Baikal: Unpaired Denoising of Fluorescence Microscopy Images using Diffusion Models

Shivesh Chaudhary[1*][0000−0002−1928−0933], Sivaramakrishnan
Sankarapandian[1*][0009−0003−7321−1087], Matt Sooknah[1][0009−0003−7145−4691],
Joy Pai[1][0009−0000−3529−6671], Caroline McCue[1][0000−0002−2304−7026], Zhenghao
Chen[1][0000−0002−9859−3005], and Jun Xu[1]

Calico Life Sciences LLC
http://www.calicolabs.com
1170 Veterans Blvd South San Francisco, CA 94080
{schaudhary, sivark, mds, joy, cmmccue, zhenghao, junxu}@calicolabs.com

Abstract. Fluorescence microscopy is an indispensable tool for biological discovery but image quality is constrained by desired spatial and temporal resolution, sample sensitivity, and other factors. Computational denoising methods can bypass imaging constraints and improve signal-to-noise ratio in images. However, current state of the art methods are commonly trained in a supervised manner, requiring paired noisy and clean images, limiting their application across diverse datasets. An alternative class of denoising models can be trained in a self-supervised manner, assuming independent noise across samples but are unable to generalize from available unpaired clean images. A method that can be trained without paired data and can use information from available unpaired high-quality images would address both weaknesses. Here, we present Baikal, a first attempt to formulate such a framework using Denoising Diffusion Probabilistic Models (DDPM) for fluorescence microscopy images. We first train a DDPM backbone in an unconditional manner to learn generative priors over complex morphologies in microscopy images. We then apply various conditioning strategies to sample from the trained model and propose an optimal strategy to denoise the desired image. Extensive quantitative comparisons demonstrate better performance of Baikal over state of the art self-supervised methods across multiple datasets. We highlight the advantage of generative priors learnt by DDPMs in denoising complex *Flywing* morphologies where other methods fail. Overall, our DDPM based denoising framework presents a new class of denoising methods for fluorescence microscopy datasets that achieve good performance without collection of paired high-quality images.

Keywords: Fluorescence Microscopy · Image Denoising · Denoising Diffusion Probabilistic Models · Unpaired Dataset

---

* These authors contributed equally to this work

1   Introduction

Fluorescence microscopy is a widely used technique to study biological phenomena. Advancements in optical techniques [5] and development of better fluorescent sensors [26] have pushed the limits of image collection. However desired imaging conditions like spatio-temporal resolution, duration for time-lapse imaging, phototoxicity and photobleaching for sensitive samples etc. constrain imaging parameters such as laser power, exposure time and frame-rate leading to acquisition of noisy images [19, 20].

Several computational denoising methods have been developed for both single images [1, 10, 12, 23] and videos [4, 11, 15] that can bypass imaging constraints and generate clean images. These methods fall into one of the two categories: 1) *paired-denoising* methods like CARE [23] offer best denoising performance, however these methods require large paired datasets of noisy and high quality images for training that are extremely time-consuming to collect. 2) *self-supervised denoising* methods such as Noise2Noise (N2N) [12] and Noise2Void (N2V) [10] can be trained using only noisy images. However, these methods assume pixel-wise independent noise, achieving worse performance in situations where spatially correlated (structured) noise is present as shown in [3]. To account for structured noise, modifications of self-supervised denoising methods have been proposed such as StructN2V [3], SSID [13] etc. Here, we consider a common scenario where an unpaired dataset of clean images is available but not paired clean data. Unpaired clean images may provide important morphological priors to denoising methods in extreme noise scenarios. Previous methods are not designed to take advantage of such unpaired images.

Our method, which we termed Baikal, leverages a Denoising Diffusion Probabilistic Model (DDPM) [8] for learning generative priors from unpaired clean images. Application of DDPMs have been shown for fluorescence microscopy [6, 14, 18] but not for denoising task. Additionally, DDPMs have been applied for denoising data from modalities such as PET, MRI, CT [7, 16, 21, 24]. However fluorescence microscopy data generation and noise characteristics differ significantly from these modalities [2]. Thus it is not clear whether DDPMs would work for denoising fluorescence microscopy images. We explore this question in this work by leveraging DDPM in a two step framework. First a diffusion model is trained only on clean images. Next, conditional sampling is used to generate clean predictions of the noisy images. In this work, we empirically evaluate several conditional samplers on multiple open-source datasets [23]. The proposed framework has several advantages over previous methods - 1) it does not require paired data as *paired-denoising* methods for training, 2) the DDPM backbone learns important morphological priors from unpaired clean images to guide denoising, 3) it is widely applicable across datasets with varying noise properties. Our contributions are as follows,

- We demonstrate the first application of Diffusion Models for denoising fluorescence microscopy images.
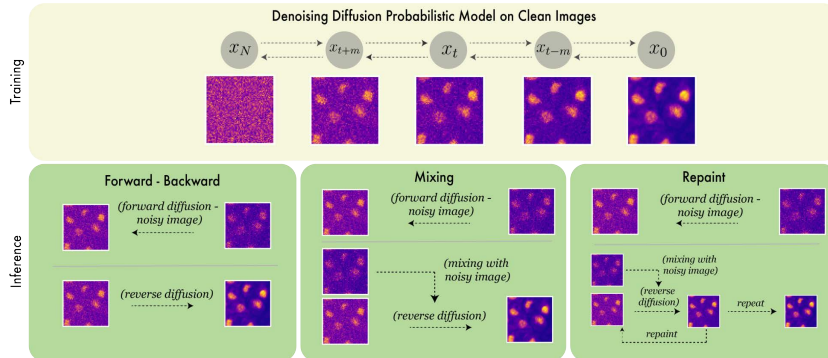
Fig. 1. Overview of the proposed method. **Training stage** - a DDPM backbone is trained using only unpaired high quality images. **Inference Stage** - image is denoised by sampling from the trained backbone. Three strategies are explored to guide the sampling process namely Forward-Backward, Mixing and Repaint.

- We systematically evaluate the applicability of several conditional sampling strategies designed for inpainting tasks for denoising and suggest an optimal strategy for fluorescence microscopy datasets.
- We evaluate if initial denoising using self-supervised methods such as N2V can better guide conditional samplers.

2   Methods

Our goal is to generate a clean image $\boldsymbol{x} \in \mathcal{X}$ given a noisy fluorescence image $\boldsymbol{y} \in \mathcal{Y}$. *Self-supervised* methods only use $\mathcal{Y}$ to generate clean images. *Paired-denoising* methods use paired clean and noisy acquisitions $\{\mathcal{X}, \mathcal{Y}\}$ to generate clean images. Here, we propose a two step method for denoising flourescence microscopy images when unpaired clean and noisy images are available (Fig. 1). First a generative backbone is trained to model clean data. Next, adapting several samplers from the inpainting literature [17] we evaluate them on three datasets of varying morphological properties. In addition, if *self-supervised* predictions are available for noisy images, we evaluate their utility on the performance of samplers.

2.1   Generative backbone on clean images

Our generative backbone is trained as a DDPM [8] on unpaired clean images $\mathcal{X}$. DDPMs are generative latent variable models that learn a model distribution $p_\theta(\boldsymbol{x})$ that approximate data distribution $q(\boldsymbol{x})$. The *forward* diffusion process applies Gaussian noise to the images until they become indistinguishable from random Gaussian noise for T steps $\boldsymbol{x}_0 \to \boldsymbol{x}_1 \to \boldsymbol{x}_2, ..., \boldsymbol{x}_T$. Variance schedule is given by $\beta_1, ..., \beta_T$.

$$q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) := \mathcal{N}(\boldsymbol{x}_t; \sqrt{1-\beta_t}\boldsymbol{x}_{t-1}, \beta_t \boldsymbol{I}) \tag{1}$$

The *reverse* process goes in the opposite direction to generate images starting from random noise $\boldsymbol{x}_T \to \boldsymbol{x}_{T-1}, ..., \boldsymbol{x}_0$ where $\boldsymbol{x}_T \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$. Reverse sampling distribution to sample $\boldsymbol{x}_{t-1}$ conditioned on $\boldsymbol{x}_t$ is given by

$$p_\theta(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t) := \mathcal{N}(\boldsymbol{x}_{t-1}; \boldsymbol{\mu}_\theta(\boldsymbol{x}_t, t), \boldsymbol{\Sigma}_\theta(\boldsymbol{x}_t, t)) \tag{2}$$

Here $\boldsymbol{\mu}_\theta$ denotes the neural network with parameters θ trained to denoise $\boldsymbol{x}_t$ at every time step t. Variances of the reverse process are also learned unless otherwise specified.

2.2 Conditional samplers as denoisers

Once the generative backbone is trained on clean images, we denoise an image $\boldsymbol{y_0} \in \mathcal{Y}$ by generating a (predicted) clean image $\boldsymbol{x_0}$ conditioned on $\boldsymbol{y_0}$. We adapt conditional sampling methods proposed in the inpainting literature [17] and evaluate their performance in the context of denoising fluorescence images. All conditional samplers in this work first forward diffuse the noisy image $\boldsymbol{y}_0$ to arbitrary time step t utilizing the closed form expression for forward process as follows,

$$q(\boldsymbol{y}_t|\boldsymbol{y}_0) = \mathcal{N}(\boldsymbol{y}_t; \sqrt{\bar{\alpha}_t}\boldsymbol{y}_0, (1 - \bar{\alpha}_t\boldsymbol{I}) \tag{3}$$

where $\alpha_t := 1 - \beta_t$ and $\bar{\alpha}_t := \prod_{s=1}^{t} \alpha_s$. To evaluate N2V predictions for their utility in our proposed denoising method, we replace $\boldsymbol{y}_0$ with $\boldsymbol{y}_0^{N2V}$ and its corresponding forward diffused as $\boldsymbol{y}_t^{N2V}$. Once noisy images $\boldsymbol{y}_0$ are forward diffused, different conditional sampling methods differ in the inputs to $\boldsymbol{\mu}_\theta(\boldsymbol{x}_t, t)$.

Forward-Backward (FB): Here, first the noisy image $\boldsymbol{y_0}$ is forward diffused to t' steps such that only the underlying signal is preserved. Next, the reverse process is run by setting $\boldsymbol{x}_{t'} = \boldsymbol{y}_{t'}$ (only at $t = t'$) to generate denoised $\boldsymbol{x}_0$ from noisy $\boldsymbol{y}_0$. The reverse step is given as follows,

$$p_\theta(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t) = \mathcal{N}(\boldsymbol{x}_{t-1}; \boldsymbol{\mu}_\theta(\boldsymbol{x}_t, t), \boldsymbol{\Sigma}_\theta(\boldsymbol{x}_t, t)) \tag{4}$$

Intuitively, since the DDPM is trained only on clean images, at the end of the reverse process, our hope is that the generated sample will approximate the denoised image (i.e. $\boldsymbol{x}_0 \approx \boldsymbol{x}$). However one disadvantage of this sampler is that the original signal destroyed during forward diffusion is lost forever and the network might not be able to recover all the components of the clean image when running the reverse process. To run on N2V predictions, we replace $\boldsymbol{x}_{t'} = \boldsymbol{y}_{t'}^{N2V}$.

Mixing: In forward-backward sampler the generation process at each time step is independent of the noisy image $\boldsymbol{y_0}$ (except for $t = t'$). Additionally, the information once destroyed is not recoverable in forward-backward denoising. Thus to feed the network all the information that was present in the original noisy image

at every time step and to further guide the generation process, we compute the weighted average of the reverse sample $\boldsymbol{x}_t$ and the forward diffused noisy image $\boldsymbol{y}_t$ and treat them as inputs to $\boldsymbol{\mu}_\theta$.

$$\boldsymbol{\mu}_\theta = \boldsymbol{\mu}_\theta((1 - w_t)\boldsymbol{x}_t + w_t\boldsymbol{y}_t, t) \tag{5}$$

Since higher weights for the noisy image in the later stages of the generative process could yield noisy generations, we monotonically decrease the weights w_t during the reverse process. We experimented with several mixing weight schedules (constant, linear, cosine and exponential). To run on N2V predictions, we replace $\boldsymbol{y}_t = \boldsymbol{y}_t^{N2V}$ when taking the weighted combination, but use forward noised $\boldsymbol{y}_0$ as $\boldsymbol{y}_t$ only for $t = t'$. A disadvantage of this sampler is that it replaces denoised pixels by noisy pixels reverting denoising, as we re-introduce noisy image at every time step by taking weighted combination.

Repaint: Although weighted combinations of $\boldsymbol{y}_t$ and DDPM outputs at each time step $\boldsymbol{x}_t$ guide the generation process, some of the denoised pixels in $\boldsymbol{x}_t$ are replaced with noisy pixels, thus reverting the denoising process. Additionally, mixing once at each time step may not be enough to ensure semantic consistency. This is because the mixing input $\boldsymbol{y}_t$ at each time step is independent of the DDPM output $\boldsymbol{x}_t$ at that time step. Thus, although the DDPM tries to predict the denoised output at the next time step using both $\boldsymbol{y}_t$ and $\boldsymbol{x}_t$, it may not be able to correct any inconsistency in the previous step. Thus, to ensure semantic consistency, similar to Repaint [17], we repeat the reverse step multiple times. Since $\boldsymbol{\mu}_\theta$ is trained as a *denoiser*, we use it for maintaining semantic consistency. Concretely, starting at a time step t, we run the the generation process for U steps, while mixing $\boldsymbol{x}_t$ and $\boldsymbol{y}_t$ at each time step, to generate $\boldsymbol{x}_{t-U}$. Next, $\boldsymbol{x}_{t-U}$ is taken as input to the time step t and generation process is repeated R times (See Algorithm 1 in Supplementary). Our strategy is different from the original Repaint, in that, empirically we found repainting every step yielded worse results compared to repainting in U intervals. Additionally, we stop mixing in the later stages of generation to avoid replacing denoised pixels by noisy pixels resulting in better results as shown in Fig.2(right).

3   Experiments

3.1   Dataset & Training

We evaluate our framework on three open source fluorescence microscopy datasets [23] with distinct morphological labelling to highlight the generalizability of Baikal. These include - 1) nuclei labeled embryos of *Tribolium castaneum*, 2) nuclei labeled embryos of the planarian *S. mediterranea* and 3) boundary labelled eplithelia of the fruit fly *Drosophila Melanogaster*. The datasets provide noisy and high quality images that were acquired at low and high laser powers, respectively. The test sets provided in the datasets were not normalized as

Datasets	Planaria			Tribolium			Flywing		
Methods	SSIM	MSE	PSNR	SSIM	MSE	PSNR	SSIM	MSE	PSNR
Noise	0.153	0.137	8.95	0.160	0.150	9.11	0.121	0.587	2.38
N2V [10]	0.296	0.055	15.58	0.208	0.074	11.55	0.174	0.576	2.46
SSID [13]	0.27	0.073	15.76	0.257	0.06	14.91	0.221	0.18	7.61
Forward Backward - Noise	0.339	**0.049**	16.71	0.260	0.052	14.71	0.260	0.209	6.94
Forward Backward - N2V	0.327	0.055	16.78	0.199	0.065	12.29	0.102	0.572	2.49
Mixing - Noise	0.317	0.056	14.25	0.229	0.058	13.68	0.241	0.220	6.69
Mixing - N2V	0.333	0.055	16.74	0.237	0.054	13.81	0.175	0.393	4.12
Repaint - Noise	**0.386**	0.050	**17.55**	**0.458**	**0.028**	**19.47**	**0.290**	**0.190**	**7.38**
Repaint - N2V	0.328	0.057	16.85	0.396	0.029	18.20	0.110	0.625	2.10
CARE* [23]	0.488	0.034	19.08	0.638	0.019	22.69	0.175	0.221	7.25

Table 1. SSIM($\uparrow$), MSE($\downarrow$) and PSNR($\uparrow$) metrics on test sets. *CARE is trained using paired dataset.
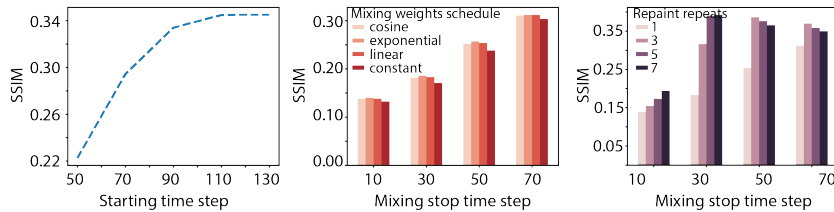
Fig. 2. Hyperparameter effects on SSIM accuracy evaluated on *Planaria* dataset. (**Left panel**) Starting time step. (**Center panel**) Mixing stop time step and mixing weights schedule. (**Right panel**) Mixing stop time step and number of repaint repeats.

the train set. Thus, for fair evaluation, we split the provided train sets in these datasets into train, eval and test sets in the ratio 80 : 10 : 10. We train diffusion backbone based on DifFace [25] codebase* using only 2D clean images from the train split. *Planaria* and *Tribolium* datasets provide noisy and clean 3D volumes whereas the *Drosophila* dataset consisted of only 2D projected clean images. In each case we train separate diffusion backbones by treating each 2D z-plane as individual data points. To quantify accuracy, we generate denoised predictions for images in the test split and compare them with the clean images in the test split. We provide all qualitative and quantitative results on test splits.

3.2   Quantitative results

We compare Baikal against state of the art methods for denoising fluorescence microscopy images. These include self-supervised methods such as N2V [10] and SpatiallyAdaptiveSSID(SSID) as well as paired-denoising method CARE [23]. We train both methods using author provided code and hyperparameters on the same train sets as used for training diffusion backbones. For the *Tribolium* dataset, we observed bad performance of N2V due to correlated noise in the

---

* https://github.com/zsyOAOA/DifFace

input images and therefore trained structN2V [3] using a square mask of size 5x5 pixels around each masked pixel. For the *Flywing* dataset, since only 2D clean images were available, we trained CARE on max-projected noisy images without using a projector network as done previously [23]. We report SSIM [22], MSE and PSNR metrics to compare performance of all methods. To be consistent with CARE, we evaluate metrics on 3D datasets by first max-projecting them across the z-plane dimension.

We observe from Table.1 that a simple conditional sampler like Forward-Backward achieves better performance than N2V across all three datasets signifying the advantage of generative priors provided by the diffusion model. Surprisingly, when N2V predictions were used as input to Forward-Backward sampler, the performance drops. We hypothesize that this could be because N2V predictions destroy underlying signal leading to sub-optimal generations. Furthermore, we observe Repaint with mixing noisy image is the best conditional sampler for all three datasets trailing only behind paired-denoising method CARE. Notably, our proposed approach is more generalizable during training time. E.g. N2V required manual tuning of the masks during training to work on *Tribolium* dataset with correlated noise in images. In addition, CARE required training an additional projector network [23] when 3D noisy and 2D clean images in *Flywing* data are available. In contrast, our proposed approach does not require any modifications while training the diffusion models across these different datasets spanning different noise profiles and shapes thus making it easy to use in practical settings.

3.3   Ablations

To observe the sensitivity of the hyperparameters in different samplers, we vary critical parameters in each sampler and report the corresponding performance achieved on *Planaria* dataset. These critical parameters include 1) Starting time step (t') denoting the time step to which noisy image y_0 is forward diffused to, 2) Mixing weights schedule (w_t in Eq.5), 3) Mixing stop time step i.e. the time step beyond which mixing is stopped while sampling in Mixing and Repaint samplers, 4) Number of repeats in Repaint sampler (parameter R in Algorithm 1 - Supplementary text). As seen from Fig.2 (left), increasing t' leads to cleaner generations for forward-backward sampler and the performance plateaus after 100 time steps. For mixing sampler Fig.2 (middle), mixing at time steps closer to $t = 0$ leads to decrease in performance as it re-introduces noisy pixels in the denoised output thus reverting denoising. Further, the schedule for the mixing weights minimally affects the results. In the Repaint sampler Fig.2 (right), increasing the number of repeats significantly improves the performance supporting our argument that semantic consistency improves by repainting. Additionally best results are obtained when mixing is stopped after $t = 50$ while repainting continues.

3.4   Qualitative results

Qualitative comparisons highlight more advantages of our proposed framework. We observe N2V predictions contain horizontal artifacts in *Planaria* dataset
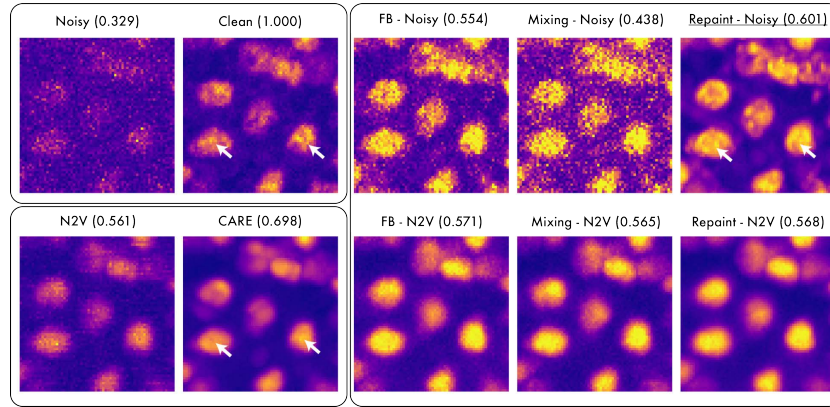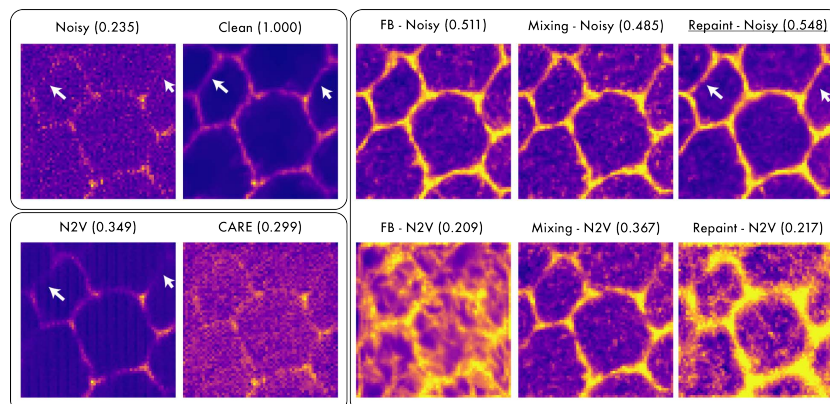
Fig. 3. Sample predictions for *Planaria*, SSIM in brackets. **(Top left panel)** Noisy and Clean images. **(Bottom left panel)** N2V and CARE predictions. **(Right panel)** Conditional diffusion denoising using noisy image and N2V prediction as inputs.
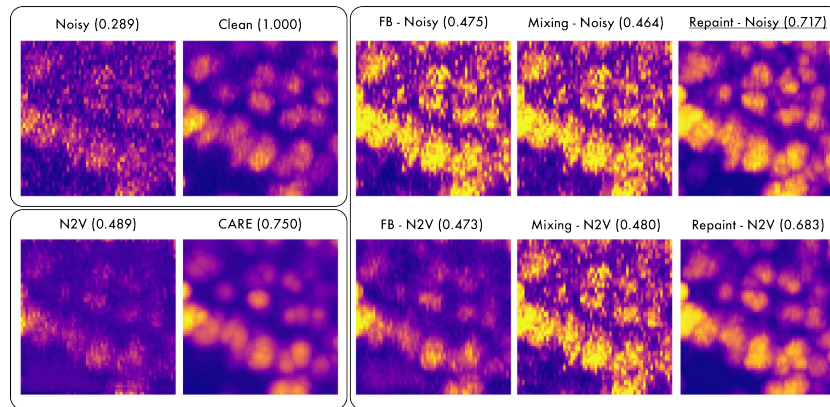
Fig. 4. Sample predictions for *Tribolium*, SSIM in brackets. (Panels are same as Fig.3)

(Fig.3). Presence of such artifacts in N2V denoised images have been reported before [9]. Further, N2V suffers from bad performance on the *Tribolium* dataset due to the presence of correlated noise (Fig.4). Finally, it fails to recover cell boundaries in *Flywing* data in extremely noisy regions (Fig.5, Supplementary Fig.1). In contrast, the best conditional sampler performs well across all scenarios. Notably in the *Flywing* dataset, Baikal is able to recover cell boundaries in extremely noisy regions (highlighted by arrows). In addition, while CARE predictions achieve high SSIM accuracy, qualitatively the cellular structures look extremely smooth, removing any subcellular features (Fig.3, Fig.4, Supplementary Fig.1). In comparison, Baikal is able to preserve such subcellular features (highlighted by arrows).

Fig. 5. Sample predictions for *Flywing*, SSIM in brackets. (Panels are same as Fig.3)

4   Conclusion

We demonstrate for the first time, to the best of our knowledge, application of Diffusion Models for denoising fluorescence microscopy images without needing paired training data. Baikal opens up future avenues for tackling other common tasks in fluorescence microscopy like de-blurring, isotropic reconstruction and super-resolution.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Batson, J., Royer, L.: Noise2self: Blind denoising by self-supervision (2019). https://doi.org/10.48550/arXiv.1901.11365
2. Belthangady, C., Royer, L.A.: Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. Nature methods **16**(12), 1215–1225 (2019)
3. Broaddus, C., Krull, A., Weigert, M., Schmidt, U., Myers, E.: Removing structured noise with self-supervised blind-spot networks. In: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI 2020). IEEE (2020). https://doi.org/10.1109/ISBI45749.2020.9098336
4. Chaudhary, S., Moon, S., Lu, H.: Fast, efficient, and accurate neuro-imaging denoising via supervised deep learning. Nature communications **13**(1), 5165 (2022)
5. Chen, B.C., Legant, W.R., Wang, K., Shao, L., Milkie, D.E., Davidson, M.W., Janetopoulos, C., Wu, X.S., Hammer, J.A., Liu, Z., English, B.P., Mimori-Kiyosue, Y., Romero, D.P., Ritter, A.T., Lippincott-Schwartz, J., Fritz-Laylin, L., Mullins, R.D., Mitchell, D.M., Bembenek, J.N., Reymann, A.C., Böhme, R., Grill, S.W., Wang, J.T., Seydoux, G., Tulu, U.S., Kiehart, D.P., Betzig, E.: Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution. Science **346**(6208), 1257998 (2014). https://doi.org/10.1126/science.1257998, https://www.science.org/doi/abs/10.1126/science.1257998

6. Eschweiler, D., Yilmaz, R., Baumann, M., Laube, I., Roy, R., Jose, A., Brückner, D., Stegmaier, J.: Denoising diffusion probabilistic models for generation of realistic fully-annotated microscopy image datasets. PLOS Computational Biology **20**(2), e1011890 (2024)

7. Gong, K., Johnson, K., El Fakhri, G., Li, Q., Pan, T.: Pet image denoising based on denoising diffusion probabilistic model. European Journal of Nuclear Medicine and Molecular Imaging pp. 1–11 (2023)

8. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models (2020). https://doi.org/10.48550/arXiv.2006.11239

9. Höck, E., Buchholz, T.O., Brachmann, A., Jug, F., Freytag, A.: N2v2-fixing noise2void checkerboard artifacts with modified sampling strategies and a tweaked network architecture. In: European Conference on Computer Vision. pp. 503–518. Springer (2022)

10. Krull, A., Buchholz, T.O., Jug, F.: Noise2void - learning denoising from single noisy images (2019). https://doi.org/10.48550/arXiv.1811.10980

11. Lecoq, J., Oliver, M., Siegle, J.H., Orlova, N., Ledochowitsch, P., Koch, C.: Removing independent noise in systems neuroscience data using deepinterpolation. Nature methods **18**(11), 1401–1408 (2021). https://doi.org/10.1038/s41592-021-01285-2

12. Lehtinen, J., Munkberg, J., Hasselgren, J., Laine, S., Karras, T., Aittala, M., Aila, T.: Noise2noise: Learning image restoration without clean data (2018). https://doi.org/10.48550/arXiv.1803.04189

13. Li, J., Zhang, Z., Liu, X., Feng, C., Wang, X., Lei, L., Zuo, W.: Spatially adaptive self-supervised learning for real-world image denoising. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9914–9924 (2023)

14. Li, R., della Maggiora, G., Andriasyan, V., Petkidis, A., Yushkevich, A., Kudryashev, M., Yakimovich, A.: Microscopy image reconstruction with physics-informed denoising diffusion probabilistic model. arXiv preprint arXiv:2306.02929 (2023)

15. Li, X., Zhang, G., Wu, J., Zhang, Y., Zhao, Z., Lin, X., Qiao, H., Xie, H., Wang, H., Fang, L., et al.: Reinforcing neuron extraction and spike inference in calcium imaging using deep self-supervised denoising. Nature Methods pp. 1–6 (2021). https://doi.org/10.1038/s41592-021-01225-0

16. Liu, X., Xie, Y., Diao, S., Tan, S., Liang, X.: A diffusion probabilistic prior for low-dose ct image denoising. arXiv preprint arXiv:2305.15887 (2023)

17. Lugmayr, A., Danelljan, M., Romero, A., Yu, F., Timofte, R., Gool, L.V.: Repaint: Inpainting using denoising diffusion probabilistic models (2022). https://doi.org/10.48550/arXiv.2201.09865

18. Pan, M., Gan, Y., Zhou, F., Liu, J., Zhang, Y., Wang, A., Zhang, S., Li, D.: Diffuseir: Diffusion models for isotropic reconstruction of 3d microscopic images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 323–332. Springer (2023)

19. Pawley, J.B.: Fundamental limits in confocal microscopy. In: Handbook of biological confocal microscopy, pp. 20–42. Springer (2006). https://doi.org/10.1007/978-0-387-45524-2

20. Scherf, N., Huisken, J.: The smart and gentle microscope. Nature biotechnology **33**(8), 815–818 (2015). https://doi.org/10.1038/nbt.3310

21. Shen, C., Yang, Z., Zhang, Y.: Pet image denoising with score-based diffusion probabilistic models. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 270–278. Springer (2023)

22. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing **13**(4), 600–612 (2004). https://doi.org/10.1109/TIP.2003.819861
23. Weigert, M., Schmidt, U., Boothe, T., Müller, A., Dibrov, A., Jain, A., Wilhelm, B., Schmidt, D., Broaddus, C., Culley, S., et al.: Content-aware image restoration: pushing the limits of fluorescence microscopy. Nature methods **15**(12), 1090–1097 (2018). https://doi.org/10.1038/s41592-018-0216-7
24. Xiang, T., Yurt, M., Syed, A.B., Setsompop, K., Chaudhari, A.: Ddm²: Self-supervised diffusion mri denoising with generative diffusion models. arXiv preprint arXiv:2302.03018 (2023)
25. Yue, Z., Loy, C.C.: Difface: Blind face restoration with diffused error contraction (2023). https://doi.org/10.48550/arXiv.2212.06512
26. Zhang, Y., Rózsa, M., Liang, Y., Bushey, D., Wei, Z., Zheng, J., Reep, D., Broussard, G.J., Tsang, A., Tsegaye, G., et al.: Fast and sensitive gcamp calcium indicators for imaging neural populations. Nature **615**(7954), 884–891 (2023)