Title: Learning how to combine internal and external denoising methods

Submitted to GCPR2013 - Paper ID: 37

Supplementary material

Contents

1 MLP vs. BM3D: More results

In the paper, we conclude that (i) external denoising is usually better on irregular and smooth images, (ii) internal denoising is usually better on regular, repeating structures, and (iii) there is no easy way to determine which is better, and our findings do not support the PatchSNR criterion [1]. In this section, we provide more evidence to support these claims. In addition, we show that external denoising using large patches is usually better than internal denoising at high noise levels, which is also contrary Mosseri et al.'s conclusion [1].

We applied BM3D (an internal method) and MLP (an external method) on 2500 test images and show the 20 images where MLP compares the most favorably compared to BM3D as well as the 20 images where MLP compares the least favorably compared to BM3D. For each image, we show (i) the noisy image itself (ii) the "preference" image, where pixels on which MLP performs better than BM3D are colored white, and (iii) the PatchSNR image, where whiter pixels mean higher PatchSNR, which in turn should indicate a higher preference for external methods according to Mosseri et al [1]. For each image, we consider six noise levels, ranging from $\sigma = 10$ to $\sigma = 170$. The test images come from the following five datasets, each containing 500 images: (i) The Berkeley segmentation dataset, (ii) the McGill dataset, (iii) the Pascal VOC 2007 test set, (iv) the Pascal VOC 2011 training set, and (v) 500 images from the ImageNet dataset. The test images were not used for training our neutal network. In this document, images have been converted to JPEG in order to save space.

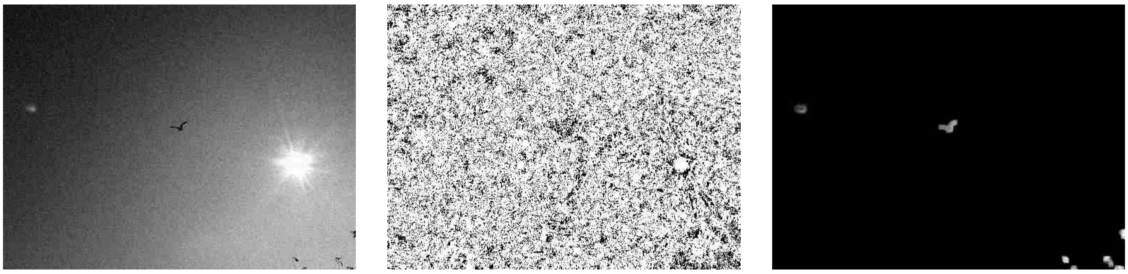
1.1 MLP beats BM3D

In image 1 we see that MLP performs better than BM3D on smooth image regions, whereas in image 17, we see that MLP performs better on irregular patterns (e.g. grass) than BM3D. In contrast, the PatchSNR predicts that internal methods such as BM3D are preferred on smooth image regions [1]. The PatchSNR also tends to predict that internal methods are to be preferred on irregular patterns. We make similar observations on the remaining images. For image image, we see that the higher the noise level, the better MLP performs against BM3D. This is contrary to the conclusions make in [1].

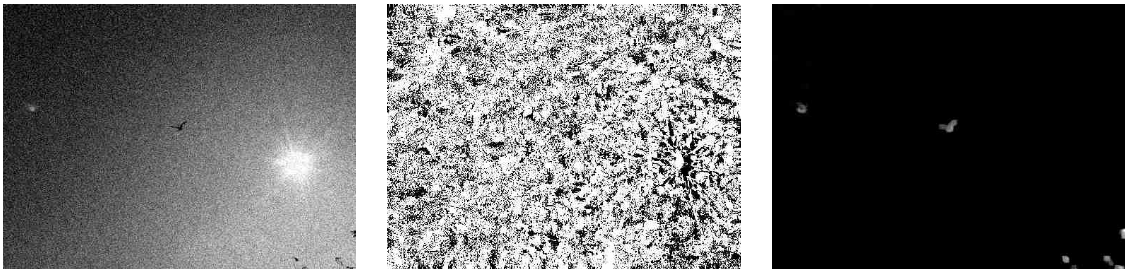
MLP beats BM3D, image 1

Notes: The image is very smooth and the MLP performs well compared to BM3D almost everywhere. However, the PatchSNR suggests the opposite: BM3D should be used almost everywhere. We also see that the MLP performs particularly well at high noise levels, whereas Mosseri et al [1] claim that external methods perform less well at high noise levels.

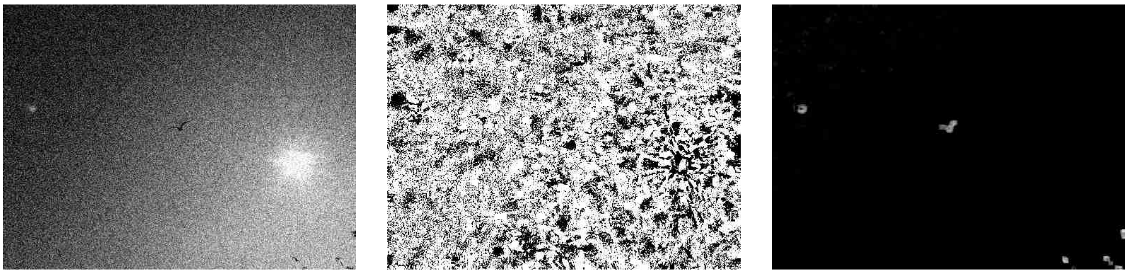
noisy image | preference image | Patch SNR



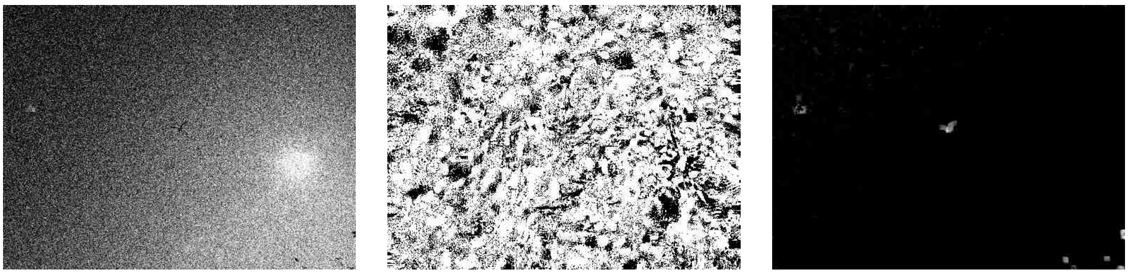
$\sigma = 10$, PSNR: 28.11dB | **MLP: 43.27dB**, BM3D: 43.20dB | between 0.00 and 6.19



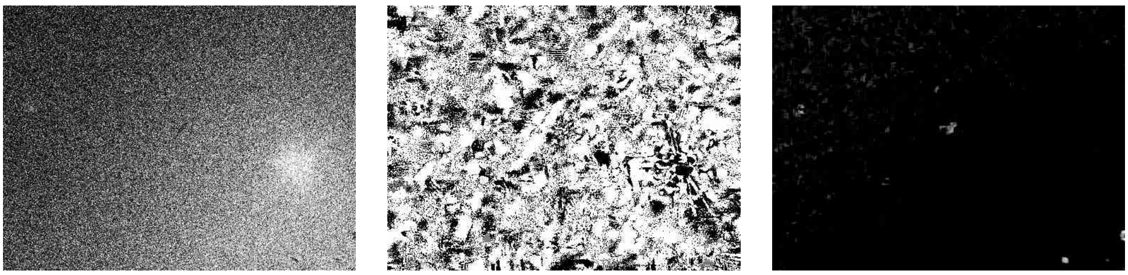
$\sigma = 25$, PSNR: 20.13dB | **MLP: 40.78dB**, BM3D: 38.86dB | between 0.00 and 2.43



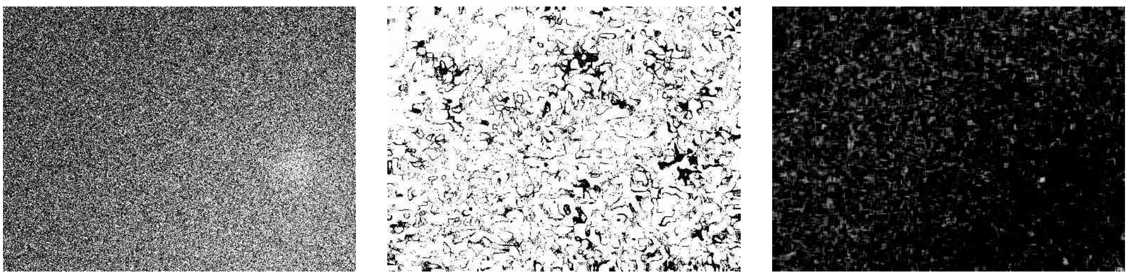
$\sigma = 35$, PSNR: 17.26dB | **MLP: 39.32dB**, BM3D: 37.36dB | between 0.00 and 1.84



$\sigma = 50$, PSNR: 14.16dB | **MLP: 37.38dB**, BM3D: 35.99dB | between 0.00 and 1.04



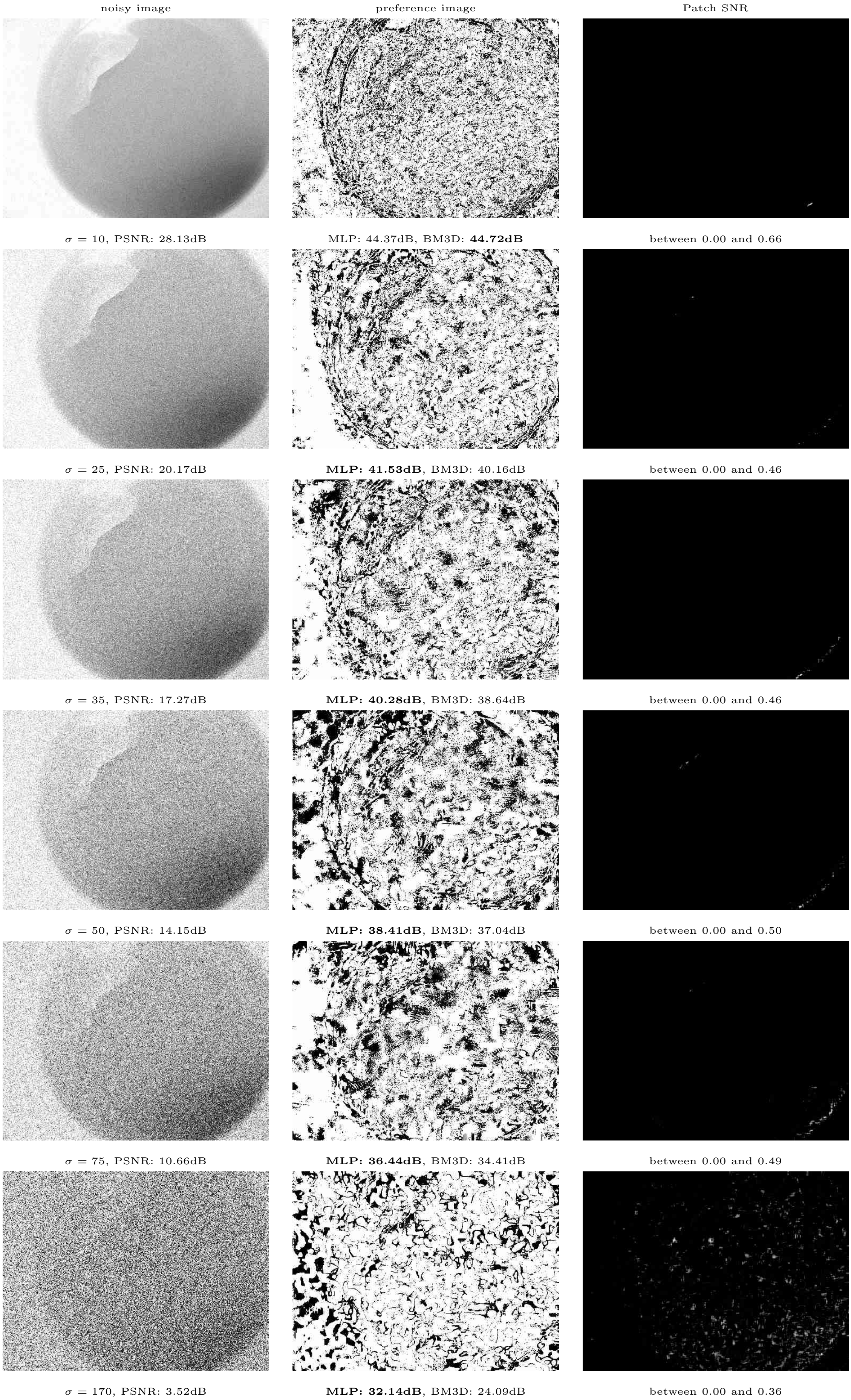
$\sigma = 75$, PSNR: 10.66dB | **MLP: 35.49dB**, BM3D: 34.07dB | between 0.00 and 0.65



$\sigma = 170$, PSNR: 3.53dB | **MLP: 30.77dB**, BM3D: 22.39dB | between 0.00 and 0.44

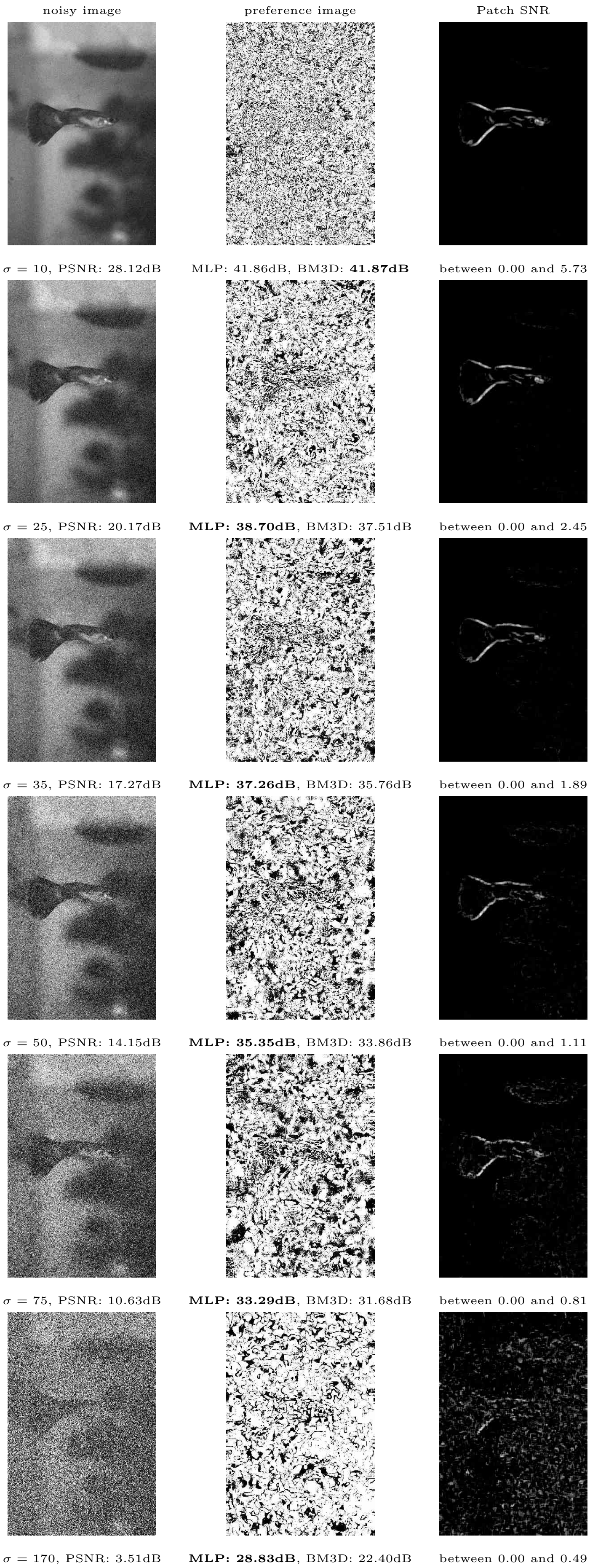
MLP beats BM3D, image 2

Notes: The image is very smooth and the MLP performs well compared to BM3D almost everywhere. However, the PatchSNR suggests the opposite: BM3D should be used almost everywhere. We also see that the MLP performs particularly well at high noise levels, whereas Mosseri et al [1] claim that external methods perform less well at high noise levels.



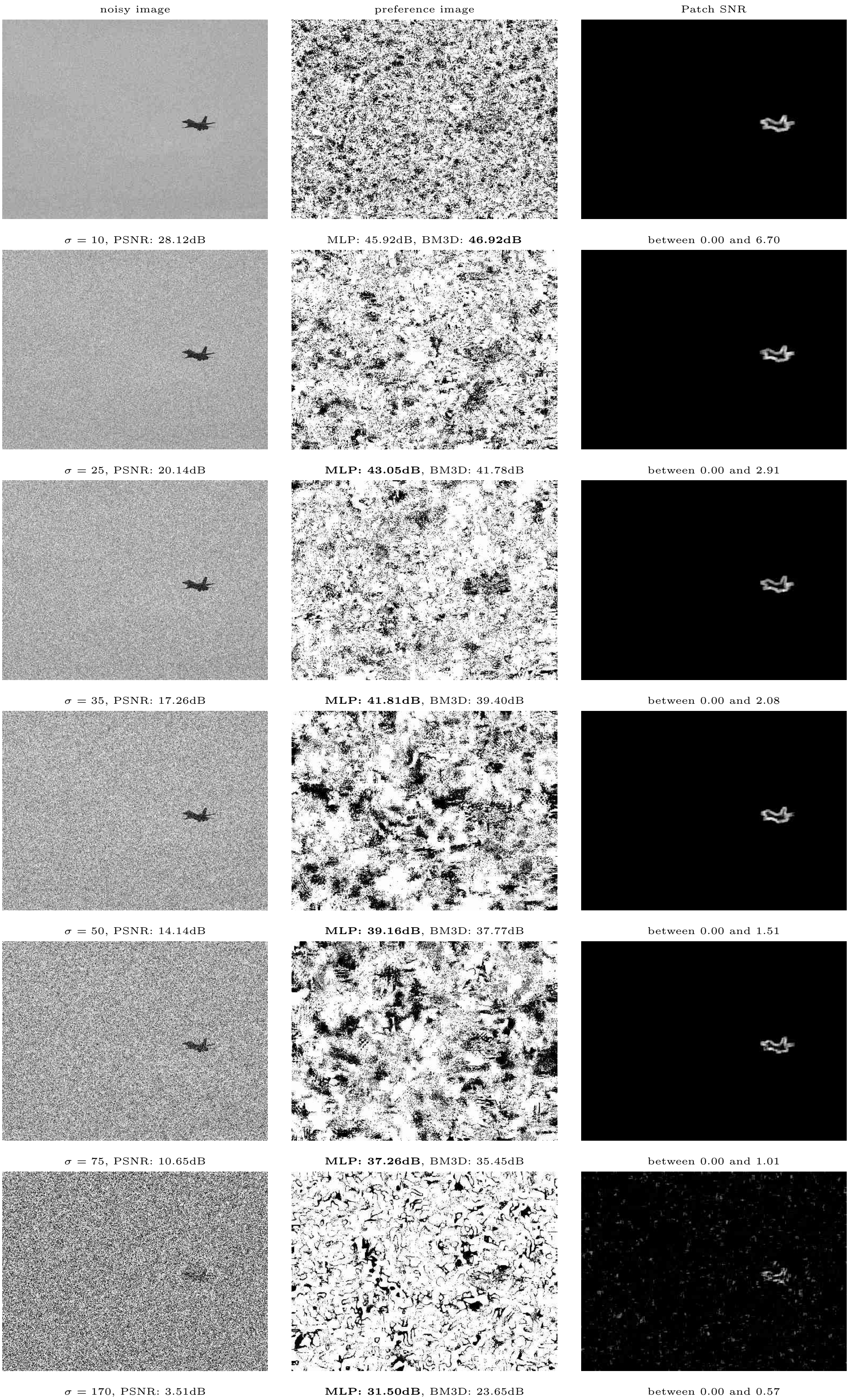
MLP beats BM3D, image 3

Notes: The image is very smooth and the MLP performs well compared to BM3D almost everywhere. However, the PatchSNR suggests the opposite: BM3D should be used almost everywhere. We also see that the MLP performs particularly well at high noise levels, whereas Mosseri et al [1] claim that external methods perform less well at high noise levels.



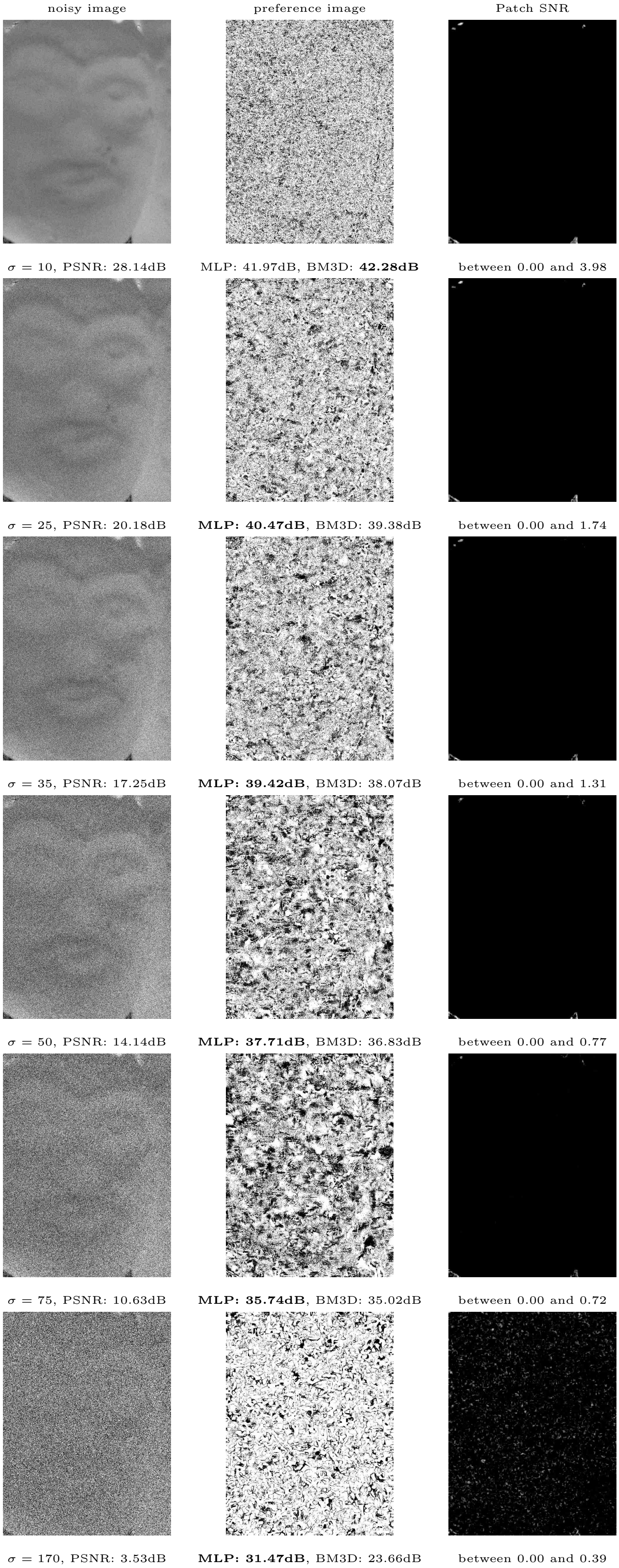
MLP beats BM3D, image 4

Notes: The image is very smooth and the MLP performs well compared to BM3D almost everywhere. However, the PatchSNR suggests the opposite: BM3D should be used almost everywhere. We also see that the MLP performs particularly well at high noise levels, whereas Mosseri et al [1] claim that external methods perform less well at high noise levels.



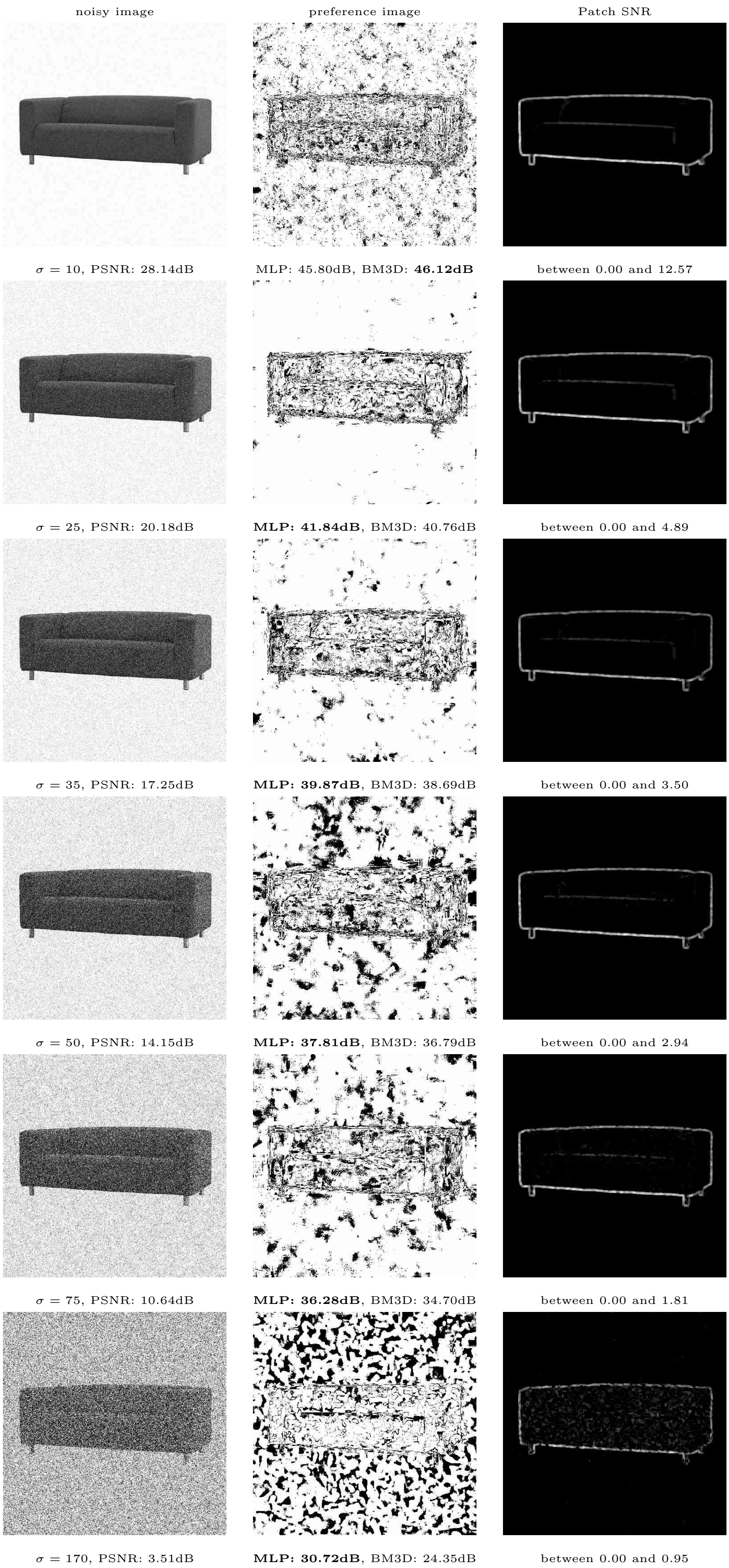
MLP beats BM3D, image 5

Notes: The image is very smooth and the MLP performs well compared to BM3D almost everywhere. However, the PatchSNR suggests the opposite: BM3D should be used almost everywhere. We also see that the MLP performs particularly well at high noise levels, whereas Mosseri et al [1] claim that external methods perform less well at high noise levels.



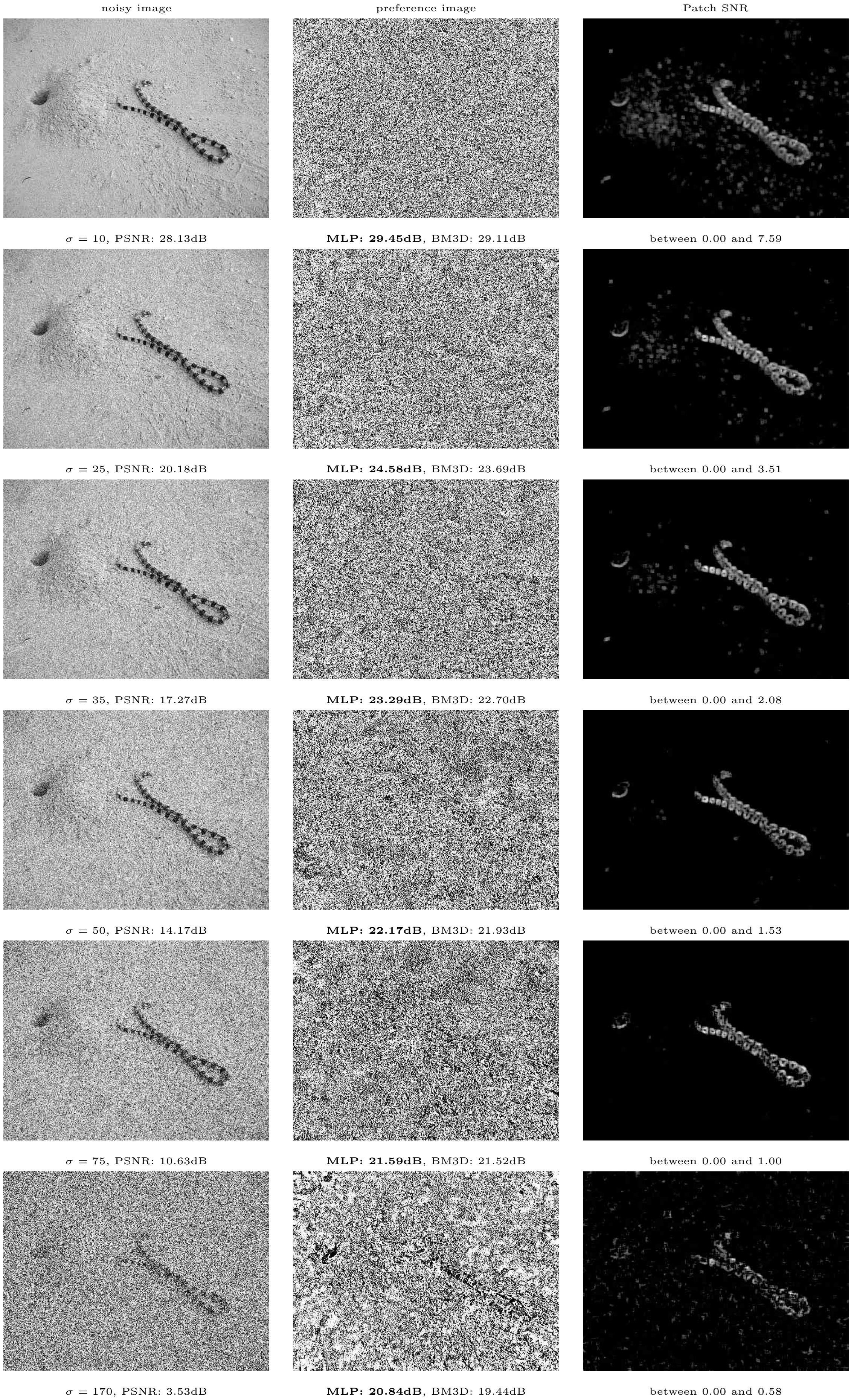
MLP beats BM3D, image 6

Notes: The MLP performs very well on the smooth regions of this image. However, the PatchSNR suggests that BM3D should be used on the smooth regions.



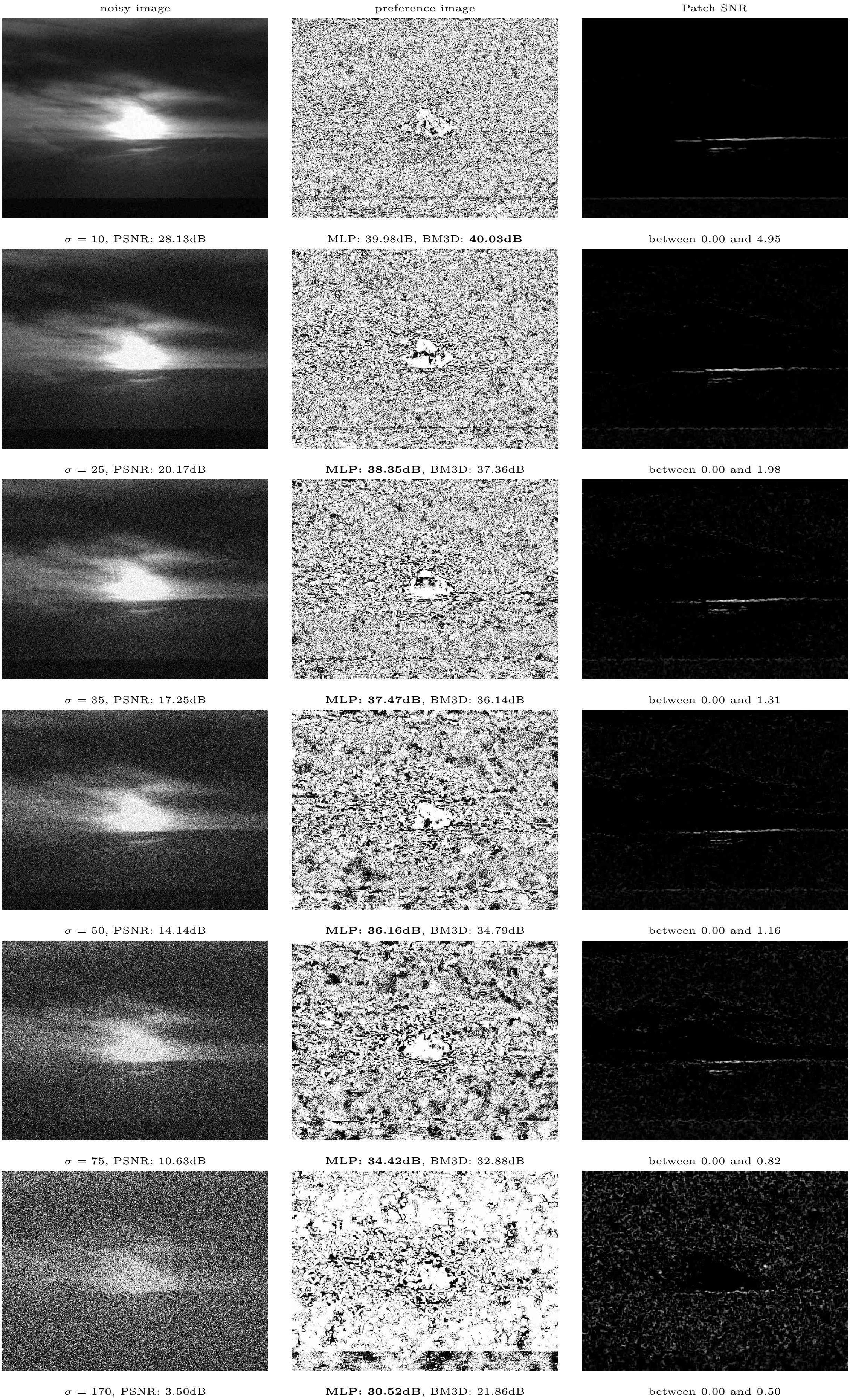
MLP beats BM3D, image 7

Notes: This is an example of an image with irregular textures (the sand). It is difficult to say where exactly the MLP performs best, but the MLP outperforms BM3D at every noise level. However, the PatchSNR suggests that BM3D should be used on the sand.



MLP beats BM3D, image 8

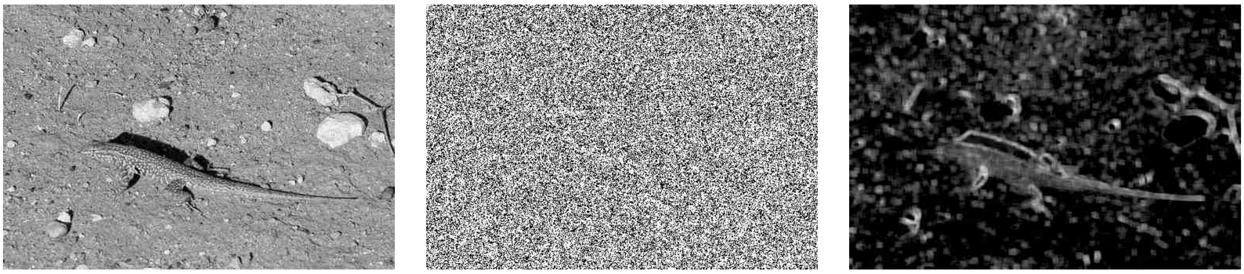
Notes: The image is very smooth and the MLP performs well compared to BM3D almost everywhere. However, the PatchSNR suggests the opposite: BM3D should be used almost everywhere. We also see that the MLP performs particularly well at high noise levels, whereas Mosseri et al [1] claim that external methods perform less well at high noise levels.



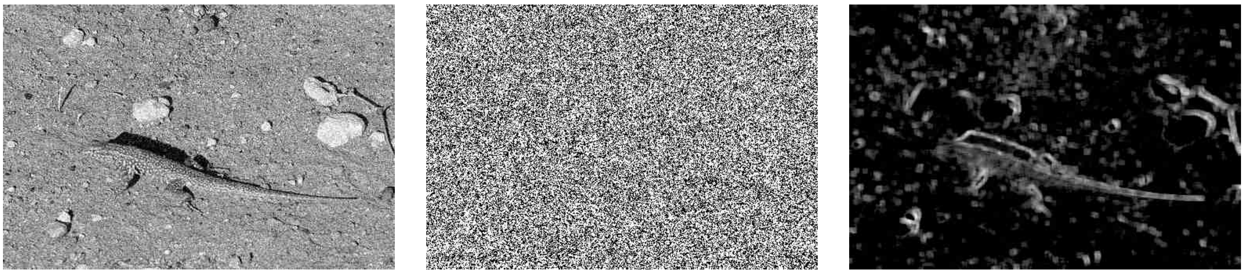
MLP beats BM3D, image 9

Notes: This is an example of an image with irregular textures (the sand). It is difficult to say where exactly the MLP performs best, but the MLP outperforms BM3D at every noise level. However, the PatchSNR suggests that BM3D should be used on the sand.

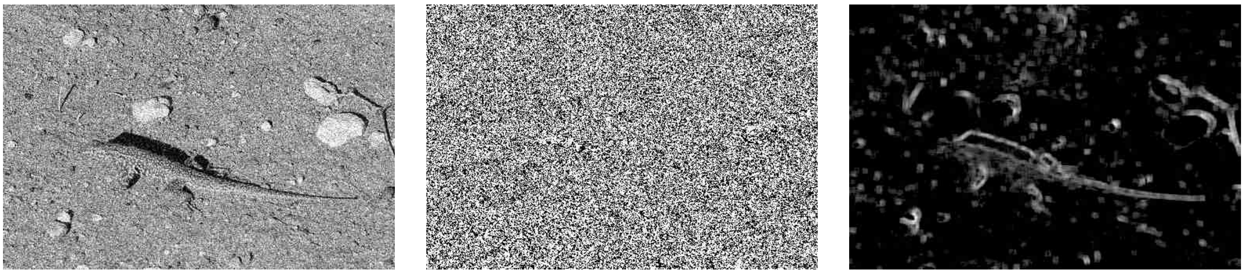
noisy image | preference image | Patch SNR



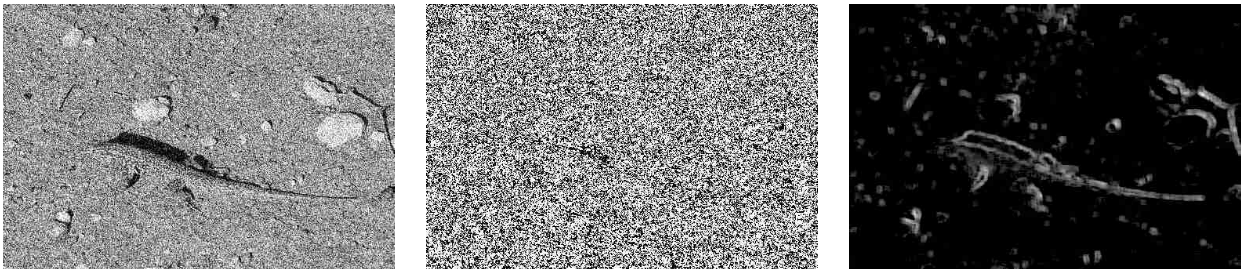
$\sigma = 10$, PSNR: 28.11dB | **MLP: 29.26dB**, BM3D: 29.05dB | between 0.00 and 8.29



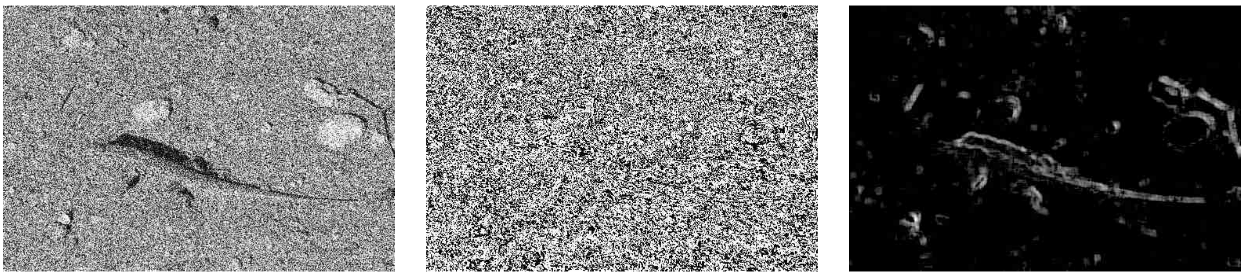
$\sigma = 25$, PSNR: 20.16dB | **MLP: 23.63dB**, BM3D: 22.79dB | between 0.00 and 3.65



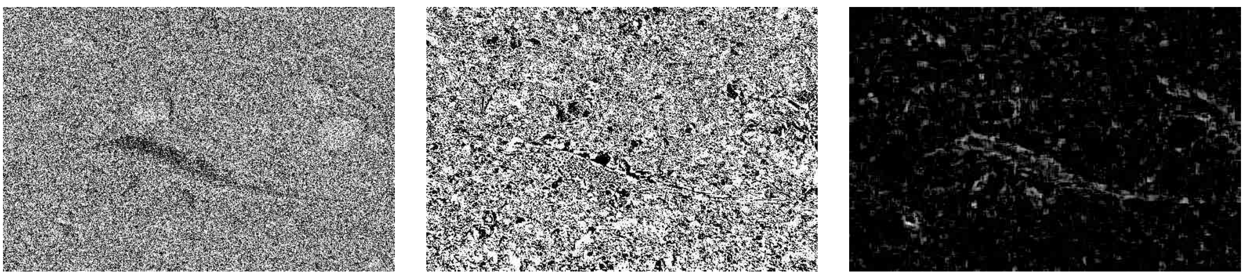
$\sigma = 35$, PSNR: 17.26dB | **MLP: 22.09dB**, BM3D: 21.30dB | between 0.00 and 2.66



$\sigma = 50$, PSNR: 14.18dB | **MLP: 20.67dB**, BM3D: 20.04dB | between 0.00 and 2.10



$\sigma = 75$, PSNR: 10.62dB | **MLP: 19.59dB**, BM3D: 19.32dB | between 0.00 and 1.15

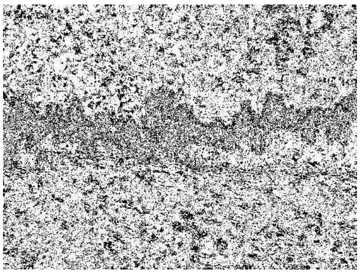
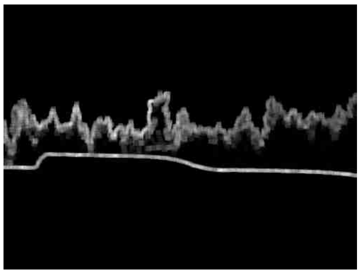


$\sigma = 170$, PSNR: 3.53dB | **MLP: 18.40dB**, BM3D: 17.71dB | between 0.00 and 0.80

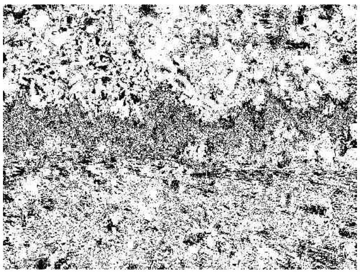
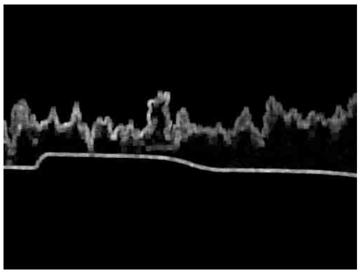
MLP beats BM3D, image 10

Notes: This image is mostly smooth. MLP performs well compared to BM3D on the smooth areas. Again, the PatchSNR suggests otherwise.

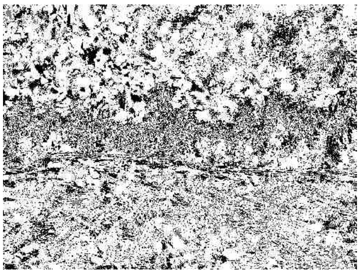
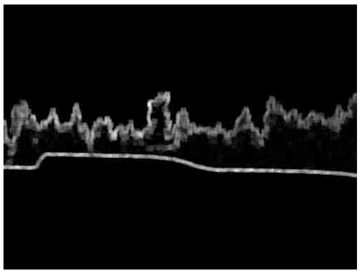
noisy image | preference image | Patch SNR

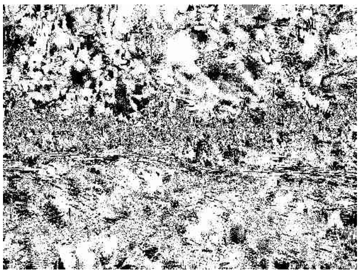
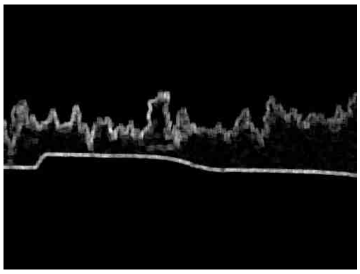
σ = 10, PSNR: 28.15dB | **MLP: 37.71dB**, BM3D: 37.57dB | between 0.00 and 8.70

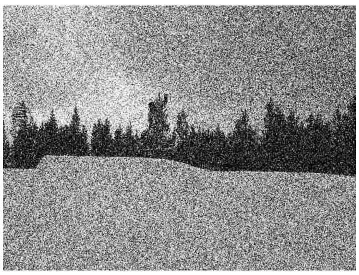
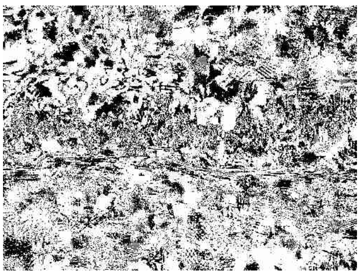
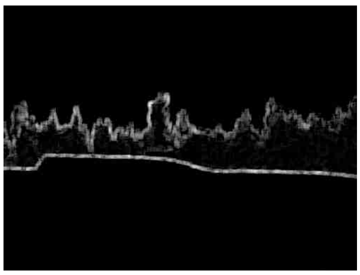
σ = 25, PSNR: 20.18dB | **MLP: 32.70dB**, BM3D: 31.86dB | between 0.00 and 3.82

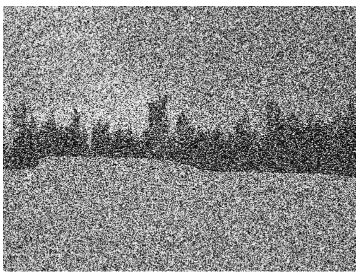
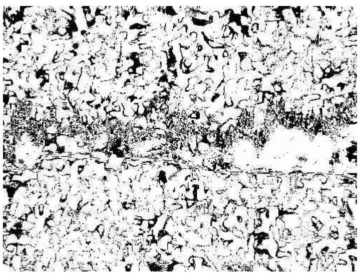
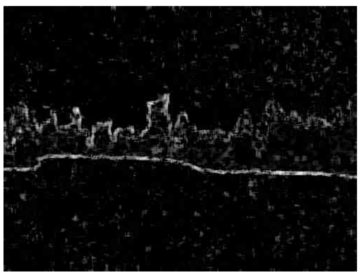
σ = 35, PSNR: 17.26dB | **MLP: 30.65dB**, BM3D: 29.79dB | between 0.00 and 2.72

σ = 50, PSNR: 14.14dB | **MLP: 28.49dB**, BM3D: 27.88dB | between 0.00 and 2.01

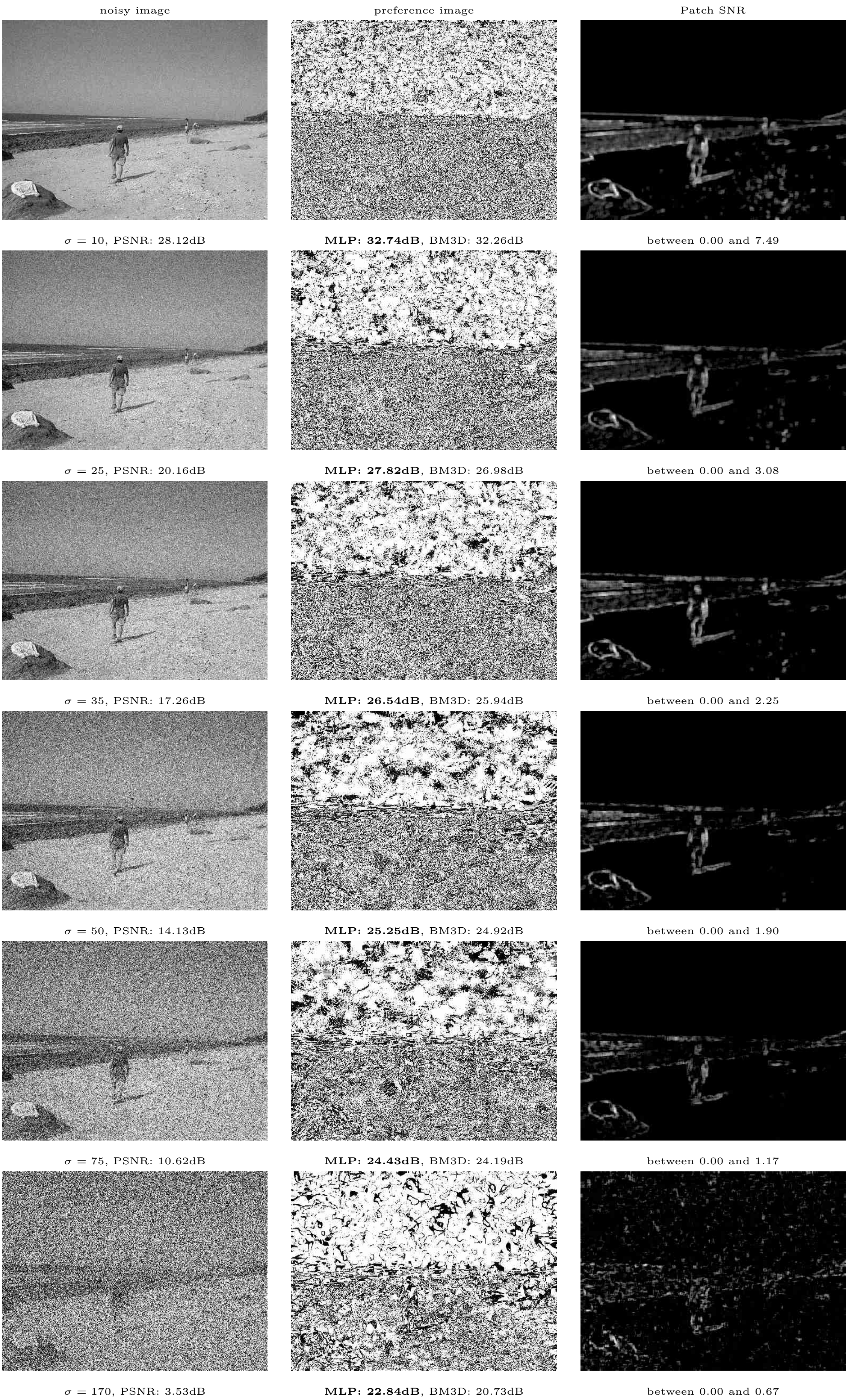
σ = 75, PSNR: 10.61dB | **MLP: 27.01dB**, BM3D: 26.49dB | between 0.00 and 1.40

σ = 170, PSNR: 3.50dB | **MLP: 24.30dB**, BM3D: 21.49dB | between 0.00 and 0.75

MLP beats BM3D, image 11

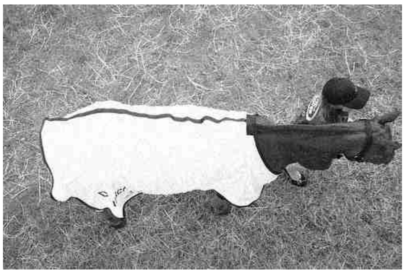
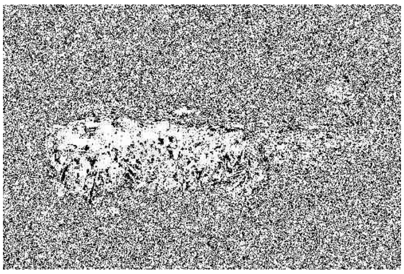
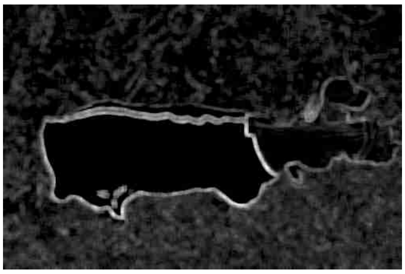
Notes: This image containts booth smooth areas and irregular textures. The MLP performs particularly well on the smooth areas. However, the PatchSNR does not seem to be correlated with preference image.



MLP beats BM3D, image 12

Notes: This image containts booth smooth areas and irregular textures. The MLP performs particularly well on the smooth areas. However, the PatchSNR indicates that BM3D should be used almost everywhere.

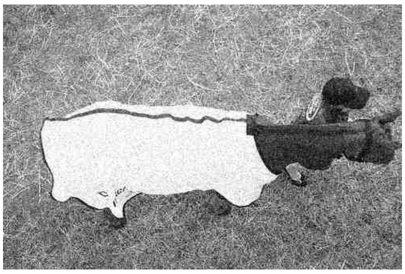
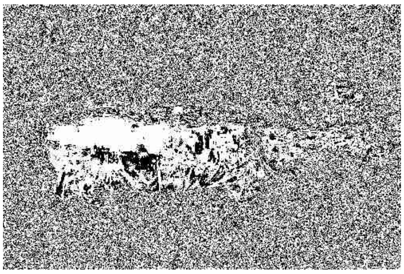
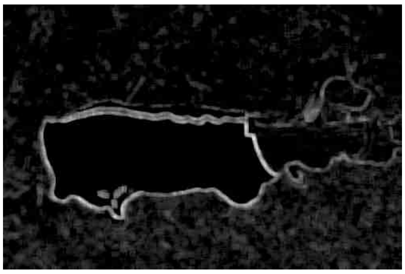
noisy image | preference image | Patch SNR

$\sigma = 10$, PSNR: 28.13dB

MLP: 29.88dB, BM3D: 29.72dB

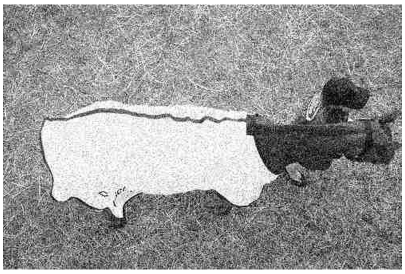
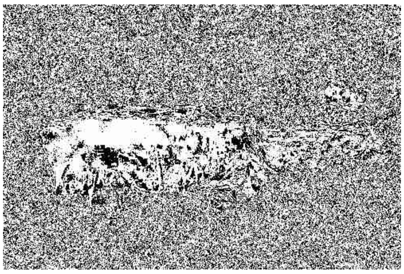
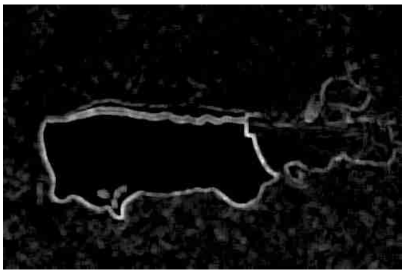
between 0.00 and 10.99

$\sigma = 25$, PSNR: 20.19dB

MLP: 24.14dB, BM3D: 23.32dB

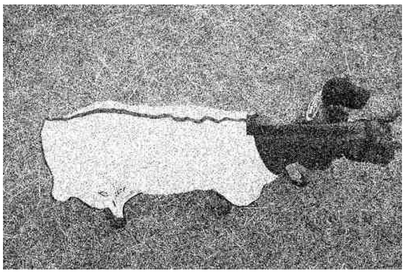
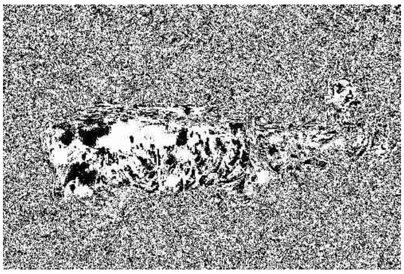
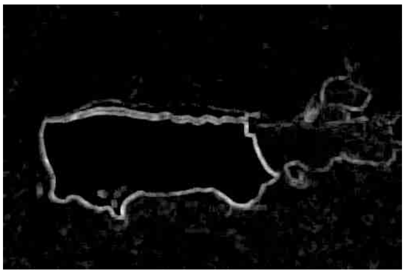
between 0.00 and 4.75

$\sigma = 35$, PSNR: 17.24dB

MLP: 22.51dB, BM3D: 21.66dB

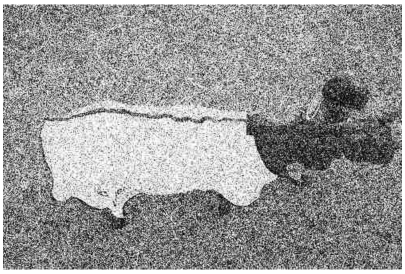
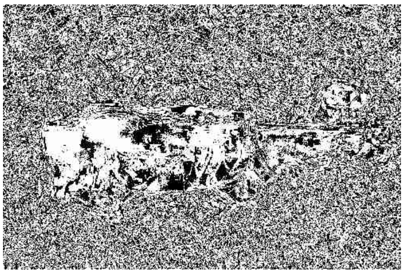
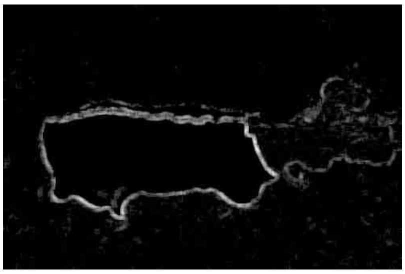
between 0.00 and 3.29

$\sigma = 50$, PSNR: 14.18dB

MLP: 21.00dB, BM3D: 20.47dB

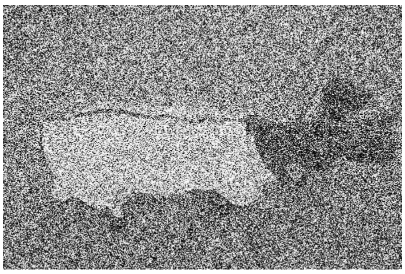
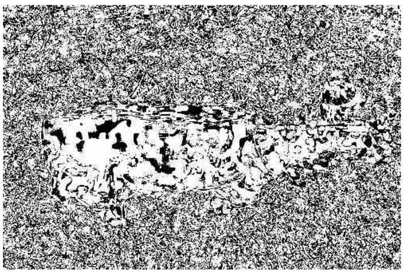
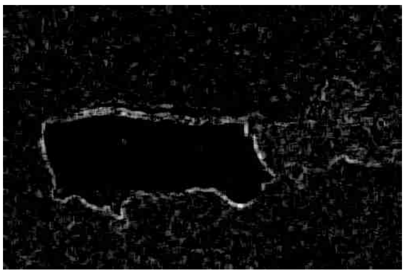
between 0.00 and 2.26

$\sigma = 75$, PSNR: 10.64dB

MLP: 20.02dB, BM3D: 19.86dB

between 0.00 and 1.69

$\sigma = 170$, PSNR: 3.51dB

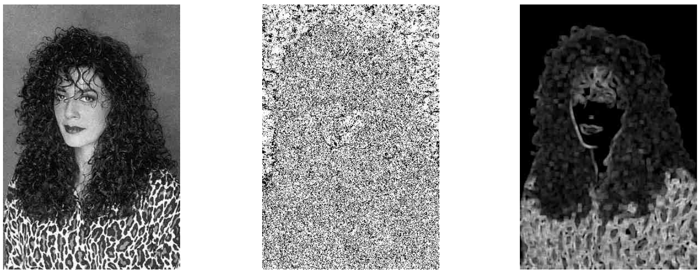
MLP: 18.96dB, BM3D: 18.13dB

between 0.00 and 0.75

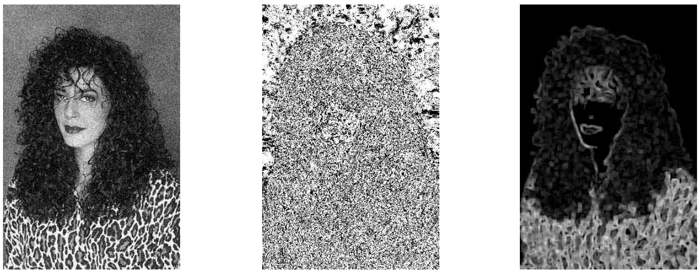
MLP beats BM3D, image 13

Notes: This image contains booth smooth areas and irregular textures. It is difficult to say where exactly one algorithm should be preferred over the other. The PatchSNR does not seem to be correlated with actual preferences.

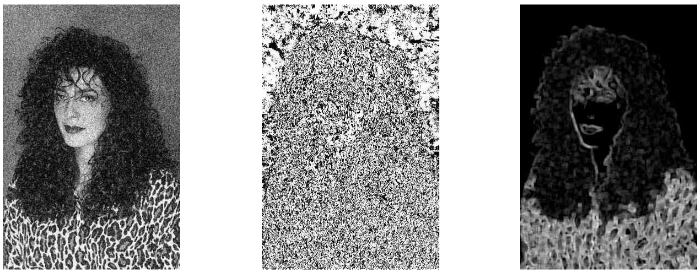
noisy image | preference image | Patch SNR



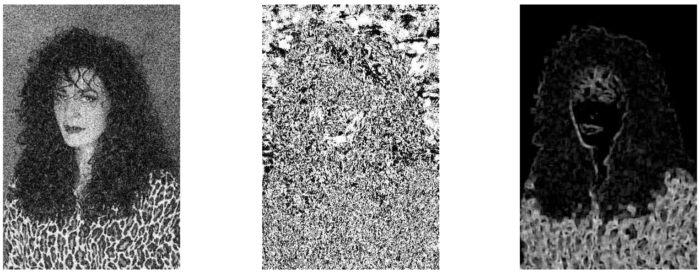
$\sigma = 10$, PSNR: 28.15dB | **MLP: 32.22dB**, BM3D: 31.69dB | between 0.00 and 11.46



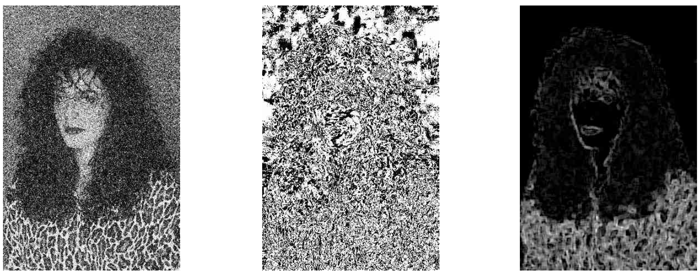
$\sigma = 25$, PSNR: 20.19dB | **MLP: 27.12dB**, BM3D: 26.27dB | between 0.00 and 4.25



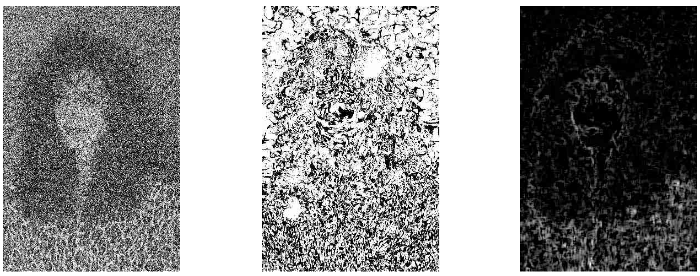
$\sigma = 35$, PSNR: 17.24dB | **MLP: 25.35dB**, BM3D: 24.49dB | between 0.00 and 3.17



$\sigma = 50$, PSNR: 14.17dB | **MLP: 23.47dB**, BM3D: 22.88dB | between 0.00 and 2.20



$\sigma = 75$, PSNR: 10.61dB | **MLP: 21.67dB**, BM3D: 21.19dB | between 0.00 and 1.62

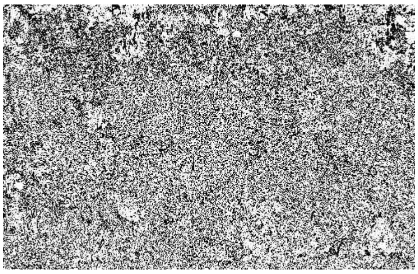
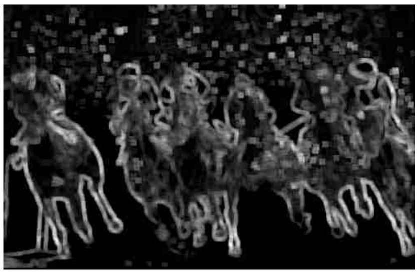


$\sigma = 170$, PSNR: 3.50dB | **MLP: 17.73dB**, BM3D: 16.69dB | between 0.00 and 0.87

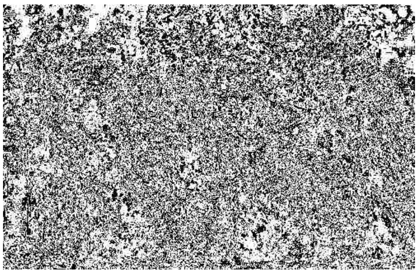
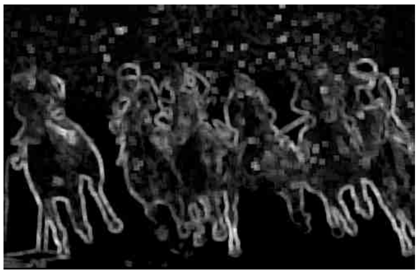
MLP beats BM3D, image 14

Notes: This image contains mostly irregular textures. The preference images look chaotic. The PatchSNR does not seem to correlate with actual preferences.

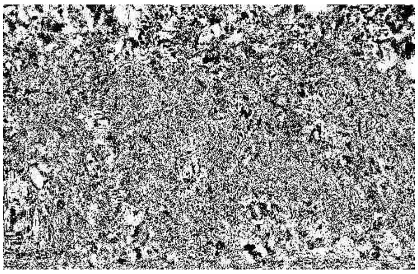
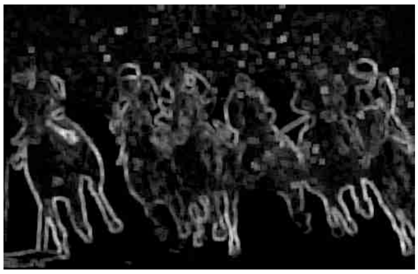
noisy image | preference image | Patch SNR

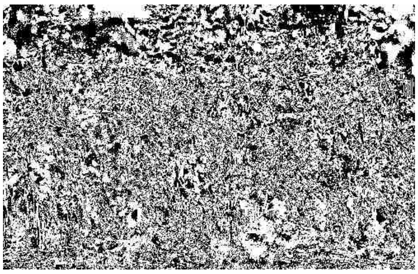
$\sigma = 10$, PSNR: 28.13dB | **MLP: 31.63dB**, BM3D: 31.27dB | between 0.00 and 11.15

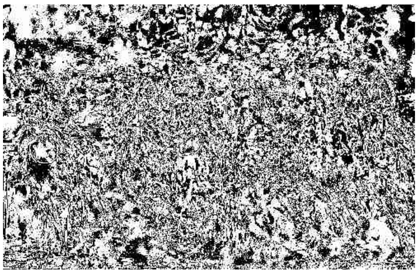
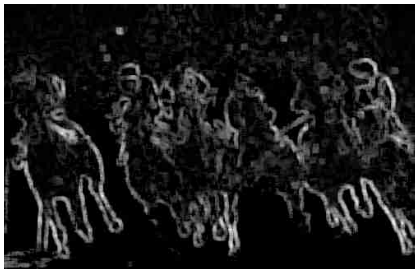
$\sigma = 25$, PSNR: 20.16dB | **MLP: 26.27dB**, BM3D: 25.49dB | between 0.00 and 5.36

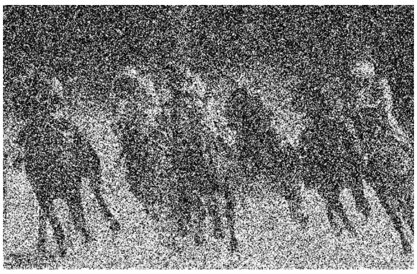
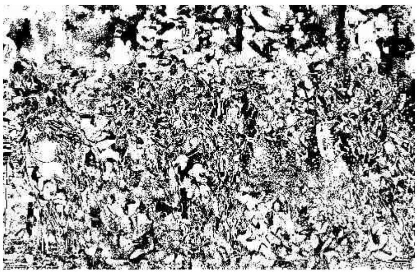
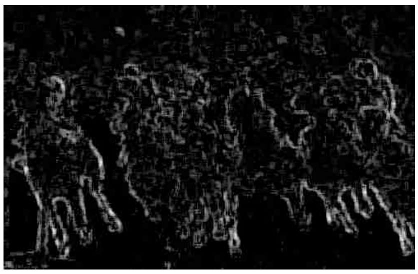
$\sigma = 35$, PSNR: 17.25dB | **MLP: 24.43dB**, BM3D: 23.74dB | between 0.00 and 3.79

$\sigma = 50$, PSNR: 14.16dB | **MLP: 22.49dB**, BM3D: 21.79dB | between 0.00 and 2.48

$\sigma = 75$, PSNR: 10.61dB | **MLP: 20.94dB**, BM3D: 20.44dB | between 0.00 and 1.83

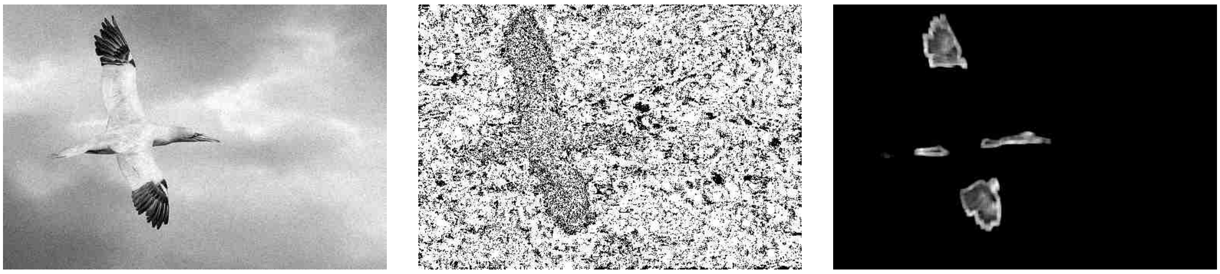
  

$\sigma = 170$, PSNR: 3.51dB | **MLP: 18.35dB**, BM3D: 17.52dB | between 0.00 and 1.06

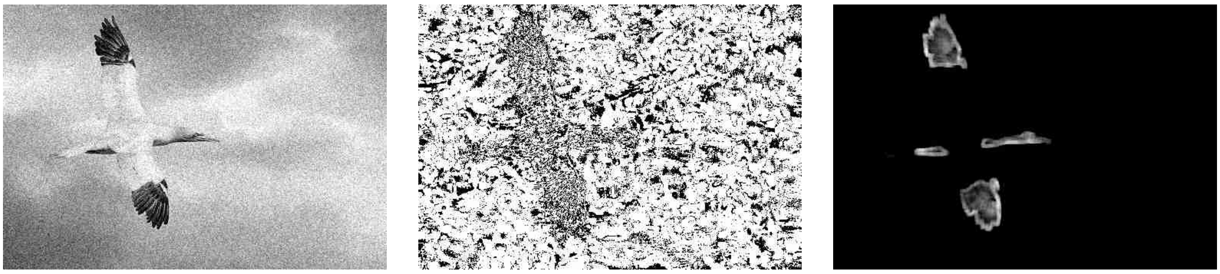
MLP beats BM3D, image 15

Notes: The MLP performs particularly well on the smooth areas of this image. However, the PatchSNR indicates that BM3D should be used in these areas. Furthermore, the MLP does not seem to perform so well on the wings of the bird. However, this is precisely the region where the PatchSNR indicates the MLP should be used.

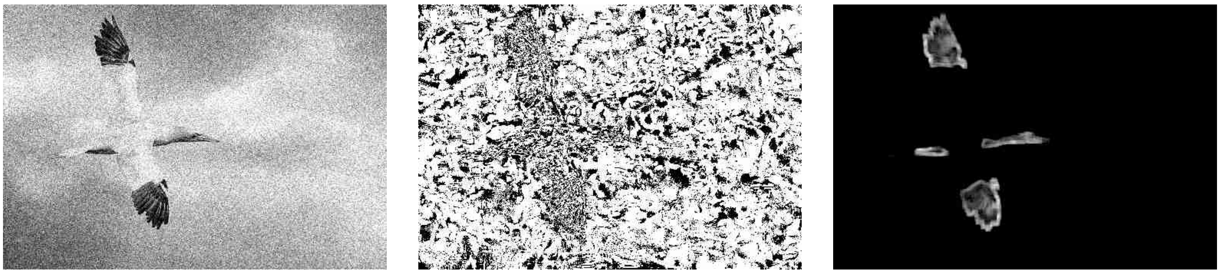
noisy image | preference image | Patch SNR



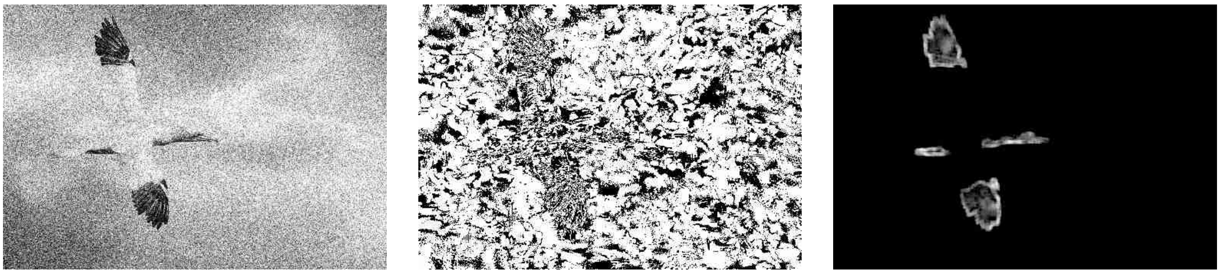
$\sigma = 10$, PSNR: 28.12dB | **MLP: 40.26dB**, BM3D: 40.13dB | between 0.00 and 8.92



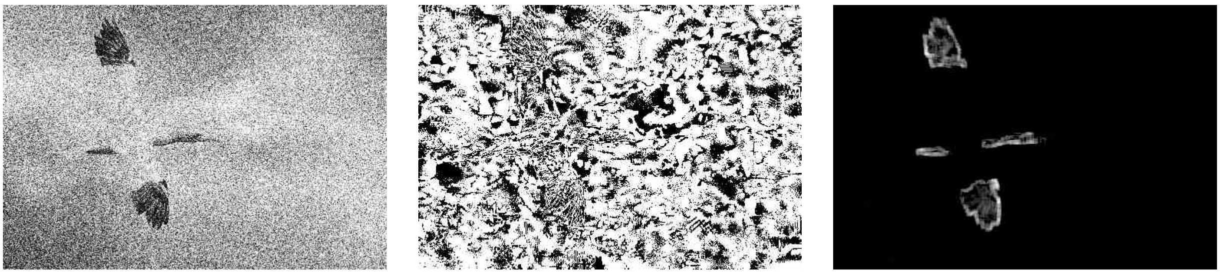
$\sigma = 25$, PSNR: 20.17dB | **MLP: 35.71dB**, BM3D: 34.75dB | between 0.00 and 4.18



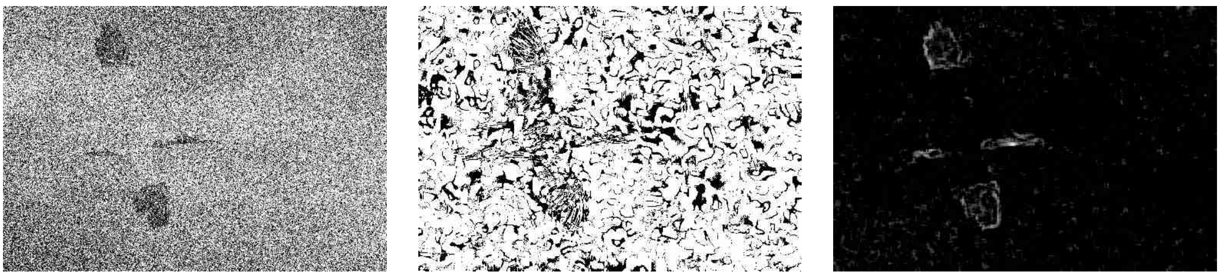
$\sigma = 35$, PSNR: 17.25dB | **MLP: 33.88dB**, BM3D: 32.92dB | between 0.00 and 2.94



$\sigma = 50$, PSNR: 14.14dB | **MLP: 31.91dB**, BM3D: 31.05dB | between 0.00 and 2.04



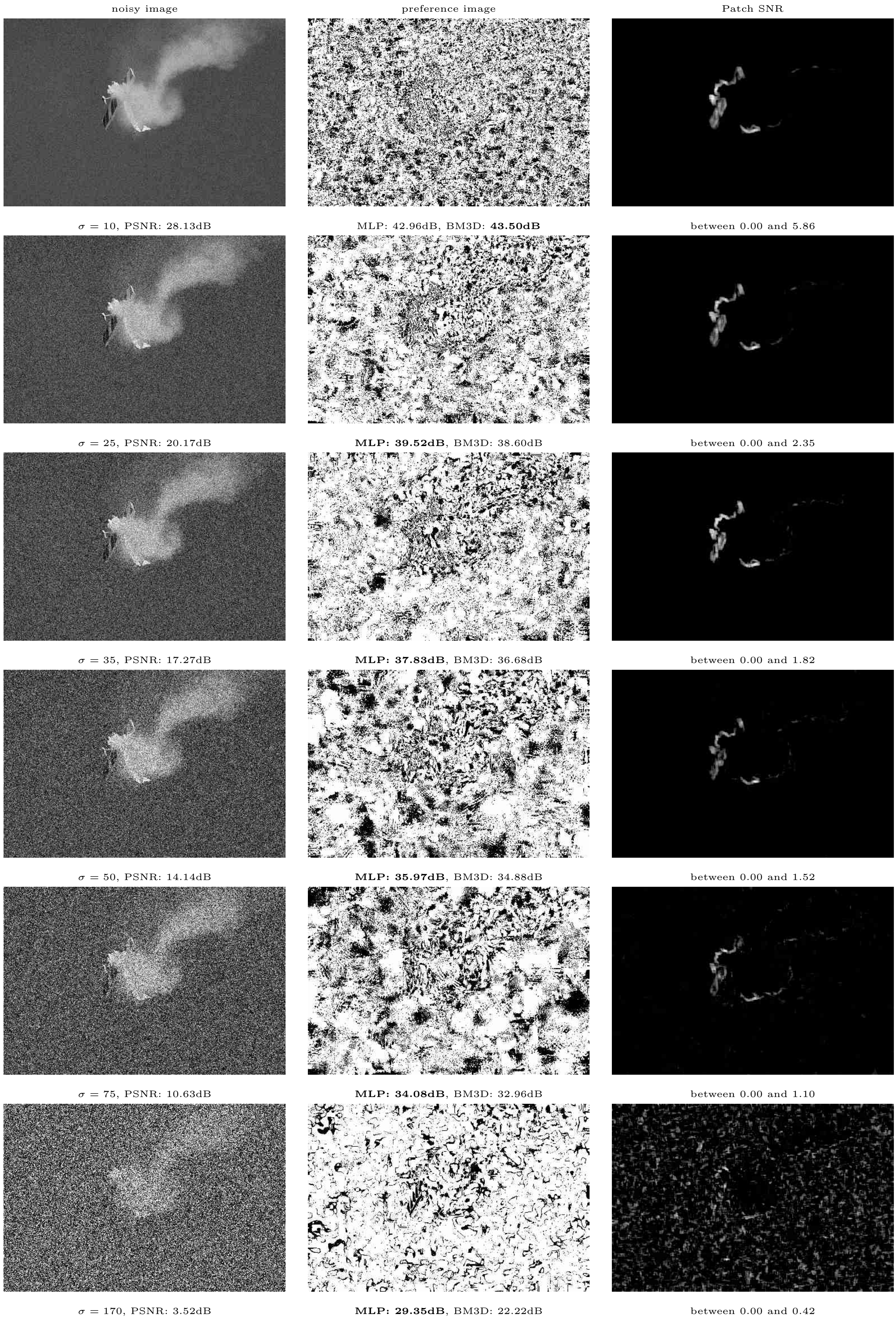
$\sigma = 75$, PSNR: 10.60dB | **MLP: 29.99dB**, BM3D: 29.12dB | between 0.00 and 1.27



$\sigma = 170$, PSNR: 3.51dB | **MLP: 26.59dB**, BM3D: 22.69dB | between 0.00 and 0.75

MLP beats BM3D, image 16

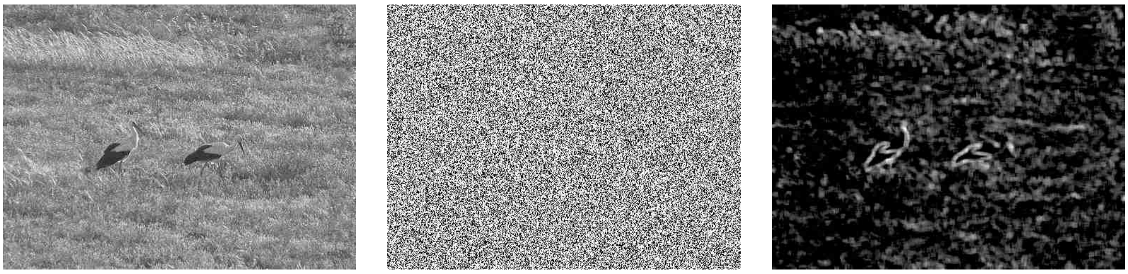
Notes: The image is very smooth and the MLP performs well compared to BM3D almost everywhere. However, the PatchSNR suggests the opposite: BM3D should be used almost everywhere. We also see that the MLP performs particularly well at high noise levels, whereas Mosseri et al [1] claim that external methods perform less well at high noise levels.



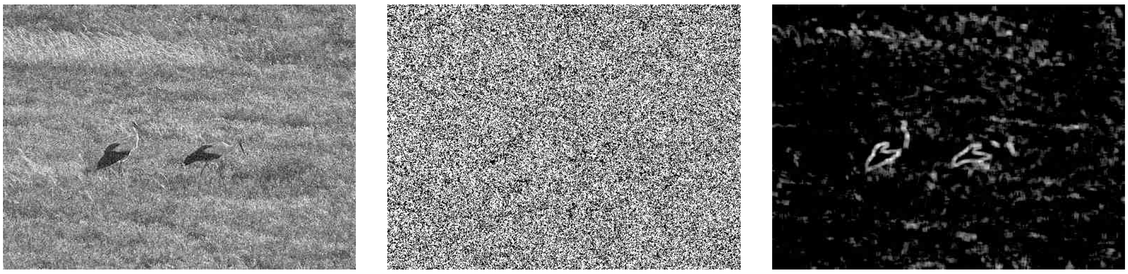
MLP beats BM3D, image 17

Notes: This image contains irregular textures. The preference images look almost random: It is difficult to tell from the image content where one method should be preferred over the other. However, the PatchSNR does not seem correlated to actual preferences.

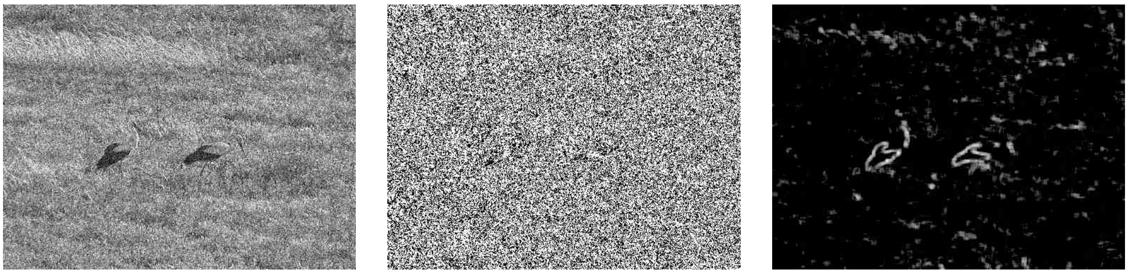
noisy image	preference image	Patch SNR



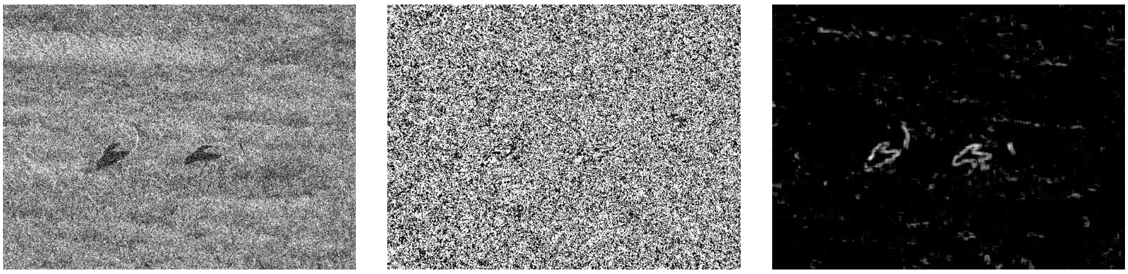
σ = 10, PSNR: 28.14dB	**MLP: 29.25dB**, BM3D: 29.10dB	between 0.00 and 5.04



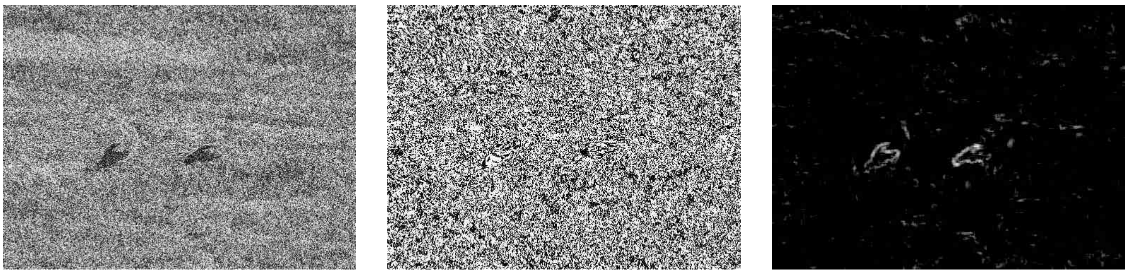
σ = 25, PSNR: 20.15dB	**MLP: 23.95dB**, BM3D: 23.16dB	between 0.00 and 2.03



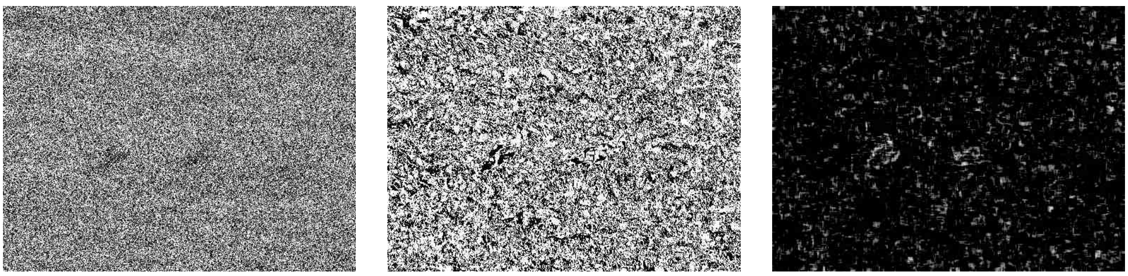
σ = 35, PSNR: 17.25dB	**MLP: 22.54dB**, BM3D: 21.90dB	between 0.00 and 1.51



σ = 50, PSNR: 14.15dB	**MLP: 21.33dB**, BM3D: 20.99dB	between 0.00 and 1.23



σ = 75, PSNR: 10.65dB	**MLP: 20.55dB**, BM3D: 20.43dB	between 0.00 and 0.81

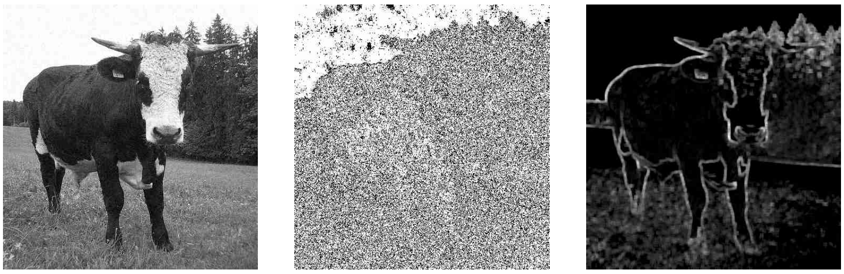


σ = 170, PSNR: 3.51dB	**MLP: 19.76dB**, BM3D: 18.79dB	between 0.00 and 0.45

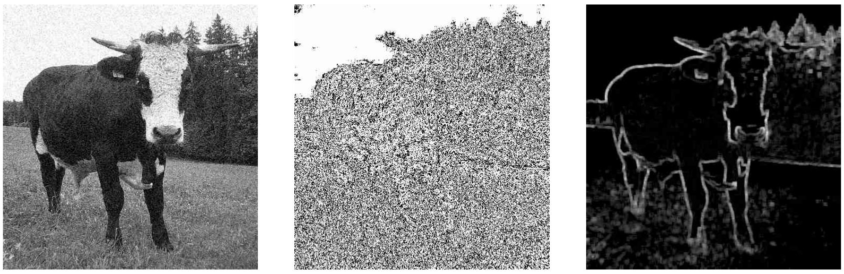
MLP beats BM3D, image 18

Notes: This image contains both irregular textures and smooth areas. On the smooth areas, the MLP outperforms BM3D on most pixels. On the irregular textures, preferences do not seem to be related to image content. The PatchSNR does not seem to be correlated to actual preferences.

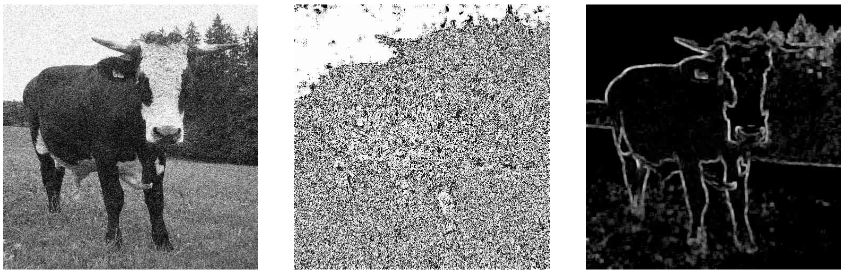
noisy image | preference image | Patch SNR



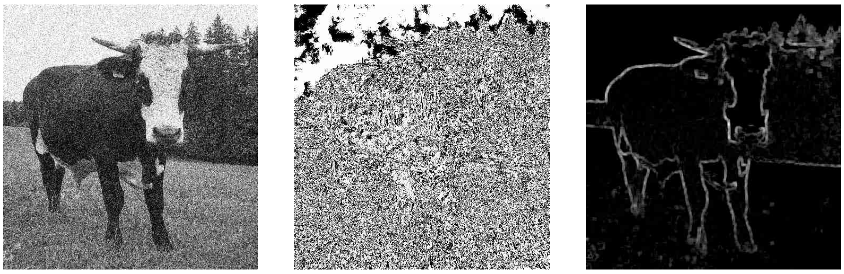
$\sigma = 10$, PSNR: 28.13dB | **MLP: 30.92dB**, BM3D: 30.60dB | between 0.00 and 10.03



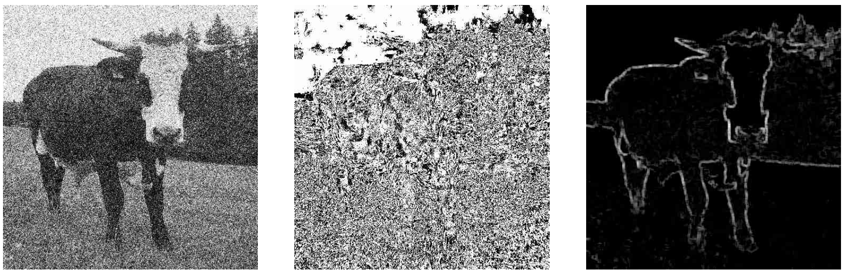
$\sigma = 25$, PSNR: 20.18dB | **MLP: 25.80dB**, BM3D: 24.99dB | between 0.00 and 3.69



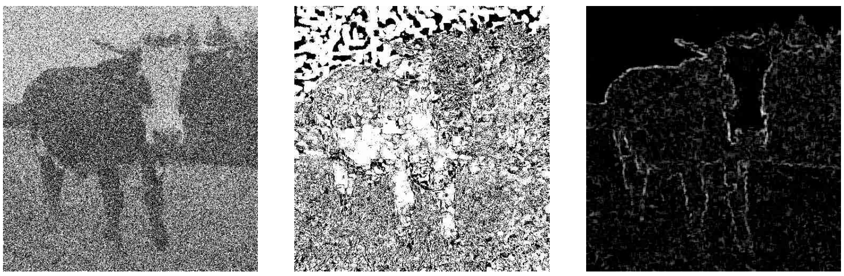
$\sigma = 35$, PSNR: 17.26dB | **MLP: 24.24dB**, BM3D: 23.58dB | between 0.00 and 3.07



$\sigma = 50$, PSNR: 14.15dB | **MLP: 22.71dB**, BM3D: 22.31dB | between 0.00 and 1.90



$\sigma = 75$, PSNR: 10.65dB | **MLP: 21.75dB**, BM3D: 21.50dB | between 0.00 and 1.41

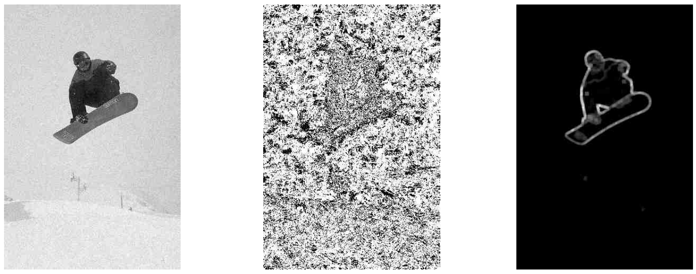


$\sigma = 170$, PSNR: 3.53dB | **MLP: 20.35dB**, BM3D: 18.96dB | between 0.00 and 0.69

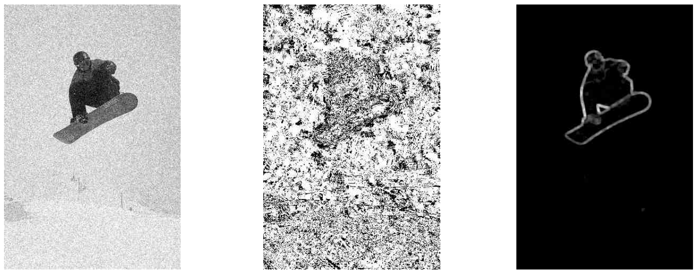
MLP beats BM3D, image 19

Notes: This image contains mostly smooth areas. The MLP performs particularly well on the smooth areas, but the PatchSNR indicates to use BM3D on these areas.

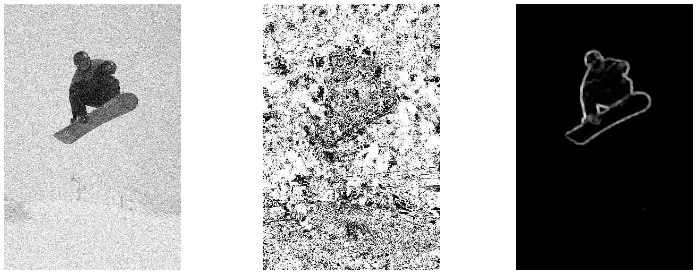
noisy image | preference image | Patch SNR



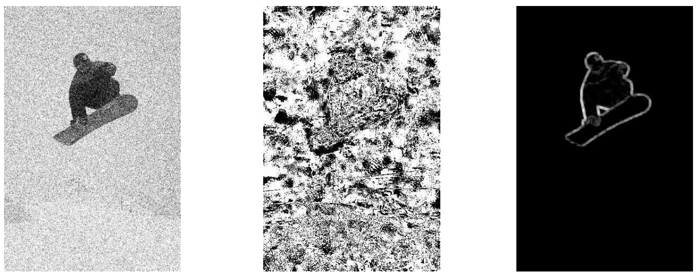
$\sigma = 10$, PSNR: 28.14dB | **MLP: 39.20dB**, BM3D: 39.17dB | between 0.00 and 8.74



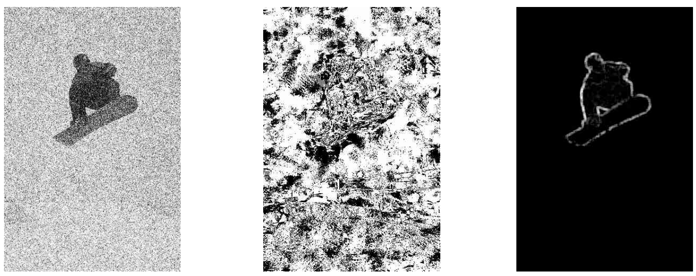
$\sigma = 25$, PSNR: 20.15dB | **MLP: 35.77dB**, BM3D: 35.02dB | between 0.00 and 4.12



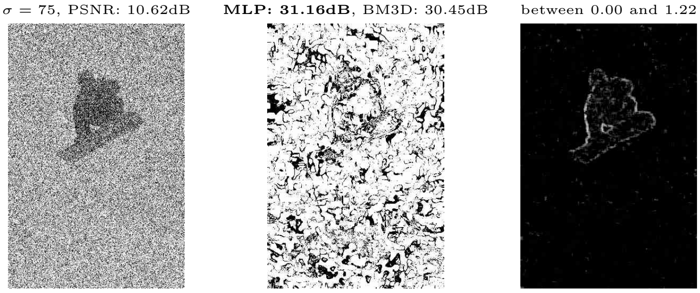
$\sigma = 35$, PSNR: 17.25dB | **MLP: 34.22dB**, BM3D: 33.55dB | between 0.00 and 3.02



$\sigma = 50$, PSNR: 14.15dB | **MLP: 32.68dB**, BM3D: 32.17dB | between 0.00 and 2.12



$\sigma = 75$, PSNR: 10.62dB | **MLP: 31.16dB**, BM3D: 30.45dB | between 0.00 and 1.22

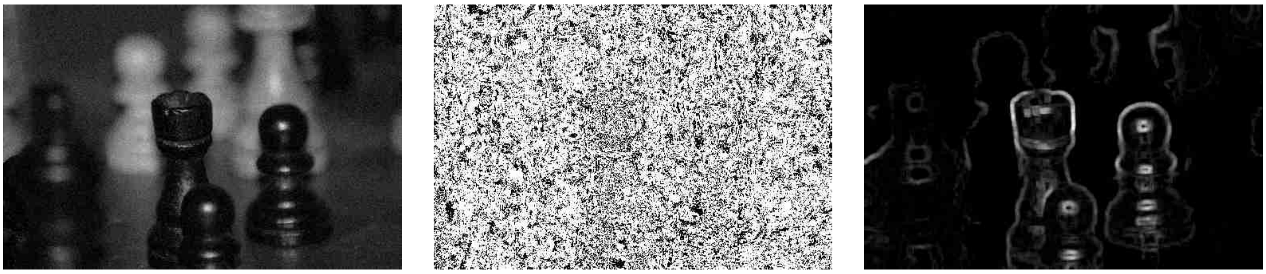


$\sigma = 170$, PSNR: 3.53dB | **MLP: 27.96dB**, BM3D: 22.93dB | between 0.00 and 0.73

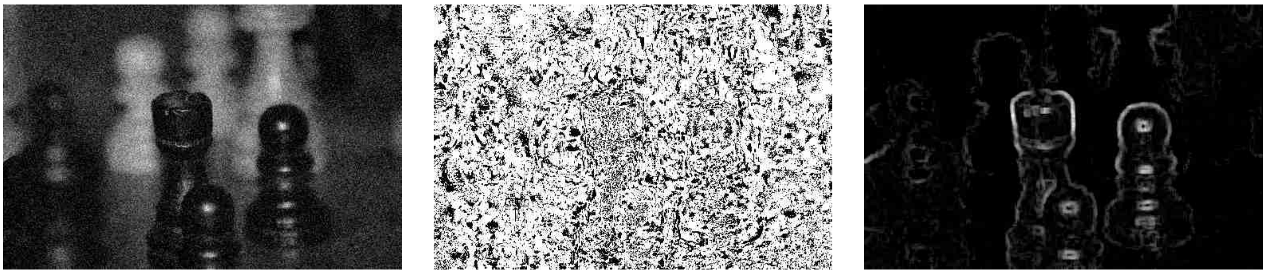
MLP beats BM3D, image 20

Notes: This image contains mostly smooth areas. The MLP performs particularly well on the smooth areas, but the PatchSNR indicates to use BM3D on these areas.

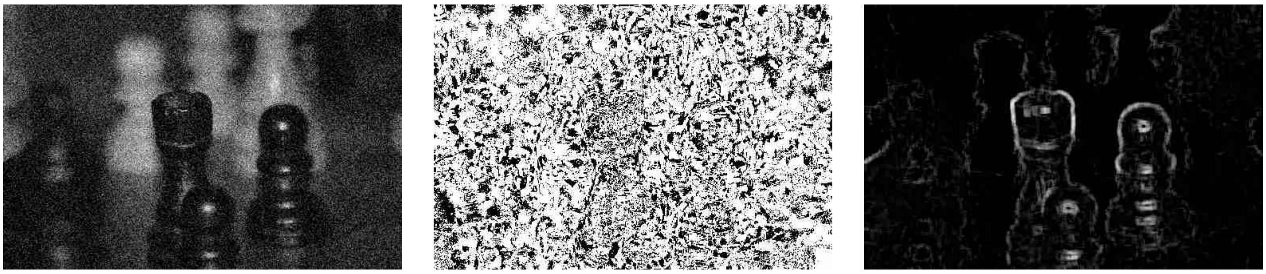
noisy image | preference image | Patch SNR



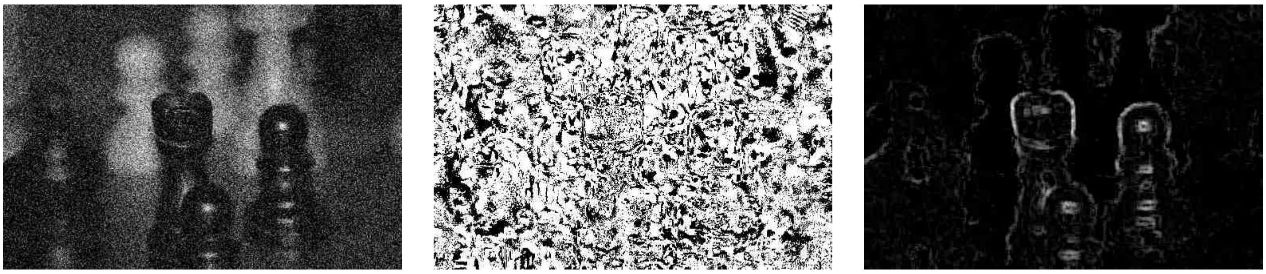
$\sigma = 10$, PSNR: 28.11dB | **MLP: 39.73dB**, BM3D: 39.55dB | between 0.00 and 6.90



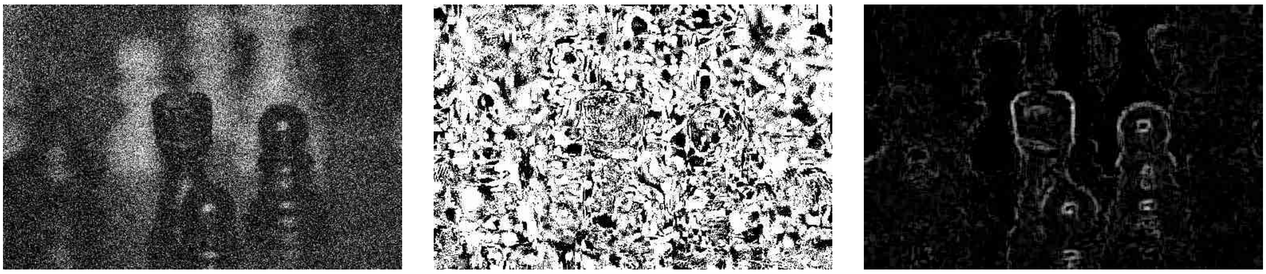
$\sigma = 25$, PSNR: 20.15dB | **MLP: 36.16dB**, BM3D: 35.38dB | between 0.00 and 2.93



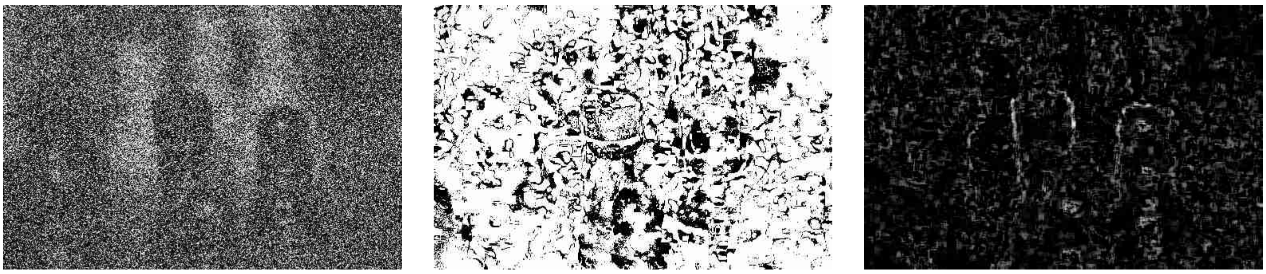
$\sigma = 35$, PSNR: 17.26dB | **MLP: 34.49dB**, BM3D: 33.61dB | between 0.00 and 2.27



$\sigma = 50$, PSNR: 14.14dB | **MLP: 32.89dB**, BM3D: 32.08dB | between 0.00 and 1.69



$\sigma = 75$, PSNR: 10.65dB | **MLP: 31.00dB**, BM3D: 30.05dB | between 0.00 and 1.07



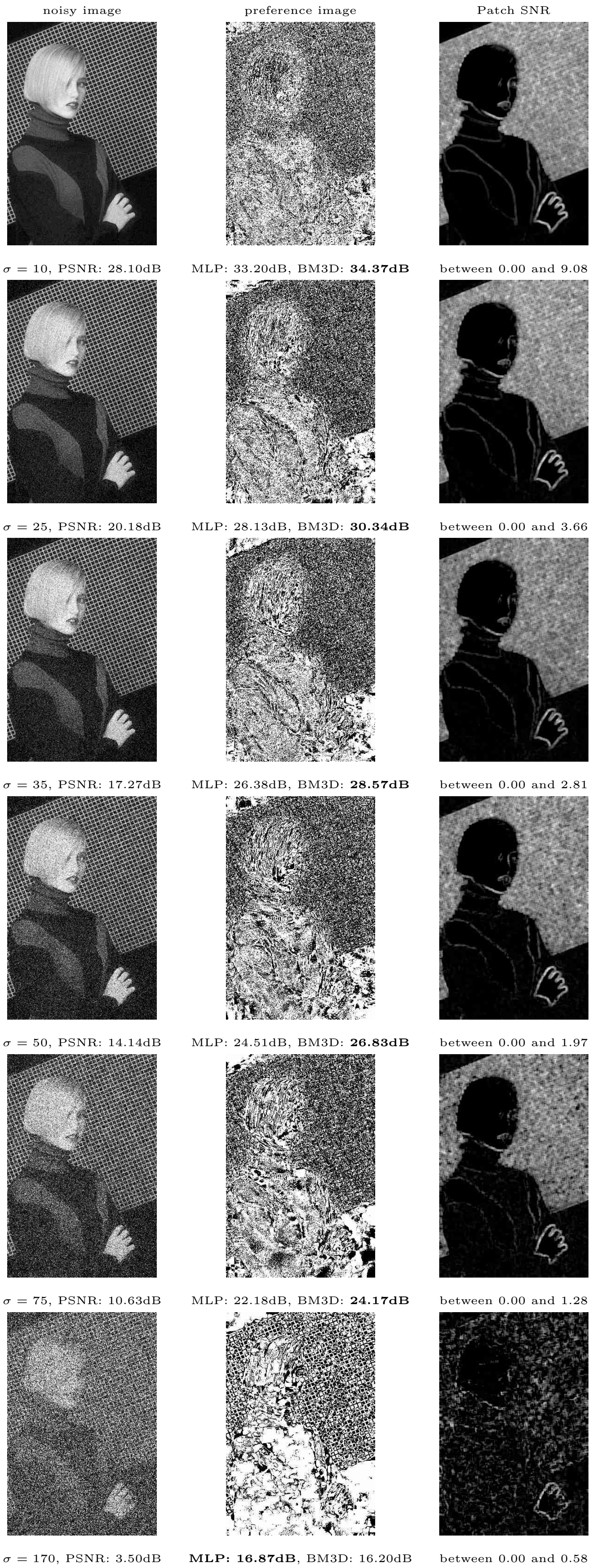
$\sigma = 170$, PSNR: 3.54dB | **MLP: 27.14dB**, BM3D: 21.68dB | between 0.00 and 0.52

1.2 BM3D beats MLP

In image 1 we see that BM3D performs well on very regular patterns, such as a grid, whereas MLP performs better in smooth regions. In the same image, it is particularly evident that the PatchSNR does not accurately predict when to prefer one algorithm over the other. Indeed, the PatchSNR predicts exactly the opposite of the actual preferences: An external method should be used on the regular grid, whereas an internal method should be used on the remaining, smooth regions of the image.
In image 4 we see that MLP outperforms BM3D on irregular patterns, whereas BM3D performs well on regular patterns. We also see that the PatchSNR is unable to predict this effect: It predicts that an external method should be used on the edges and that an internal method should be used for the rest of the image. However, these predictions do not seem to correlate with actual preferences. We observe the same effects as for image 1 and 4 on the remaining 18 images.
Finally, we observe the MLP always outperforms BM3D for $\sigma = 170$. MLP also sometimes outperforms BM3D for $\sigma = 75$, and on image 18 also for $\sigma = 50$. Hence we conclude that MLP tends to become better (compared to BM3D) at higher noise levels, which is the opposite of the conclusion made by [1].

BM3D beats MLP, image 1

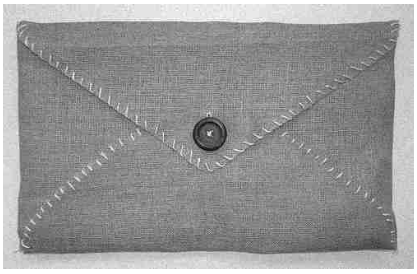
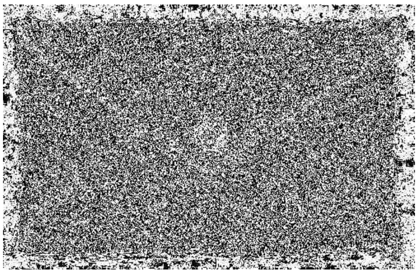
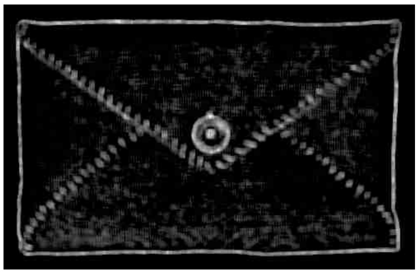
Notes: The MLP is at a disadvantage compared to BM3D on the regular grid in the background, but performs better on most of the rest of the image. However, the PatchSNR indicates the opposite: BM3D should be used in areas where the MLP performs well, and an external method should be used on the regular grid (where BM3D performs well).



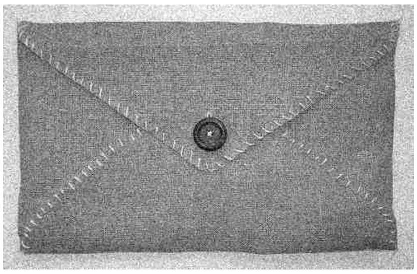
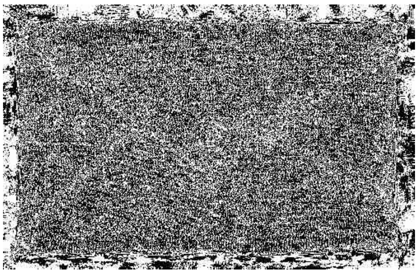
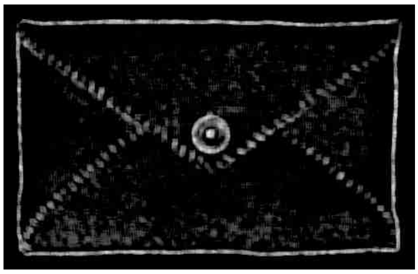
BM3D beats MLP, image 2

Notes: BM3D performs particularly well on regular, repeating texture of the wallet. This is indeed correctly predicted by the PatchSNR. However, the PatchSNR also indicates that BM3D should be used on the smooth rim of the image, where MLP actually performs better.

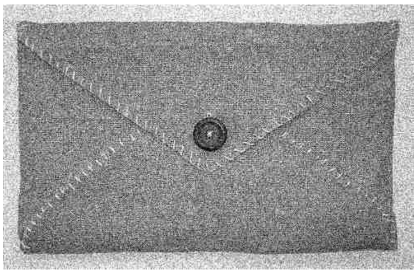
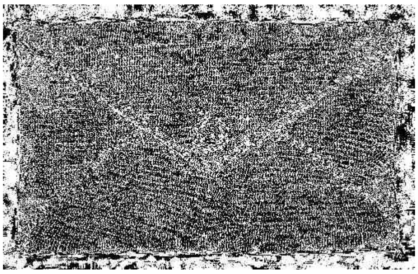
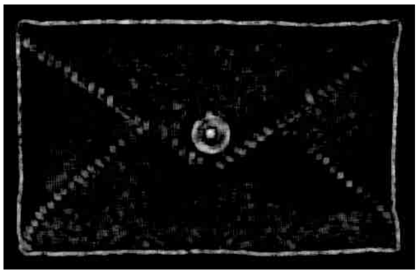
noisy image | preference image | Patch SNR

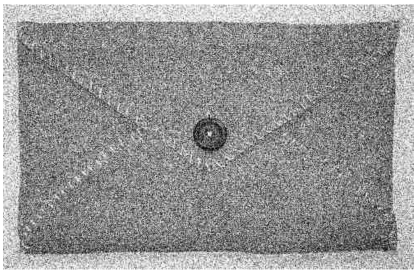
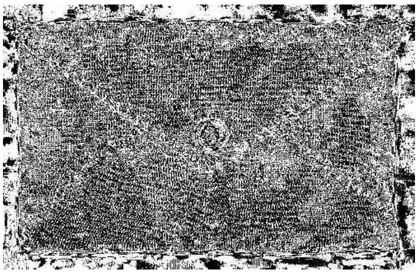
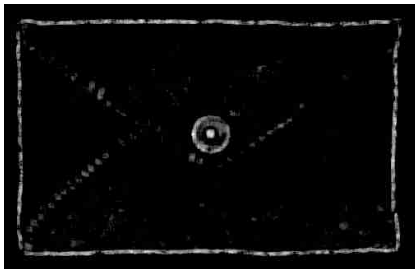
$\sigma = 10$, PSNR: 28.11dB | MLP: 30.55dB, BM3D: **31.46dB** | between 0.00 and 5.84

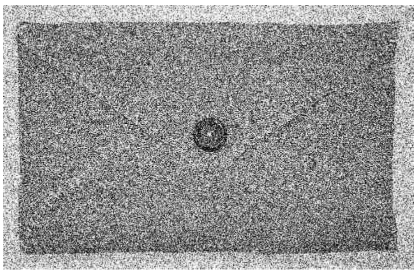
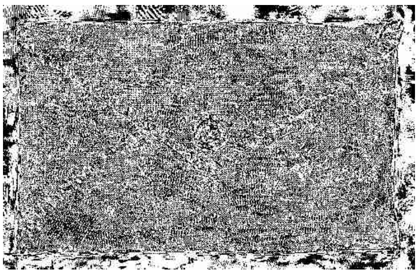
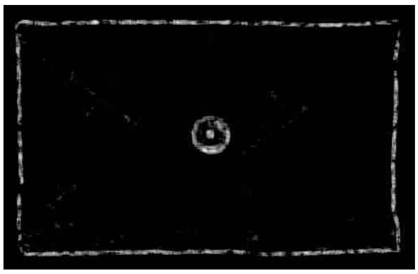
$\sigma = 25$, PSNR: 20.17dB | MLP: 25.50dB, BM3D: **27.03dB** | between 0.00 and 2.17

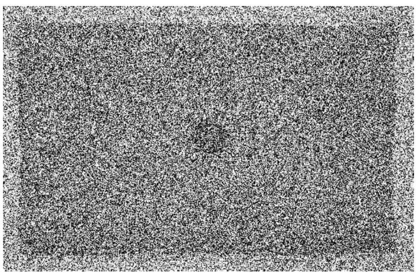
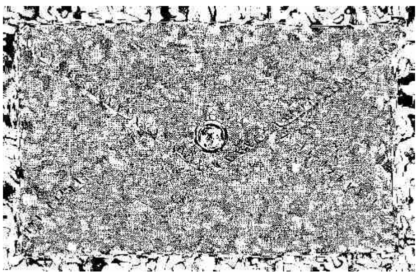
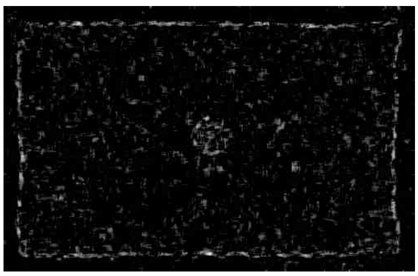
$\sigma = 35$, PSNR: 17.25dB | MLP: 24.10dB, BM3D: **25.40dB** | between 0.00 and 1.66

$\sigma = 50$, PSNR: 14.16dB | MLP: 23.35dB, BM3D: **23.85dB** | between 0.00 and 1.24

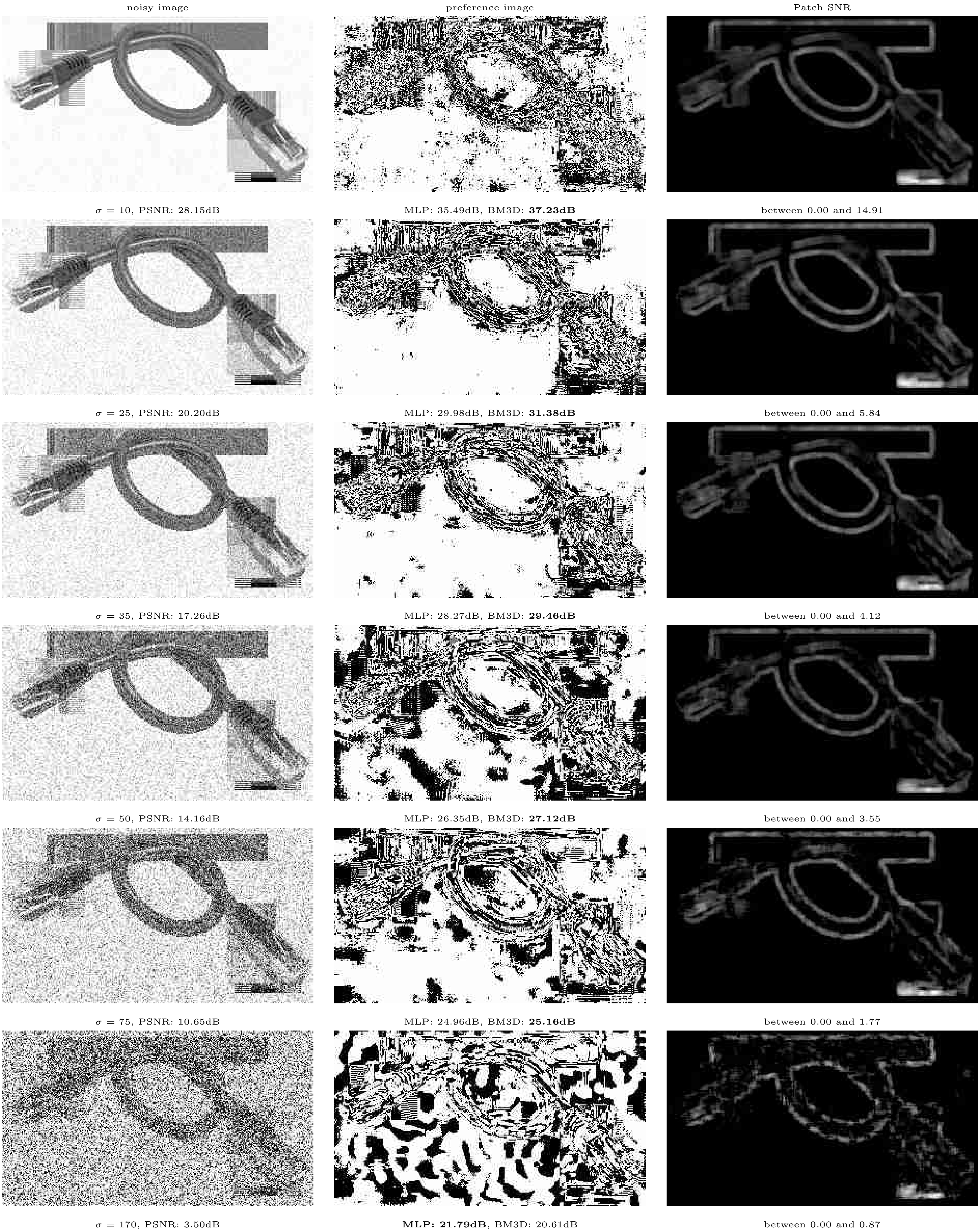
$\sigma = 75$, PSNR: 10.63dB | MLP: 22.90dB, BM3D: **22.95dB** | between 0.00 and 0.91

$\sigma = 170$, PSNR: 3.52dB | **MLP: 22.08dB**, BM3D: 20.18dB | between 0.00 and 0.50

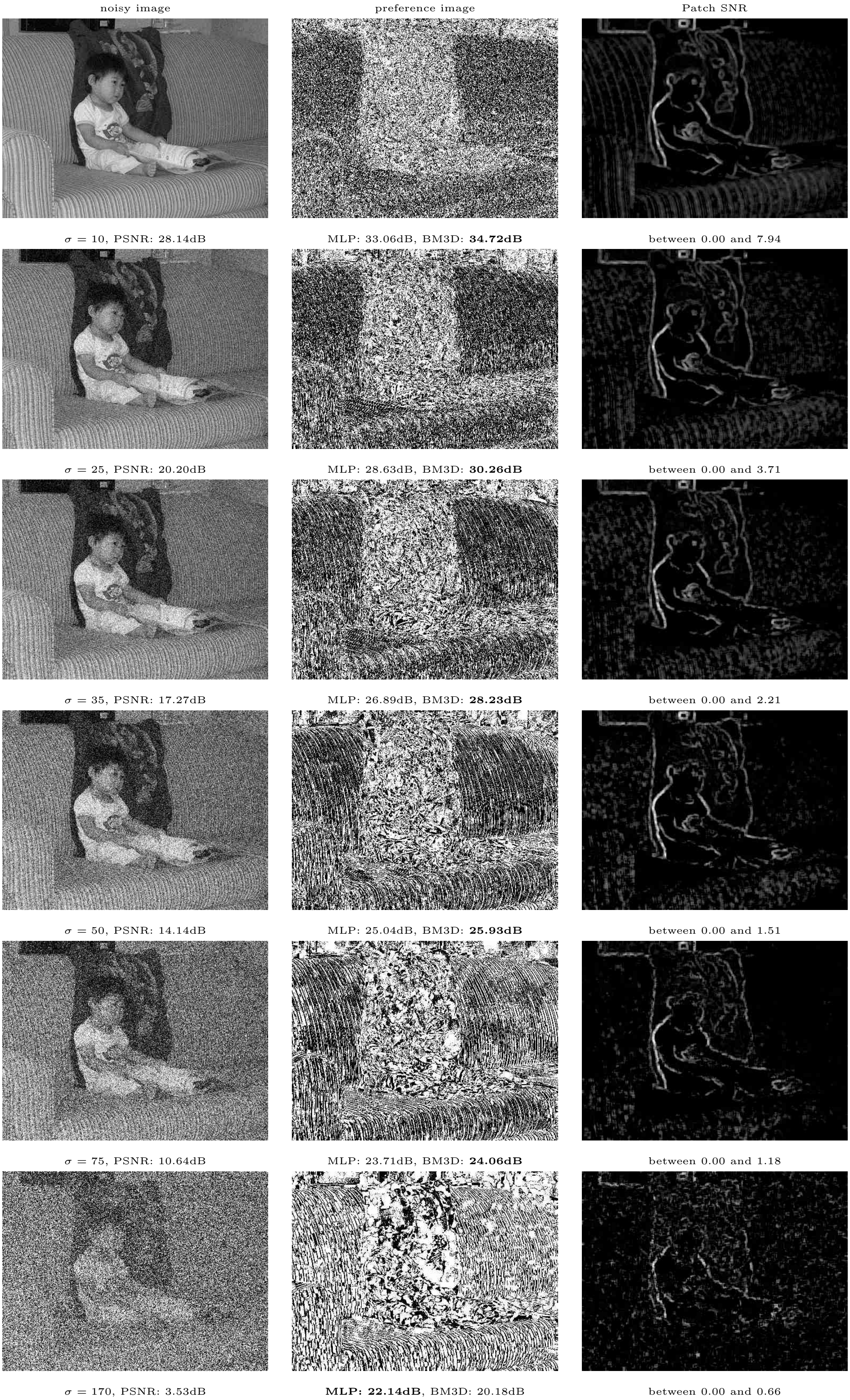
BM3D beats MLP, image 3

Notes: Parts of this image contain geometric patterns, which are better denoised with BM3D than with MLP. The smooth background is better denoised with MLP. However, the PatchSNR indicates that BM3D should be used almost everywhere.



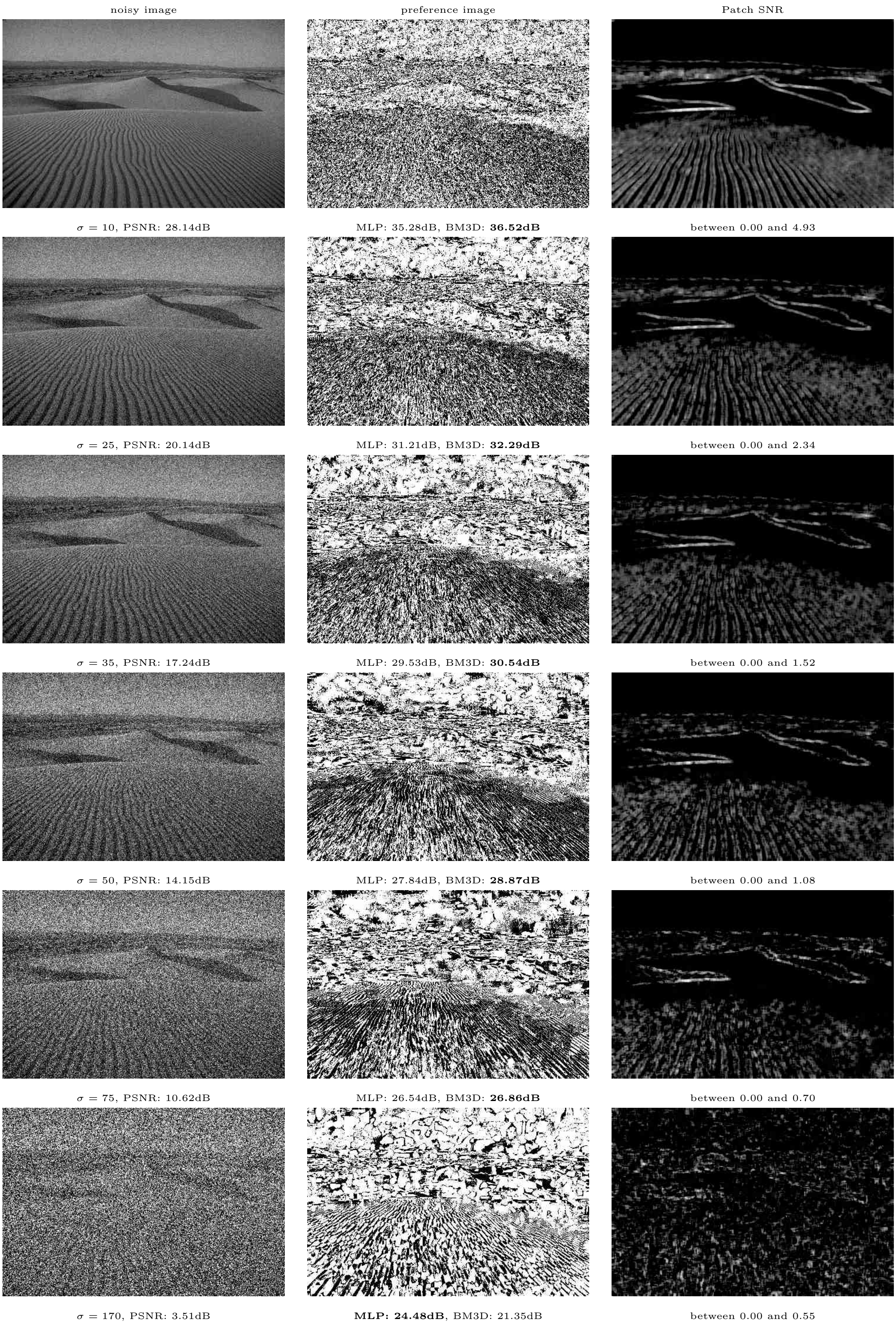
BM3D beats MLP, image 4

Notes: The couch contains regular, repeating patterns and is better denoised with BM3D. The center part of the image with the child is better denoised with MLP because it contains irregular patterns. However, the PatchSNR does not correctly predict this effect.



BM3D beats MLP, image 5

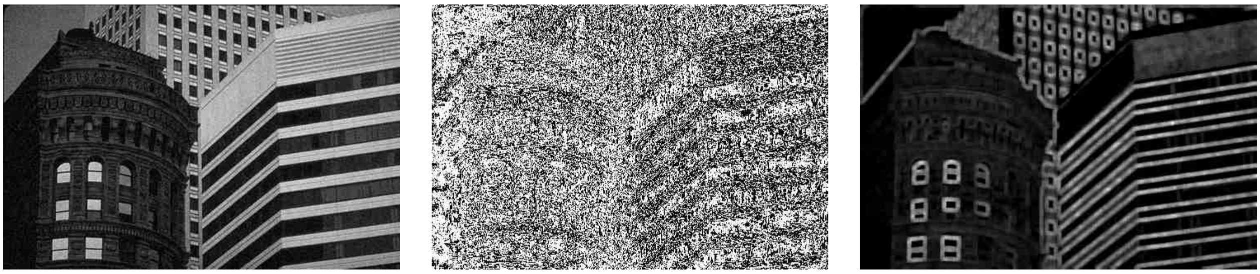
Notes: The repeating patterns of the dunes are better denoised with BM3D, but the sky and the smoother dunes in the background are better denoised with MLP. The PatchSNR does not correctly predict this effect.



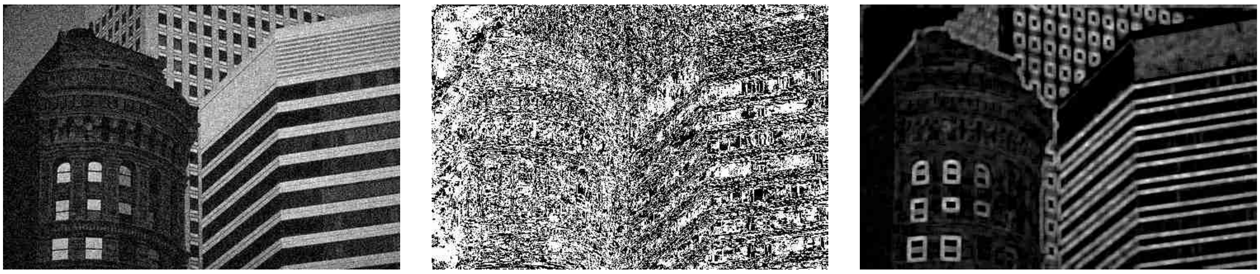
BM3D beats MLP, image 6

Notes: BM3D performs better on the straight lines contained in this image. This effect is not correctly predicted by the PatchSNR.

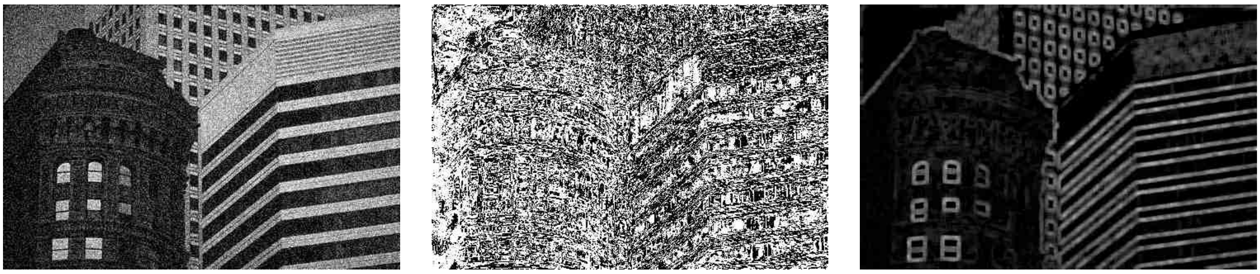
noisy image | preference image | Patch SNR



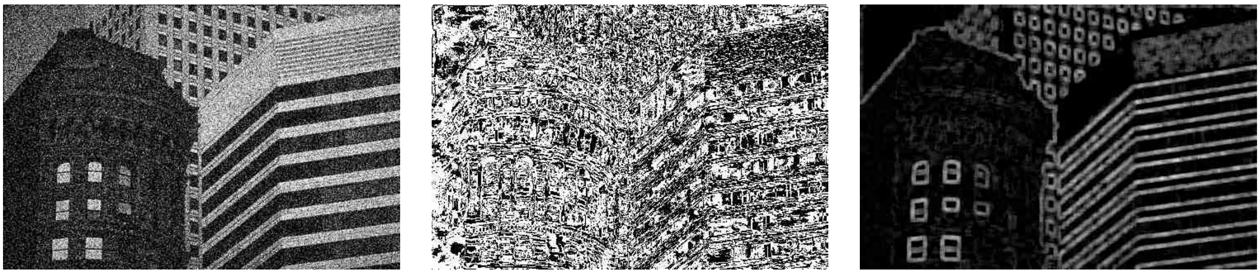
$\sigma = 10$, PSNR: 28.15dB | MLP: 33.64dB, BM3D: **34.63dB** | between 0.00 and 9.49



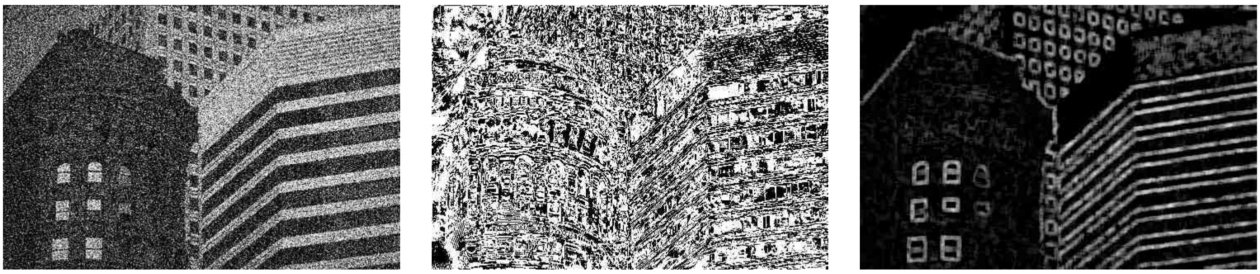
$\sigma = 25$, PSNR: 20.19dB | MLP: 28.89dB, BM3D: **29.89dB** | between 0.00 and 4.19



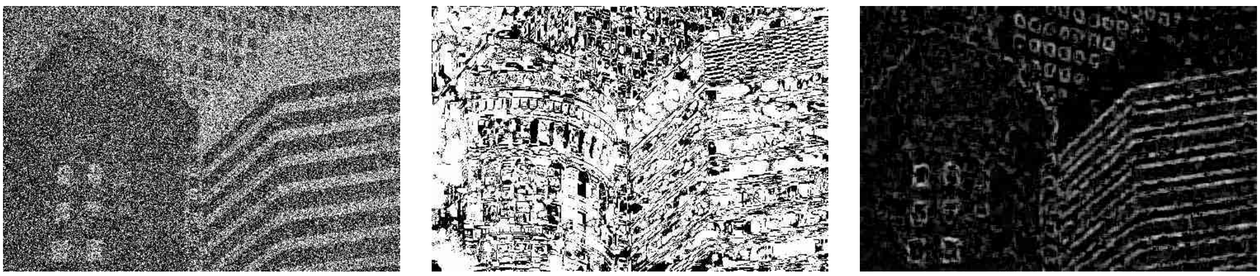
$\sigma = 35$, PSNR: 17.25dB | MLP: 27.25dB, BM3D: **28.12dB** | between 0.00 and 3.36



$\sigma = 50$, PSNR: 14.17dB | MLP: 25.50dB, BM3D: **25.97dB** | between 0.00 and 2.31



$\sigma = 75$, PSNR: 10.63dB | MLP: 23.90dB, BM3D: **24.06dB** | between 0.00 and 1.61

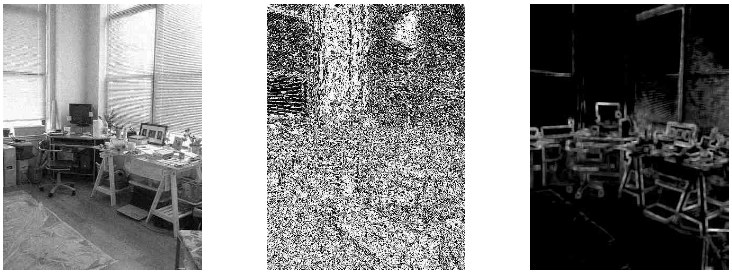


$\sigma = 170$, PSNR: 3.51dB | **MLP: 20.22dB**, BM3D: 18.43dB | between 0.00 and 0.92

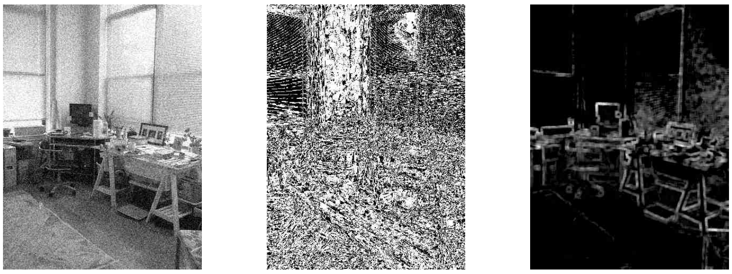
BM3D beats MLP, image 7

Notes: The window blinds are very regular and therefore better denoised by BM3D. The smoother areas of this image are better denoised with MLP. The PatchSNR does not predict this effect: It suggests to use BM3D almost everywhere.

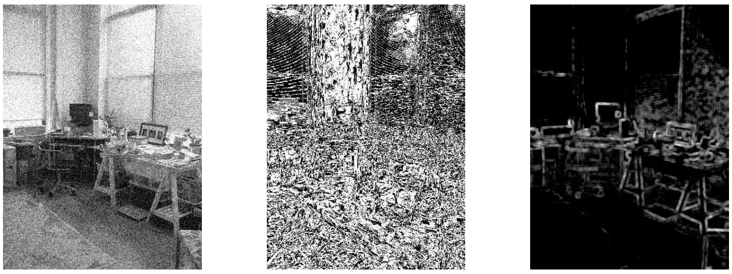
noisy image | preference image | Patch SNR



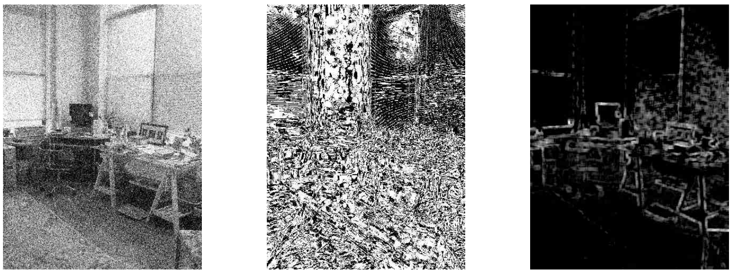
$\sigma = 10$, PSNR: 28.13dB | MLP: 33.61dB, BM3D: **34.58dB** | between 0.00 and 9.56



$\sigma = 25$, PSNR: 20.15dB | MLP: 28.96dB, BM3D: **29.83dB** | between 0.00 and 3.71

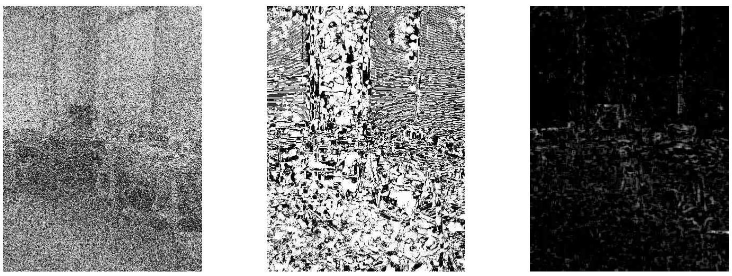


$\sigma = 35$, PSNR: 17.26dB | MLP: 27.17dB, BM3D: **28.10dB** | between 0.00 and 2.44



$\sigma = 50$, PSNR: 14.15dB | MLP: 25.07dB, BM3D: **26.23dB** | between 0.00 and 2.06

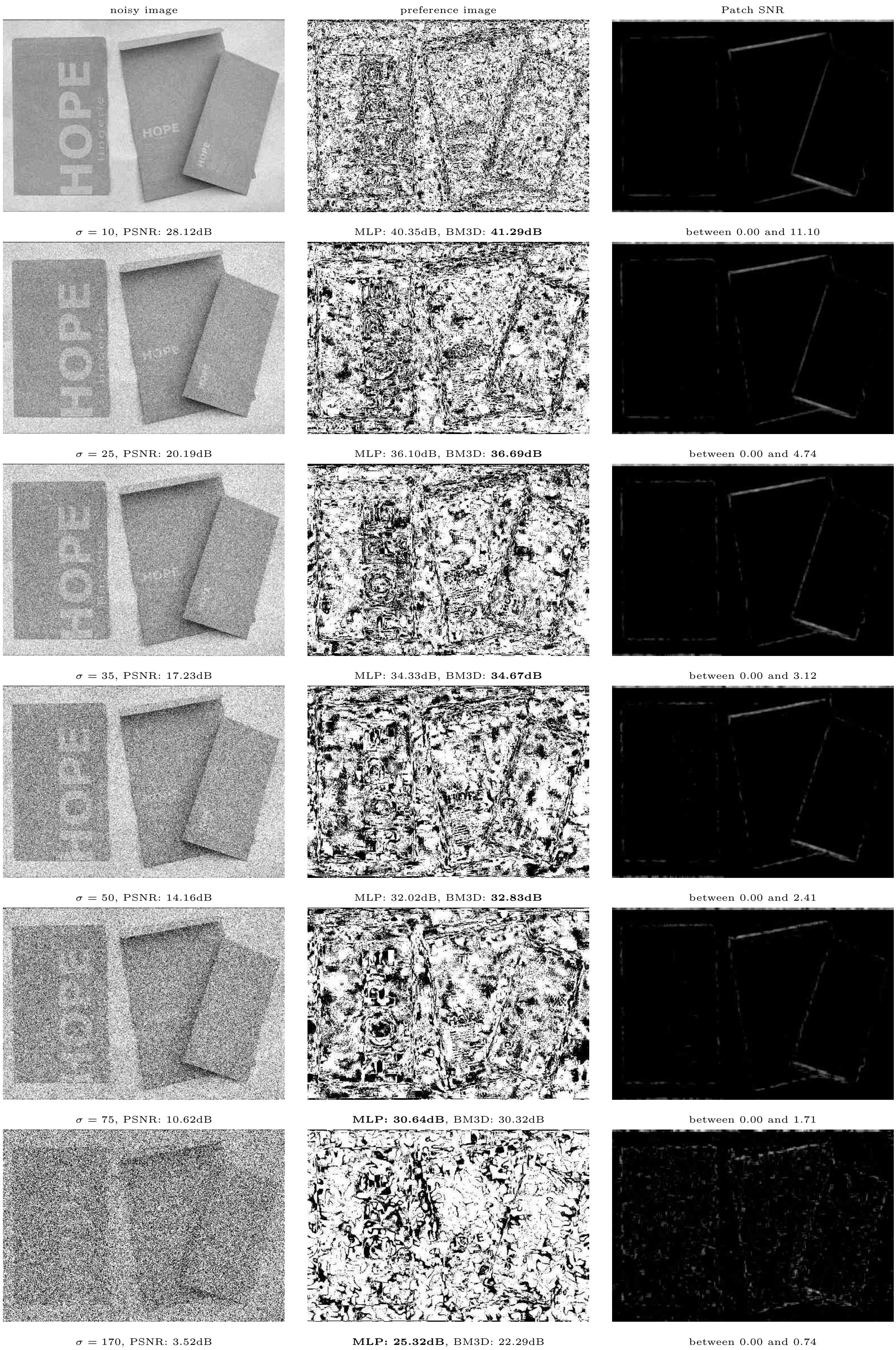
$\sigma = 75$, PSNR: 10.62dB | MLP: 23.26dB, BM3D: **24.36dB** | between 0.00 and 1.24



$\sigma = 170$, PSNR: 3.54dB | **MLP: 21.07dB**, BM3D: 19.70dB | between 0.00 and 0.68

BM3D beats MLP, image 8

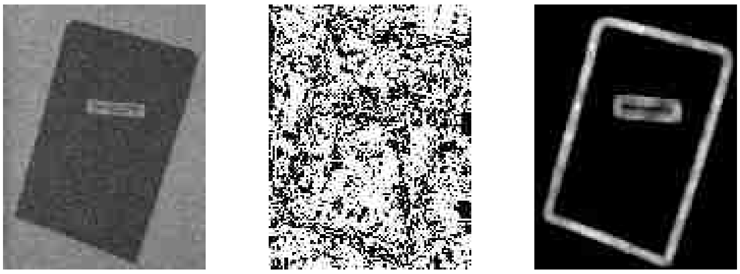
Notes: BM3D performs well on the straight edges contained in this image. However, the PatchSNR predicts the contrary effect: An external method should be used for the edges.



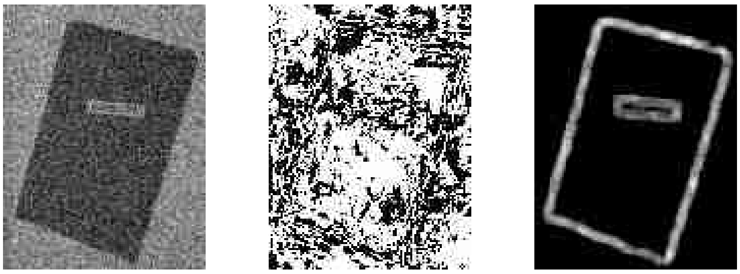
BM3D beats MLP, image 9

Notes: As in the previous image, BM3D performs well on the straight edges contained in this image. However, the PatchSNR predicts the contrary effect: An external method should be used for the edges.

noisy image | preference image | Patch SNR

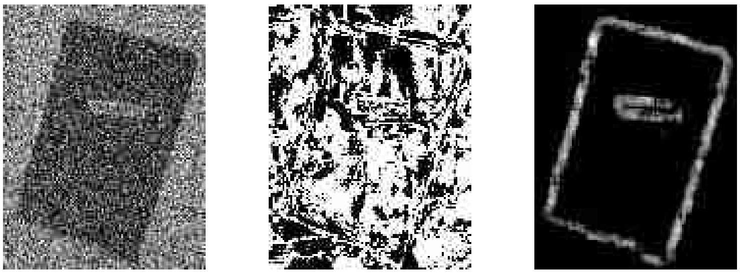


$\sigma = 10$, PSNR: 28.12dB | MLP: 38.90dB, BM3D: **40.10dB** | between 0.00 and 4.18



$\sigma = 25$, PSNR: 20.18dB | MLP: 34.62dB, BM3D: **35.59dB** | between 0.00 and 1.63

$\sigma = 35$, PSNR: 17.25dB | MLP: 33.36dB, BM3D: **34.30dB** | between 0.00 and 1.32



$\sigma = 50$, PSNR: 14.10dB | MLP: 30.99dB, BM3D: **31.30dB** | between 0.00 and 0.99

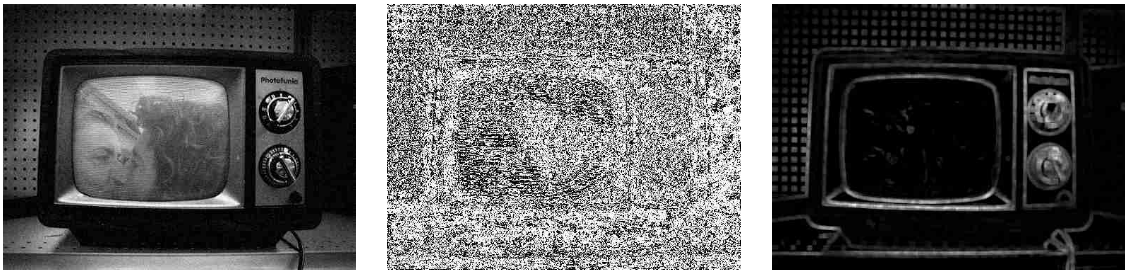
$\sigma = 75$, PSNR: 10.65dB | **MLP: 28.95dB**, BM3D: 28.87dB | between 0.00 and 0.68

$\sigma = 170$, PSNR: 3.56dB | **MLP: 24.70dB**, BM3D: 21.15dB | between 0.00 and 0.47

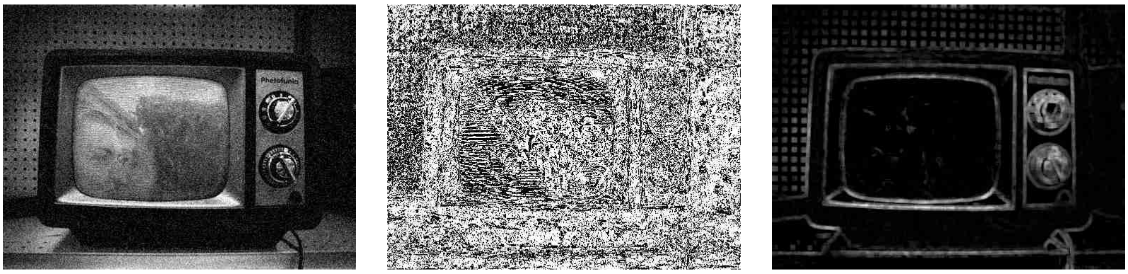
BM3D beats MLP, image 10

Notes: This image contains regular patterns (e.g. the lines in the TV image), which are better denoised with BM3D. However, the smooth areas are better denoised with MLP. The PatchSNR does not correctly predict this effect.

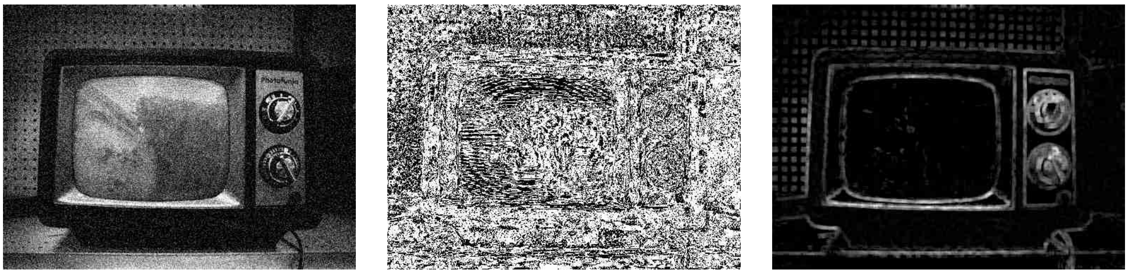
noisy image | preference image | Patch SNR



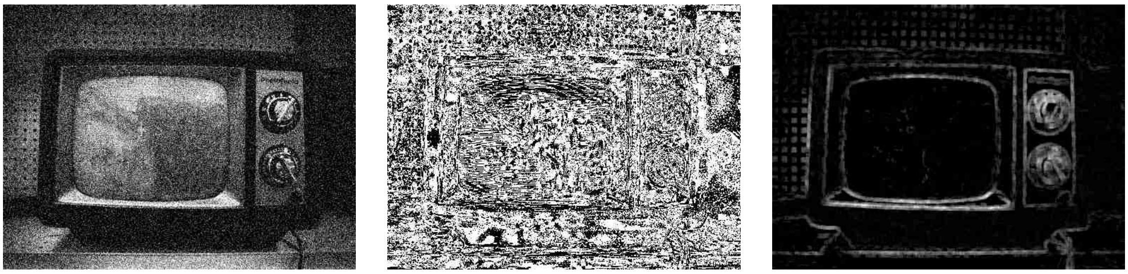
$\sigma = 10$, PSNR: 28.13dB | MLP: 34.57dB, BM3D: **35.03dB** | between 0.00 and 10.14



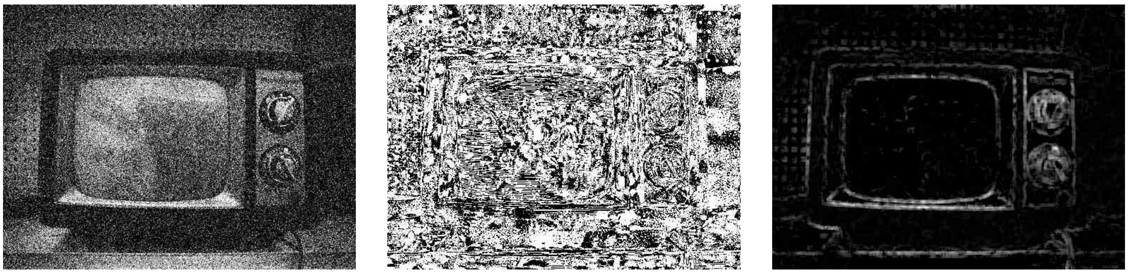
$\sigma = 25$, PSNR: 20.18dB | MLP: 30.43dB, BM3D: **31.17dB** | between 0.00 and 4.41



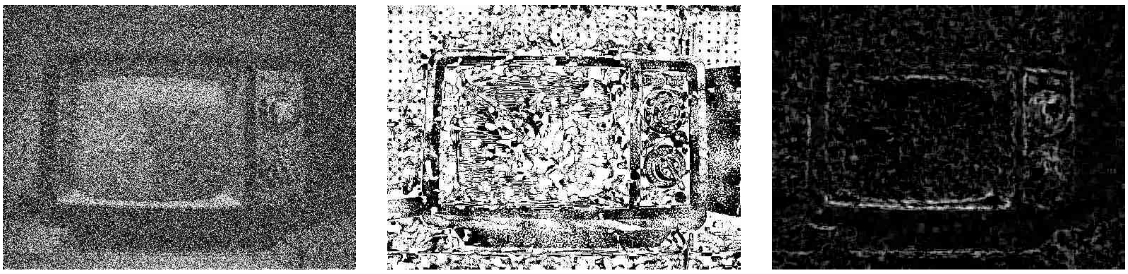
$\sigma = 35$, PSNR: 17.25dB | MLP: 28.84dB, BM3D: **29.44dB** | between 0.00 and 3.07



$\sigma = 50$, PSNR: 14.15dB | MLP: 27.44dB, BM3D: **27.54dB** | between 0.00 and 2.25



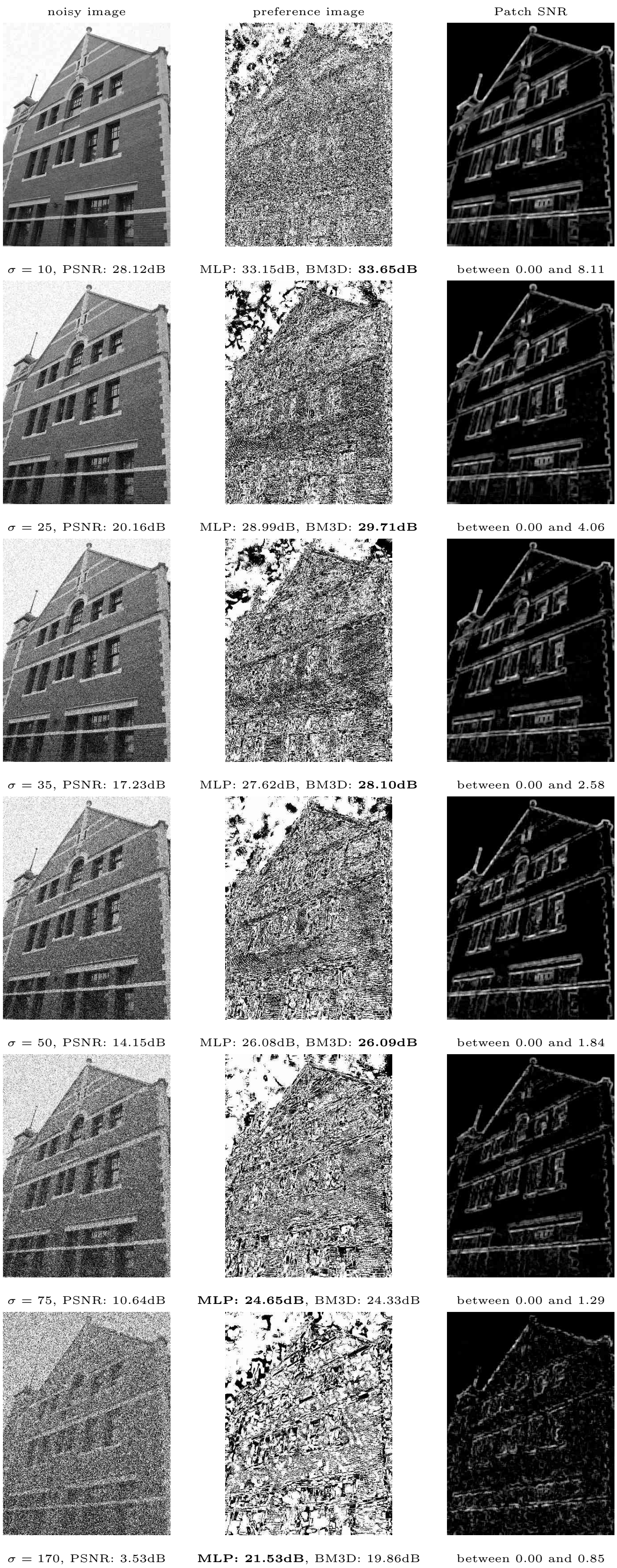
$\sigma = 75$, PSNR: 10.62dB | **MLP: 26.03dB**, BM3D: 25.67dB | between 0.00 and 1.68



$\sigma = 170$, PSNR: 3.52dB | **MLP: 23.14dB**, BM3D: 20.49dB | between 0.00 and 0.85

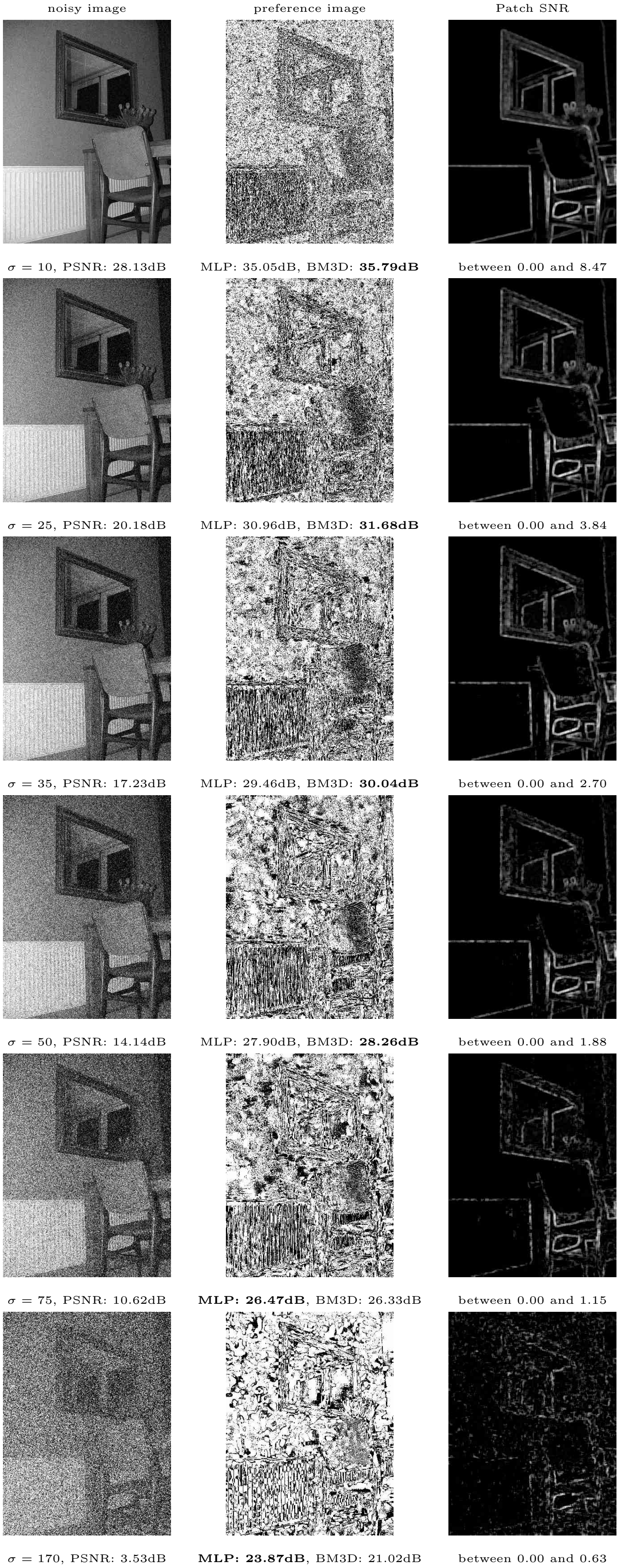
BM3D beats MLP, image 11

Notes: The brick wall contains regular patterns that are better denoised with BM3D. The sky should be preferentially denoised with MLP. The PatchSNR does not correctly predict this effect.



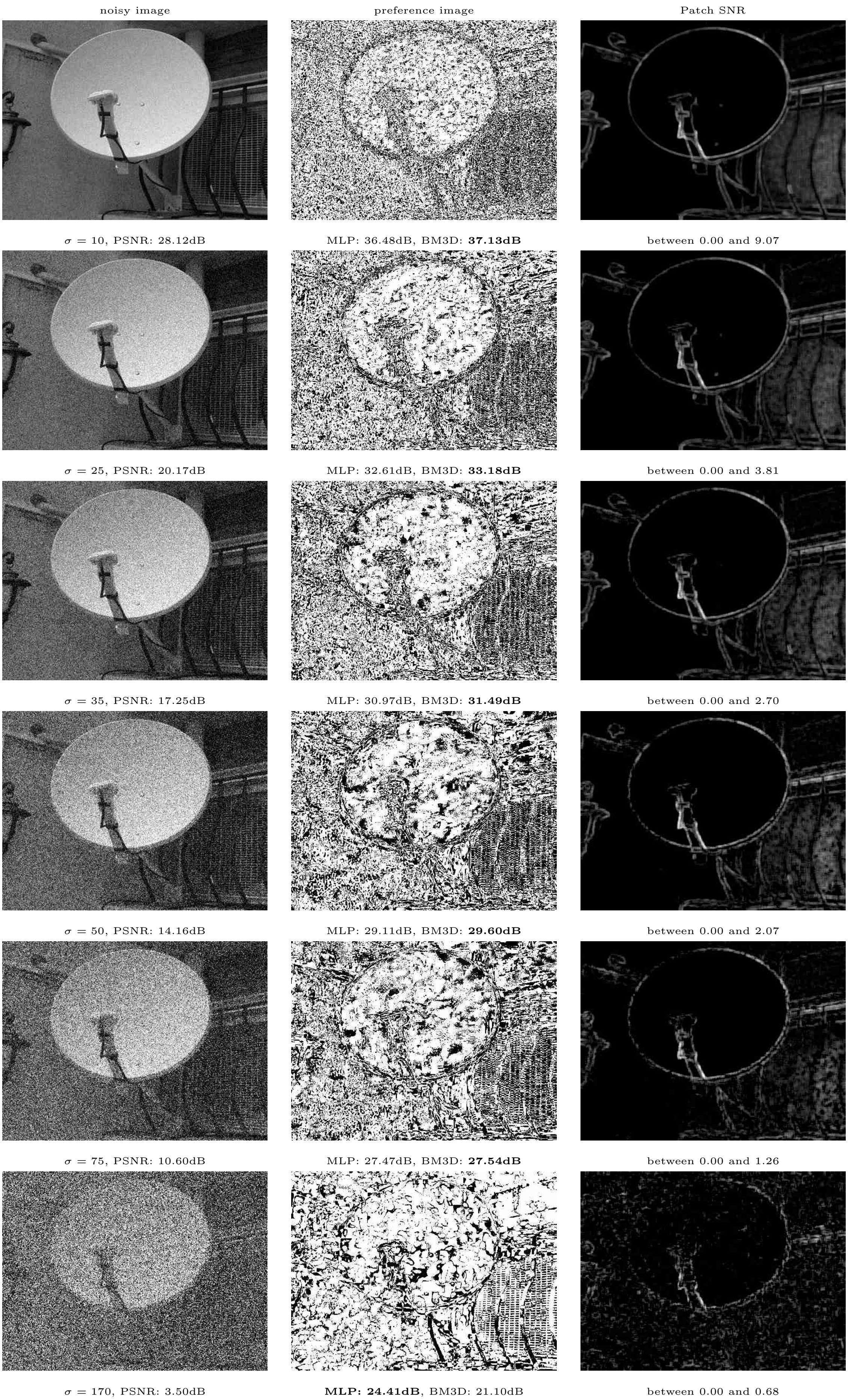
BM3D beats MLP, image 12

Notes: The radiator, part of the chair, and the rim of the mirror contain regular patterns, which should be denoised with BM3D. The rest of the image is rather smooth and is better denoised with MLP. The PatchSNR does not correctly predict this effect.



BM3D beats MLP, image 13

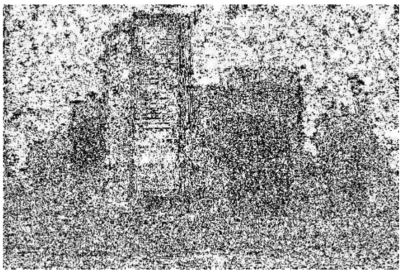
Notes: The vent of the aircondioner contains regular, grid-like patterns that are better denoised with BM3D, whereas the rest of the image is rather smooth and therefore better denoised with MLP. However, the PatchSNR predicts the opposite.



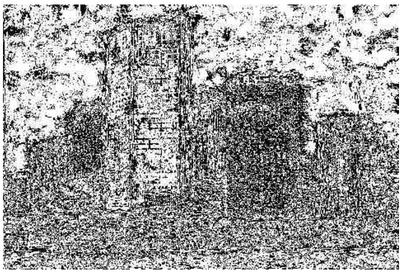
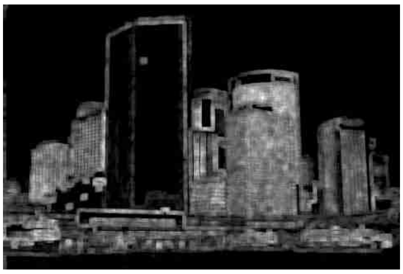
BM3D beats MLP, image 14

Notes: Some of the building contain regular patterns, which are better denoised with BM3D. The sky and some other buildings are smooth and therefore better denoised with BM3D. The PatchSNR predicts the opposite.

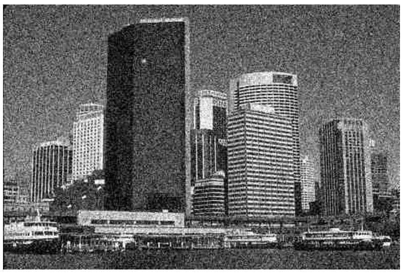
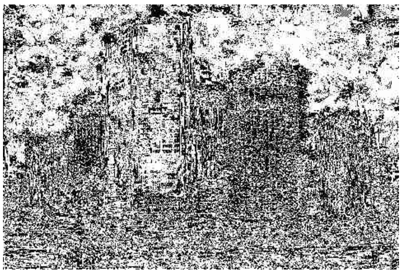
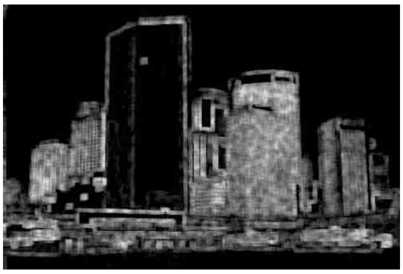
noisy image | preference image | Patch SNR

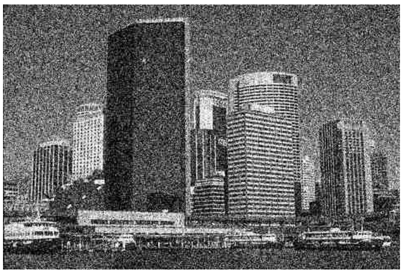
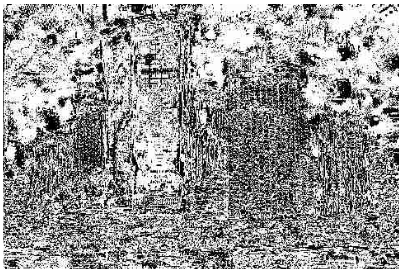
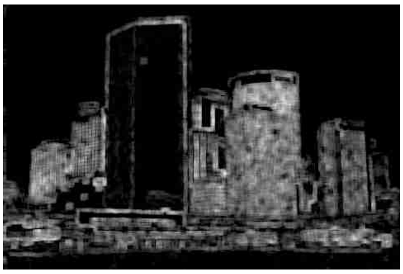

$\sigma = 10$, PSNR: 28.11dB | MLP: 33.08dB, BM3D: **33.67dB** | between 0.00 and 9.93

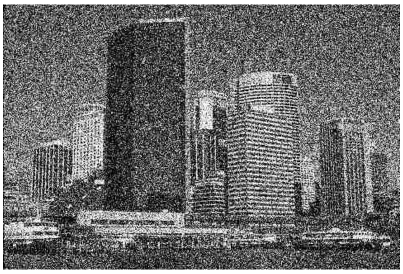
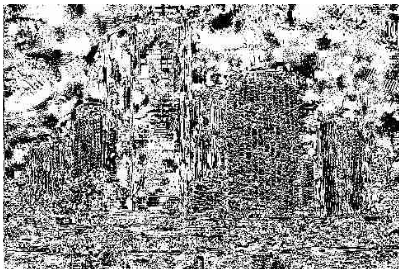
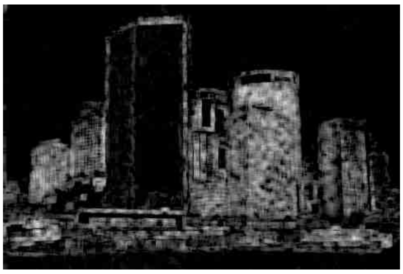
$\sigma = 25$, PSNR: 20.13dB | MLP: 27.84dB, BM3D: **28.52dB** | between 0.00 and 4.70

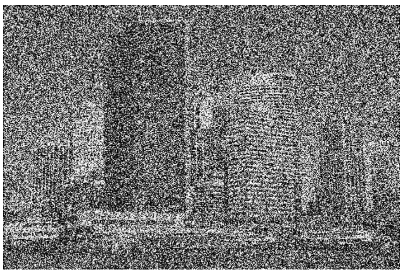
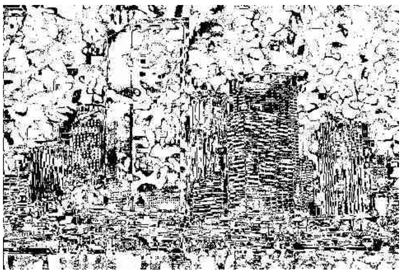
$\sigma = 35$, PSNR: 17.26dB | MLP: 26.07dB, BM3D: **26.57dB** | between 0.00 and 3.21

$\sigma = 50$, PSNR: 14.14dB | MLP: 24.25dB, BM3D: **24.57dB** | between 0.00 and 2.90

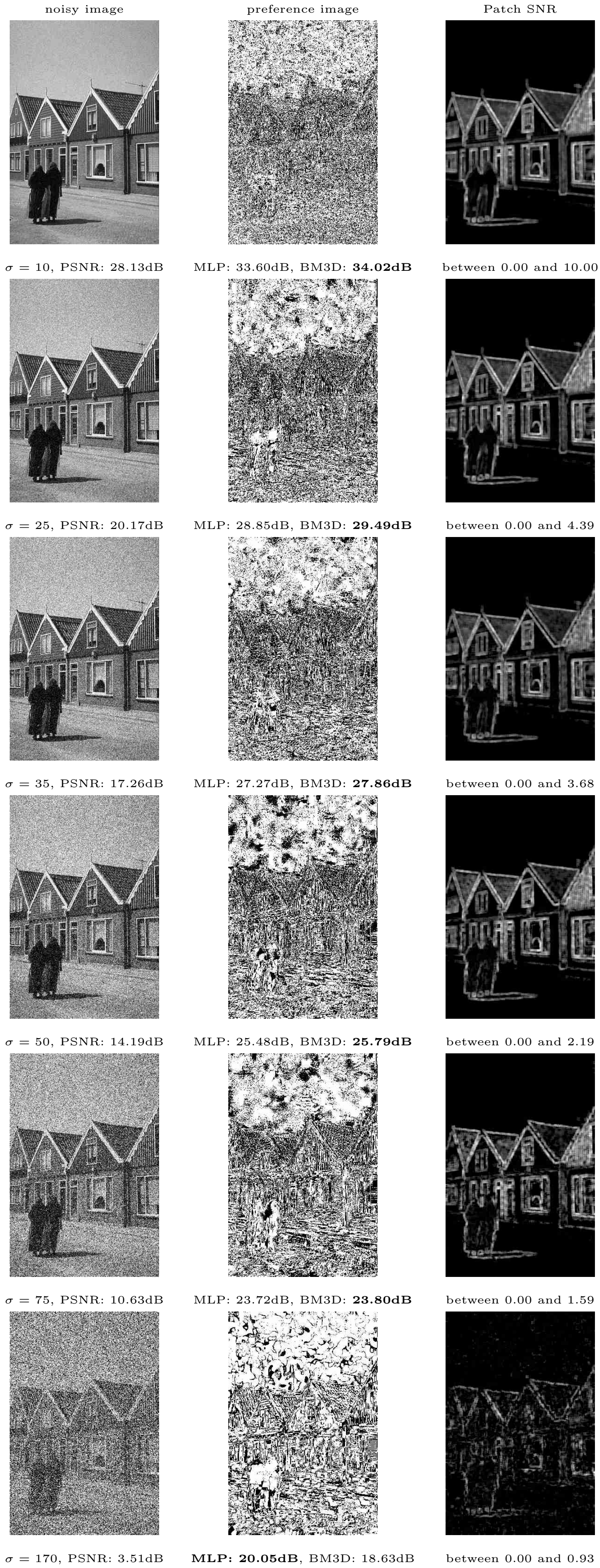
$\sigma = 75$, PSNR: 10.63dB | MLP: 22.42dB, BM3D: **22.64dB** | between 0.00 and 1.65


$\sigma = 170$, PSNR: 3.51dB | **MLP: 18.05dB**, BM3D: 17.10dB | between 0.00 and 0.98

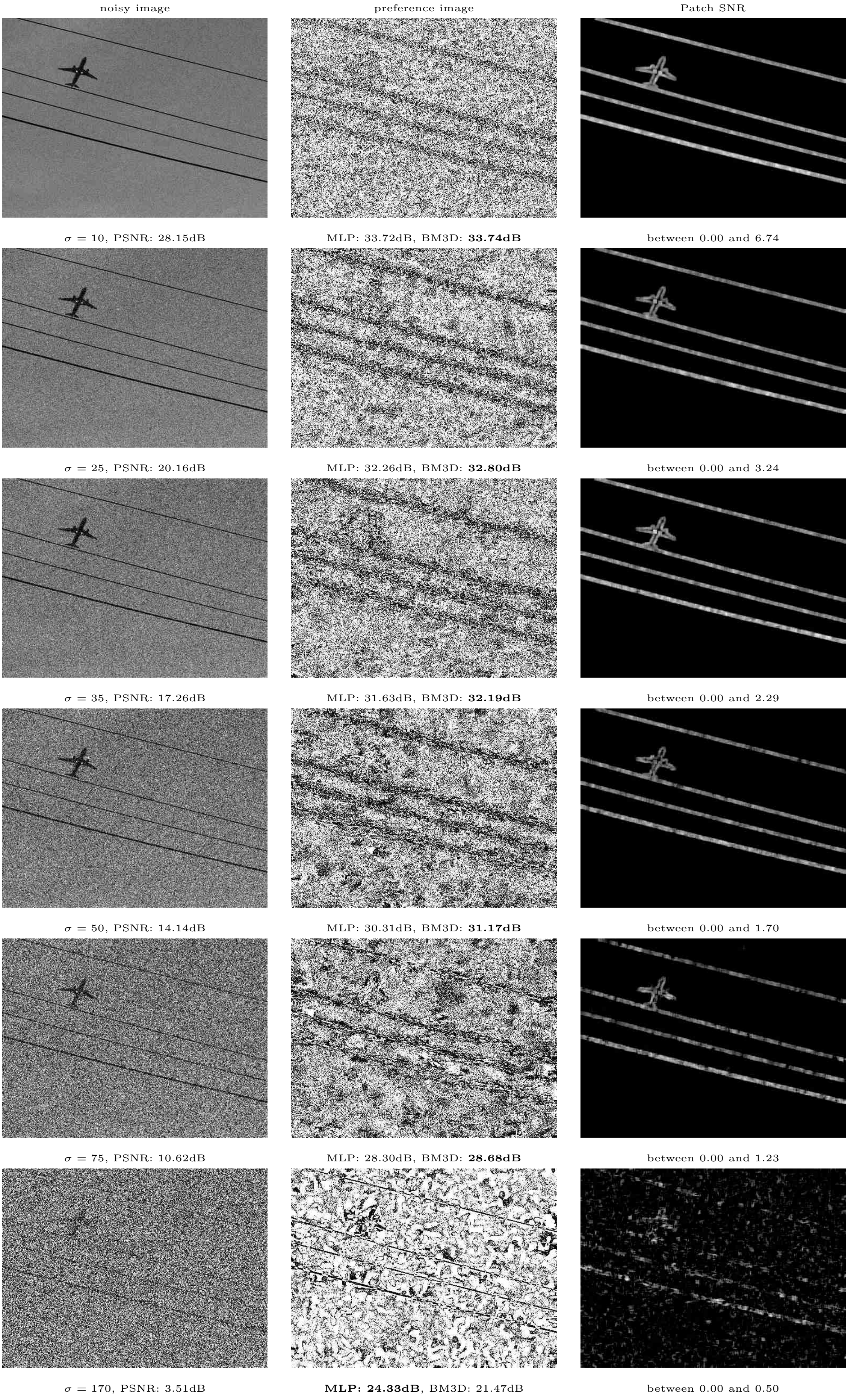
BM3D beats MLP, image 15

Notes: The roofs and other parts of the houses contain repeating patterns which are better denoised with BM3D, whereas the sky is smooth and better denoised with MLP. The PatchSNR predicts the opposite effect.



BM3D beats MLP, image 16

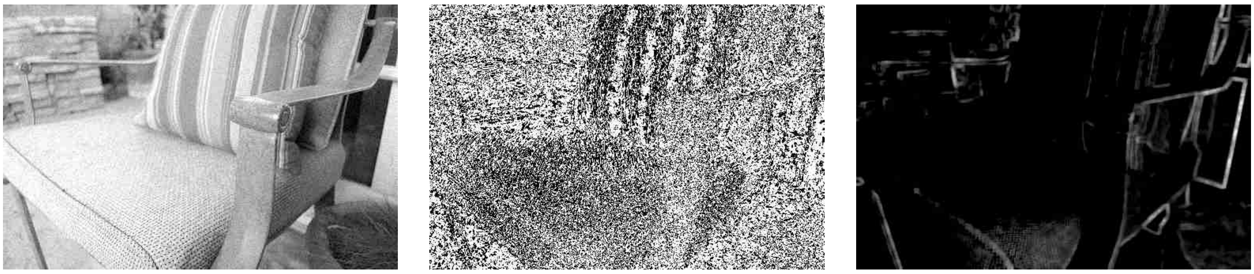
Notes: The electricity lines are regular and therefore better denoised with BM3D, whereas the sky is better denoised with MLP. The PatchSNR predicts the opposite effect.



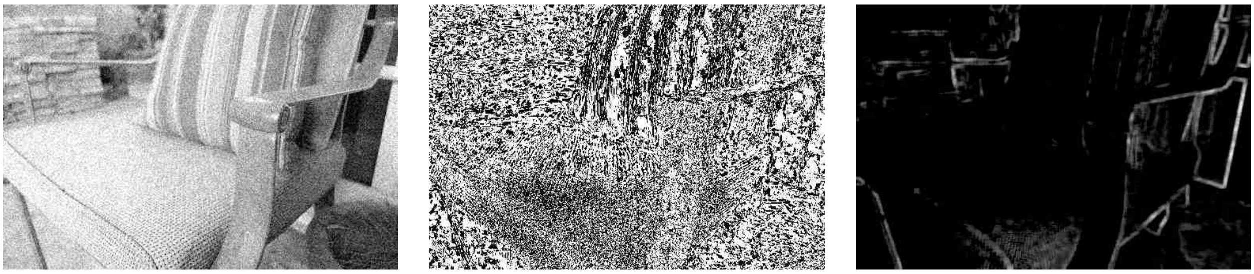
BM3D beats MLP, image 17

Notes: The chair contains a regular pattern that is better denoised with BM3D. The PatchSNR does not predict this effect.

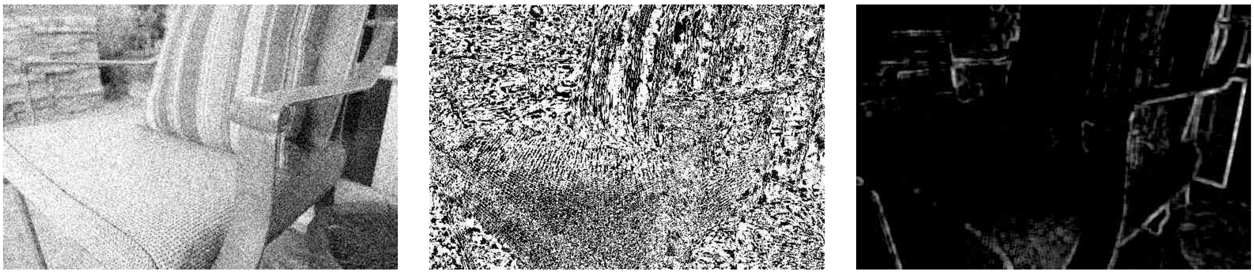
noisy image | preference image | Patch SNR



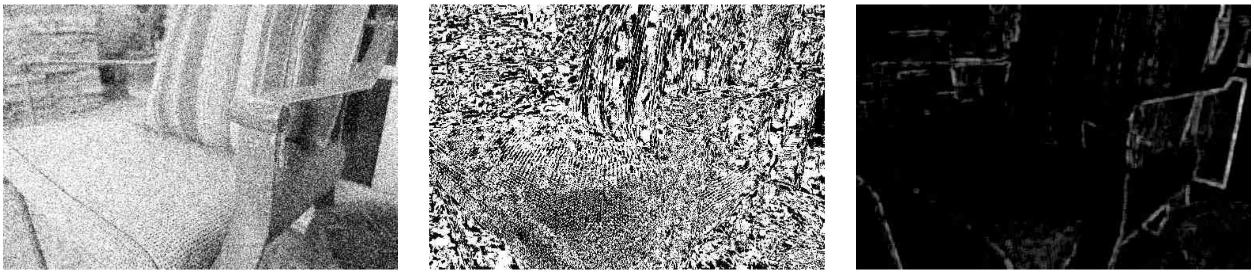
$\sigma = 10$, PSNR: 28.14dB | MLP: 33.46dB, BM3D: **34.01dB** | between 0.00 and 9.02



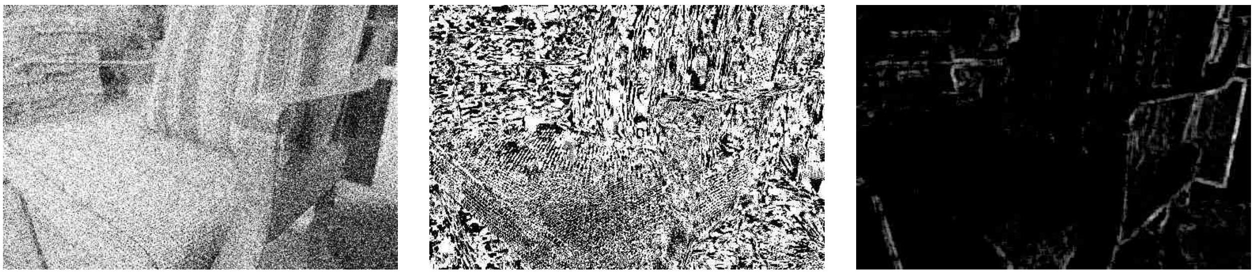
$\sigma = 25$, PSNR: 20.17dB | MLP: 29.07dB, BM3D: **29.66dB** | between 0.00 and 3.92



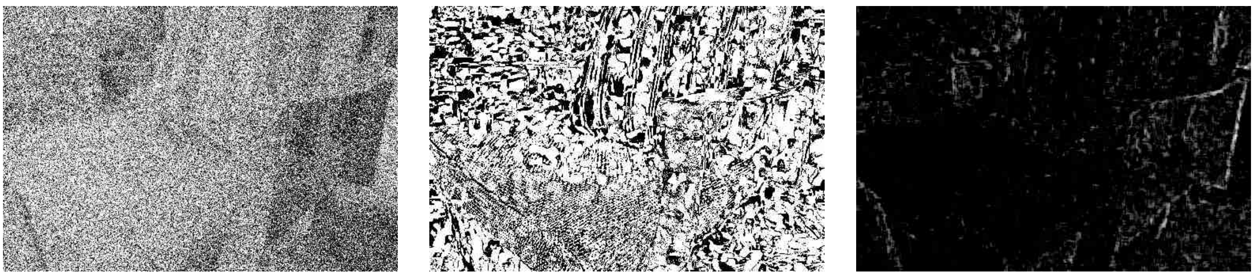
$\sigma = 35$, PSNR: 17.24dB | MLP: 27.59dB, BM3D: **28.03dB** | between 0.00 and 2.75



$\sigma = 50$, PSNR: 14.13dB | MLP: 26.10dB, BM3D: **26.43dB** | between 0.00 and 2.09



$\sigma = 75$, PSNR: 10.62dB | **MLP: 24.86dB**, BM3D: 24.83dB | between 0.00 and 1.43

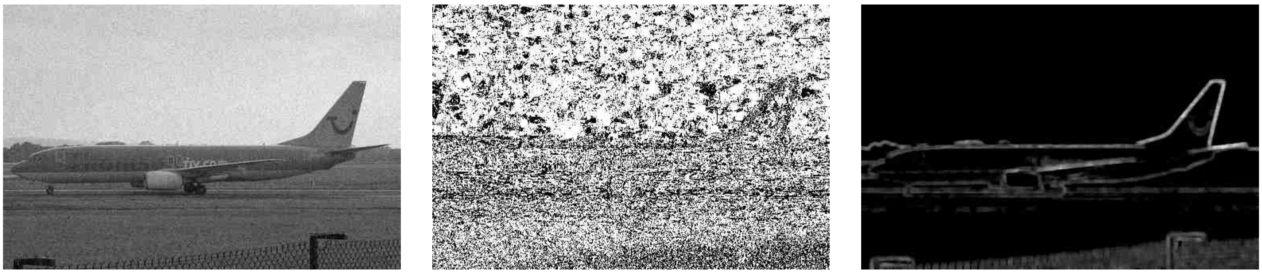


$\sigma = 170$, PSNR: 3.53dB | **MLP: 22.51dB**, BM3D: 20.82dB | between 0.00 and 0.82

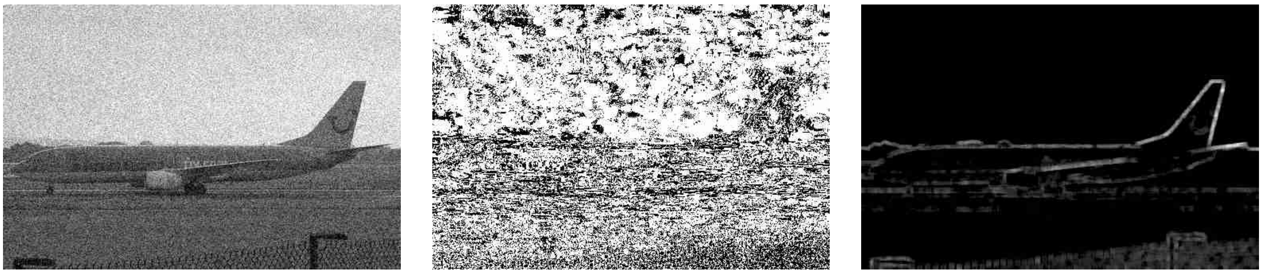
BM3D beats MLP, image 18

Notes: The fence is regular and hence better denoised with BM3D. The PatchSNR predicts the opposite effect.

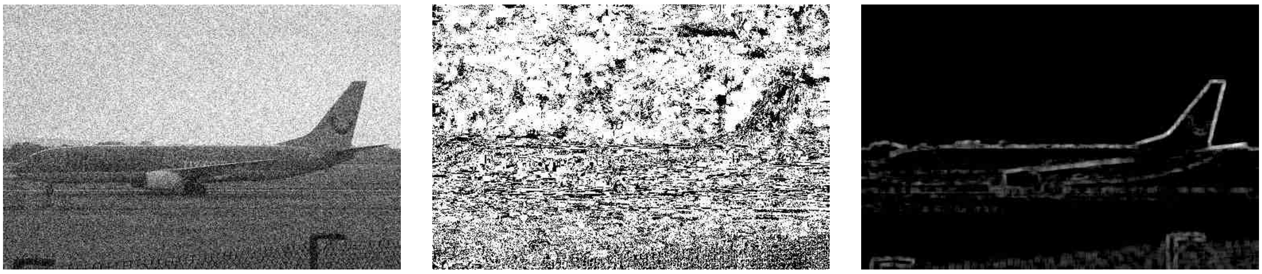
noisy image	preference image	Patch SNR



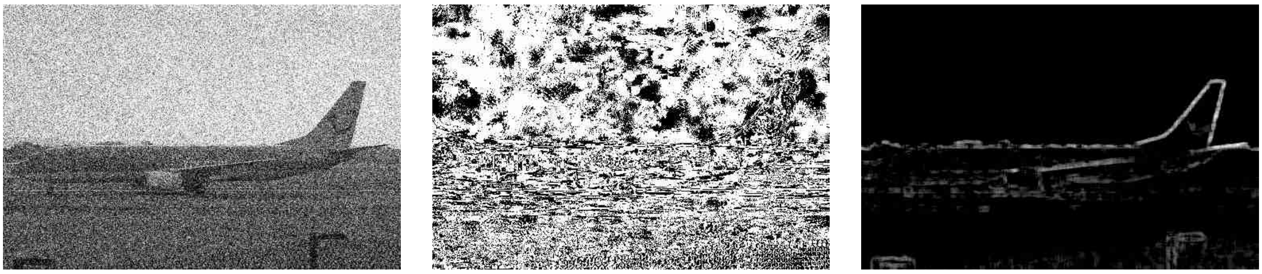
$\sigma = 10$, PSNR: 28.13dB | MLP: 36.28dB, BM3D: **36.73dB** | between 0.00 and 7.27



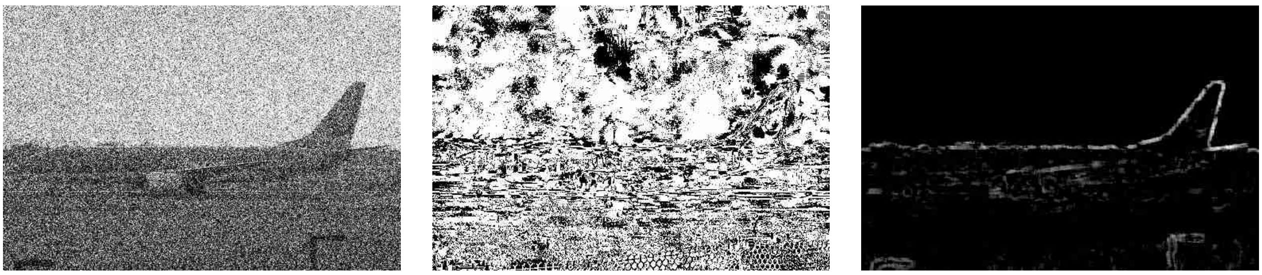
$\sigma = 25$, PSNR: 20.17dB | MLP: 32.35dB, BM3D: **32.83dB** | between 0.00 and 2.77



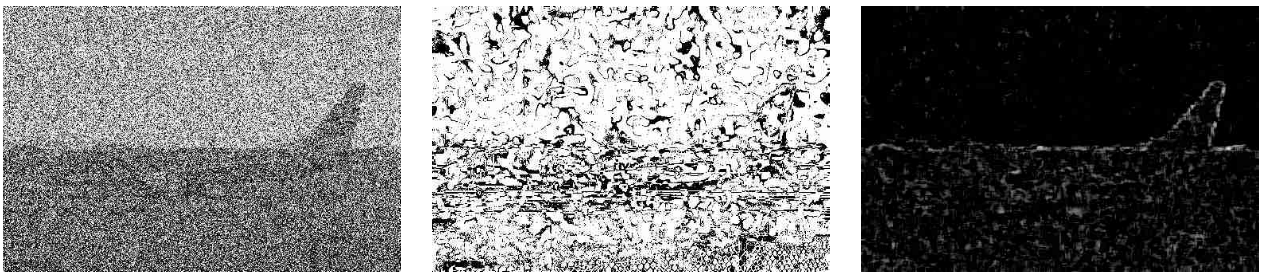
$\sigma = 35$, PSNR: 17.27dB | MLP: 30.97dB, BM3D: **31.06dB** | between 0.00 and 2.30



$\sigma = 50$, PSNR: 14.15dB | **MLP: 29.75dB**, BM3D: 29.57dB | between 0.00 and 1.78



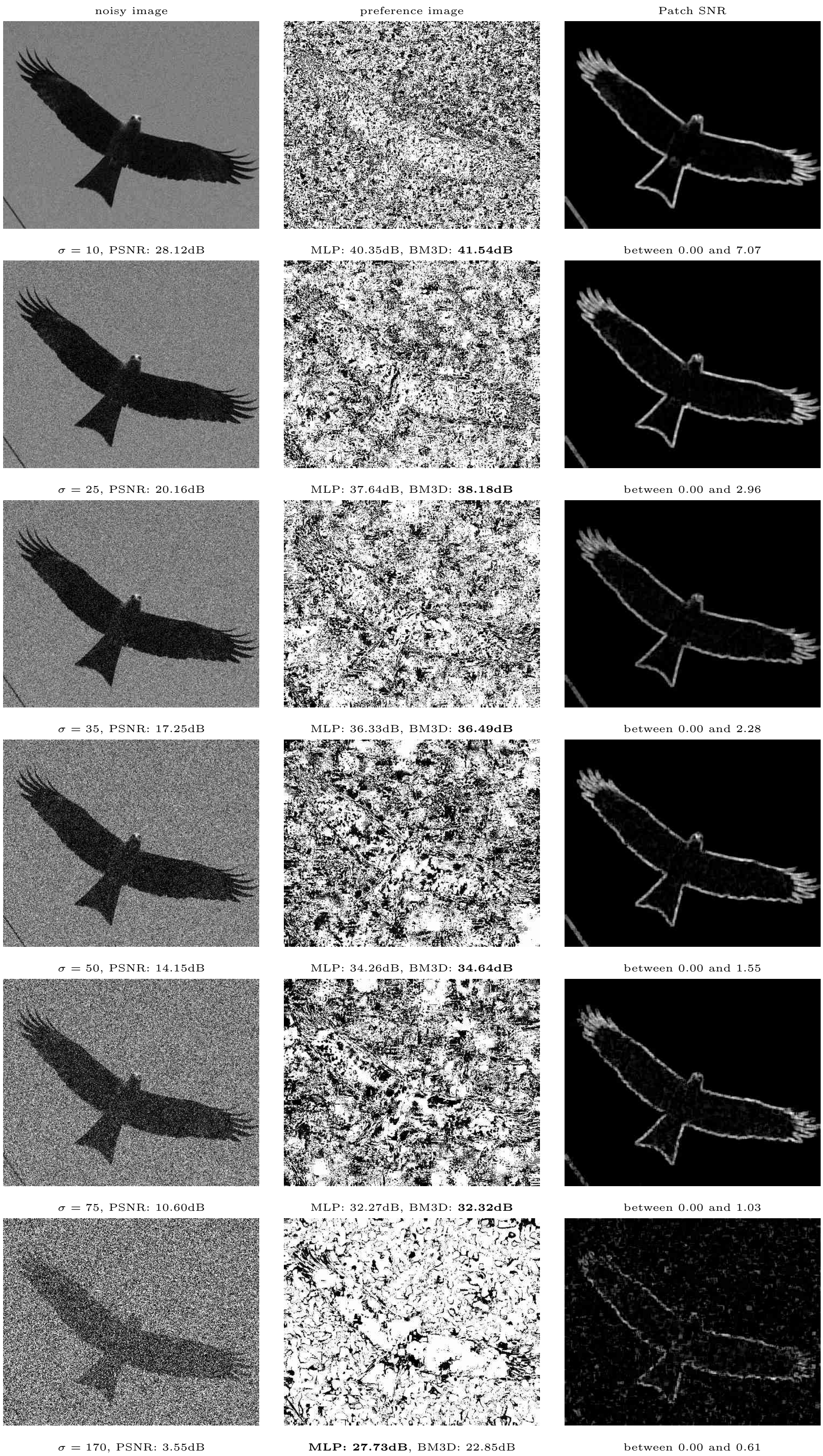
$\sigma = 75$, PSNR: 10.60dB | **MLP: 28.64dB**, BM3D: 28.04dB | between 0.00 and 1.35



$\sigma = 170$, PSNR: 3.51dB | **MLP: 26.25dB**, BM3D: 22.09dB | between 0.00 and 0.64

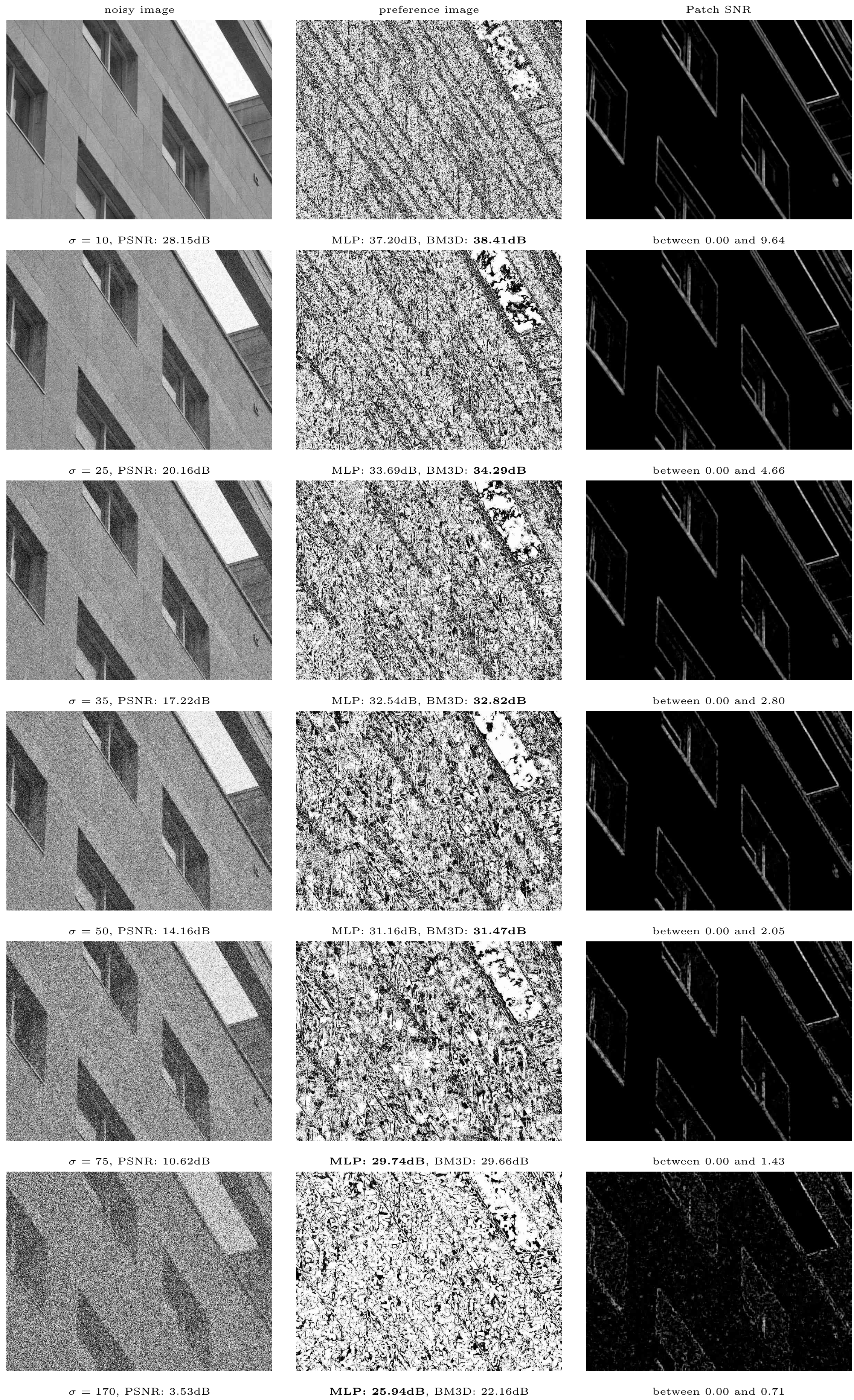
BM3D beats MLP, image 19

Notes: BM3D performs better than MLP on the electricity line and on the contour of the bird. The PatchSNR predicts the opposite effect.



BM3D beats MLP, image 20

Notes: This image contains very regular lines which are better denoised with BM3D. The PatchSNR predicts the opposite effect.

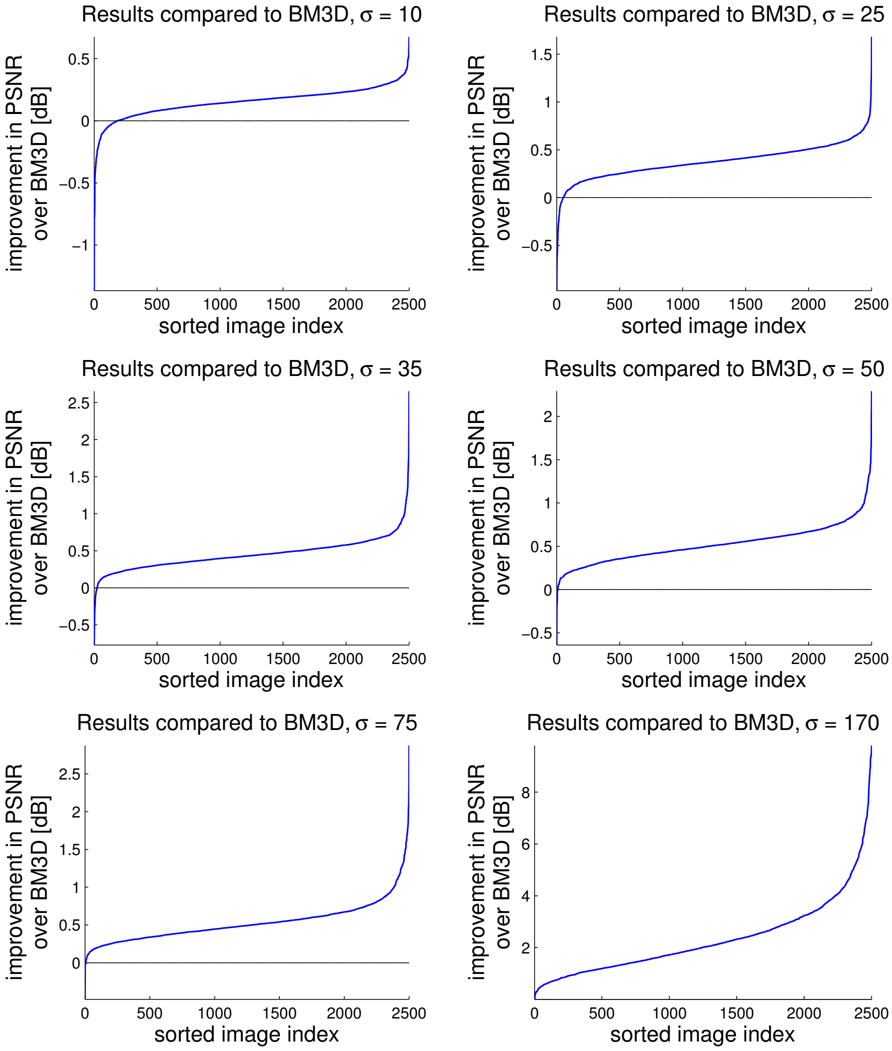


2 Performance profiles

In this section, we provide further results to show that our method indeed achieves the best denoising results reported in literature. We compare our method to both BM3D and MLP (the best internal and external methods, respectively) on a large test set of images on six different noise levels. We present the results using a "performance profile", where values below 0 indicate that our method performs worse than the competitor, whereas values above 0 indicate that it performs better.

2.1 Compared to BM3D

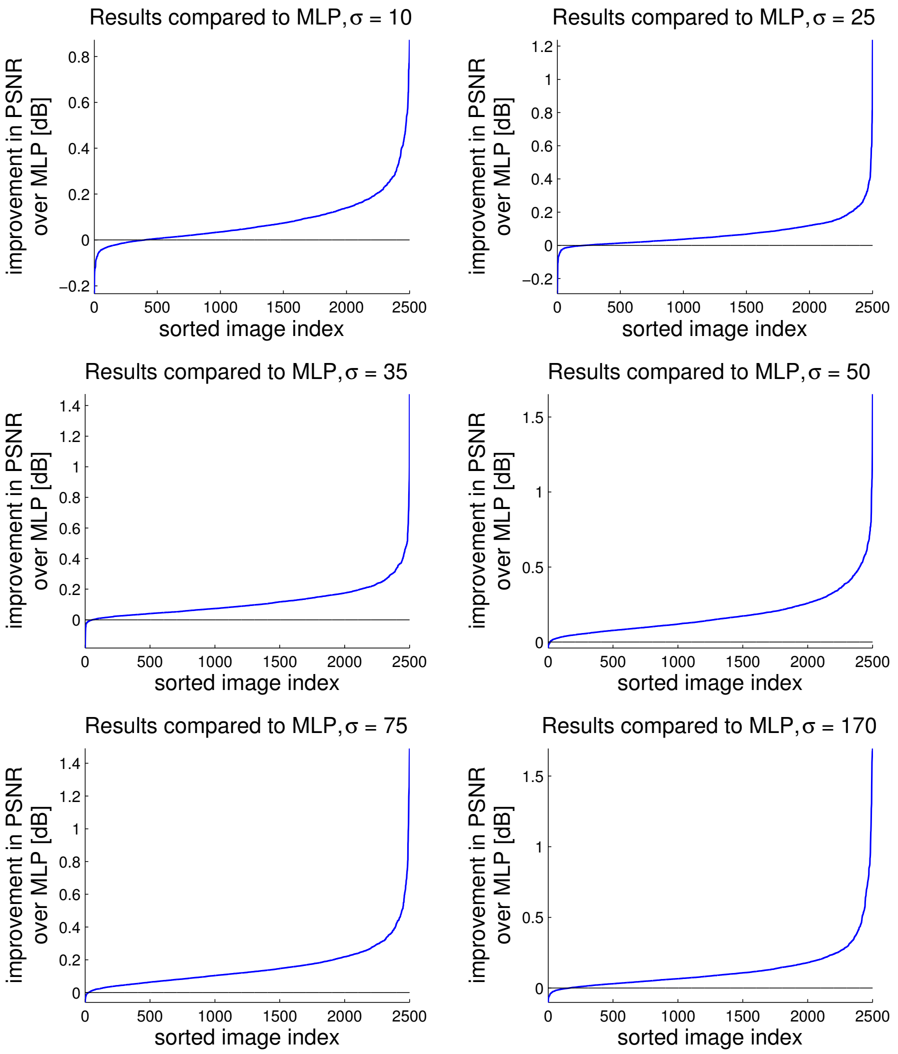
Our method outperforms BM3D on 92.6%, 97.92%, 99.04%, 99.72%, 99.84%, and 100% of the images, on the noise levels $\sigma = 10, 25, 35, 50, 75$, and 170, respectively:



Sometimes, our method outperforms BM3D by several dB. The higher the noise level, the better are our results compared to BM3D. At $\sigma = 170$, there is no image on which BM3D performs better than our approach. The average improvements on the same noise levels are 0.15dB, 0.38dB, 0.45dB, 0.52dB, 0.53dB, and 2.32dB.

2.2 Compared to MLP

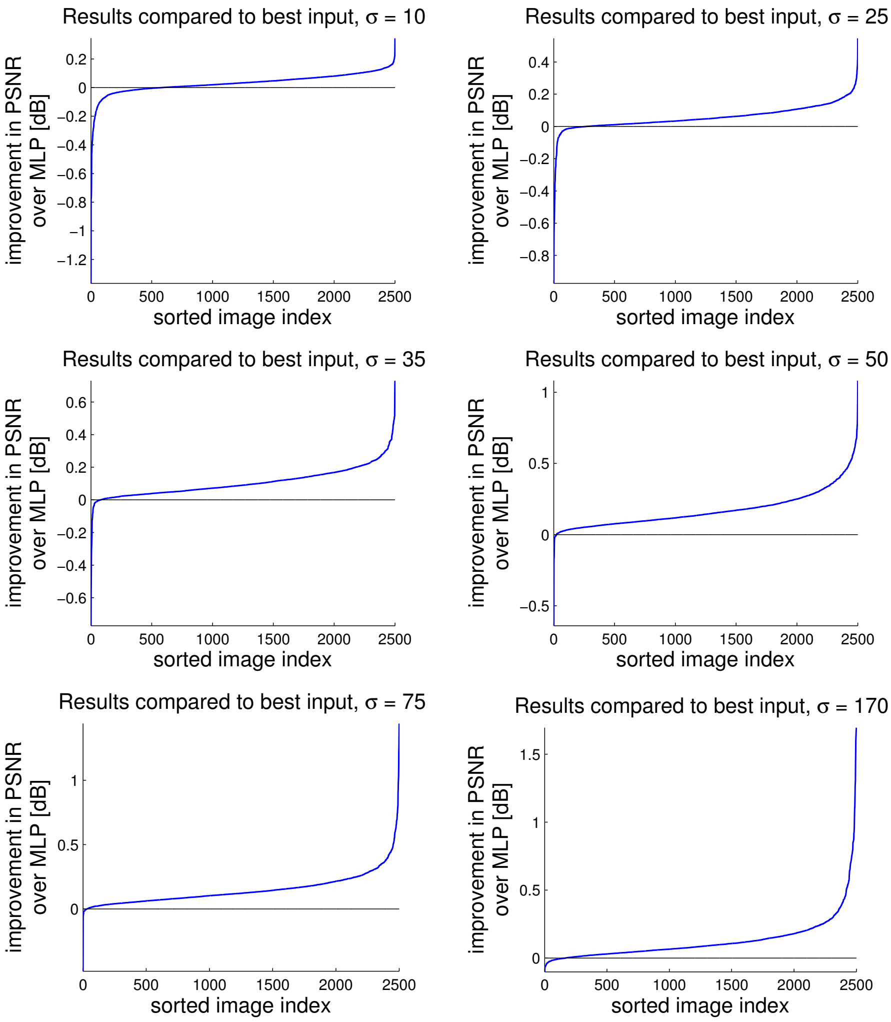
Our method outperforms MLP on 84.08%, 91.12%, 97.76%, 99.4%, 98.96%, and 93.48% of the images, on the noise levels $\sigma = 10, 25, 35, 50, 75$, and 170, respectively:



The improvement of our method over MLP is sometimes above 1dB. The average improvements on the same noise levels are 0.08dB, 0.07dB, 0.12dB, 0.18dB, 0.15dB, 0.13dB.

2.3 Compared to the best of the inputs

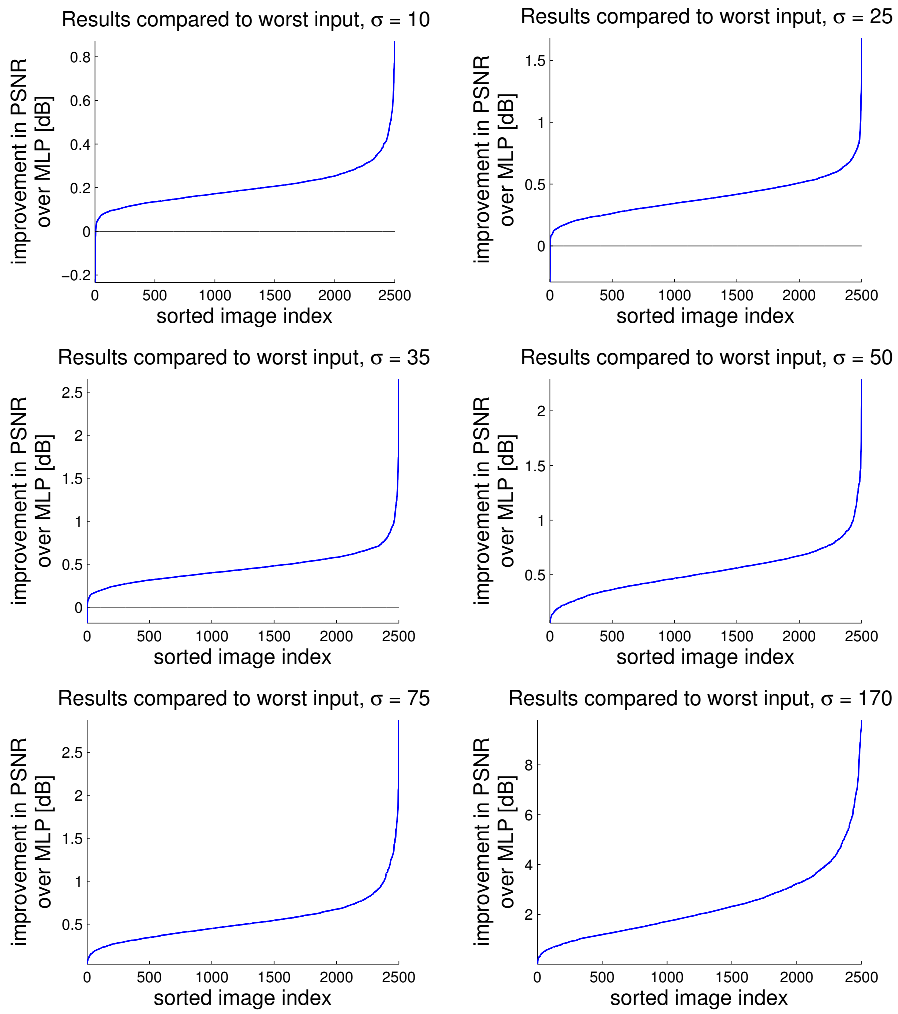
How often does our method outperform both BM3D and MLP? This happens for 76.92%, 89.12%, 96.92%, 99.12%, 98.8% and 93.48% of the images on the noise levels $\sigma = 10, 25, 35, 50, 75$, and 170, respectively:



This means that we are usually better than the best of the two inputs methods (BM3D and MLP). The average improvements on the same noise levels are 0.03dB, 0.06dB, 0.11dB, 0.17dB, 0.15dB, and 0.13dB.

2.4 Compared to the worst of the inputs

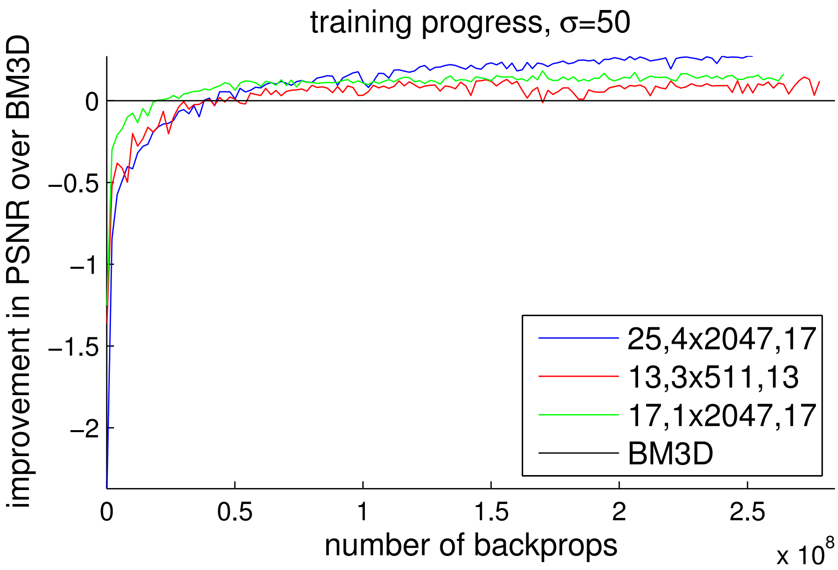
How often does our method outperform at least of the two methods (BM3D or MLP)? This is the case on 99.76%, 99.92%, 99.88%, 100.00%, 100.00%, and 100.00% of the images on the noise levels $\sigma = 10, 25, 35, 50, 75$, and 170, respectively:



This means that we are almost always at least as good as the worst of the input methods (BM3D or MLP). The average improvements on the same noise levels are 0.20dB, 0.39dB, 0.46dB, 0.53dB, 0.54dB, and 2.32dB.

3 Other architectures

Are the results we achieve due to our specific choice of architecture? To answer this question, we train different architectures. We periodically interrupt the training procedure and evaluate the different architectures on a test set of 11 images. The figure below plots the average result achieved on the test images at different stages of the training procedure.



The notation $25, 4 \times 2047, 17$ refers to input patches of size 25×25, four hidden layers, each of size 2047, and output patches of size 17. We see that all architectures we tried achieve improvements over BM3D. We achieve the best results with the architecture with the most and largest hidden layers as well as the largest patch sizes.

4 The whitening transform

In our paper, we noted that the BM3D and MLP images look similar, justifying the use of a whitening step. Here, we explain the whitening step in more detail.

The whitening matrix is learned by performing a PCA on a training set of noisy and denoised images. For the noise level $\sigma = 35$, we obtained the following matrix:

$$W = \begin{pmatrix} 0.0053 & 0.0048 & 0.0048 \\ 0.0270 & -0.0150 & -0.0148 \\ 0.0009 & 0.2504 & -0.2514 \end{pmatrix}.$$

The three whitened images are obtained by right-multiplying W with the three inputs images. The first row approximately creates an average of the three inputs. The second row approximately creates the difference image between the noisy image and the two denoised images, and the third row approximately creates the difference image between BM3D and MLP.

The whitening matrix also handles the fact that the difference between BM3D and MLP is quite small (i.e. has small variance). The three whitened images all have approximately variance 1 (averaged on a large dataset of images).

The effect of the whitening transform is illustrated in the images below, where the images in the top row are the inputs and the images in the bottom row are the outputs of the whitening transform.

References

[1] Inbar Mosseri, Maria Zontak, and Michal Irani. Combining the power of internal and external denoising. In *International Conference on Computational Photography (ICCP)*, pages 1–9. IEEE, 2013. 2, 3, 4, 5, 6, 7, 8, 11, 19, 24