

Deep-inverse correlography: towards real-time high-resolution non-line-of-sight imaging: supplementary material

CHRISTOPHER A. METZLER^{1,4,*}, FELIX HEIDE², PRASANA RANGARAJAN³, MURALIDHAR MADABHUSHI BALAJI³, APARNA VISWANATH³, ASHOK VEERARAGHAVAN⁴, AND RICHARD G. BARANIUK⁴

¹Department of Electrical Engineering, Stanford University, 450 Serra Mall, Stanford, CA 94305, USA

²Department of Computer Science, Princeton University, Princeton, NJ 08544, USA

³Department of Electrical Engineering, Southern Methodist University, 6425 Boaz Lane, Dallas, TX 75205, USA

⁴Department of Electrical and Computer Engineering, Rice University, 6100 Main Street, Houston, TX 77005, USA

*Corresponding author: cmetzler@stanford.edu

Published 16 January 2020

This document provides supplementary information to “Deep-inverse correlography: towards real-time high-resolution non-line-of-sight imaging,” <https://doi.org/10.1364/optica.374026>. It includes a derivation the autocorrelation estimate, a comparison of various training losses, a comparison between correlography and previous correlation-imaging methods, a radiometric analysis of the experimental NLoS setup, and additional experimental results.

1. DERIVATION OF AUTOCORRELATION ESTIMATE

In this supplement, we derive Eq. (8) and Eq. (9) from the paper, which relate the autocorrelation of the hidden object’s albedo with the measurements I_1, I_2, \dots, I_n that we capture with the camera. To keep notation succinct, we denote the field propagating from the object, $E_{O_{out}}$, using the variable f and index the hidden object’s spatial domain using x rather than x_O . Recall from Section 3 that $E_{O_{out}}$ follows a circular Gaussian distribution with autocorrelation function $\sigma^2 r \delta(\Delta x)$. Without loss of generality, we assume $\sigma^2 = 1$.

This analysis is almost identical to that of standard, line-of-sight correlography [1].

A. Useful Properties

The following two properties will prove useful in our derivations.

- The power spectral density of a signal is equal to the Fourier transform of its autocorrelation.

$$|F(u)|^2 = \overline{F(u)} \circ F(u) = F(u_{\text{time reversed}} * u) = F(u \star u) \quad (\text{S1})$$

where $\overline{(\cdot)}$ denotes complex conjugation, \circ denotes a Hadamard (elementwise) product, $*$ denotes convolution, and \star denotes autocorrelation.

- If random variables G_1, G_2, G_3 , and G_4 are jointly Gaussian, we

have that

$$\begin{aligned} \mathbb{E}[G_1 G_2 G_3 G_4] &= \mathbb{E}[G_1 G_2] \mathbb{E}[G_3 G_4] + \\ &\quad \mathbb{E}[G_1 G_3] \mathbb{E}[G_2 G_4] + \mathbb{E}[G_1 G_4] \mathbb{E}[G_2 G_3]. \end{aligned} \quad (\text{S2})$$

B. Proof of Eq. (8)

In this appendix, we derive that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N |F^{-1}(I_n)|^2(\Delta x) = r \star r(\Delta x) + \delta(\Delta x) \left[\int_{x_1=-\infty}^{\infty} r(x_1) dx_1 \right]^2.$$

Following the law of large numbers

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N |F^{-1}(I_n)|^2 &= \mathbb{E}[|F^{-1}(I)|^2] \\ &= \mathbb{E}[|F^{-1}(|F(f)|^2)|^2] \\ &= \mathbb{E}[|F^{-1}(F(f \star f))|^2] \\ &= \mathbb{E}[|(f \star f)|^2]. \end{aligned} \quad (\text{S3})$$

$$\begin{aligned}
& \mathbb{E}[|(f \star f)|^2](\Delta x) \\
&= \mathbb{E}[\overline{(f \star f)}(f \star f)](\Delta x) \\
&= \mathbb{E}\left[\left(\int_{x_1=-\infty}^{\infty} f(x_1 + \Delta x)\bar{f}(x_1)dx_1\right)\left(\int_{x_2=-\infty}^{\infty} f(x_2 + \Delta x)\bar{f}(x_2)dx_2\right)\right] \\
&= \mathbb{E}\left[\int_{x_1=-\infty}^{\infty} \int_{x_2=-\infty}^{\infty} \bar{f}(x_1 + \Delta x)f(x_1)f(x_2 + \Delta x)\bar{f}(x_2)dx_1dx_2\right].
\end{aligned} \tag{S4}$$

Integration is linear which implies the following.

$$\begin{aligned}
& \mathbb{E}[|(f \star f)|^2](\Delta x) \\
&= \int_{x_1=-\infty}^{\infty} \int_{x_2=-\infty}^{\infty} \mathbb{E}[\bar{f}(x_1 + \Delta x)f(x_1)f(x_2 + \Delta x)\bar{f}(x_2)]dx_1dx_2.
\end{aligned} \tag{S5}$$

Taking advantage of Eq. (S2), we have

$$\begin{aligned}
& \mathbb{E}[|(f \star f)|^2](\Delta x) \\
&= \int_{x_1=-\infty}^{\infty} \int_{x_2=-\infty}^{\infty} \mathbb{E}[\bar{f}(x_1 + \Delta x)f(x_1)]\mathbb{E}[f(x_2 + \Delta x)\bar{f}(x_2)]dx_1dx_2 \\
&+ \int_{x_1=-\infty}^{\infty} \int_{x_2=-\infty}^{\infty} \mathbb{E}[\bar{f}(x_1 + \Delta x)f(x_2 + \Delta x)]\mathbb{E}[f(x_1)\bar{f}(x_2)]dx_1dx_2 \\
&+ \int_{x_1=-\infty}^{\infty} \int_{x_2=-\infty}^{\infty} \mathbb{E}[\bar{f}(x_1 + \Delta x)\bar{f}(x_2)]\mathbb{E}[f(x_1)f(x_2 + \Delta x)]dx_1dx_2.
\end{aligned} \tag{S6}$$

Assuming that the speckle is fully realized (the autocorrelation function of f is a Dirac delta) we have

$$\begin{aligned}
& \mathbb{E}[\bar{f}(x_1 + \Delta x)f(x_1)] = r(x_1)\delta(\Delta x) \\
& \mathbb{E}[f(x_2 + \Delta x)\bar{f}(x_2)] = r(x_2)\delta(\Delta x) \\
& \mathbb{E}[\bar{f}(x_1 + \Delta x)f(x_2 + \Delta x)] = r(x_1 + \Delta x)\delta(x_2 - x_1) \\
& \mathbb{E}[f(x_1)\bar{f}(x_2)] = r(x_1)\delta(x_2 - x_1) \\
& \mathbb{E}[\bar{f}(x_1 + \Delta x)\bar{f}(x_2)] = \mathbb{E}[\bar{f}(x_2)^2]\delta(x + x_1 - x_2) \\
& \mathbb{E}[f(x_1)f(x_2 + \Delta x)] = \mathbb{E}[f(x_1)^2]\delta(x + x_2 - x_1).
\end{aligned}$$

Note that, because the real and imaginary parts of f are independent, have zero mean, and have equal variances, we have that

$$\begin{aligned}
& \mathbb{E}[\bar{f}(x_2)^2] \\
&= \mathbb{E}[\mathcal{R}(f(x_2))^2 + 2i\mathcal{R}(f(x_2))\mathcal{I}(f(x_2)) - \mathcal{I}(f(x_2))^2] = 0
\end{aligned} \tag{S7}$$

and

$$\mathbb{E}[f(x_1)^2] = \mathbb{E}[\mathcal{R}(f(x_1))^2 + 2i\mathcal{R}(f(x_1))\mathcal{I}(f(x_1)) - \mathcal{I}(f(x_1))^2] = 0$$

where $\mathcal{R}(\cdot)$ and $\mathcal{I}(\cdot)$ take the real and imaginary parts of their respective arguments.

Substituting the above equalities into Eq. (S6), we have that

$$\begin{aligned}
& \mathbb{E}[|(f \star f)|^2](\Delta x) \\
&= \int_{x_1=-\infty}^{\infty} \int_{x_2=-\infty}^{\infty} r(x_1)\delta(\Delta x)r(x_2)\delta(\Delta x)dx_1dx_2 \\
&+ \int_{x_1=-\infty}^{\infty} \int_{x_2=-\infty}^{\infty} r(x_1 + \Delta x)\delta(x_2 - x_1)r(x_1)\delta(x_2 - x_1)dx_1dx_2 \\
&= \delta(\Delta x) \int_{x_1=-\infty}^{\infty} \int_{x_2=-\infty}^{\infty} r(x_1)r(x_2)dx_1dx_2 \\
&+ \int_{x_1=-\infty}^{\infty} r(x_1 + \Delta x)r(x_1)dx_1 \\
&= \delta(\Delta x) \left[\int_{x_1=-\infty}^{\infty} r(x_1)dx_1 \right]^2 + r \star r(\Delta x).
\end{aligned}$$

C. Proof of Eq. (9)

In this appendix, we derive that

$$\lim_{N \rightarrow \infty} \left| \frac{1}{N} \sum_{n=1}^N F^{-1}(I_n) \right|^2(\Delta x) = \delta(\Delta x) \left[\int_{x_1=-\infty}^{\infty} r(x_1)dx_1 \right]^2.$$

Following the law of large numbers

$$\begin{aligned}
\lim_{N \rightarrow \infty} \left| \frac{1}{N} \sum_{n=1}^N F^{-1}(I_n) \right|^2 &= |\mathbb{E}[F^{-1}(I)]|^2 \\
&= |\mathbb{E}[F^{-1}(|F(f)|^2)]|^2 \\
&= |\mathbb{E}[F^{-1}(F(f \star f))]|^2 \\
&= |\mathbb{E}[f \star f]|^2.
\end{aligned}$$

$$\begin{aligned}
|\mathbb{E}[f \star f]|^2 &= |\mathbb{E} \int_{x_1=-\infty}^{\infty} f(x_1 + \Delta x)\bar{f}(x_1)dx_1|^2 \\
&= \left| \int_{x_1=-\infty}^{\infty} \mathbb{E}[f(x_1 + \Delta x)\bar{f}(x_1)]dx_1 \right|^2 \\
&= \left| \int_{x_1=-\infty}^{\infty} r(x_1)\delta(\Delta x)dx_1 \right|^2 \\
&= \delta(\Delta x) \left| \int_{x_1=-\infty}^{\infty} r(x_1)dx_1 \right|^2 \\
&= \delta(\Delta x) \left[\int_{x_1=-\infty}^{\infty} r(x_1)dx_1 \right]^2.
\end{aligned} \tag{S8}$$

The last of the above equalities holds because the albedo r is non-negative.

2. COMPARISON OF TRAINING LOSSES

Figure S1 compares the training losses associated with the $\|\hat{r} - r\|_1$, $\|\hat{r} - r\|_2$, $\|\hat{r} \star \hat{r} - r \star r\|_1$, $\|\hat{r} \star \hat{r} - r \star r\|_2$, $\| |\mathcal{F}\hat{r}| - |\mathcal{F}r| \|_1$, and $\| |\mathcal{F}\hat{r}| - |\mathcal{F}r| \|_2$ losses. Surprisingly, the $\|\hat{r} - r\|_1$ and $\|\hat{r} - r\|_2$ losses can be minimized; the network can memorize the locations of the training data. Among the translation-invariant losses, the $\|\hat{r} \star \hat{r} - r \star r\|_1$ loss converges the fastest.

Figure S2 compares reconstructions produced by the various networks on unseen experimental test data of a hidden ‘‘t’’. Although they can minimize the training loss, the $\|\hat{r} - r\|_1$ and $\|\hat{r} - r\|_2$ networks do not generalize to unseen data. In contrast, all four networks trained with translation-invariant losses produce reasonable reconstructions.

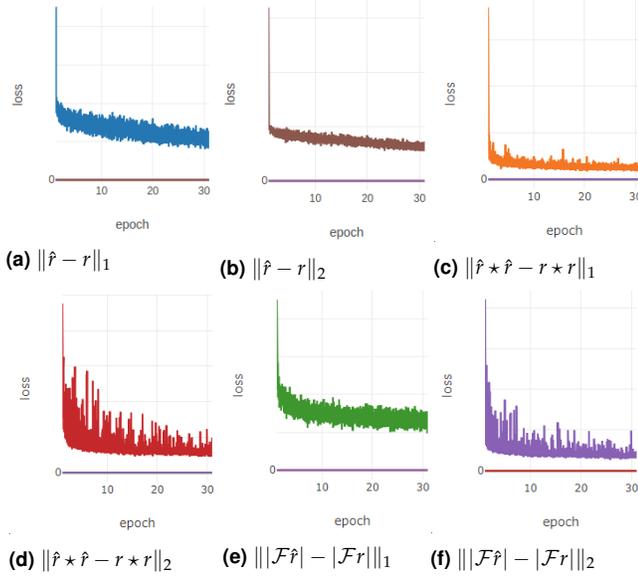


Fig. S1. Training loss versus number of training epochs with various loss functions. An ℓ_1 loss with respect to the autocorrelation converged the fastest.

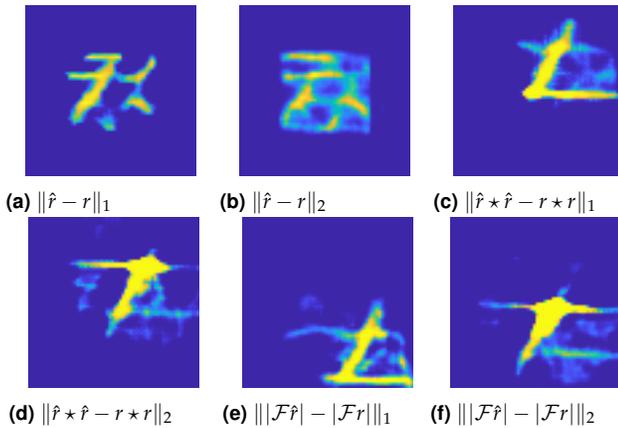


Fig. S2. Reconstructions of a hidden “t” with various training losses. All the networks trained with translation-invariant losses eventually produce reasonable results.

3. CORRELOGRAPHY VS. CORRELATION IMAGING

At first blush, NLoS correlography seems very similar to correlation imaging [2]. Both reconstruct a hidden object by estimating its autocorrelation from speckle reflected off a virtual source. However, they have one consequential difference; correlation imaging is based on temporally coherent, spatially incoherent illumination whereas NLoS correlography relies upon temporally and spatially coherent illumination. Both methods are temporally coherent, and so will form speckle. However, under the former model speckle intensities add while under the latter model speckle fields add. This has several distinct advantages, the most important of which are increased contrast and robustness to ambient illumination.

In this subsection, we attempt to recover the hidden object’s autocorrelation using spatially incoherent illumination, i.e., the technique used in [2]. Instead of using a rotating diffusers, we destroy the temporal coherence of our light source by quickly sweeping the laser’s focus across the wall.

The estimated autocorrelation is shown in Figure S4. Unlike the estimate formed using a spatially coherent light source, the estimate from the incoherent light source presents no discernible features. Attempts at reconstructing r from this estimate of $r * r$ fail.

4. RADIOMETRIC ANALYSIS

Let ρ_{VS} , ρ_{obj} , and ρ_{VD} denote the albedos of the virtual source, hidden object, and virtual detector respectively. Let \vec{n}_{VS} , \vec{n}_{obj} , and \vec{n}_{VD} denote the surface normals of the virtual source, hidden object, and virtual detector respectively. Let \vec{SO} , \vec{OD} , and \vec{DC} denote the vectors from the virtual source to the hidden object, the hidden object to the virtual detector, and the virtual detector to the camera, respectively. Let A_{VS} , A_{obj} , and A_{pix} denote the surface areas of the virtual source, the hidden object, and a single pixel on the camera. Let D_{Ap} denote the camera’s aperture diameter. Let N denote the number of pixels at the detector. Let m denote magnification between the virtual detector and the camera.

The radiometric throughput of our NLoS imaging system, defined as the flux at the camera divided by the input power is given by

$$\text{Throughput} = \left(\frac{\rho_{VS}\rho_{obj}\rho_{VD}}{\pi^3} \right) \times \quad (\text{S9})$$

$$\left(\|\vec{SO}\|^{-2} \|\vec{DC}\|^{-2} \|\vec{OD}\|^{-2} \right) \times \quad (\text{S10})$$

$$\left(|\vec{n}_{VS}^t \vec{SO}| \times |\vec{n}_{obj}^t \vec{SO}|^2 \times |\vec{n}_{VD}^t \vec{OD}| \right) \times \quad (\text{S11})$$

$$\left(\frac{1_{obj}}{A_{beam}} \times \frac{\pi}{4} D_{Ap}^2 \times N \times (m^{-2} A_{pix}) \right) \quad (\text{S12})$$

where line Eq. (S9) corresponds to the Albedo and Lambertian BRDF loss, line Eq. (S10) corresponds to the propagation loss, line Eq. (S11) corresponds to the foreshortening loss, and line Eq. (S12) corresponds to the area loss.

In our experimental setup, an image of which can be seen in Figure 8 of the main text, the hidden object is at 45° with respect to the virtual source’s surface normal, the virtual source is at 0° with respect to the hidden object’s surface normal, and the hidden object and camera are at roughly 70° and 0° with respect to the virtual detector’s surface normal. The virtual source is .5 meters from the hidden object, the hidden object is 1 meter from the virtual detector, and the virtual detector is .8 meters from the camera. Altogether, this results in a radiometric throughput of -182 dB.

5. TESTING MORE ALGORITHMS

In this section we test out the alternating minimization (Alt) [3], median truncated Wirtinger flow (MTWF) [4], truncated amplitude flow (TAF) [5], and truncated Wirtinger flow (TWF) [6] PR algorithms on both the simulated and experimental data from Section 6. We also test alternating direction method of multipliers with a total variation prior (ADMM-TV) [7]. All algorithms were initialized with HIO (spectral initializers, e.g., [8], fail with Fourier measurement operators) and used their default parameters.

Figures S5 and S6 demonstrate that the CNN-based method consistently produces the best results with both simulated and experimental data.

6. IMAGING WITH FEW SPECKLE IMAGES

In this section, we test how the algorithms handle finite-sample approximation error by estimating the hidden object’s autocorrelation using small subsets of the total detector area. These subsets and the associated noisy autocorrelation estimates are illustrated in Figure S7.

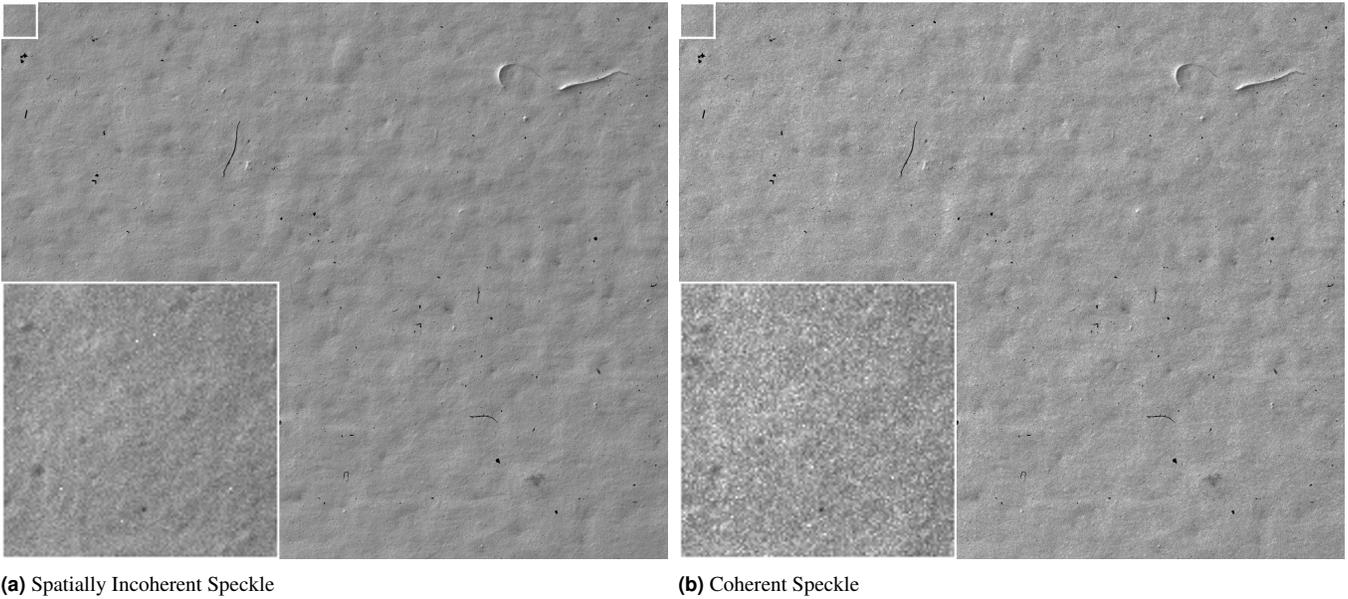
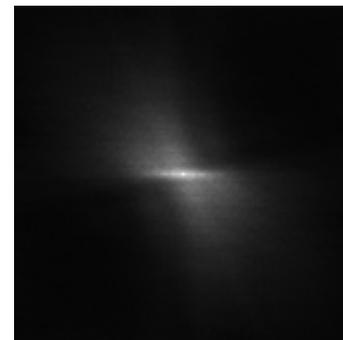


Fig. S3. Examples of a speckle images formed using spatially incoherent illumination (left) and spatially coherent illumination (right). The image captured using spatially coherent illumination has significantly more speckle contrast.

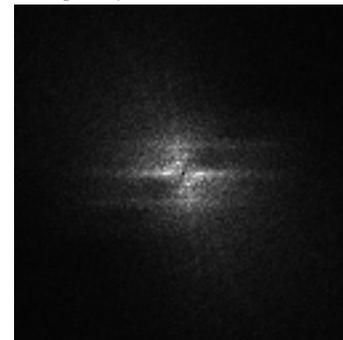
Recall that the finite-sample-approximation error presents as shot-noise-like, spatially-varying noise in the estimates of $\widehat{r \star r}$, not the Gaussian background noise that results from low flux measurements. Figure S8 demonstrates that, despite being trained on a very different noise distribution, the CNN-based method is still robust to this form of noise and offers improved reconstructions across all operating regimes.

REFERENCES

1. P. S. Idell, J. R. Fienup, and R. S. Goodman, “Image synthesis from nonimaged laser-speckle patterns,” *Opt. letters* **12**, 858–860 (1987).
2. O. Katz, P. Heidmann, M. Fink, and S. Gigan, “Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations,” *Nat. photonics* **8**, 784 (2014).
3. P. Netrapalli, P. Jain, and S. Sanghavi, “Phase retrieval using alternating minimization,” in *Advances in Neural Information Processing Systems*, (2013), pp. 2796–2804.
4. H. Zhang, Y. Chi, and Y. Liang, “Provable non-convex phase retrieval with outliers: Median truncated wirtinger flow,” in *International conference on machine learning*, (2016), pp. 1022–1031.
5. G. Wang, G. B. Giannakis, and Y. C. Eldar, “Solving systems of random quadratic equations via truncated amplitude flow,” *IEEE Transactions on Inf. Theory* **64**, 773–794 (2018).
6. Y. Chen and E. Candes, “Solving random quadratic systems of equations is nearly as easy as solving linear systems,” in *Advances in Neural Information Processing Systems*, (2015), pp. 739–747.
7. F. Heide, S. Diamond, M. Nießner, J. Ragan-Kelley, W. Heidrich, and G. Wetzstein, “Proximal: Efficient image optimization using proximal algorithms,” *ACM Transactions on Graph. (TOG)* **35**, 84 (2016).
8. E. J. Candes, X. Li, and M. Soltanolkotabi, “Phase retrieval via wirtinger flow: Theory and algorithms,” *IEEE Transactions on Inf. Theory* **61**, 1985–2007 (2015).



(a) Spatially Incoherent $r \star r$ Estimate



(b) Coherent $r \star r$ Estimate

Fig. S4. Autocorrelation estimates from incoherently illuminated hidden “F” (a) and coherently illuminated hidden “F” (b). The estimate formed with coherent illumination has distinct features that do not show up in the estimate formed with spatially incoherent illumination.

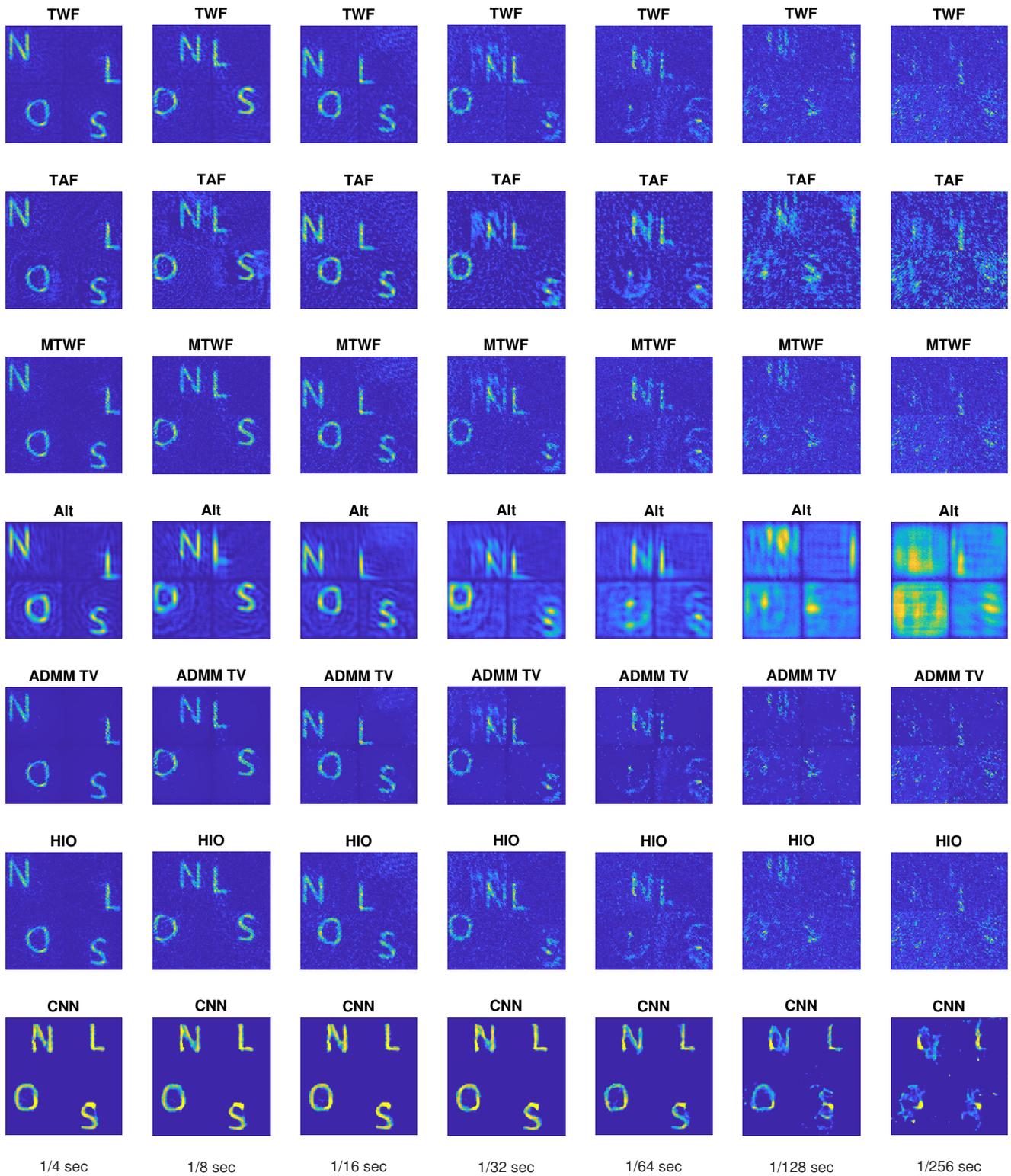


Fig. S5. Simulated results. Reconstructions from simulated short-exposure measurements. The CNN-based methods consistently produce the best reconstructions.

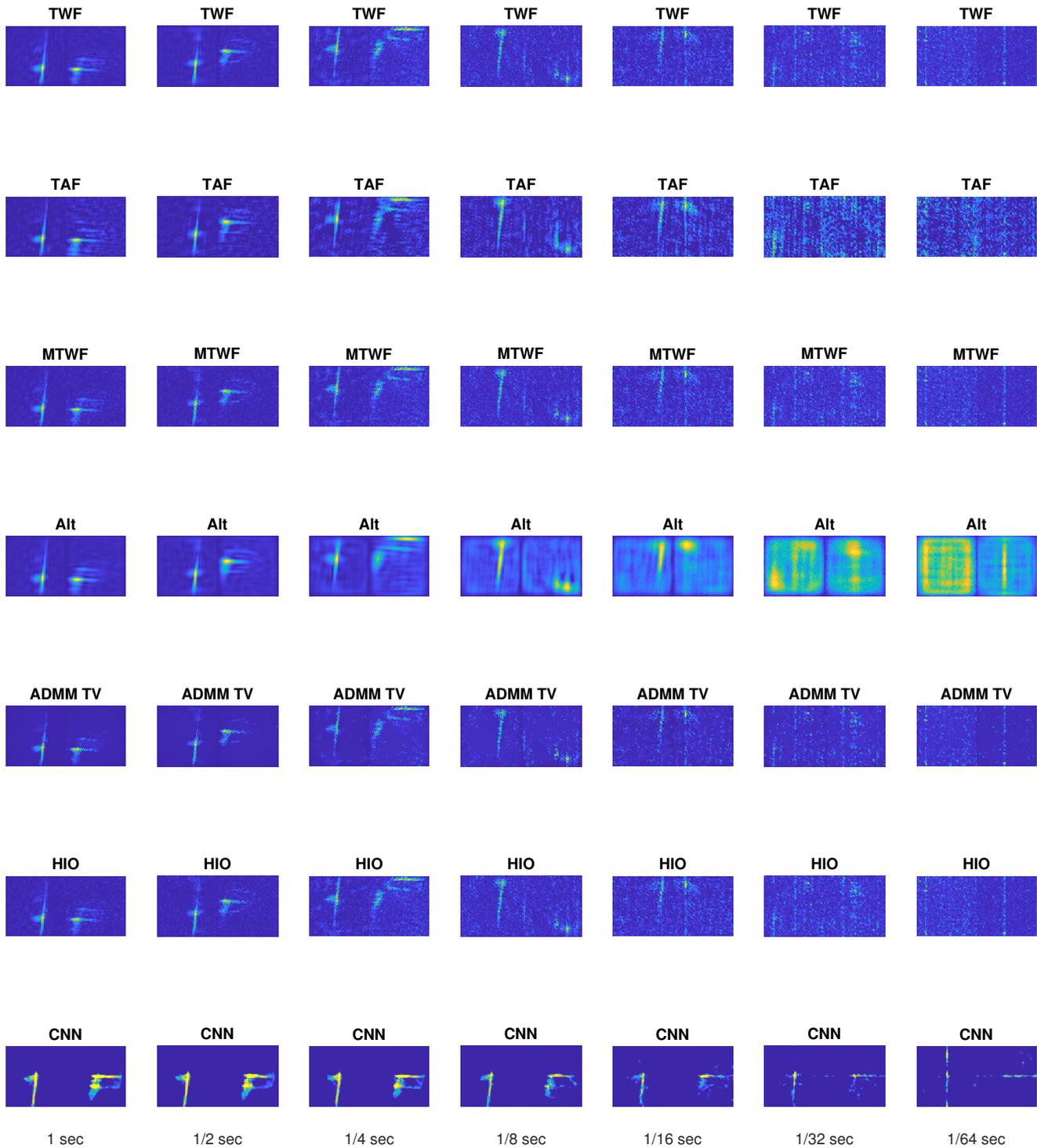


Fig. S6. Short-exposure experimental results. Reconstructions from experimental short-exposure measurements. The CNNs form recognizable reconstructions with less light.

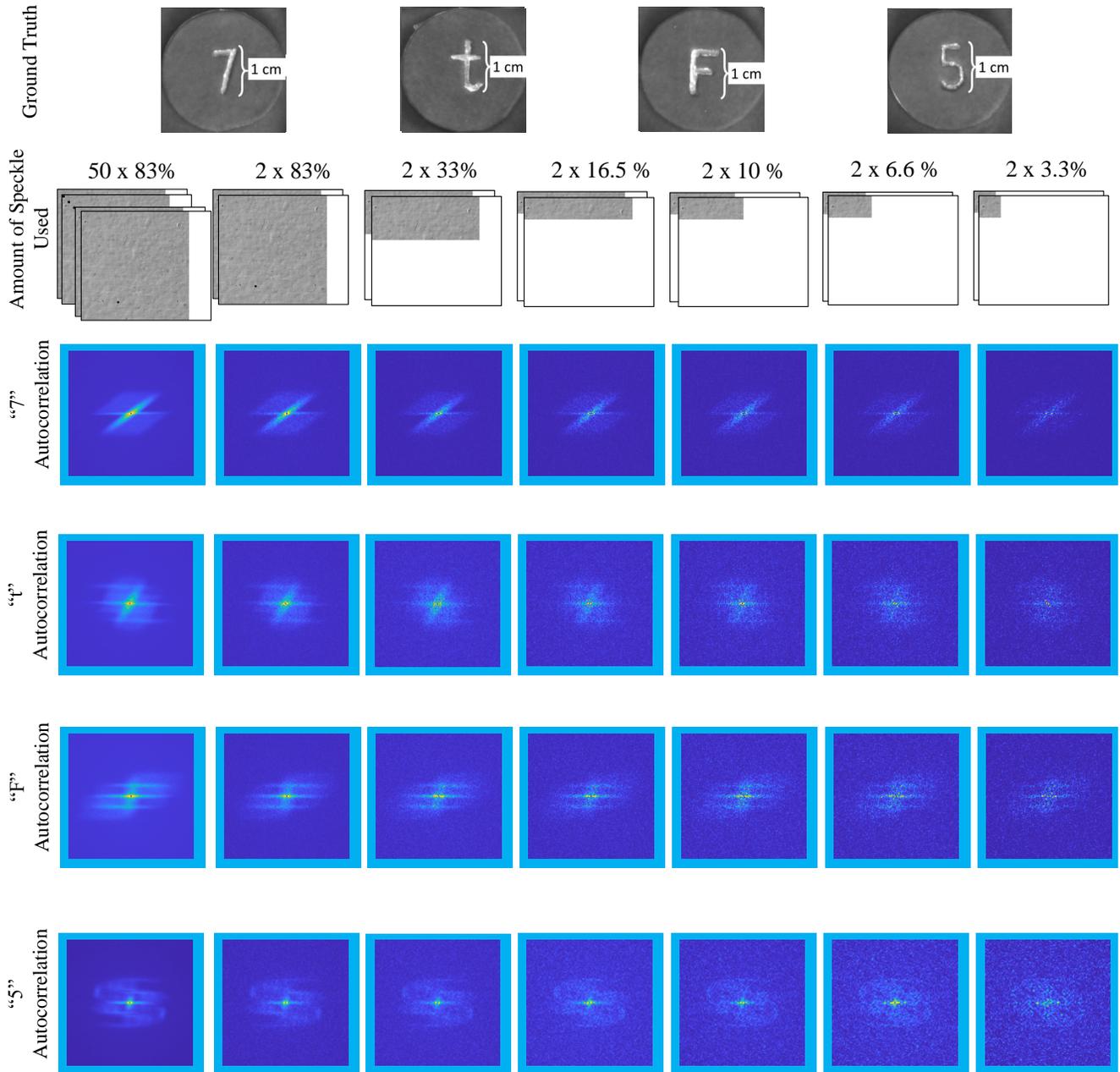


Fig. S7. Subsampled speckle regions. We estimated the autocorrelation from sub-sampled 1 second exposure-length speckle images. The resulting autocorrelation estimates display shot-noise-like finite-sample-approximation error.

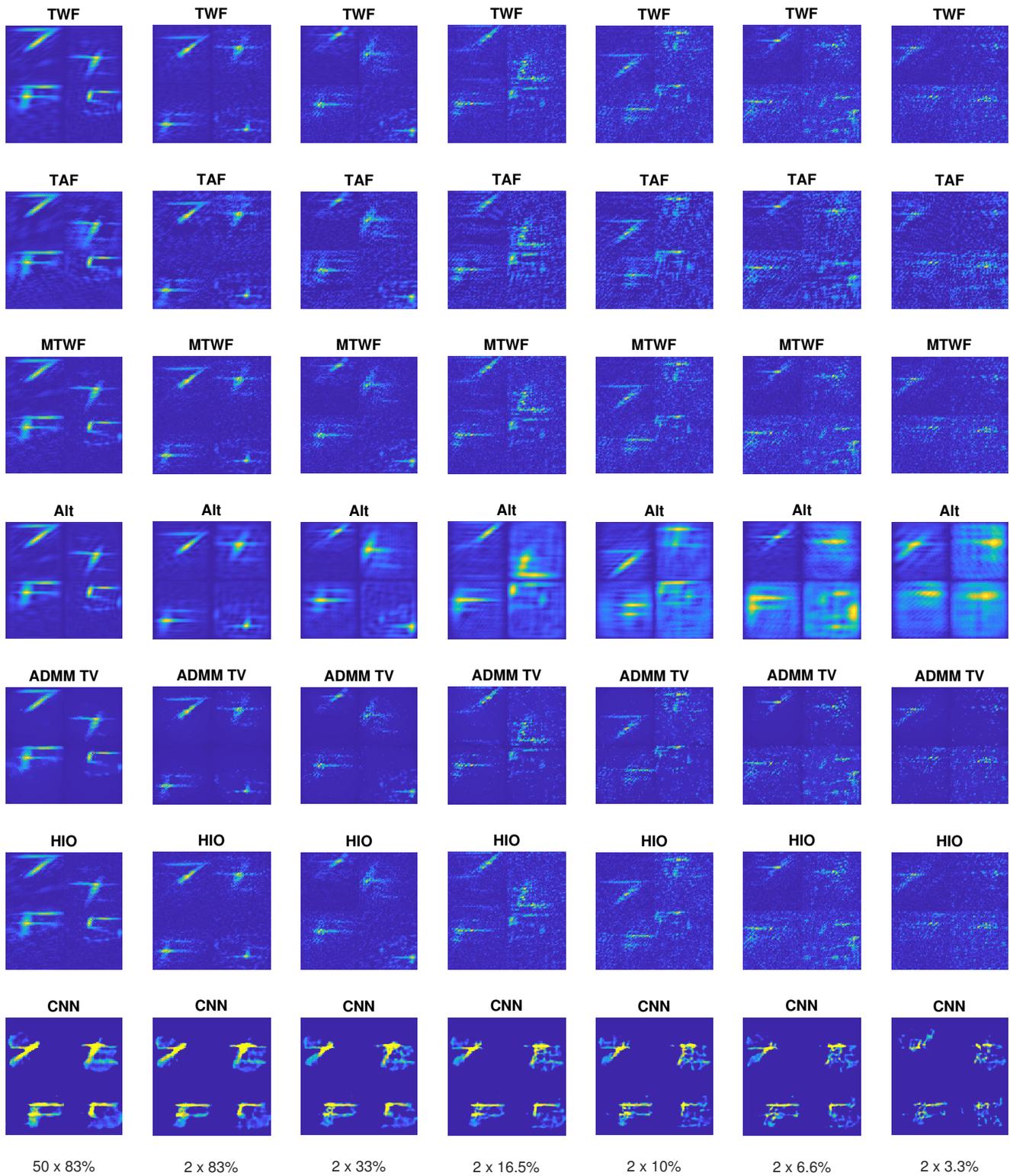


Fig. S8. Sub-sampled speckle experimental results. Reconstructions from subsets of experimental data, which are subject to finite-sample-approximation error. The CNN-based method can reconstruct using fewer measurements.