Ref. Ares(2026)9854608 - 20/10/2026

BE-LIGHT

Improving biomedical diagnosis through light–based technologies and machine learning

Doctoral Networks (MSCA-DN)

HORIZON - MSCA - 2022 - DN

Grant Agreement Nº 101119924

WP3: *Advanced microscopy for the analysis of cells and tissues involved in the development of diseases*

Deliverable D3.3

Confocal/Superresolution images

Ioannis Georvasilis, FSJD

Mònica Roldán, FSJD

Date: 14 10 2026

Type: Data

Dissemination Level: PU-Public

Funded by the European Union

DELIVERABLE IDENTIFICATION SHEET

Document version	Date	Editor	Change(s)
1	14-10-2026	Mònica Roldán	First version
2	15-10-2026	Ioannis Georvasilis	Data review, references, Data Availability and FAIR section, Annex A.
3	18-10-2026	Mònica Roldán	Last revision

Grant Agreement Nº	101119924
Call Identifier	HORIZON - MSCA - 2022 – DN
Acronym	BE-LIGHT
Full title	Improving biomedical diagnosis through light-based technologies and machine learning
Project URL	http://www.belightproject.eu/

Deliverable Nº and Name	D3.3
Work Package	WP3 Advanced microscopy for the analysis of cells and tissues involved in the development of diseases
Lead Beneficiary	UMG
Type	DATA - data sets
Dissemination Level	PU-Public
Due date	Month 30

Authors (Partner)	Mònica Roldán, FSJD Ioannis Georvasilis, FSJD			
Responsible Author	Mònica Roldán, FSJD Ioannis Georvasilis, FSJD		**Email**	monica.roldan@sjd.es ioannis.georvasilis@sjd.es
	Partner	FSJD	**Phone**	93-2532100

Abstract (for dissemination)	Fluorescence microscopy enables researchers to study cells and understand how rare diseases affect their structure and function. However, images may contain missing information or areas of poor quality, making it difficult to identify important biological features. These challenges are particularly relevant in rare disease research, where the availability of biological samples and high-quality reference images is often limited. This study evaluates a computational framework designed to improve fluorescence microscopy images by recovering information from

	missing or degraded image channels. The framework is assessed using three complementary datasets comprising a total of **501 multichannel images**. An external reference dataset includes **240 images of breast cancer cells**, with only 10% used as high-quality reference data to simulate the limited availability of curated images. The framework is further evaluated using two real-world datasets from Hospital Sant Joan de Déu: **184 multichannel skeletal muscle mosaics** from the Col6a1-KI dataset, from which 113 high-quality image regions were selected as references, and **77 skin cell images** from the Dnm1L-Mito dataset, including 73 with complete fluorescence channel acquisitions. By evaluating the framework across different biological samples and imaging conditions, this study investigates its ability to improve image quality and recover missing information when high-quality references are scarce. Ultimately, this approach aims to support more reliable fluorescence microscopy image analysis and facilitate the study of biological mechanisms underlying rare diseases.
Keywords	Fluorescence microscopy; Rare diseases; Image restoration; Image quality enhancement; Biomedical image analysis; Deep learning; Missing data reconstruction; Mitochondrial disorders; Skeletal muscle imaging; Computational imaging.

Table of Contents

DELIVERABLE IDENTIFICATION SHEET 2
EXECUTIVE SUMMARY 5
INTRODUCTION 5
1 DATA DESCRIPTION 6
1.1 Overview 6
1.2 Data Organisation 7
1.2.1. External Reference Dataset 7
1.2.2. Col 601-K1 Skeletal Muscle Dataset 7
1.2.3. Dnm1L-Mito Skin Fibroblast Dataset 8
1.2.4. Fluorescence Channel and Spatial Structure 8
1.2.5. Reference Data Organisation 8
1.3 Evaluation Protocol 9
1.3.1. Controlled Evaluation 9
1.3.2. Real-World Evaluation 9
2 DATA ACQUISITION 10
2.1 Experimental Set-up 10
2.2 Data Processing 11
3 DATA AVAILABILITY AND FAIR COMPLIANCE 12
4 SUMMARY 13
5. REFERENCES 13
6. Annex A 15

EXECUTIVE SUMMARY

This deliverable presents the evaluation of a computational framework for improving fluorescence microscopy images in rare disease research. The framework aims to recover information from missing or degraded fluorescence channels, addressing common challenges associated with variable image quality and the limited availability of high-quality reference data. The evaluation is based on three complementary datasets comprising 501 multichannel images: 240 images of breast cancer cells from an external reference dataset, 184 skeletal muscle mosaics from the Col6a1-KI dataset, and 77 skin cell images from the Dnm1L-Mito dataset. To simulate the limited availability of curated data in rare disease research, only 10% of the external dataset is used as high-quality reference data. In addition, 113 high-quality image regions from the Col6a1-KI dataset and 73 images with complete fluorescence channel acquisitions from the Dnm1L-Mito dataset are retained as reference samples.

This work contributes directly to the objectives of the BE-LIGHT project, particularly Task 3.2, which focuses on developing computer-aided tools for rare disease diagnosis. By evaluating image restoration under heterogeneous acquisition conditions and limited reference availability, this work supports the development of computational methods for more reliable fluorescence microscopy image analysis. These methods may facilitate quantitative assessment of subcellular structures in fibroblasts and muscle tissues, contributing to research on rare diseases, including muscular dystrophies and mitochondrial diseases, and supporting future advances in image-based diagnostic approaches.

INTRODUCTION

Multichannel fluorescence microscopy is a key tool in biological and medical research, enabling the simultaneous visualization of multiple cellular structures and molecular processes. By combining different fluorescence channels, researchers can obtain complementary biological information that supports the study of cellular mechanisms and disease-related changes.

Despite these advantages, fluorescence microscopy datasets, particularly those from rare disease research, often contain missing or degraded channels due to limited sample availability, variable acquisition conditions, and differences in imaging protocols. These limitations can hinder image analysis and reduce the amount of data available for computational studies, especially when high-quality reference images are scarce.

This work addresses these challenges by introducing RM-DDPM [1], a computational framework designed to reconstruct missing or degraded information in multichannel fluorescence microscopy images. By using a reliability mask to identify trustworthy and unreliable image regions, the framework aims to handle different types of image degradation within a single model. This approach is evaluated using both an external reference dataset and real-world rare disease microscopy data, with the goal of supporting more robust image analysis under heterogeneous acquisition conditions.

1 DATA DESCRIPTION

1.1 Overview

The dataset collection comprises three complementary multichannel fluorescence microscopy datasets used to evaluate a computational framework for reconstructing missing or degraded image information. The datasets include an external reference collection of breast cancer cells and two real-world rare-disease datasets involving skeletal muscle tissue and skin fibroblasts.

All datasets contain three fluorescence channels (C = 3), although the corresponding biological markers differ. Together, they comprise 501 original multichannel images and mosaics, providing a range of acquisition conditions and biological structures for evaluating image reconstruction.

The external dataset provides fully observed images for controlled quantitative evaluation, whereas the two rare-disease datasets represent real-world scenarios characterised by heterogeneous acquisition quality, incomplete or degraded channels, and limited availability of high-quality reference data.

Key Dataset Characteristics:

- **Total number of original images and mosaics:** 501.
- **External reference dataset:** 240 multichannel images of MCF-7 breast cancer cells.
- **Col6a1-KI dataset:** 184 skeletal muscle mosaics, subdivided into 1,877 tiles.
- **Dnm1L-Mito dataset:** 77 skin fibroblast images.
- **Number of fluorescence channels:** 3 per dataset.
- **External image dimensions:** 1024 × 1280 pixels.
- **Col6a1-KI tile dimensions:** 1024 × 1024 pixels.
- **Dnm1L-Mito image dimensions:** 1024 × 1024 pixels.
- **High-quality reference data:** 10% of the external dataset, 113 tiles from Col6a1-KI, and 73 images from Dnm1L-Mito.

1.2. Data Organisation

The data collection is organised into three datasets according to their origin, biological content, and role in the evaluation. Each dataset contains three fluorescence channels, providing complementary information about cellular structures and biological processes.

The datasets differ in their acquisition formats and spatial dimensions. BBBC021v1 consists of individual multichannel images, Dnm1L-Mito contains images acquired directly at tile resolution, and Col6a1-KI comprises larger mosaics that are subdivided into fixed-size tiles for quality assessment and restoration evaluation.

- **1.2.1. External Reference Dataset: BBBC021v1**

The external dataset corresponds to the Week10_40111 subset of BBBC021v1 [2, 3]. It comprises 240 multichannel fluorescence microscopy images of MCF-7 breast cancer cells, with a spatial resolution of 1024 × 1280 pixels.

Each image contains three fluorescence channels:

- **DAPI:** nuclear staining.
- **Tubulin:** cytoskeletal structures.
- **Actin:** cellular actin structures.

The complete availability of the three channels enables controlled experiments in which channels are artificially removed or degraded and subsequently reconstructed against their original reference images.

To simulate the limited availability of curated data commonly encountered in rare-disease research, a randomly selected 10% of the dataset is used as the high-quality reference subset across the experiments. This provides a controlled setting for evaluating image generation, missing-channel imputation, degraded-channel restoration, and the joint reconstruction of missing and degraded channels.

- **1.2.2. Col6a1-KI Skeletal Muscle Dataset**

The Col6a1-KI dataset [4] was acquired by the Neuromuscular Unit at Hospital Sant Joan de Déu. It comprises 184 multichannel skeletal muscle mosaics containing three fluorescence channels:

- **DAPI:** nuclear staining.
- **Perlecan:** extracellular matrix structures.
- **Collagen VI:** collagen VI distribution in muscle tissue.

The original mosaics were subdivided into fixed-size tiles of 1024 × 1024 pixels to enable quality assessment at the regional level. This approach accounts for the fact that individual mosaics may contain both high-quality and degraded regions.

A total of 1,877 tiles were evaluated using channel-specific quality criteria, resulting in 113 high-quality reference tiles originating from 23 mosaics. To prevent subject-level data leakage, the mosaics contributing reference tiles for model training are disjoint from those providing non-reference tiles for restoration evaluation.

This dataset is used to assess degraded-channel restoration under real-world acquisition conditions, particularly when the target channel does not have a paired, clean reference image. The quality assessment protocol and selection criteria are detailed in Annex A.

- **1.2.3. Dnm1L-Mito Skin Fibroblast Dataset**

The Dnm1L-Mito dataset was acquired by the Synaptic Metabolism and Personalized Therapies Laboratory at Hospital Sant Joan de Déu. It comprises 77 multichannel fluorescence microscopy images of primary skin fibroblasts obtained from patients with DNM1L-related disorders, patients with other mitochondrial disorders, and matched controls.

Each image has a spatial resolution of 1024 × 1024 pixels and contains three fluorescence channels:

- **DAPI:** nuclear staining.
- **α-tubulin:** cytoskeletal structures.
- **TOMM20:** mitochondrial network.

The dataset is used to characterise mitochondrial network morphology, an established phenotypic marker relevant to mitochondrial disorders and DNM1L-related mitochondrial fission deficiency [5, 6].

Unlike Col6a1-KI, Dnm1L-Mito is acquired directly at tile resolution rather than as larger mosaics. Of the 77 images, 73 contain complete acquisitions of all three fluorescence channels and are retained as high-quality reference samples.

This dataset is used to evaluate missing-channel imputation in real-world samples where paired, clean ground-truth references are not available for all acquisitions.

- **1.2.4. Fluorescence Channel and Spatial Structure**

All three datasets contain three fluorescence channels, although the markers differ according to the biological specimen. Each channel provides complementary information about specific cellular or tissue structures.

The spatial organisation also differs between datasets. BBBC021v1 images have dimensions of 1024 × 1280 pixels, whereas Dnm1L-Mito images have dimensions of 1024 × 1024 pixels. Col6a1-KI mosaics are subdivided into 1024 × 1024-pixel tiles to support region-level quality assessment and image reconstruction.

The use of fixed-size tiles in Col6a1-KI enables individual regions to be assessed independently, accounting for spatial variability in image quality within the same mosaic. It also allows high-quality reference regions to be separated from degraded regions while maintaining the required separation between training and evaluation data.

- **1.2.5. Reference Data Organisation**

High-quality reference data are defined separately for each dataset according to its acquisition characteristics and quality assessment procedure.

For BBBC021v1, a randomly selected 10% of the 240 images is used as the high-quality reference subset. For Col6a1-KI, 113 high-quality tiles originating from 23 mosaics are selected from the 1,877 tiles evaluated. For Dnm1L-Mito, 73 images with complete acquisitions of all three fluorescence channels are retained as reference samples.

These subsets provide high-quality examples of complete multichannel observations for model training and evaluation. The remaining data allow the framework to be assessed under more challenging conditions involving missing or degraded fluorescence channels and heterogeneous acquisition quality.

- **1.3. Evaluation Protocol**

The proposed framework is evaluated under two complementary settings: controlled experiments using BBBC021v1 and real-world experiments using the Col6a1-KI and Dnm1L-Mito datasets.

- **1.3.1. Controlled Evaluation**

The BBBC021v1 dataset is used to evaluate four tasks:

- **Unconditional image generation:** generating multichannel fluorescence microscopy images and assessing their similarity to the reference data distribution.
- **Missing-channel imputation:** reconstructing a fluorescence channel that has been withheld while using the remaining observed channels as input.
- **Degraded-channel restoration:** recovering a fluorescence channel affected by simulated optical blur, signal attenuation, background offset, or noise.
- **Joint imputation–restoration:** reconstructing two channels simultaneously when one is missing and another is degraded, using the remaining clean channel as input.

Since the original channels are available, reconstructed images can be compared against paired reference images. Reconstruction quality is assessed using the Structural Similarity Index Measure (SSIM) [7], Learned Perceptual Image Patch Similarity (LPIPS) [8], and object-level Dice scores derived from Cellpose-SAM [9]. For image generation, distributional quality is assessed using Improved Precision and Recall (iF1) [10], Fréchet Inception Distance (FID) [11], and Kernel Inception Distance (KID) [12].

The framework is compared with baseline diffusion models and task-specific methods, including approaches based on image restoration, denoising, and deep-learning-based reconstruction.

- **1.3.2. Real-World Evaluation**

The two rare-disease datasets are used to assess the framework under heterogeneous acquisition conditions, where paired, clean reference images are unavailable for every sample. Consequently, evaluation relies on surrogate measures of structural fidelity, biological realism, and image quality rather than direct pixel-wise comparison against ground-truth images.

For **Dnm1L-Mito**, missing-channel imputation is evaluated by reconstructing the α-tubulin channel. The assessment includes channel-correlation shift, tubulin realism statistics, fibril-likeness [13], and cellular network topology. Because these surrogate metrics require reliably segmented tubulin networks, they are computed on an exploratory subset of four images.

For **Col6a1-KI**, degraded-channel restoration is evaluated on tiles where the Collagen VI channel does not meet the high-quality reference criteria. The assessment considers fibril-like structures, branch-point counts, network topology, and spatial correlation between the restored Collagen VI channel and the unchanged DAPI channel.

This evaluation strategy enables the framework to be assessed both quantitatively under controlled conditions and in realistic rare-disease research scenarios with limited high-quality reference data.

2 DATA ACQUISITION

2.1. Experimental Set-up

Fluorescence microscopy images were acquired using confocal microscopy systems with different configurations depending on the biological specimen and imaging requirements. The acquisition protocols enabled the visualization of complementary cellular and tissue structures through three fluorescence channels.

Dnm1L-Mito skin fibroblast imaging. Fluorescence imaging was performed using a Leica TCS SP8 confocal microscope equipped with a white-light laser and hybrid spectral detectors (Leica Microsystems GmbH, Mannheim, Germany). Images were acquired using an HC PL APO 100×/1.4 oil-immersion objective and HyD detectors, with a zoom factor of 2.5 for colocalization imaging. DAPI was excited using a 405 nm diode laser, with emission collected between 425 and 480 nm. TOMM20 was excited using a 488 nm argon laser, with emission detected between 500 and 550 nm. Drp1 and tubulin were excited using a 555 nm white-light laser, with emission collected between 575 and 610 nm. Z-stacks comprising 13–15 optical sections were acquired at intervals of 0.25–0.30 µm across the sample thickness.

Col6a1-KI skeletal muscle imaging. Automated tissue mapping was performed using a Leica TCS SP8 STED 3X confocal microscopy system equipped with a white-light laser, HyVolution mode, and hybrid detectors (Leica Microsystems, Wetzlar, Germany). A high-precision motorized stage controlled by LAS X Navigator software was used to acquire large-scale three-dimensional mosaics of tissue sections. Imaging was performed at low magnification using an HC PL APO CS2 10×/0.4 dry objective. The software automatically determined the optimal stage positions based on image dimensions and the overlap required between adjacent tiles, ensuring coverage of the selected tissue region.

Individual image tiles were acquired at 1024 × 1024 pixels, with 12-bit intensity resolution and a z-step of 5 µm. Between 4 and 20 stacks were collected for each extended image. Muscle tissue was sequentially excited at 405 nm for DAPI, 488 nm for Collagen VI (COL6), and 594 nm for Perlecan. The corresponding emission detection ranges were 420–465 nm, 500–560 nm, and 610–755 nm, respectively. Ten optical sections were acquired at 5 µm intervals across the tissue thickness, and maximum-intensity projections were generated using LAS AF software (Leica Microsystems, Heidelberg, Germany).

These acquisition protocols produced multichannel fluorescence images suitable for studying cellular structures, mitochondrial networks, and extracellular matrix organisation. However, differences in acquisition conditions and image quality, together with incomplete or degraded channels, motivated the development of computational methods for image reconstruction.

2.2. Data Processing

The image processing pipeline combines quality assessment, data normalisation, image augmentation, and diffusion-based model training to support the reconstruction of missing or degraded fluorescence channels.

Image quality assessment and reference selection. For the Col6a1-KI dataset, quality control was performed at the tile level because individual mosaics could contain both high-quality and degraded regions. The original 184 mosaics were divided into 1,877 fixed-size tiles of 1024 × 1024 pixels. Channel-specific quality criteria were applied independently, resulting in 113 high-quality reference tiles originating from 23 mosaics. To prevent subject-level data leakage, mosaics contributing reference tiles for training were kept separate from those used for restoration evaluation. For Dnm1L-Mito, 73 of the 77 images, all containing complete acquisitions of the three fluorescence channels, were retained as reference samples. For the external BBBC021v1 dataset, a randomly selected 10% of the 240 images was used as the high-quality reference subset.

Image normalisation and augmentation. Each fluorescence channel was normalised independently using global percentile bounds calculated from the high-quality reference subset. Specifically, the median of the per-image 1st and 99th intensity percentiles was used to define the lower and upper bounds. Intensities were then rescaled to the range [−1, 1]. During training, random 256 × 256 pixels crops were extracted from the BBBC021v1 and Col6a1-KI images and augmented using random scale transformations to increase variability in the training data. Dnm1L-Mito images were instead resized to 256 × 256 pixels.

Diffusion model implementation and training. The proposed Reliability-Mask-Conditioned Denoising Diffusion Probabilistic Model (RM-DDPM) [14] uses a U-Net architecture [15] with six resolution blocks, two residual layers per block, and one self-attention block. The diffusion process comprises 1,000 timesteps and uses a cosine variance schedule [16]. An auxiliary reconstruction loss is included with a weighting factor of $\lambda x0 = 1.0$.

During training, reliability masks are sampled from a uniform mixture of reconstruction scenarios, enabling the model to learn from different combinations of observed, missing, and degraded fluorescence information. A separate model is trained for each dataset using its corresponding high-quality reference subset. Once trained, the same model can be used for image generation, missing-channel imputation, degraded-channel restoration, and joint imputation–restoration without task-specific retraining.

Optimisation is performed using AdamW [17] with a base learning rate of 1×10^{-4} and a cosine learning-rate schedule with warmup. Training runs for 80,000 iterations with a batch size of 8. An exponential moving average of the model weights, with a decay factor of 0.999, is maintained throughout training and used for the reported results. The final checkpoint is used without validation-based model selection.

The model is implemented in PyTorch [18] and trained on an NVIDIA RTX 5000 Ada Generation GPU, with an average training time of approximately four hours per experiment.

Inference and reconstruction. At inference, the model receives a reliability mask, the available image context, and a starting diffusion timestep. These inputs determine which observations are preserved and which regions or channels are reconstructed. For degraded-channel restoration, the starting timestep is selected according to the dataset and degradation severity. A value of 500 is used for Col6a1-KI and for joint imputation–restoration. For BBBC021v1, the starting timestep is set to 200 and 250 for mild and severe DAPI degradation, respectively, and to 400 and 500 for mild and severe

Tubulin or Actin degradation. Resampling, consistency mechanisms, and initial-noise injection are enabled during inference.

Baseline implementation. The proposed framework is compared with established diffusion-based and task-specific reconstruction methods. Publicly available implementations are used where possible, retaining their default architectures and training procedures unless otherwise specified. All comparison methods are trained on the same high-quality reference subsets and for the same number of iterations to ensure consistent access to training data. The StyleGAN2-ADA [19] implementation is used for image generation, while the CARE-based U-Net [20], MAXIM [21] and SIMIF [22] are used for missing-channel imputation, and Noise2Void [23] and Cellpose3 [24] are used for degraded-channel restoration, with CARE serving as an approximate upper bound. The SIMIF method is reproduced from its published description because an official implementation was unavailable. Baseline DDPM [14] and DDIM [25] models use the same backbone and training configuration as RM-DDPM but without reliability-mask conditioning.

3 Data Availability and FAIR Compliance

In line with the open-science obligations of Horizon Europe, the data underlying this deliverable are managed according to the FAIR principles [26] and the principle "as open as possible, as closed as necessary". The availability of each dataset is summarised below.

Findable. The three datasets are described in Section 1, including their origin, biological content, fluorescence markers, number of images, spatial dimensions and high-quality reference subsets. The methodology and results are reported in the associated preprint, which has a persistent identifier (doi: 10.64898/2026.10.07.753438) [1]. The Col6a1-KI model is further described in the corresponding publication [4].

Accessible. BBBC021v1 is publicly available from the Broad Bioimage Benchmark Collection (https://bbbc.broadinstitute.org/BBBC021) [2, 3]. The Col6a1-KI and Dnm1L-Mito datasets contain fluorescence microscopy images derived from patient-related biological material collected at Hospital Sant Joan de Déu. In accordance with the data-protection policy of the hospital and the General Data Protection Regulation (GDPR), these datasets cannot be deposited in a public repository. They are stored on secure institutional infrastructure at FSJD, and any request for access must be addressed to the responsible authors and is subject to approval by Hospital Sant Joan de Déu.

Interoperable. The images are multichannel confocal fluorescence images stored in standard microscopy image formats. The acquisition metadata (microscope, objectives, excitation wavelengths, emission ranges, z-sampling and bit depth) are reported in Section 2.1, and the processing steps in Section 2.2.

Reusable. The source code implementing RM-DDPM, the baselines and the evaluation protocol is publicly available at https://github.com/igeor/ddpm4het. This allows the methodology to be reproduced on the public BBBC021v1 dataset and applied to other multichannel fluorescence microscopy datasets. Data provenance, quality-control criteria and processing steps are documented in Sections 1 and 2 and in Annex A.

4 SUMMARY

This deliverable presents the evaluation of RM-DDPM, a computational framework for reconstructing missing or degraded information in multichannel fluorescence microscopy images. The framework is evaluated using three datasets comprising 501 original images and mosaics: 240 breast cancer cell images from BBBC021v1, 184 skeletal muscle mosaics from Col6a1-KI, and 77 skin fibroblast images from Dnm1L-Mito. The evaluation considers scenarios with limited high-quality reference data, reflecting the challenges commonly encountered in rare-disease research.

RM-DDPM unifies image generation, missing-channel imputation, and degraded-channel restoration within a single model per dataset. Results show that the framework is competitive with specialised methods, particularly under severe degradation and limited reference availability. It also generates pixel-wise uncertainty maps to help identify potentially unreliable reconstructions.

This work contributes to the BE-LIGHT project, particularly Task 3.2, by supporting the development of computational tools for more reliable fluorescence microscopy analysis and quantitative assessment of biological structures relevant to rare diseases.

REFERENCES

[1] Georvasilis I, Frías Nestares M, Badosa C, Musokhranova U, Oyarzábal A, Jimenez-Mallebrera C, Rumiński J, Roldán M. Diffusion-based inference for heterogeneous multichannel fluorescence imaging. bioRxiv. 2026. doi: 10.64898/2026.10.07.753438.

[2] Caie PD, Walls RE, Ingleston-Orme A, Daya S, Houslay T, Eagle R, Roberts ME, Carragher NO. High-content phenotypic profiling of drug response signatures across distinct cancer cells. Mol Cancer Ther. 2010;9(6):1913–1926.

[3] Ljosa V, Sokolnicki KL, Carpenter AE. Annotated high-throughput microscopy image sets for validation. Nat Methods. 2012;9(7):637. doi: 10.1038/nmeth.2083.

[4] López-Márquez A, Badosa C, Enjuanes-Ruiz L, Hernández-Carabias P, Sánchez-Martín M, Cadot B, Guesmia Z, Georvasilis I, Balsells S, Blanco-Ramos M, et al. Col6a1 knock-in mice provide a promising preclinical model for collagen VI-related dystrophies. Dis Model Mech. 2026;19(1):dmm052460.

[5] Waterham HR, Koster J, van Roermund CWT, Mooyer PAW, Wanders RJA, Leonard JV. A lethal defect of mitochondrial and peroxisomal fission. N Engl J Med. 2007;356(17):1736–1741. doi: 10.1056/NEJMoa064436.

[6] Vanstone JR, Smith AM, McBride S, Naas T, Holcik M, Antoun G, Harper ME, Michaud J, Sell E, Chakraborty P, et al. DNM1L-related mitochondrial fission defect presenting as refractory epilepsy. Eur J Hum Genet. 2016;24(7):1084–1088.

[7] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process. 2004;13(4):600–612.

[8] Zhang R, Isola P, Efros AA, Shechtman E, Wang O. The unreasonable effectiveness of deep features as a perceptual metric. In: Proc IEEE Conf Comput Vis Pattern Recognit (CVPR). 2018. p. 586–595.

[9] Pachitariu M, Rariden M, Stringer C. Cellpose-SAM: superhuman generalization for cellular segmentation. bioRxiv. 2025. doi: 10.1101/2025.04.28.651001.

[10] Kynkäänniemi T, Karras T, Laine S, Lehtinen J, Aila T. Improved precision and recall metric for assessing generative models. Adv Neural Inf Process Syst. 2019;32.

[11] Heusel M, Ramsauer H, Unterthiner T, Nessler B, Hochreiter S. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. Adv Neural Inf Process Syst. 2017;30.

[12] Bińkowski M, Sutherland DJ, Arbel M, Gretton A. Demystifying MMD GANs. In: International Conference on Learning Representations (ICLR). 2018.

[13] Frangi AF, Niessen WJ, Vincken KL, Viergever MA. Multiscale vessel enhancement filtering. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer; 1998. p. 130–137.

[14] Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. Adv Neural Inf Process Syst. 2020;33.

[15] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer; 2015. p. 234–241.

[16] Nichol AQ, Dhariwal P. Improved denoising diffusion probabilistic models. In: Proceedings of the 38th International Conference on Machine Learning (ICML). PMLR; 2021. p. 8162–8171.

[17] Loshchilov I, Hutter F. Decoupled weight decay regularization. In: International Conference on Learning Representations (ICLR). 2019.

[18] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: an imperative style, high-performance deep learning library. Adv Neural Inf Process Syst. 2019;32:8024–8035.

[19] Karras T, Aittala M, Hellsten J, Laine S, Lehtinen J, Aila T. Training generative adversarial networks with limited data. Adv Neural Inf Process Syst. 2020;33:12104–12114.

[20] Weigert M, Schmidt U, Boothe T, Müller A, Dibrov A, Jain A, et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. Nat Methods. 2018;15(12):1090–1097.

[21] Shaban M, Lassoued W, Canubas K, Bailey S, Liu Y, Allen C, et al. Novel method (MAXIM) uses deep learning model to impute missing stains in multiplex images (mIF). J Clin Oncol. 2024;42(16_suppl):2580. doi: 10.1200/JCO.2024.42.16_suppl.2580.

[22] Li X, Rana P, Gide TN, Adegoke NA, Mao Y, Long G, et al. Stain imputation in multiplex immunofluorescence imaging (SIMIF) based on random channel-wise masking. In: 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI). IEEE; 2025. p. 1–5.

[23] Krull A, Buchholz TO, Jug F. Noise2Void – learning denoising from single noisy images. In: Proc IEEE/CVF Conf Comput Vis Pattern Recognit (CVPR). 2019. p. 2129–2137.

[24] Stringer C, Pachitariu M. Cellpose3: one-click image restoration for improved cellular segmentation. Nat Methods. 2025;22(3):592–599.

[25] Song J, Meng C, Ermon S. Denoising diffusion implicit models. In: International Conference on Learning Representations (ICLR). 2021.

[26] Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3:160018. doi: 10.1038/sdata.2016.18.

ANNEX

Annex A. Quality assessment and selection of high-quality reference tiles (Col6a1-KI)

For the Col6a1-KI dataset, channel-specific image quality metrics (IQMs) and thresholds were derived through an expert-guided calibration procedure. For each fluorescence channel, experts selected representative examples with clearly high-quality and clearly low-quality signal. These examples were used only to calibrate the quality control and were not used to train the diffusion model.

A broad set of candidate IQMs was computed for these images: intensity statistics (mean, standard deviation, low and high percentiles, percentile spread and interquartile range), distribution descriptors (coefficient of variation, robust coefficient of variation, skewness and kurtosis), signal descriptors (signal-to-background ratio, signal-to-noise proxy, saturation rate, low-signal rate, zero fraction, maximum-value fraction and dynamic-range usage) and focus measures (Laplacian-variance sharpness, Tenengrad sharpness and Brenner focus).

For each channel, the expert-labelled high- and low-quality examples were treated as two classes, and a shallow decision tree was fitted on the candidate IQMs as an interpretable tool to identify the metrics and threshold values that best separated the two groups. The resulting criteria were then applied automatically to the 1,877 tiles. A tile was retained as a high-quality reference only if all criteria were satisfied for all required channels, which yielded 113 reference tiles from 23 mosaics.

The selected criteria reflect the main degradation modes of the Col6a1-KI dataset: for DAPI, unstable background, loss of clear nuclear sparsity or high-frequency artefact-like texture; for Perlecan, low contrast and poor focus; and for Collagen VI, weak signal or low dynamic range. Overall, the quality-control rules target residual degradation related to weak fluorescence signal, reduced contrast, blur and unstable background.