We thank the reviewers for constructive and detailed comments. We will revise the paper according to the above.

We will discuss in detail and visualize the learned features for various noise levels in the revised paper.

Our PSNR results are 25.57 (15%), 24.01 (20%), 21.79 (30%) on (R4) Improvement of the proposed method in Fig. 3.

This is the method that removes the Wiener deconvolution in our model, thus (R4) “Ours w/o Wiener” is not clear.

The robustness and effectiveness of the SNR estimation have been demonstrated in (R4) Robustness to noise & inaccurate kernels vs. competing methods.

As we use the same training dataset as [49], we do not retrain (R3, R4) Some recent approaches are not retrained.

The motivation to use Wiener deconvolution is that we can use it to derive an effective feature-based deconvolution (R1, R3) Contribution and motivation.

Tabs. 3 and 5 are evaluated on [19] and [16], respectively, with (R1) Why PSNRs in Tabs. 3 and 5 are so different?

The derivations of the feature-based Wiener (R1, R3) Parameters of the feature refinement.

blocks (except the first), the number of features are the same and thus the parameters can be shared.

First, the feature-based Wiener (R4) Why feature-based Wiener deconvolution is better than image-based method.

deconvolution can utilize more useful information beyond the image intensity to better constrain the deconvolution process (L110–114). Second, finer-scale detail is better modeled in the feature space. Especially, the deep feature extractor in our end-to-end network can adaptively learn useful features for the final image restoration (Figs. 3 and 6).

(R4) Robustness to noise & inaccurate kernels vs. competing methods.

First, benefitting from the proposed feature-based Wiener deconvolution, our method can adaptively estimate the SNR from blurry features. Second, our end-to-end network facilitates the feature extractor learning useful features for deconvolution with fewer artifacts and benefits the feature refinement in handling deconvolved features to reconstruct clearer images. The robustness is also evaluated on real-world images in Figs. 1, 5, 17–19 (suppl.), where the noise is unknown and estimated kernels are inaccurate.

(R4) Effect of SNR estimation.

The robustness and effectiveness of the SNR estimation have been demonstrated in Tabs. 1–2, 9 (suppl.) and Figs. 4–5, 11–19 (suppl.). We further carry out a sensitivity analysis w.r.t. the SNR estimation using [19] by adding 0–20% perturbation to our estimated SNR (no retraining). The PSNRs differ no more than 0.06dB, suggesting the robustness of our method to the SNR estimate. In addition, as we do not model spatially-variant noise, our method may not be so robust to non-stationary noise. This is worth further study.

(R4) “Ours w/o Wiener” is not clear.

This is the method that removes the Wiener deconvolution in our model, thus is not guided by the blur kernel. We use this baseline to illustrate the importance of the Wiener deconvolution on the end-to-end network for non-blind image deblurring, which can effectively incorporate the kernel information.

(R4) Improvement of the proposed method in Fig. 3.

As both the methods in Fig. 3 use the same multi-scale feature refinement, the improvement is due to the feature-based Wiener deconvolution module.

We thank the reviewers for constructive and detailed comments. We will revise the paper according to the above responses and add all other suggested references & experimental results and correct the unclear statements as suggested.