

# **AI-based automated active learning for discovery of hidden dynamic processes: A use case in light microscopy**

Nils Friederich<sup>1,2</sup> Angelo Yamachui Sitcheu<sup>1</sup>, Oliver Neumann<sup>1</sup>,  
Süheyla Eroğlu-Kayıkçı<sup>2</sup>, Roshan Prizak<sup>2</sup>, Lennart Hilbert<sup>2,3</sup>,  
Ralf Mikut<sup>1</sup>

<sup>1</sup> Institute for Automation and Applied Informatics

Karlsruhe Institute of Technology

Hermann-von-Helmholtz-Platz 1

76344 Eggenstein-Leopoldshafen

<sup>2</sup> Institute of Biological and Chemical Systems

Karlsruhe Institute of Technology

Hermann-von-Helmholtz-Platz 1

76344 Eggenstein-Leopoldshafen

<sup>3</sup> Zoological Institute

Karlsruhe Institute of Technology

Fritz-Haber-Weg 4

76131 Karlsruhe

## **Abstract**

In the biomedical environment, experiments assessing dynamic processes are primarily performed by a human acquisition supervisor. Contemporary implementations of such experiments frequently aim to acquire a maximum number of relevant events from sometimes several hundred parallel, non-synchronous processes. Since in some high-throughput experiments, only one or a few instances of a given process can be observed simultaneously, a strategy for planning and executing an efficient acquisition paradigm is essential. To address this problem, we present two new methods in this paper. The first method,

Encoded Dynamic Process (EDP), is Artificial Intelligence (AI)-based and represents dynamic processes so as to allow prediction of pseudo-time values from single still images. Second, with Experiment Automation Pipeline for Dynamic Processes (EAPDP), we present a Machine Learning Operations (MLOps)-based pipeline that uses the extracted knowledge from EDP to efficiently schedule acquisition in biomedical experiments for dynamic processes in practice. In a first experiment, we show that the pre-trained State-Of-The-Art (SOTA) object segmentation method Contour Proposal Networks (CPN) works reliably as a module of EAPDP to extract the relevant object for EDP from the acquired three-dimensional image stack.

## 1 Introduction

For the imaging-based assessment of dynamic processes in biomedical settings, objects of interest must be identified and relevant events that characterize the dynamic process must be recorded during their time of occurrence. Commonly, a human operator controls the imaging instrument to examine a biomedical sample using a microscope and relevant objects are found in the sample by inspection. Alternatively, the operator estimates for each object of interest the time at which an event of interest is expected to occur based on previous experience and triggers the recording of the event at that time. Nevertheless, many contemporary experiments provide several hundred relevant objects that can, in principle, be imaged in parallel. Events of interest, however, are non-synchronous and the estimation of future event times requires extensive human effort, is prone to error, and not necessarily time-efficient. These obstacles can result in unnecessarily large amounts of irrelevant data, unnecessary experimental repeats, or experimental biases inflicted by additional light exposure of the sample [18]. To address these obstacles, we present two new methods for the automated, real-time planning and execution of such experiments.

**EDP.** The traditional method of capturing all data or relying on human experience to predict future events is outdated and inefficient. Instead of relying on human experience, an accurate and comprehensive model of the dynamic

process should be created. This model should be capable of uniquely identifying a relevant object state through a process known as fingerprinting, similar to how humans do it. In biomedicine, it is crucial that this fingerprint remains consistent despite contextual changes such as noise, brightness changes, or affine transformations. By modeling the relationship between the fingerprints, the dynamic process can be represented orderly. This representation suits as an approximation of relative progress within the dynamic process. This relative progress can also be interpreted as a relative time, known as pseudo-time. By adopting this method, we can achieve efficiency in predicting future events and a deeper understanding of the dynamic process.

**EAPDP** To unlock the full potential of the EDP, a well-designed pipeline is an absolute must. Such a pipeline should be able to recognize specific states in the real world, identify relevant objects, and then calculate a pseudo-time for those states using the EDP. Armed with this knowledge, the pipeline can automatically plan and execute a new state capture for any significant event that occurs. Due to the uncertain nature of the EDP predictions, the pipeline must be able to respond to unsuccessful recordings and learn from them. This is where MLOps [1] comes in. By retraining an existing production model in accordance with the live context, MLOps ensures that the pipeline always uses a current and accurate model, resulting in a potentially better outcome in real-world experiments.

## 2 Related Work

**Object Extraction.** Basically, there are different possibilities in Computer Vision (CV) like object detection and object segmentation, to identify individual objects in an image [10] and thus extract them. In the biomedical context, many methods mostly focus on segmentation [42] with SOTA methods like StarDist [38] and CPN [46].

**Pseudo-time predictions.** A first approach for pseudo-time predictions with classical, non Deep Learning (DL) methods was presented in [14]. For extract-

ing relevant objects from the acquired image, thresholding is used as a classical CV segmentation method. Then the object's fingerprint is generated linearly with a Principal Component Analysis (PCA) [19, 20]. However, biological processes are usually not linear [5]. Therefore, recently, non-linear encodings of the dynamic processes using DL methods have become popular [17, 9, 21, 32]. For example, [21] and [9] present DL approaches to encode cell cycles and derive predefined cell phases. However, this classification-based approach does not allow for deriving continuous relations like the pseudo-time to each other directly. This continuous relation was modeled with DeepCycle in [32]. The training of DeepCycle is performed supervised. For this purpose, virtual labels are calculated based on the fluorescence intensity in specifically labeled channels of the cells. These classes can then be used as anchor points during training to determine a (relative) cell state as a pseudotime. It is important to note that the assumption that a correlation of fluorescence intensity to cell phase can be used is not always true in the biological context. A DL method that follows a comparable pseudo-time approach to the given constraints, in this paper, was presented in [17]. In [17], an autoencoder (AE) approach for pseudo-time approximation is used as a Self-Supervised Learning (SSL) approach. A DL model is used as the Variational Autoencoder (VAE) [23] encoding, from whose Hierarchical Agglomerative Clustering (HAC) and Minimum Spanning Tree (MST) code the pseudo-time is then determined. However, this approach also has a few limitations. First, a recording necessarily contains exactly one relevant object in one acquired image. Second, the entire dataset was acquired under comparable acquisition conditions, which also only contain identical positioned and oriented objects and are not able to learn affine transformations [3] between objects. Both constraints are generally not satisfied for microscopic images, such as in [27, 36]. Furthermore, this pseudo-time method was not designed as an End-to-end (E2E) model, which deprives the DL model of the ability to internally bind affine transformed objects.

**AutoEncoder.** Autoencoders are SSL methods to learn a representation from a given suitability, such as an image [49, 47]. For example, the autoencoder can be represented by a Conventional AutoEncoder (CAE) [49] and/or a VAE [23]. Especially recently, masked autoencoders (MAEs) [15] have become more



popular than CAE because of their ability for a better visual representation learning [49], either using a transformer-based approach [15] or the Convolutional Neural Network (CNN)-based approach [47]. However, since in [47] the higher efficiency of ConvNeXt V2 is shown, this model is chosen for this work. In addition to the pure learning of a visual representation in the form of a fingerprint, relations between the fingerprints can also be learned, e.g. with VAE [17].

**SSL** To train the DL model in a supervised manner, labeled data are generally rare in the biomedical domain [13]. There are DataBases (DBs) like BioMed-Image.io [30] or several challenges with own datasets [26, 2, 28]. However, especially for biological datasets with sometimes hundreds of relevant objects in an image, the datasets are often limited to the 2D case. Furthermore, in the context of this work, labeling the relevant events with pseudo-time stamps is only approximate, demanding, error-prone and time-consuming. Therefore, unsupervised or SSL methods are often used in the biomedical context [43]. Thereby, Active Learning (AL) [41] is used to selectively integrate expert knowledge into the learning process. Since AL aims to keep the number of interactions to a minimum [7, 33], data-efficient learning is preferable. For example, existing datasets from a related context can be identified and leveraged to train more robust models in transfer learning [11, 29, 48]. In order to be able to use possibly directly existing pre-trained models from similar contexts, a new concept was developed in [48].

## 3 Methodology

### 3.1 EDP

For the modeling of the dynamic process, a new concept is introduced with EDP. The new concept of EDP is based on an AE and is visualized in Figure 1. The basic idea of the new concept is to separate the generation of the fingerprint from the learning of the relation as a representation between all states. The fingerprint generation is done using a MAE as an evolution of CAE.

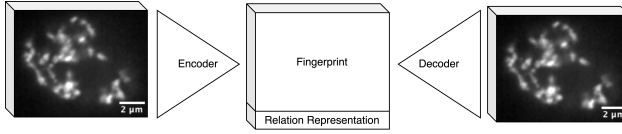


Figure 1: Visualization of the EDP model. A 3D object is transformed by an encoder into a fingerprint (like MAE) and into a relation representation between the fingerprints (like VAE). The example object is a recording of the DNA channel of a nucleus from an internal zebrafish embryo dataset. A scale bar of  $2\mu\text{m}$  is indicated at the bottom of the input/output image.

Specifically, the SOTA MAE-based method ConvNeXt V2 is chosen. During the learning process, a maximum recovery of the encoded image is aimed in accordance with SSL. According to the challenge posed by biological objects, a context-independent representation is required. For this purpose, the images can be modified during the learning process through Data Augmentation [16] techniques like Rotations, reflections, contrast adjustments or noise additions. In addition to the fingerprint, the relation must also be learned as the actual modeling of the dynamic process. For this purpose, a VAE-like modeling is used by learning the uncertainty  $v$  in addition to the circle angle  $\alpha$ . The assumption is that objects succeeded each other in the dynamic process with the relative distance  $t$  corresponding to this relative distance and differ in the same ratio in the circle representation. Such an exemplary circle representation is shown in Figure 2 using a cell division process of the zebrafish embryo. The state of the cell after cell division is visualized at 00 o'clock and up to the state of the cell just before cell division at 11 o'clock. This corresponds to a relative distance of  $\sim 0.92$  (normalized between  $[0,1)$ ). This temporal difference must also be valid in reality for the temporal distance according to the model statement.

### 3.2 EAPDP

The new EDP module is integrated as a module into the new MLOps-based pipeline EAPDP. The pipeline concept is visualized in Figure 3 and contains nine other modules besides the EDP module. Each of these ten modules is

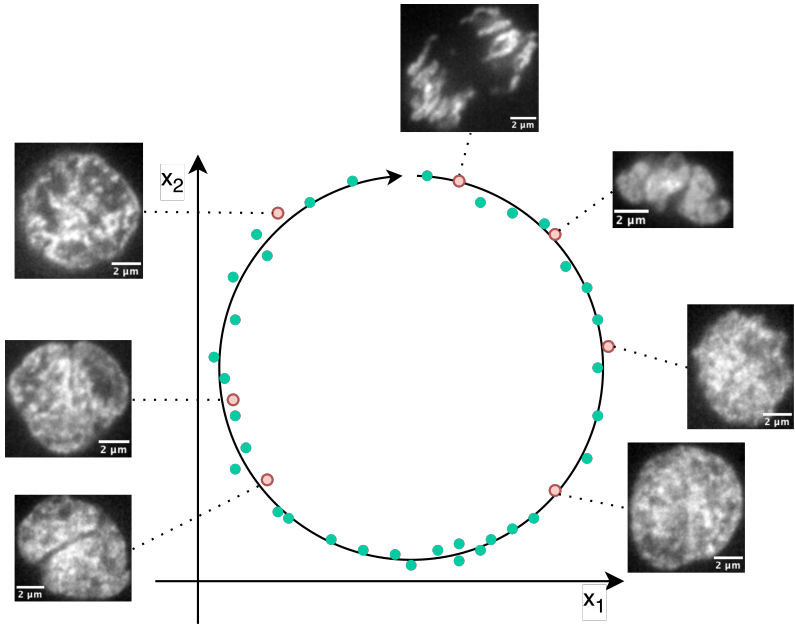


Figure 2: Example of a 2D feature space representation for an encoded dynamic process. Each point represents an encoded image. The circle serves as an estimation for the positioning of points in its vicinity. It's worth noting that due to the presence of uncertainty, the points may not always be precisely on the circle, but rather in its proximity. For the seven red dots, example images of cell nuclei from zebrafish embryos at various stages of cell division are shown. A scale bar of  $2\ \mu m$  is indicated at the bottom of each example image. The images are from an internal dataset.

briefly described below. The explanation of the modules and their relationships to each other is based on the pipeline visualization of Figure 3.

**Microscope setup.** In the EAPDP, the microscope is used as an actuator to the real-world environment represented by a biomedical sample. For this purpose, all microscope components relevant to image acquisition and the microscope accessories, such as lasers in the case of a laser scanning microscope, must be controllable via appropriate interfaces of the specific microscope setup. In addition, the microscope must be able to react on given com-

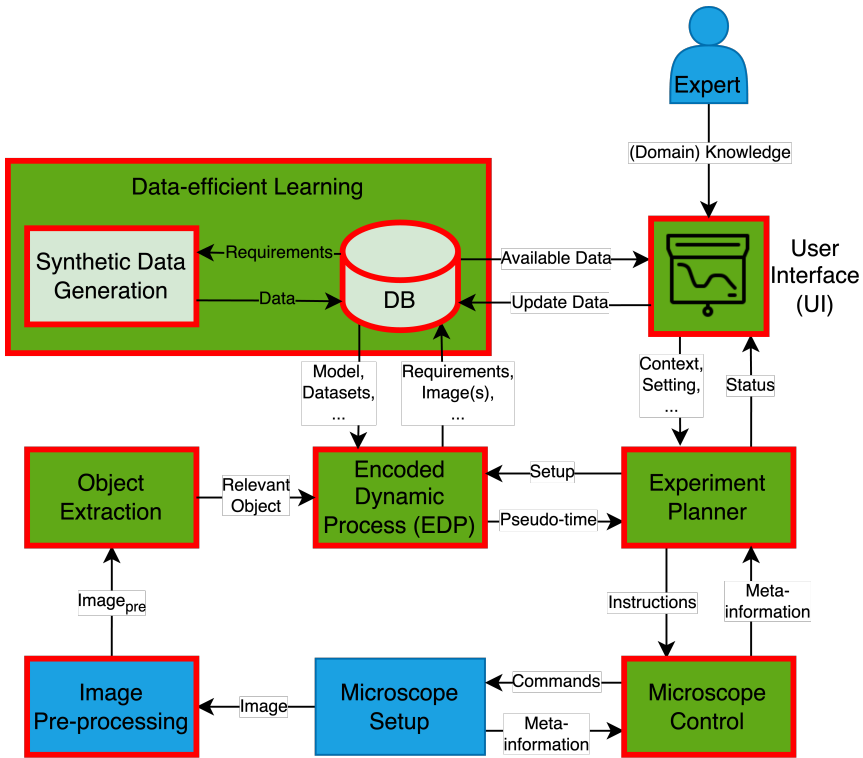


Figure 3: MLOps pipeline with the new EDP module. AI-based MLOps modules are marked with a green background, non AI-based ones with a blue background. Additionally, all modules marked with a red border must be newly developed or only partially adapted from existing methods.

mands like image acquisition or requested meta-information like the objective position in a standardized way.

**Image Pre-processing.** To optimize the analysis of dynamic processes in the biomedical environment, the raw images acquired through experiments must be pre-processed according to the microscope setup and the context of the targeted event. This may involve methods such as cropping, contrast adjustment, or denoising. Various libraries, such as Albumentations [4], offer pre-processing

methods that can be used to improve the quality of the images and optimize their analysis by other modules in the machine learning operation pipeline.

**Object Extraction.** During an acquisition, the relevant object and the surrounding context are captured. In order to better analyze the object, it is necessary to separate it from the surrounding context. The extraction from the whole image is done via pre-trained segmentation methods. To find a suitable method, we compare the SOTA cell segmentation algorithms StarDist and CPN using a microscopic dataset in a first experiment. Importantly, the actual pseudo-time determination cannot be performed if both methods' segmentation is insufficient. Therefore, this submodule is of particular importance. Because (well) labeled data are generally scarce in the biological context, this work evaluates generalization performance during inference with already pre-trained models on new, unknown images. Since there are only pre-trained weights for 2D segmentation for both methods, the dataset was split into 2D images along the z-axis.

**EDP.** The EDP module gets the extracted object and should pass the pseudo-time to the experiment planner. In order to do this, the module is equipped by the Experiment planner before with the appropriate experiment setup. With the setup, the EDP model can then query according to its existing knowledge like pre-trained models in the context of data-efficient learning. If no weights are available, training can also be done with/without AL as specified by the expert. After successful training, the model is passed to the DB with the appropriate required metadata for possible further use. Then, when the inference with the original extracted relevant object has been determined, the results are passed to the expert planner accordingly. The recorded inference image is also sent to the Data-efficient Learning module and stored in its DB.

**Experiment planner.** The Experiment planner is the central module of the experiment automatization. As input it gets the pseudo-times for the recorded objects. Based on the experiment context, including interesting events, the Experiment planner can plan future experiments with utmost precision. Once

the plan is set, the Experiment planner gives the microscope the command to ensure that the image captures the object's state at the right time, leaving no room for errors. Additionally, it can query the state of the microscope to ensure that there was no hardware drift, such as when moving to the object position. All the information about the experiment's state is then passed on to the User Interface (UI), ensuring that all aspects of the experiment are under control.

**UI.** The UI is the interface between the expert and the MLOps pipeline. On the one hand, simple interactions can be provided, such as displaying meta-information or adapting the experimental context, e.g. the cell classes that occur. On the other hand, much more complex interactions such as result justifications of DL models can be represented through Explainable Artificial Intelligence (XAI) [34] or expert knowledge can be brought into the pipeline within the context of AL. With XAI, the expert should be able to understand better the processes in the DL models used and why decisions were made, e.g. for event detection. This helps the expert to eliminate potential errors like unfavorable experiment settings at an early stage. Such XAI methods can be realized using a library like PyTorch Captum [25]. For AL, only if the expert can capture the actual state in the best possible way, the expert can transfer his domain knowledge to the method in the best possible way and support the method. For example, points, boxes, or entire segmented regions can be passed to the method as hints. For this purpose, a custom segmentation module can be developed based on existing AL labeling platforms like ObiWanMicrobi [40] or Karlsruhe Image Data Annotation (KaIDA) [37].

**Expert (Domain) Knowledge.** The domain knowledge contributed by the expert to the MLOps pipeline can take several forms. For example, the context of the experiment with a specific cell class can improve a more efficient event detection module. Furthermore, knowledge can be injected, e.g. by labeling in the context of AL. For this, the expert must ensure the quality of the injected domain knowledge with maximum correctness. Incorrect information can affect the learning processes in the network.

**Data-efficient Learning.** To minimize AL interactions with the human expert, as much existing knowledge as possible is reused. To this end, building on [48], a new AE-based fingerprinting approach for datasets and Machine Learning (ML) models is being implemented to reuse as much knowledge as possible. For this, from a DB, context-given requirements can query existing knowledge. If no data is available, it can be created synthetically e.g. with biophysical simulations like Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [44].

**Microscope control.** In order for the planned experiment to be automated and performed in real-time, a corresponding software library is needed to control the microscope. Since the first release in 2010,  $\mu$ Manager has been used for this purpose as one of the SOTA open-source solutions [6, 22, 31, 45]. Therefore, this is also used in this work.

### 3.3 Exemplary Use Cases

Example use cases for the presented EAPDP with the EDP are presented using record extracts in Figure 4 below. A first use case is shown in Figure 4a and represents the temporal sorting of RiboNucleic Acid (RNA) Polymerase II (Pol II) clusters that occur in the nuclei of pluripotent zebrafish embryos. A method for this use case has already been presented in [14]. A comparison of the pipeline based on classical ML methods with our DL-based EDP method allows a direct statement about limitations or improvements of our approach. The second use-case in Figure 4b is the recording of cell divisions in pluripotent zebrafish embryos, where the time of reaching a new division stage and thus the regions of an event of interest need to be extrapolated. A final biological application from the field of microbiology is presented in Figure 4c. In this example, one interesting event could be the state at which  $n$  microbes reach the recording region. For this purpose, a modeling of the cell division process with EDP can be used to plan the experiment accordingly and automatically record the event of interest at a time  $t$ . The modeling of the cell division process with EDP can be used for this purpose.

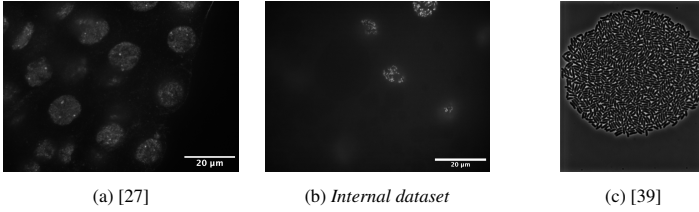


Figure 4: Three exemplary images from biological datasets for dynamic processes. Figure 4a shows cell nuclei of zebrafish embryos with marked Pol II Ser5P clusters. Figure 4b shows the DeoxyriboNucleic Acid (DNA) of zebrafish embryos nuclei. In these first two images, a scale bar of  $20\mu\text{m}$  is shown in the lower left. The last Figure 4c shows a microbial cell division state.

In addition to these biological use cases, other use cases are also possible, e.g. in medicine. For example, by modeling a tumor accordingly, a prediction can be made about the relative stage. Consequently, a therapy concept such as surgery or medication can be tailored to the patient.

## 4 Experiments

The comparison of segmentation algorithms was performed on Helmholtz AI Compute REsources (HAICORE) resources equipped with Intel Xeon Platinum 8368 Central Processing Units (CPUs) and an Nvidia A100-40 Graphics Processing Unit (GPU) [24]. The operating system utilized was Red Hat Enterprise Linux (RHEL) version 8.6.

### 4.1 Dataset

The internal microscope dataset from Figure 4b is used to compare the segmentation algorithms. This dataset was chosen over the other two example datasets from Figures 4a and 4c because of the challenging, frayed structure of the nuclei as the relevant image object. This is because the fibrillar structure of the nuclei sometimes deviates strongly from their typical ellipsoidal shape as in Figure 4a due to individually advanced cytokinesis. This poses a challenge because contiguous pixel regions are not trivially identifiable and correct



boundary segmentation is a challenge. With microbeSEG [35], a working SOTA solution for microbes like in Figure 4c also already exists.

For the dataset in Figure 4b, zebrafish embryo DNA was imaged. DNA was stained with 1:10000 5'-TMR Hoechst in TDE or glycerol. Confocal z-sections were obtained using a commercial instant SIM microscope (iSIM, VisiTech). A Nikon 100x oil immersion objective (NA 1.49, SR HO Apo TIRF 100xAC Oil) and a Hamamatsu ORCA-Quest camera were used for image acquisition. In accordance with a common problem in biology, no labels exist for this dataset. According to the desired 2D segmentation, the 3D images are split into 2D images along the z-axis.

## 4.2 Object extraction

These 2D images were then segmented using each of the two methods. In the following, the results are evaluated qualitatively because of the non-existent labels. Therefore, the results are shown in Figure 5. The comparison of the original image in Figure 5a with the StarDist prediction in Figure 5b, shows that StarDist cannot well segment semantically related objects as the nucleus in the upper area. For the method designs with center prediction, StarDist focused primarily on segmenting ellipsoidal objects from [12] and was trained only on these. The cell detection was designed to be more flexible and additionally trained on a more heterogeneous set of non-elliptical cells such as *MCF7* from the dataset [8]. This leads to better generalization and results in qualitatively evaluated good initial segmentation performance on this most challenging of our datasets from Figure 4.

Thus, we could show that CPN is a good pre-trained SOTA approach for extracting the relevant objects from the 2D decomposition. The 2D segmentations can be reassembled back to 3D segmentations in post-processing, e.g. using Nearest Neighbor. Based on this, the further submodules of EDP can be developed in future work and the presented MLOps pipeline can be built upon it.

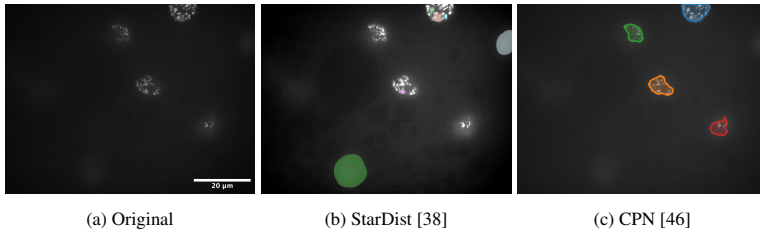


Figure 5: Comparison of segmentation predictions for the two SOTA methods StarDist [38] and CPN [46]. Figure 5a represents the original image, duplicated from Figure 4b. In Figure 5b and Figure 5c, the predictions of pre-trained StarDist or CPN are then shown. The predictions are highlighted differently for better visual differentiation depending on the method used.

## 5 Conclusion and Further Work

In this work, we motivated that due to the large number of parallel non-synchronous dynamic processes, a novel concept for automated planning and execution of two novel DL-based approaches is essential. First, the EDP was introduced to model dynamic processes and derivate a pseudo-time for a given object state. The pseudo-time prediction can then be used with the EAPDP for real-time experiment automation. We explained the EDP realized within the MLOps pipeline by an AE and trained using SSL with AL. At the same time, the key advantage of higher execution speed and lower human cost while minimizing user interactions with data-efficient learning was highlighted. Finally, as a first Proof of Concept (PoC), we showed the necessary pre-processing step for the EDP to extract the relevant objects based on good inference results of CPN.

However, the lack of pre-trained weights for 3D segmentation was a drawback of the segmentation experiments. However, since the fragmented objects are partially reconnected along the z-axis, this could simplify the problem and improve accuracy. This will be done as soon as appropriate weights are available. In addition, a suitable affine-invariant 3D AE needs to be developed for use within the EDP method. In this context, further research is needed to investigate whether the ConvNeXt V2 is suitable for 3D segmentation, also

from an efficiency perspective. Of course, the modules of the MLOps pipeline must be implemented accordingly.

## Acknowledgments

This project is funded by the Helmholtz Association under the Program Natural, Artificial and Cognitive Information Processing (NACIP) and the Helmholtz Association's Initiative & Networking Fund through Helmholtz AI. All experiments were performed on the HAICORE. We sincerely thank all of them for supporting our research.

The authors have accepted responsibility for the entire content of this manuscript and approved its submission. We describe the individual contributions of N. Friederich (NF), A. Yamachui Sitcheu (AYS), O. Neumann (ON), S. Eroglu-Kayıkçı (SEK), R. Prizak (RP), L. Hilbert (LH), R. Mikut (RM): Conceptualization: NF, LH, RM; Methodology: NF, LH, RM; Software: NF; Investigation: NF; Resources: SEK, RP, LH; Writing – Original Draft: NF; Writing – Review & Editing: NF, AYS, ON, SEK, RP, LH, RM; Supervision: LH, RM; Project administration: RM; Funding Acquisition: LH, RM

## References

- [1] S. Alla, S. K. Adari, S. Alla, and S. K. Adari. What is MLOps? *Beginning MLOps with MLFlow: Deploy Models in AWS SageMaker, Google Cloud, and Microsoft Azure*, pages 79–124, 2021.
- [2] M. Aubreville, C. Bertram, K. Breininger, S. Jabari, N. Stathonikos, and M. Veta. MITosis DOmain Generalization Challenge 2022. <https://doi.org/10.5281/zenodo.6362337>, 03 2022.
- [3] E. Begelfor and M. Werman. Affine Invariance Revisited. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, volume 2, pages 2087–2094, 2006.

- [4] A. V. Buslaev, A. Parinov, E. Khvedchenya, V. I. Iglovikov, and A. A. Kalinin. Albumentations: fast and flexible image augmentations. *CoRR*, abs/1809.06839, 2018.
- [5] J. Carballido-Landeira and B. Escribano. *Nonlinear dynamics in biological systems*. Springer, 2016.
- [6] C. Conrad, A. Wünsche, T. H. Tan, J. Bulkescher, F. Sieckmann, F. Verissimo, A. Edelstein, T. Walter, U. Liebel, R. Pepperkok, et al. Micropilot: automation of fluorescence microscopy-based imaging for systems biology. *Nature Methods*, 8(3):246–249, 2011.
- [7] A. M. Das, G. Bhatt, M. M. Bhalerao, V. R. Gao, R. Yang, and J. Bilmes. Continual Active Learning. <https://openreview.net/forum?id=GC5MsCxrU->, 2023.
- [8] C. Edlund, T. R. Jackson, N. Khalid, N. Bevan, T. Dale, A. Dengel, S. Ahmed, J. Trygg, and R. Sjögren. LIVECell—A large-scale dataset for label-free live cell segmentation. *Nature Methods*, 18(9):1038–1045, 2021.
- [9] P. Eulenberg, N. Köhler, T. Blasi, A. Filby, A. E. Carpenter, P. Rees, F. J. Theis, and F. A. Wolf. Reconstructing cell cycle and disease progression using deep learning. *Nature Communications*, 8(1):463, 2017.
- [10] X. Feng, Y. Jiang, X. Yang, M. Du, and X. Li. Computer vision algorithms and hardware implementations: A survey. *Integration*, 69:309–320, 2019.
- [11] P. Godau and L. Maier-Hein. Task Fingerprinting for Meta Learning inBiomedical Image Analysis. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part IV 24*, pages 436–446. Springer, 2021.
- [12] A. Goodman, A. Carpenter, E. Park, jlefman nvidia, Josette-BoozAllen, Kyle, Maggie, Nilofer, P. Sedivec, and W. Cukierski. 2018 Data Science Bowl. <https://kaggle.com/competitions/data-science-bowl-2018>, 2018.

- [13] H. A. Gündüz, M. Binder, X.-Y. To, R. Mreches, B. Bischl, A. C. McHardy, P. C. Münch, and M. Rezaei. A self-supervised deep learning method for data-efficient training in genomics. *Communications Biology*, 6(1):928, 2023.
- [14] H. Hajiabadi, I. Mamontova, R. Prizak, A. Pancholi, A. Koziolk, and L. Hilbert. Deep-learning microscopy image reconstruction with quality control reveals second-scale rearrangements in RNA polymerase II clusters. *PNAS nexus*, 1(3):pgac065, 2022.
- [15] K. He, X. Chen, S. Xie, Y. Li, P. Dollár, and R. Girshick. Masked autoencoders are scalable vision learners. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16000–16009, 2022.
- [16] A. Hernandez-Garcia and P. König. Data augmentation instead of explicit regularization. *arXiv preprint arXiv:1806.03852*, 2018.
- [17] J. Hong, S. K. Kang, I. Alberts, J. Lu, R. Sznitman, J. S. Lee, A. Rominger, H. Choi, K. Shi, and A. D. N. Initiative. Image-level trajectory inference of tau pathology using variational autoencoder for Flortaucipir PET. *European Journal of Nuclear Medicine and Molecular Imaging*, 49(9):3061–3072, 2022.
- [18] J. Icha, M. Weber, J. C. Waters, and C. Norden. Phototoxicity in live fluorescence microscopy, and how to avoid it. *BioEssays*, 39(8):1700003, 2017.
- [19] J. E. Jackson. *A user’s guide to principal components*. John Wiley & Sons, 2005.
- [20] I. T. Jolliffe. *Principal component analysis for special types of data*. Springer, 2002.
- [21] A. Jose, R. Roy, D. Moreno-Andrés, and J. Stegmaier. Automatic Detection of Cell-cycle Stages using Recurrent Neural Networks. *bioRxiv*, pages 2023–02, 2023.

- [22] R. A. Kellogg, R. Gómez-Sjöberg, A. A. Leyrat, and S. Tay. High-throughput microfluidic single-cell analysis pipeline for studies of signaling dynamics. *Nature Protocols*, 9(7):1713–1726, 2014.
- [23] D. P. Kingma and M. Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.
- [24] KIT. HAICORE - Hardware Overview. <https://www.nhr.kit.edu/userdocs/haicore/hardware/>, Sep 2022.
- [25] N. Kokhlikyan, V. Miglani, M. Martin, E. Wang, B. Alsallakh, J. Reynolds, A. Melnikov, N. Kliushkina, C. Araya, S. Yan, and O. Reblitz-Richardson. Captum: A unified and generic model interpretability library for PyTorch. *CoRR*, abs/2009.07896, 2020.
- [26] K. Lee, H. Byun, and H. Shim. Cell Segmentation in Multi-modality High-Resolution Microscopy Images with Cellpose. In *Proceedings of The Cell Segmentation Challenge in Multi-modality High-Resolution Microscopy Images*, volume 212 of *Proceedings of Machine Learning Research*, pages 1–11. PMLR, 28 Nov–09 Dec 2023.
- [27] I. Mamontova, A. Pancholi, R. Prizak, and L. Hilbert. Microscopy data: interaction of the gene *klf2b* with RNA polymerase II clusters during early zebrafish embryo development. <https://doi.org/10.5281/zenodo.5268833>, 08 2021.
- [28] M. Maška, V. Ulman, P. Delgado-Rodriguez, E. Gómez-de Mariscal, T. Nečasová, F. A. Guerrero Peña, T. I. Ren, E. M. Meyerowitz, T. Scherr, K. Löffler, et al. The Cell Tracking Challenge: 10 years of objective benchmarking. *Nature Methods*, pages 1–11, 2023.
- [29] M. Molina-Moreno, M. P. Schilling, M. Reischl, and R. Mikut. Automated Style-Aware Selection of Annotated Pre-Training Databases in Biomedical Imaging. In *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, pages 1–5, 2023.
- [30] W. Ouyang, F. Beuttenmueller, E. Gómez-de Mariscal, C. Pape, T. Burke, C. Garcia-López-de Haro, C. Russell, L. Moya-Sans, C. de la Torre-Gutiérrez, D. Schmidt, et al. BioImage Model Zoo: A

Community-Driven Resource for Accessible Deep Learning in BioImage Analysis. *bioRxiv*, pages 2022–06, 2022.

- [31] H. Pinkard, N. Stuurman, K. Corbin, R. Vale, and M. F. Krummel. Micro-Magellan: open-source, sample-adaptive, acquisition software for optical microscopy. *Nature Methods*, 13(10):807–809, 2016.
- [32] L. Rappez, A. Rakhlin, A. Rigopoulos, S. Nikolenko, and T. Alexandrov. DeepCycle reconstructs a cyclic cell cycle trajectory from unsegmented cell images using convolutional neural networks. *Molecular Systems Biology*, 16(10):e9474, 2020.
- [33] P. Ren, Y. Xiao, X. Chang, P.-Y. Huang, Z. Li, B. B. Gupta, X. Chen, and X. Wang. A Survey of Deep Active Learning. *ACM Computing Surveys (CSUR)*, 54(9):1–40, 2021.
- [34] W. Saeed and C. Omlin. Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities. *Knowledge-Based Systems*, 263:110273, 2023.
- [35] T. Scherr, J. Seiffarth, B. Wollenhaupt, O. Neumann, M. P. Schilling, D. Kohlheyer, H. Scharr, K. Nöh, and R. Mikut. microbeSEG: A deep learning software tool with OMERO data management for efficient and accurate cell segmentation. *PLOS ONE*, 17(11):1–14, 11 2022.
- [36] T. Scherr, J. Seiffarth, B. Wollenhaupt, O. Neumann, M. P. Schilling, D. Kohlheyer, H. Scharr, K. Nöh, and R. Mikut. microbeSEG dataset. <https://doi.org/10.5281/zenodo.6497715>, 04 2022.
- [37] M. P. Schilling, S. Schmelzer, L. Klinger, and M. Reischl. KaIDA: a modular tool for assisting image annotation in deep learning. *Journal of Integrative Bioinformatics*, 19(4):20220018, 2022.
- [38] U. Schmidt, M. Weigert, C. Broaddus, and G. Myers. Cell Detection with Star-Convex Polygons. In *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018 - 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II*, pages 265–273, 2018.

- [39] J. Seiffarth, L. Blöbaum, K. Löffler, T. Scherr, A. Grünberger, H. Scharr, R. Mikut, and K. Nöh. Data for - Tracking one in a million: Performance of automated tracking on a large-scale microbial data set. <https://doi.org/10.5281/zenodo.7260137>, 10 2022.
- [40] J. Seiffarth, T. Scherr, B. Wollenhaupt, O. Neumann, H. Scharr, D. Kohlheyer, R. Mikut, and K. Nöh. ObiWan-Microbi: OMERO-based integrated workflow for annotating microbes in the cloud. *bioRxiv*, pages 2022–08, 2022.
- [41] B. Settles. *Active Learning Literature Survey*. University of Wisconsin-Madison Department of Computer Sciences, 2009.
- [42] P. Shrestha, N. Kuang, and J. Yu. Efficient end-to-end learning for cell segmentation with machine generated weak annotations. *Communications Biology*, 6(1):232, 2023.
- [43] R. Sun, J. Wu, Y. Miao, L. Ouyang, and L. Qu. Progressive 3D biomedical image registration network based on deep self-calibration. *Frontiers in Neuroinformatics*, 16:932879, 2022.
- [44] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comp. Phys. Comm.*, 271:108171, 2022.
- [45] S. Tosi, A. Lladó, L. Bardia, E. Rebollo, A. Godo, P. Stockinger, and J. Colombelli. AutoScanJ: A Suite of ImageJ Scripts for Intelligent Microscopy. *Frontiers in Bioinformatics*, 1:627626, 2021.
- [46] E. Upschulte, S. Harmeling, K. Amunts, and T. Dickscheid. Contour proposal networks for biomedical instance segmentation. *Medical Image Analysis*, 77:102371, 2022.
- [47] S. Woo, S. Debnath, R. Hu, X. Chen, Z. Liu, I. S. Kweon, and S. Xie. ConvNeXt V2: Co-designing and Scaling ConvNets with



Masked Autoencoders. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16133–16142, 2023.

- [48] A. Yamachui Sitcheu, N. Friederich, S. Baeuerle, O. Neumann, M. Reischl, and R. Mikut. MLOps for Scarce Image Data: A Use Case in Microscopic Image Analysis. In *Proceedings 33. Workshop Computational Intelligence*, volume 1, 11 2023. Submitted manuscript.
- [49] C. Zhang, C. Zhang, J. Song, J. S. K. Yi, K. Zhang, and I. S. Kweon. A survey on masked autoencoder for self-supervised learning in vision and beyond. *arXiv e-prints*, pages arXiv–2208, 2022.