nature portfolio

Corresponding author(s): Paul Bogdan
Last updated by author(s): 01-27-2026

Machine Learning Checklist V 1.1

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form is intended to provide structure for consistency and transparency in reporting of works using or developing Machine Learning models. Some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

1. Availability and reproducibility of Code and Data

Please select all that apply regarding the availability of the data and code used in the study.

☑ Code will be included in a CodeOcean capsule.

☑ The source code is included in the submission or available in a public repository:
https://github.com/Belis0811/Unet

☐ A compiled standalone version of the software is included in the submission or available in a public repository:
URL

☑ A test dataset and instructions/scripts for replicating the results are included in the submission or available in a public repository:
https://github.com/Belis0811/Unet

☑ A Readme file with instructions for installing and running the code is included in the submission or available in a public repository:
https://github.com/Belis0811/Unet

☑ The code is made available to reviewers during review.

☑ Pretrained models are used in the study and accessible through:
Uhttps://github.com/Belis0811/Unet

☐ Pretrained models are used in the study and are not accessible.

☑ The paper contains information on how to obtain code and data after publication.

2. Datasets

A. All data sources are listed in the paper.
   ☑ Yes
   ☐ No

B. The train, test and validation datasets are publicly available, and links/accession numbers have been provided in the manuscript or supplementary materials.
   ☐ Yes
   ☑ No

C. We have reported and discussed potential dataset biases in the paper. Where applicable, appropriate mitigation strategies were used.

☑ Yes Methods : Data Preparation (p.12) describes acquisition across multiple microscopes (Zeiss/Leica),

☐ No Explain why this information is not reported/not relevant

D. The data cleaning and preprocessing steps are clearly and fully described, either in text or as a code pipeline.

☑ Yes Methods : Data Preparation (p.12) details imaging protocol, staining, and ground-truth label generation via

☐ No Explain why this information is not reported/not relevant

E. Instances of combining data from multiple sources are clearly identified, and potential issues mitigated.

☑ Yes Methods : labels are produced by combining aligned fluorescence channels (DAPI/FITC) with SLIM

☐ No Explain why this information is not reported/not relevant

3. Model and training

A. What model architecture is the current model based on? UNet

B. A Model Card is provided[1].

☐ Yes

☑ No

C. The model clearly splits data into different sets for training (model selection), validation (hyperparameter optimization), and testing (final evaluation).

☑ Yes

☐ No

D. The method of data splitting (e.g. random, cluster- or time-based splitting, forward cross-validation) is clearly stated.

☑ Yes Methods : Datasets splits (p.13): sequences are split at the experimental-well level into train/val/test;

☐ No Explain why this information is not reported/not relevant

E. The data splitting mimics anticipated real-world applications.

☑ Yes Methods : Datasets splits (p.13): splitting by experimental wells ensures test data comes from unseen

☐ No Explain why this information is not reported/not relevant

F. The data splitting procedure has been chosen to avoid data leakage.

☑ Yes Methods : Datasets splits (p.13) explicitly states leakage control: “to avoid information leakage…partition…

☐ No Explain why this information is not reported/not relevant

[1] https://huggingface.co/docs/hub/model-cards

G. The interpretability of the model has been studied and clearly validated.

☑ Yes Results : Fig. 3 provides qualitative visual examples (inputs + ground truth) for interpreting

☐ No Explain why this information is not reported/not relevant

4. Evaluation

A. The performance metrics used are described and justified in the paper.

☑ Yes Appendix : Model Comparison and Metrics defines IoU/AUC/F1/precision/recall and states identical test

☐ No Explain why this information is not reported/not relevant

B. Cross-validation of the results is included.

☐ Yes

☑ No

C. Community-accepted benchmark datasets/tasks are used for comparisons.

☑ Yes Results : comparisons are reported on both your dataset and the E-UNet dataset, using standard cell

☐ No Explain why this information is not reported/not relevant

D. Baseline comparisons to simple/trivial models (for example, 1-nearest neighbour, random forest, most frequent class) are provided.

☑ Yes Appendix : baseline UNet learning curves (Fig. S1) and Table S2 compare against UNet; Main paper also

☐ No Explain why this information is not reported/not relevant

E. Benchmarks with current state-of-the-art are provided.

☑ Yes Results : benchmark against SAM2, E-UNet, and transformer-based baselines (µSAM, CelloType); the

☐ No Explain why this information is not reported/not relevant

F. Ablation experiments are included.

☑ Yes Appendix : UNet vs Self-Attention UNet training curves (Figs. S1–S2) and Table S2 (model complexity

☐ No Explain why this information is not reported/not relevant

G. The model has been tested on a fully independent dataset.

☑ Yes

☐ No

5. Computational resources

A. The paper contains information on hardware/computing resources that were used.

☐ Yes

☑ No

B. The paper includes information on the computational costs in terms of computation time, parallelization or carbon footprints estimates.

☑ Yes

☐ No