Efficient 2D Image Reconstruction with Modified CM²Net: Enhancing Fluorescence Microscopy for Large-Scale Biological Dynamics

Jamin Xie[1,2], Qianwan Yang[2], Lei Tian[2]

Valley Christian High School, 100 Skyway Dr, San Jose, CA 95111[1]
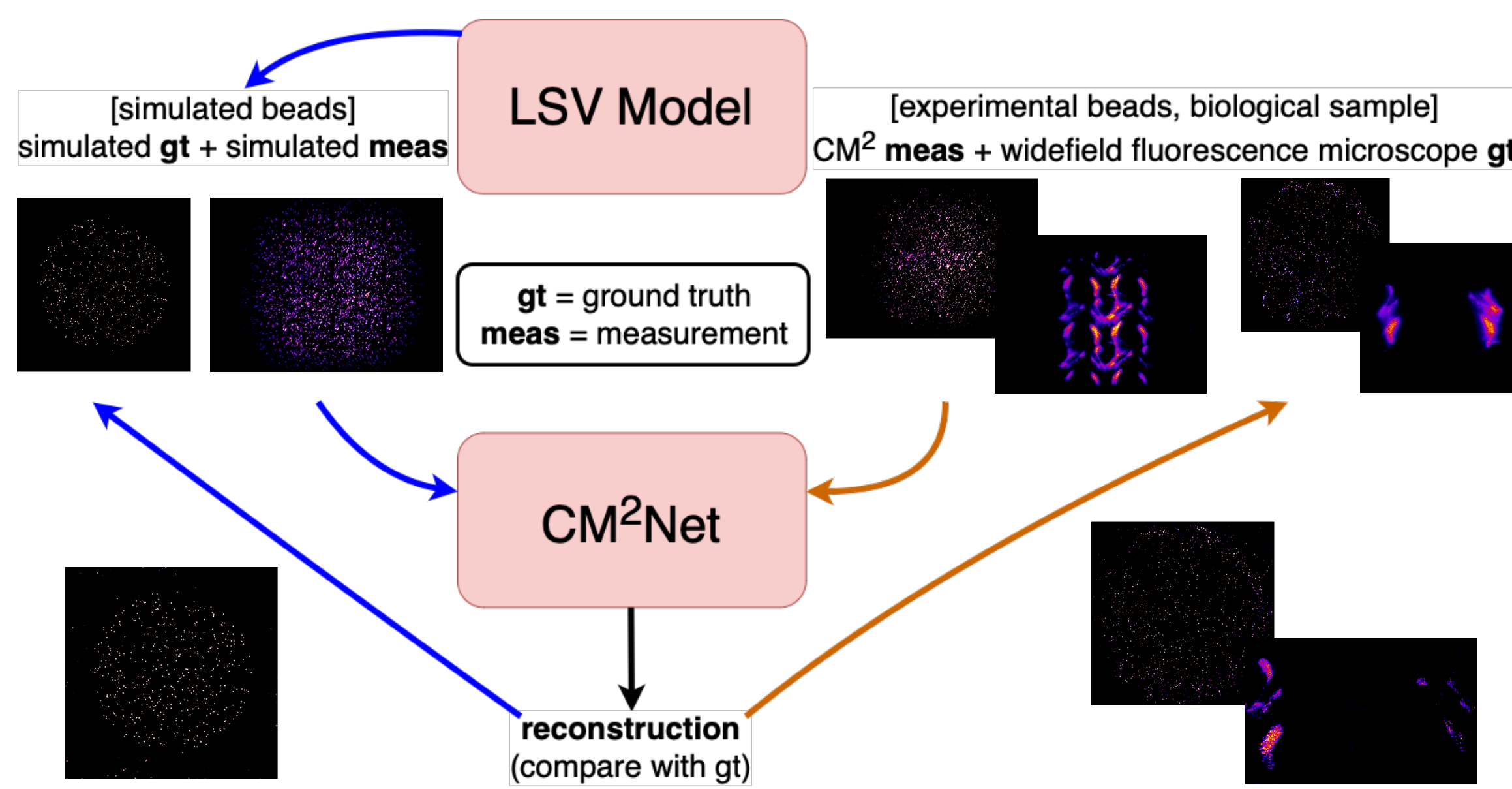Department of Electrical and Computer Engineering, Boston University, 8 St. Mary's Street, Boston, MA 02215[2]

Introduction

- Fluorescence microscopy is indispensable for studying biological structures
- Conventional optics suffers from the **trade-off between field of view (FOV), resolution, and miniaturization**
- Solution: combine miniature optics and advanced computational algorithms with **Computational Miniature Mesoscope (CM²)** and **CM²Net** for high-quality 3D reconstruction
- Linear shift-variant (LSV) model[1] characterizes realistic shift-variant point spread functions (PSFs) and generates large-scale fluorescent bead data
- Present **Modified CM²Net** for accurate and efficient **2D image reconstruction**

[simulated beads]
simulated gt + simulated meas

LSV Model

gt = ground truth
meas = measurement

CM²Net

[experimental beads, biological sample]
CM² meas + widefield fluorescence microscope gt

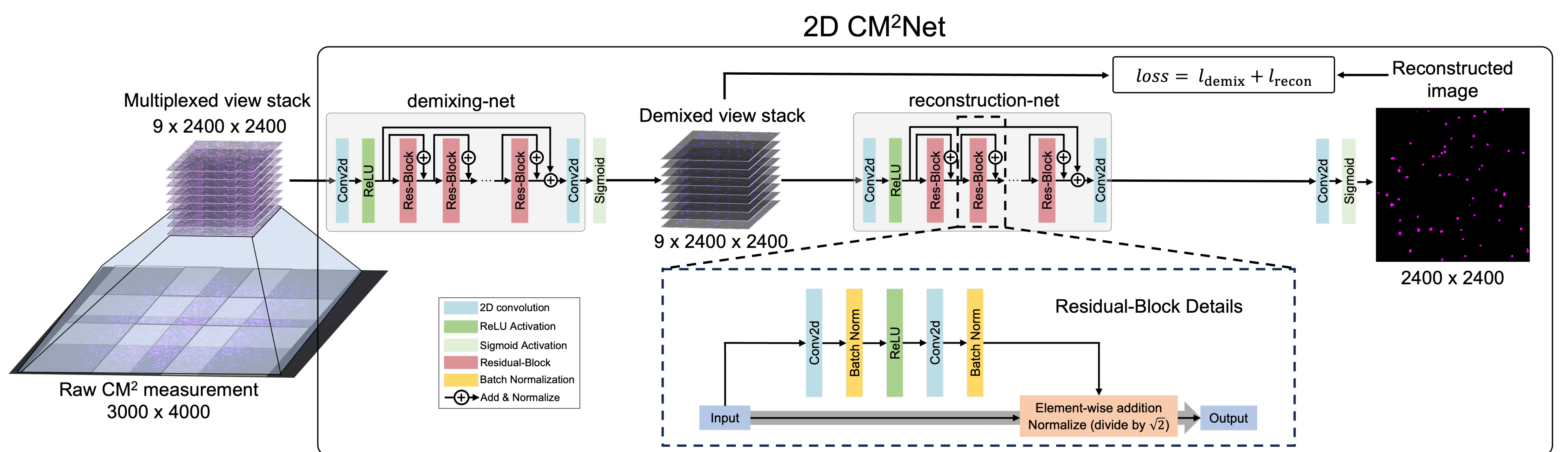reconstruction
(compare with gt)

Methods

CM² V2[2]
- 3x3 Microlens array (MLA) to acquire multiple views
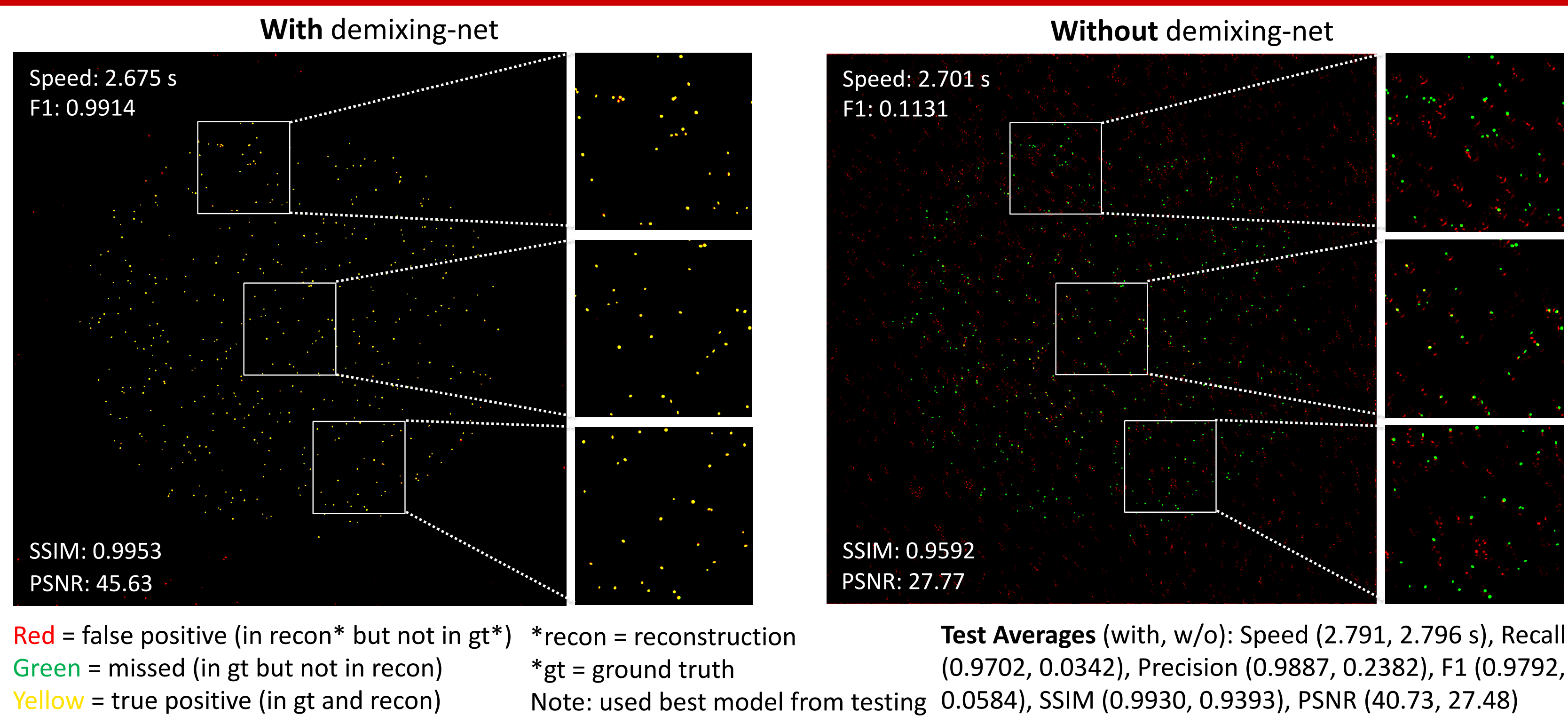- Results in strong multiplexing between neighboring views
- Enables large FOV
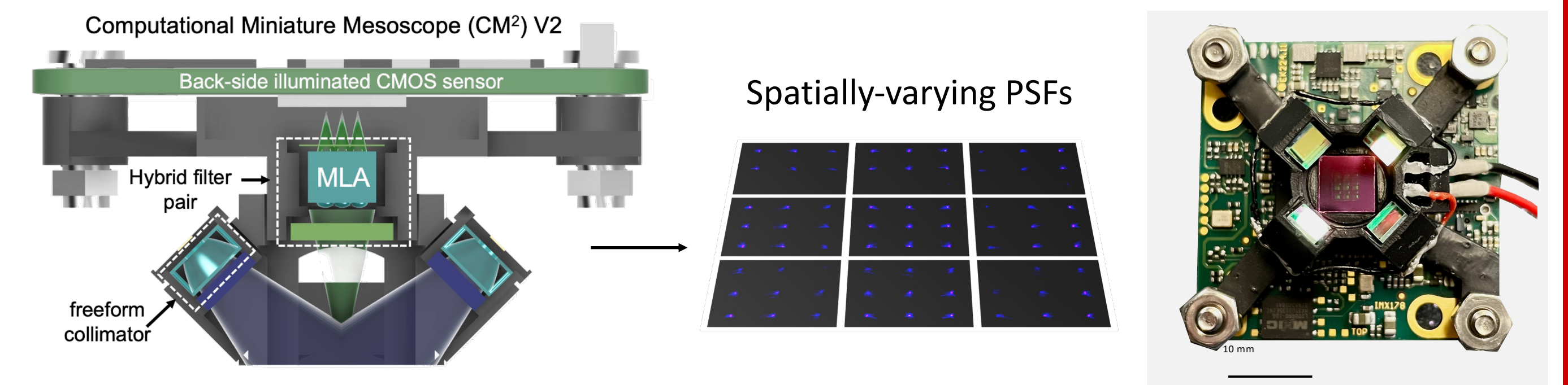
CM²Net (2D)
- 20 Res-Blocks per net: use normalization for variance stabilization to speed up training convergence[3]
- Training was conducted for for a max of 150 epochs or 48 hours
- Hyperparameter tuning

- Data Augmentation (training)
  - Random patch sampling (256x256): reduce computational complexity
  - Mixed Poisson-Gaussian noise

Computational Miniature Mesoscope (CM²) V2
Back-side illuminated CMOS sensor
Hybrid filter pair
MLA
freeform collimator

Spatially-varying PSFs

2D CM²Net
Multiplexed view stack 9 x 2400 x 2400
demixing-net
Demixed view stack 9 x 2400 x 2400
reconstruction-net
$loss = l_{demix} + l_{recon}$
Reconstructed image 2400 x 2400
Raw CM² measurement 3000 x 4000

2D convolution / ReLU Activation / Sigmoid Activation / Residual-Block / Batch Normalization / Add & Normalize

Residual-Block Details
Input → Conv2d → Batch Norm → ReLU → Conv2d → Batch Norm → Element-wise addition Normalize (divide by √2) → Output

Demixing-net Ablation Study

With demixing-net
Speed: 2.675 s
F1: 0.9914
SSIM: 0.9953
PSNR: 45.63

Without demixing-net
Speed: 2.701 s
F1: 0.1131
SSIM: 0.9592
PSNR: 27.77

Red = false positive (in recon* but not in gt*)
Green = missed (in gt but not in recon)
Yellow = true positive (in gt and recon)

*recon = reconstruction
*gt = ground truth

Test Averages (with, w/o): Speed (2.791 s, 2.796 s), Recall (0.9702, 0.0342), Precision (0.9887, 0.2382), F1 (0.9792, 0.0584), SSIM (0.9930, 0.9393), PSNR (40.73, 27.48)
Note: used best model from testing

Two-Sample T-test for Difference in Mean PSNR

Let μ be the population mean PSNR. α = 0.05
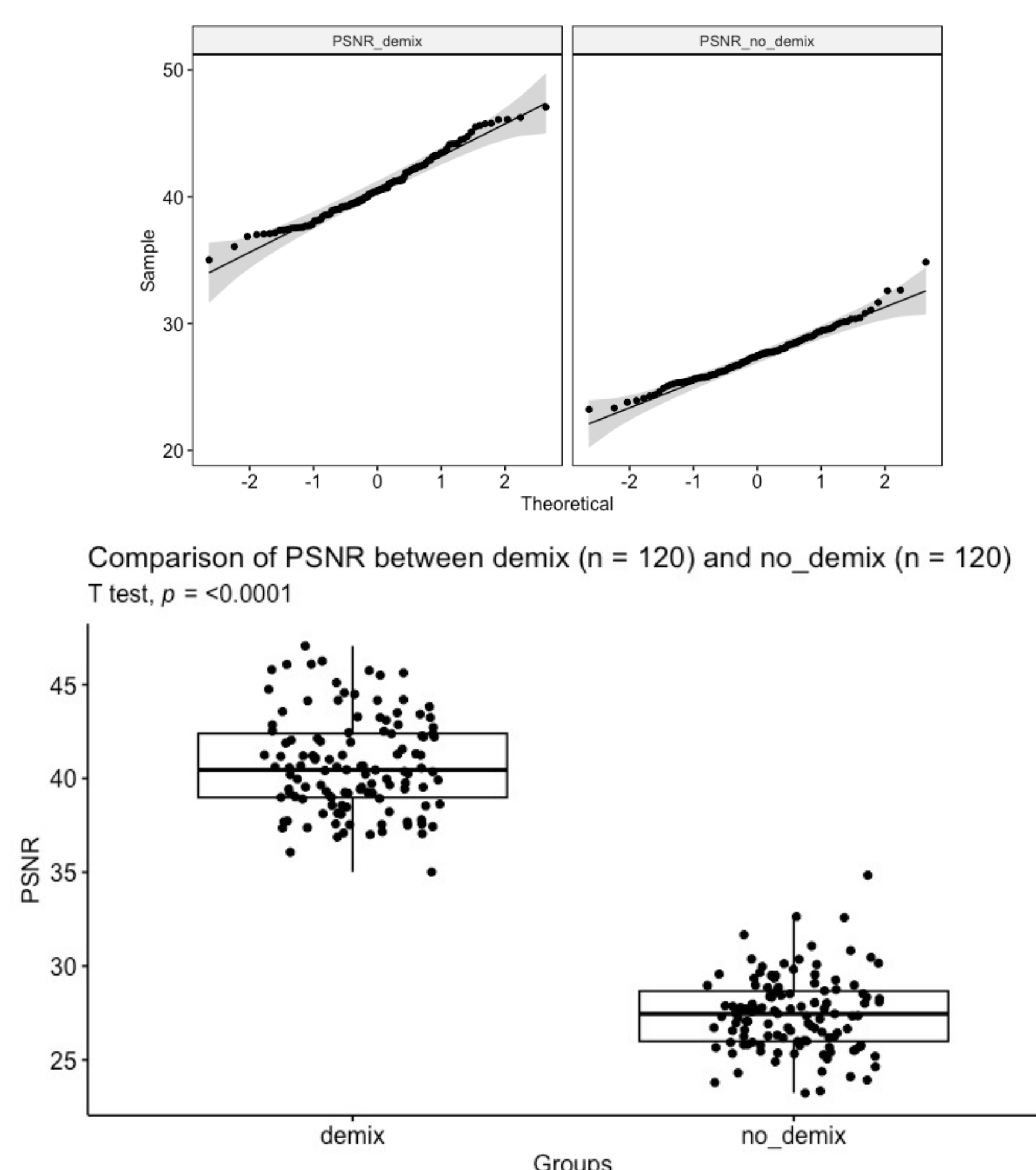
H_0: $\mu_{demix} = \mu_{no\_demix}$ H_a: $\mu_{demix} > \mu_{no\_demix}$
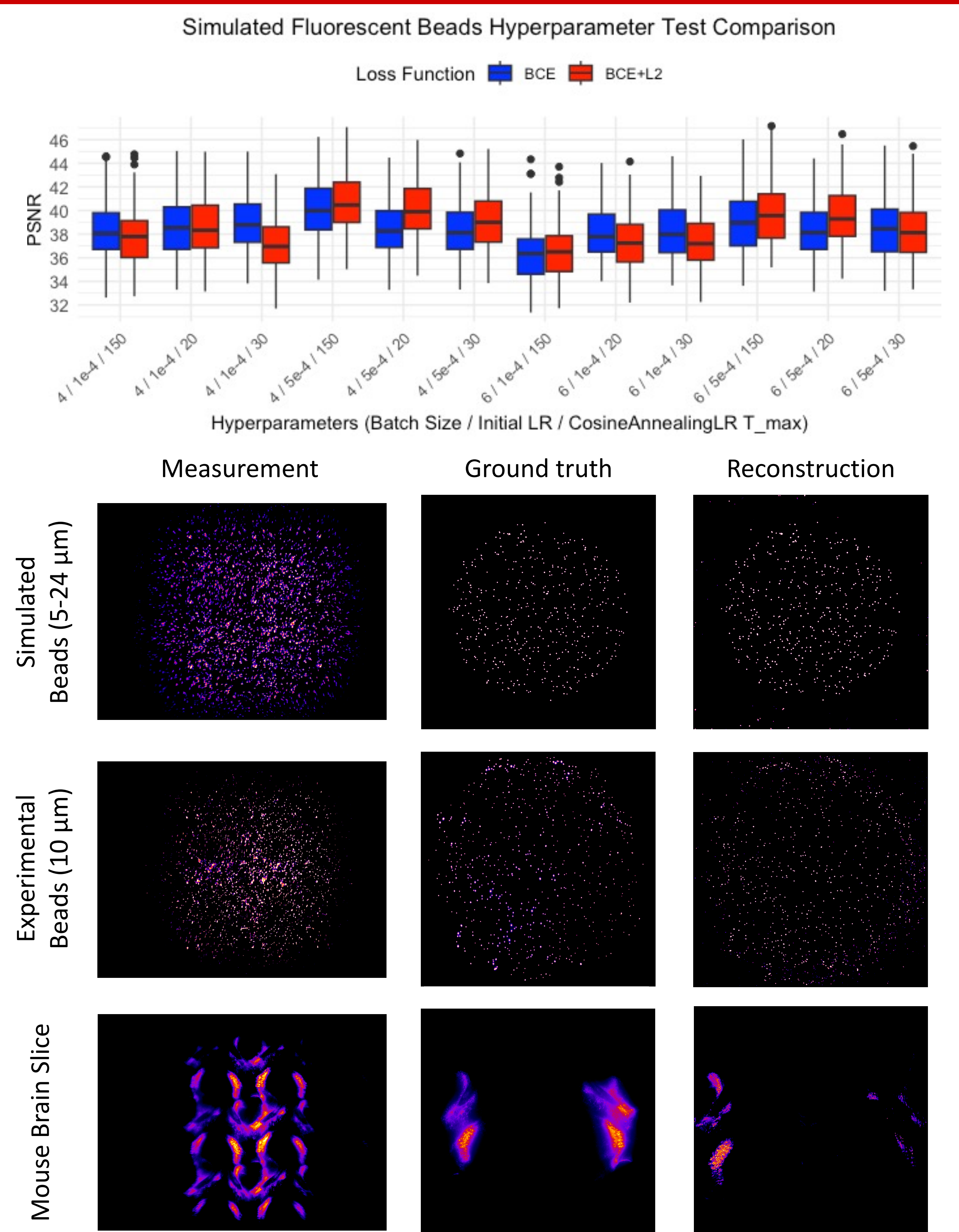
Assumptions:
- ✓ Independence of the observations
- ✓ No extreme outliers
- ✓ Normality, checking Shapiro-Wilk test and QQ plot (normal probability plot)
  - demix: p = 0.0392 < 0.05 -> check QQ plot approx. linear
  - no_demix: p = 0.0633 > 0.05 -> assume normality
- ✓ Homogeneity of variances not satisfied -> Welch's t-test
  - Levene's test: p = 0.0041 < 0.05 -> variances significantly different

Conclusion:
- Since Welch's t-test p < 0.0001 < 0.05, reject H_0. Sufficient statistical evidence that $\mu_{demix} > \mu_{no\_demix}$ i.e. demixing-net significantly improves model reconstruction.

PSNR_demix   PSNR_no_demix

Comparison of PSNR between demix (n = 120) and no_demix (n = 120)
T test, p = <0.0001

demix   no_demix
Groups

Results

Simulated Fluorescent Beads Hyperparameter Test Comparison
Loss Function: BCE / BCE+L2
PSNR
Hyperparameters (Batch Size / Initial LR / CosineAnnealingLR T_max)

Measurement / Ground truth / Reconstruction

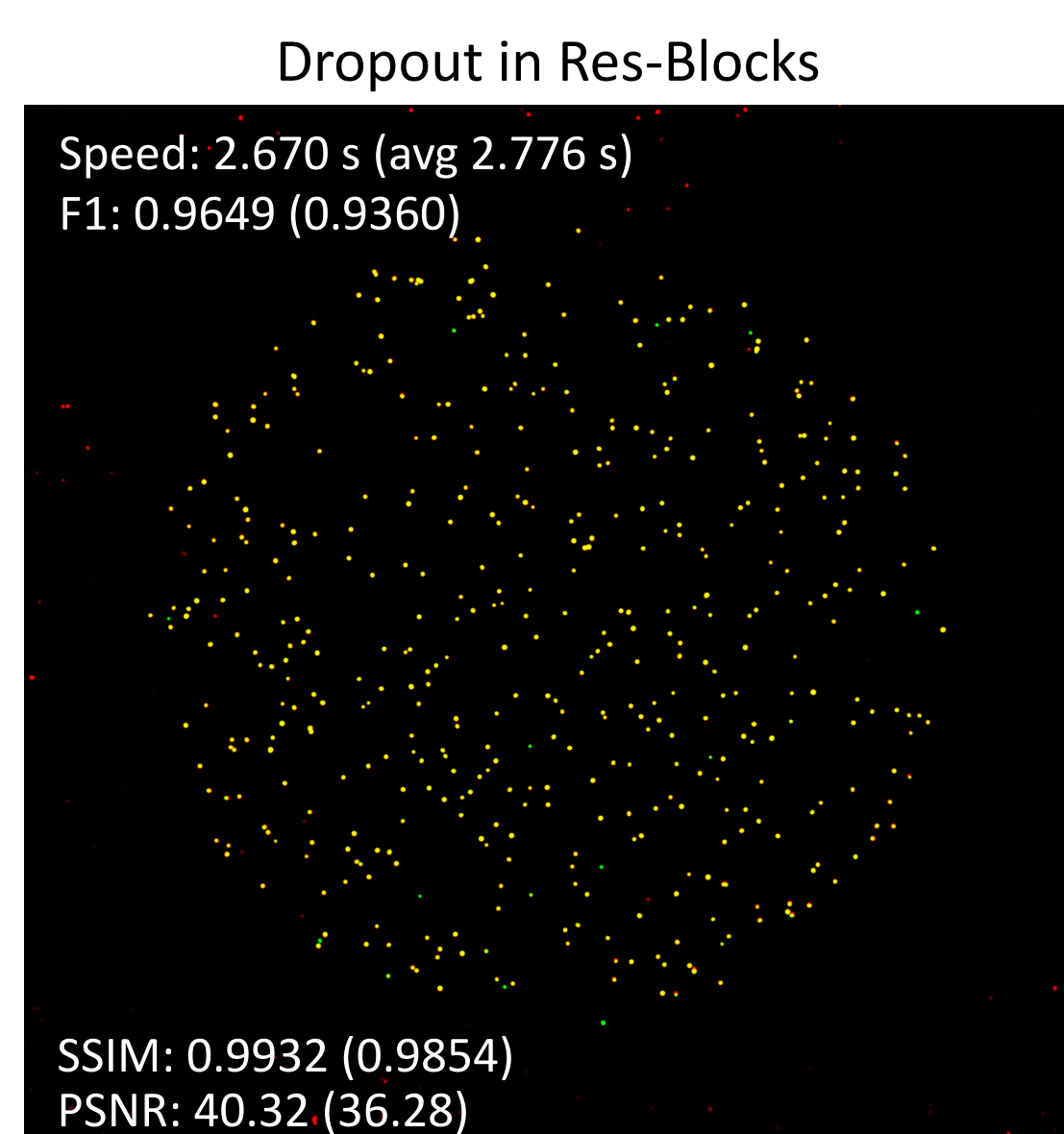Simulated Beads (5–24 μm)
Experimental Beads (10 μm)
Mouse Brain Slice

Discussion

- 2D CM²Net achieves robust, high-resolution, wide FOV 2D reconstruction
- LSV model generates large-scale realistic simulated fluorescent emitter data for model training
- Ablation study shows that the demixing-net module significantly improves the true positive rate, F1 score, SSIM, and PSNR by removing crosstalk artifacts between adjacent views
- Trained network performs exceptionally well on various unseen samples
- 2D CM²Net performs reasonably well on mouse brain slices, which has different statistics than beads

Future directions:
- Explore different network structures: dropout, Deformable ConvNet[4]
- Sparsity restrictions of simulated data must be accounted for by transfer learning on various biological samples
- Enhanced 2D CM²Net can be potentially used for recording large-scale biological dynamics

Dropout in Res-Blocks
Speed: 2.670 s (avg 2.776 s)
F1: 0.9649 (0.9360)
SSIM: 0.9932 (0.9854)
PSNR: 40.32 (36.28)

References

[1] Sroubek, F.; Kamenicky, J.; Lu, Y. M. Decomposition of Space-Variant Blur in Image Deconvolution. *IEEE Signal Processing Letters* **2016**, *23* (3), 346–350. DOI:10.1109/lsp.2016.2519764.

[2] Xue, Y.; Yang, Q.; Hu, G.; Guo, K.; Tian, L. Deep-Learning-Augmented Computational Miniature Mesoscope. *Optica* **2022**, *9* (9), 1009. DOI:10.1364/optica.464700.

[3] Alido, J.; Greene, J.; Xue, Y.; Hu, G.; Li, Y.; Monk, K. J.; DeBenedicts, B. T.; Davison, I. G.; Tian, L. Robust Single-Shot 3D Fluorescence Imaging in Scattering Media with a Simulator-Trained Neural Network. DOI:10.48550/arXiv.2303.12573.

[4] Zhu, X.; Hu, H.; Lin, S.; Dai, J. Deformable ConvNets V2: More Deformable, Better Results. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* **2019**. DOI:10.1109/cvpr.2019.00953.

Acknowledgements