Deep Burst Denoising
Supplementary Material

Clément Godard[1,*] Kevin Matzen[2] Matt Uyttendaele[2]
[1]University College London [2]Facebook

1. Network description

In Table 1 we detail our network architecture. All the convolutional layers are followed by a ReLU except the last layers conv8_m and conv8_s. We experimented with Batch Normalization [3] but did not find it useful given the increased memory cost at training time.

SFD				**MFD**			
layer	**input**	**in**	**out**	**layer**	**input**	**in**	**out**
conv1_s^t	N^t	3	64	conv1_m^t	$\text{conv1}_s^t, \text{conv1}_m^{t-1}$	128	64
conv2_s^t	conv1_s^t	64	64	conv2_m^t	$\text{conv2}_s^t, \text{conv2}_m^{t-1}, \text{conv1}_m^t$	196	64
conv3_s^t	conv2_s^t	64	64	conv3_m^t	$\text{conv3}_s^t, \text{conv3}_m^{t-1}, \text{conv2}_m^t$	196	64
conv4_s^t	conv3_s^t	64	64	conv4_m^t	$\text{conv4}_s^t, \text{conv4}_m^{t-1}, \text{conv3}_m^t$	196	64
conv5_s^t	conv4_s^t	64	64	conv5_m^t	$\text{conv5}_s^t, \text{conv5}_m^{t-1}, \text{conv4}_m^t$	196	64
conv6_s^t	conv5_s^t	64	64	conv6_m^t	$\text{conv6}_s^t, \text{conv6}_m^{t-1}, \text{conv5}_m^t$	196	64
conv4_s^t	conv6_s^t	64	64	conv7_m^t	$\text{conv7}_s^t, \text{conv7}_m^{t-1}, \text{conv6}_m^t$	196	64
conv8_s^t	conv7_s^t	64	3	conv8_m^t	$\text{conv8}_s^t, \text{conv8}_m^{t-1}, \text{conv7}_m^t$	70	3

Table 1: Our network architecture

2. Sensor Noise Modelling

In order to use our denoiser on real digital photographs, we need to train it with a realistic noise model. Following [1] we model the noise in each pixel as a clipped mixture of Poisson noise and additive Gaussian noise. In order to generate realistic noise, we need to simulate the digital imaging process.

Digital cameras make use of a colored Bayer filter to generate color images, which means that RAW photographs are actually single channel images where each pixel only accounts for one primary color. We simulate this process by keeping the green, red and blue channel for respectively half and a quarter of the pixels. We then convert this single channel image to linear space by using the inverse of the sRGB gamma curve. We then add noise to each pixel intensity $x \in [0, 1]$ following:

$$z = a x_P + x_N \quad \text{where} \quad x_P \sim \mathcal{P}(x/a), \quad x_N \sim \mathcal{N}(0, \sigma) \tag{1}$$

The resulting noisy image is then converted back to sRGB and demoisaiced using bilinear interpolation. We used 10 different values of a from 0.001 to 0.01 in linear increments and used a fixed value of $\sigma = 0.0001$. Although possible, we did not estimate the noise level in test images and reported the best looking image.

3. Additional Gaussian Denoising results

In Figures 1 and 2 we show additional results of our denoiser on, respectively stabilized and unstabilized, live photo data with additive white Gaussian noise with $\sigma = 50$.

4. Additional Super-Resolution 4× results

In Figure 3 we show additional results of our super-resolution $4\times$ model on live photos data for stabilized sequences.

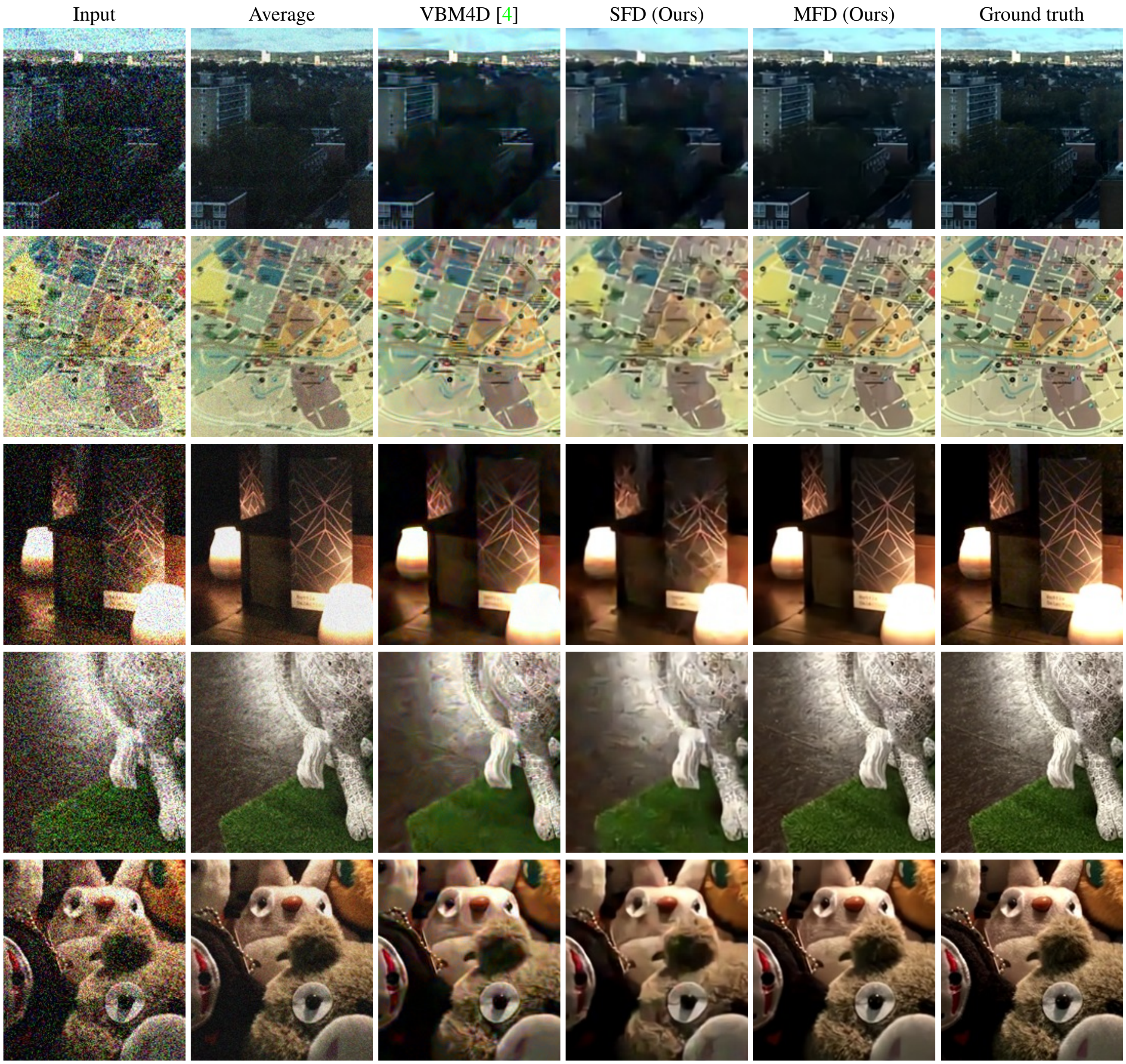

Figure 1: **Multiframe Gaussian denoising on stabilized Live Photo test data with $\sigma = 50$.** Our multiframe denoising approach is able to resolve more details than other methods.

5. Additional FlexISP results

In Figure 4 we show additional results of our denoiser on a real burst from the FlexISP dataset [2].

References

[1] A. Foi. Clipped noisy images: Heteroskedastic modeling and practical denoising. *Signal Processing*, 89(12):2609–2629, 2009. 1

[2] F. Heide, M. Steinberger, Y.-T. Tsai, M. Rouf, D. Pajak, D. Reddy, O. Gallo, J. Liu, W. Heidrich, K. Egiazarian, et al. Flexisp: A flexible camera image processing framework. *ACM Transactions on Graphics (TOG)*, 33(6):231, 2014. 2, 5

[3] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International Conference on Machine Learning*, pages 448–456, 2015. 1

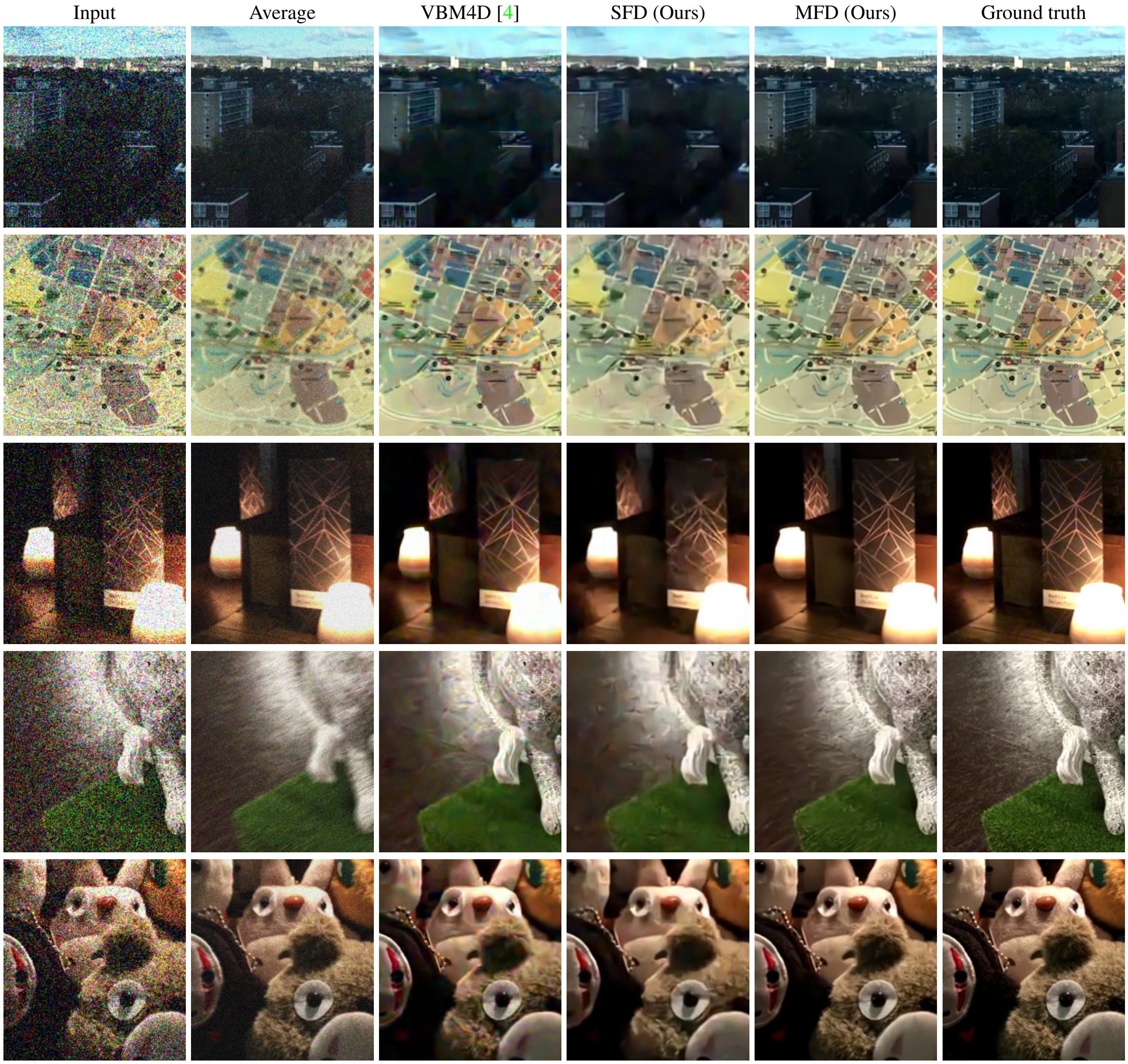

Figure 2: **Multiframe Gaussian denoising on unstabilized Live Photo test data with $\sigma = 50$.** We can see that the average image can be very blurry and that only our multiframe denoising approach is able to resolve details.

[4] M. Maggioni, G. Boracchi, A. Foi, and K. O. Egiazarian. Video denoising using separable 4d nonlocal spatiotemporal transforms. In *Image Processing: Algorithms and Systems*, page 787003, 2011. 2, 3

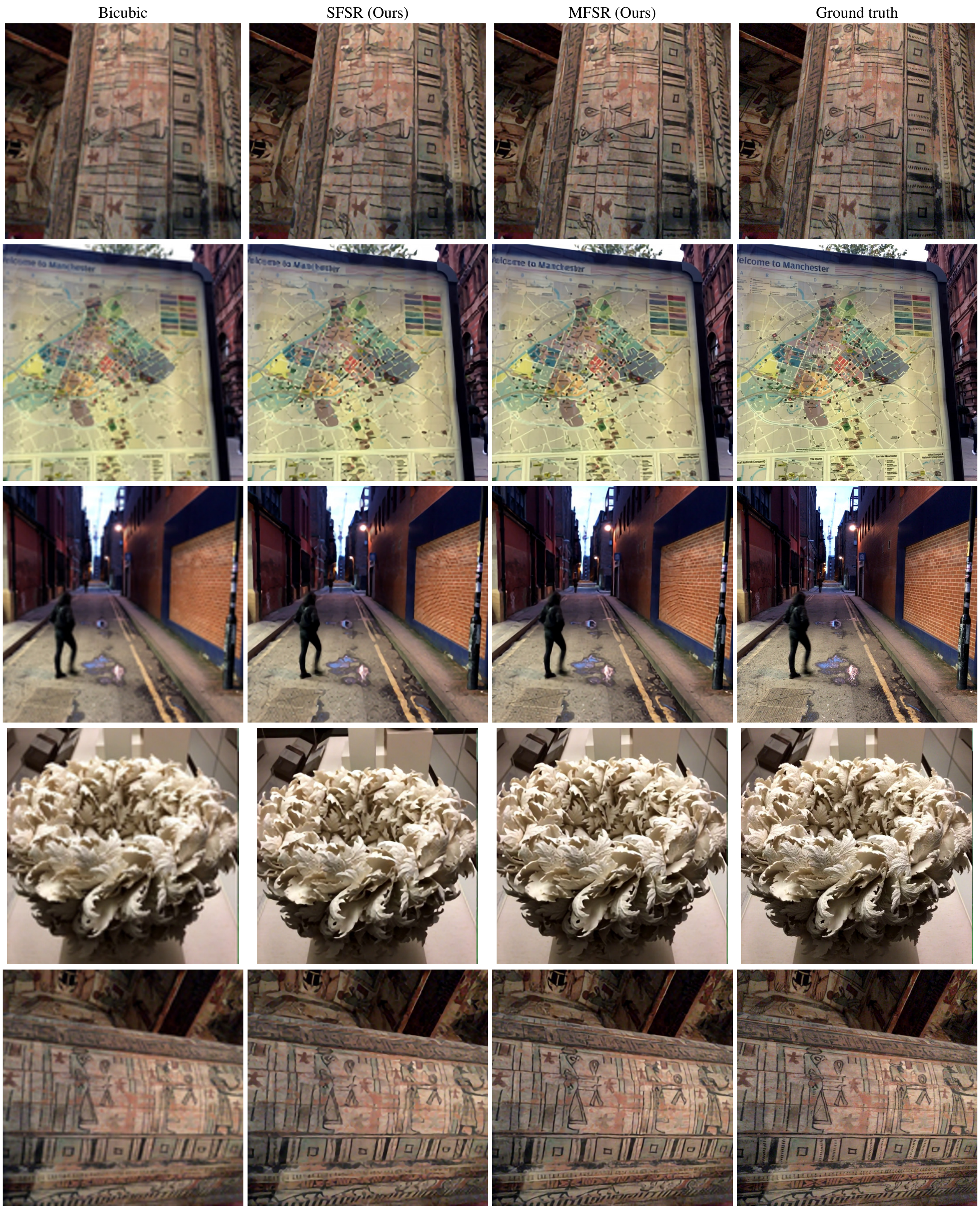

Figure 3: **Multiframe** $4\times$ **super-resolution on stabilized Live Photo test data.** While our single frame model achieves a good upsampling, the increase in sharpness from our multiframe approach brings a significant quality improvement.

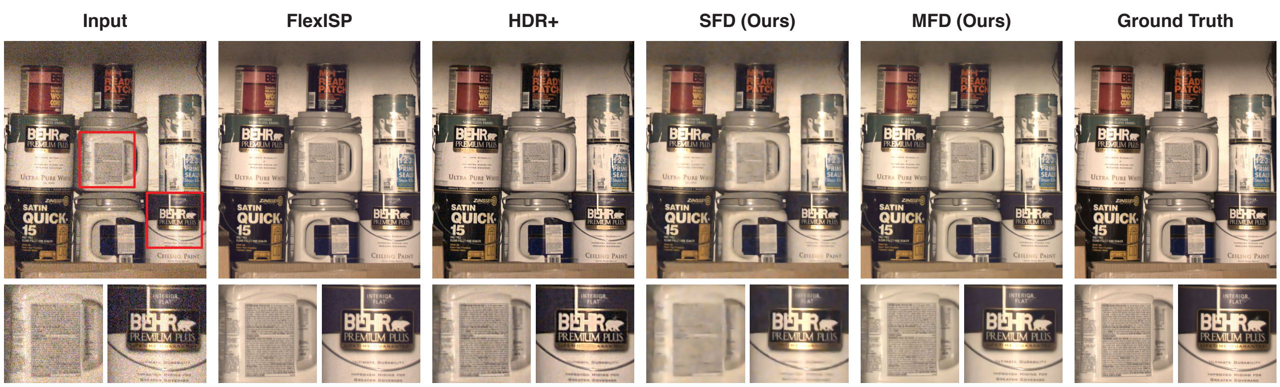

Figure 4: **Denoising results on a real and burst from the FlexISP dataset** [2].