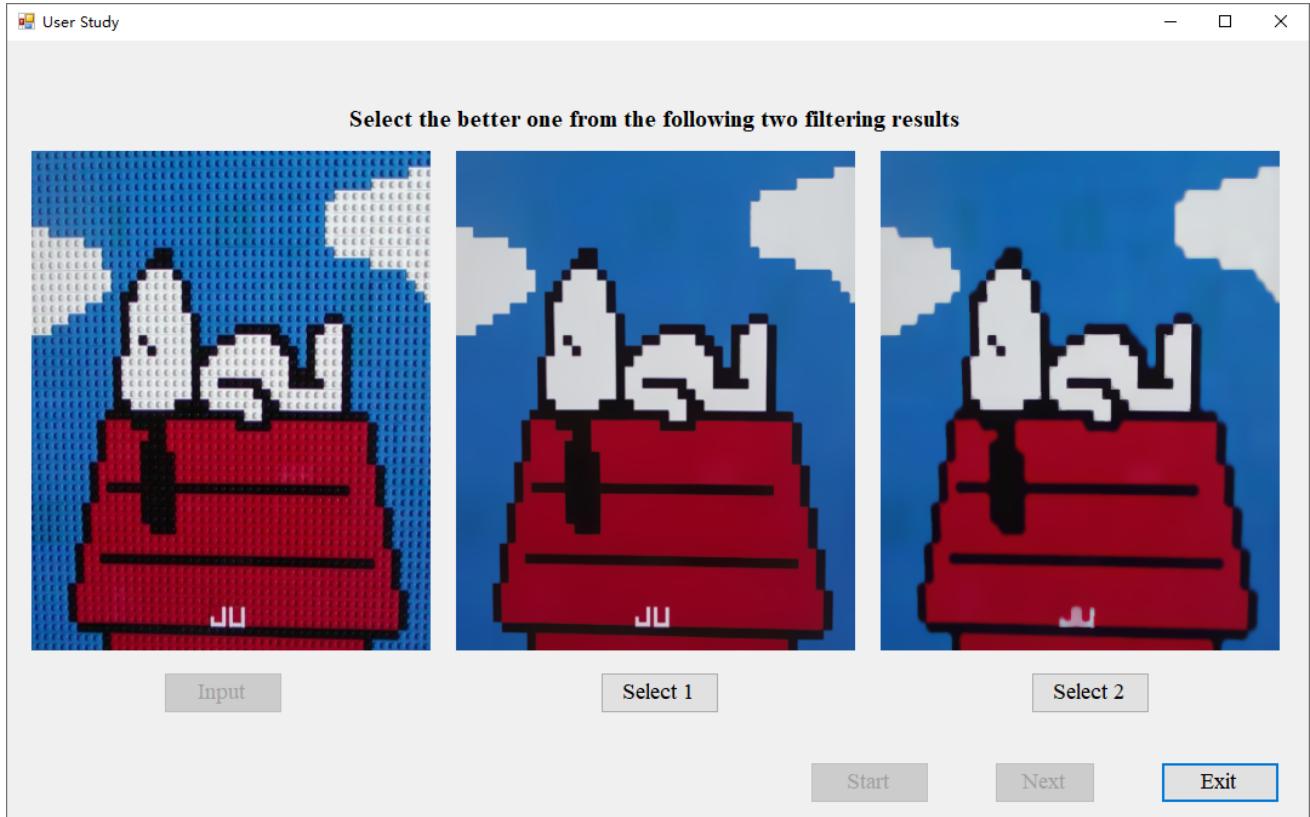


Fig. 21: A snapshot of our interactive interface for user study.

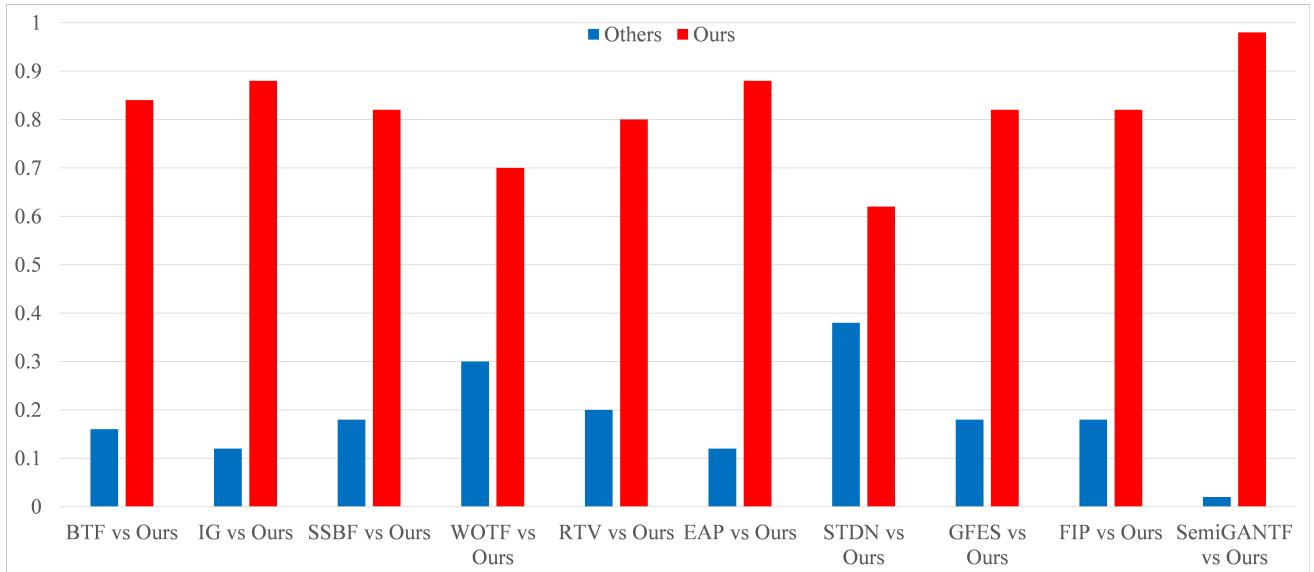


Fig. 22: Statistics for one-by-one comparison of the filtering results. The blue and red bars represent the ratios of the preferred images generated by the compared methods and ours, respectively.