DeepMuCS: A Framework for Co-culture Microscopic Image Analysis: From Generation to Segmentation

Nabeel Khalid*, Mohammadmahdi Koochali*, Vikas Rajashekar*, Mohsin Munir*, Christoffer Edlund†, Timothy R Jackson¶, Johan Trygg†‡, Rickard Sjögren†‡, Andreas Dengel*§, Sheraz Ahmed*
*German Research Center for Artificial Intelligence (DFKI) GmbH, Kaiserslautern 67663, Germany
†Sartorius Corporate Research, Sweden
‡Computational Life Science Cluster (CLiC), Umeå University, Sweden
§Technische Universität Kaiserslautern, Kaiserslautern 67663, Germany
¶Arthrex California Technology, Santa Barbara, CA, USA

***Abstract*—Discrimination between cell types in the co-culture environment with multiple cell lines can assist in examining the interaction between different cell populations. Identifying different cell cultures in addition to cell segmentation in co-culture is essential for understanding the cellular mechanisms associated with disease states. In drug development, biologists are more interested in co-culture models because they replicate the tumor environment in vivo better than the monoculture models. Additionally, they have a measurable effect on cancer cell response to treatment. Co-culture models are critical for designing a drug with maximum efficacy on cancer while minimizing harm to the rest of the body. In the past, there existed minimal progress related to cell-type aware segmentation in the monoculture and no development whatsoever for the co-culture. The introduction of the LIVECell dataset has allowed us to perform experiments for cell-type-aware segmentation. However, it is composed of microscopic images in a monoculture environment. This paper presents a framework for co-culture microscopic image data generation, where each image can contain multiple cell cultures. The framework also presents a pipeline for culture-dependent cell segmentation in co-culture microscopic images. The extensive evaluation revealed that it is possible to achieve cell-type aware segmentation in co-culture microscopic images with good precision.**

***Index Terms*—biomedical, healthcare, deep learning, cell segmentation, co-culture**

I. Introduction

In drug development, biologists are interested in co-culture models because they have a measurable effect on cancer cell response to treatment. In the study [1], prostate cancer cell proliferation is inhibited by several different drugs in monoculture, but to a lesser extent in co-culture with stromal cells. This suggests that monoculture conditions alone are insufficient to select a suitable treatment for prostate cancer, making co-culture much more relevant for drug discovery.

Deep learning-based approaches are showing promising results in microscopic image analysis [2]–[4] and datasets to train instance segmentation models in monoculture environments, such as EVICAN60 [5] and LIVECell [3], are now available to further explore deep learning-based approaches for cell-type aware segmentation. In the cell biology domain, there is no co-culture microscopic dataset and method available. To tackle that issue, this paper presents a pipeline to generate multiple subsets of synthetic co-culture microscopic images using the LIVECell dataset.

Compared to multi-class segmentation in natural images [6], segmenting multiple cell types in microscopic images suffer from certain challenges including low contrast as well as irregularly shaped and overlapping cells of different cell types, making segmentation challenging. Multiple cell types in the same microscopic image can possess comparable properties like low contrast, overlapping cells, and unclear boundaries. In the proposed pipeline, parameters for features extraction and anchor sizes are proposed to detect, segment, and classify different cell cultures with good precision using variations of the LIVECell [3] dataset. We propose DeepMuCS: a framework for co-culture microscopic image analysis. DeepMuCS framework is further divided into two modules i.e., for synthetic co-culture microscopic data generation (DeepMuCS-Generation) and to perform cell type-aware segmentation in co-culture microscopic images (DeepMuCS-Segmentation). The main contributions of this study are as follows:

- A framework to generate co-culture microscopic images, and deep learning based cell type-aware segmentation in co-culture environment.
- Extensive evaluation of proposed method using different variations of the LIVECell [3] dataset.

II. DeepMuCS-Genration: Proposed Co-culture Microscopic Data generation Pipeline

Fig. 1 provides a system overview of the proposed co-culture microscopic data generation pipeline. Synthetic co-culture images are generated using the original LIVECell dataset. Initially, cell instances from different cell cultures are extracted using their segmentation annotations. Backgrounds

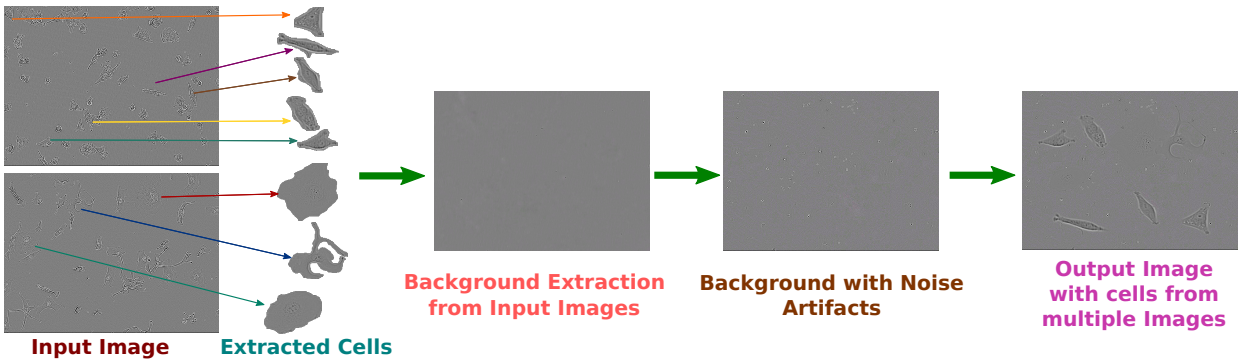

Fig. 1: System overview of DeepMuCS-Generation. Images from different cell cultures are passed to the proposed pipeline, which outputs the co-culture images.

are generated by initially filtering out the cells from the original images. Subsequently, noise artifacts are extracted from the original images and pasted on the filtered image on a random basis to make the synthetic images realistic. As a last step, extracted cell instances from different cell cultures are pasted on the generated background images to obtain the synthetic co-culture images.

III. DeepMuCS-Segmentation: Proposed Cell Segmentation Pipeline for Co-culture Microscopic Images

Fig. 2 provides a system overview of the proposed cell segmentation pipeline for co-culture microscopic images. The pipeline is divided into three blocks.

A. Feature Extraction

The purpose of this block is to extract feature maps from the input image at different scales. The feature extraction module of our proposed methodology is composed of Feature Pyramid Network (FPN) [7] along with ResNeSt-200 [8]. FPN combines the low resolution, semantically strong features with high resolution, semantically weak features. It takes a single-scale image as an input and outputs feature maps of proportional size at multiple levels by operating on a bottom-up pathway, top-down pathway, and lateral connections. The bottom-up pathway uses a normal feed-forward CNN architecture to compute a hierarchy of features consisting of feature maps at various scales. The output of each CNN layer is used later in the top-down pathway via lateral connections. We are using ResNeSt-200 [8] with deformable convolution [9] as a feed-forward CNN architecture in the bottom-up pathway of our approach. The output of each convolution layer of ResNeSt is used in the top-down pathway which constructs higher resolution layers from the semantic rich layer. As the final task, the FPN applies a 3x3 convolution operation on each merged map to overcome the aliasing effect after the upsampling to generate the final feature map.

B. Object Region Detection and Groundtruth Matching

Multi-scale features from the backbone network are passed onto a Regional Proposal Network (RPN), which detects regions that contain objects and matches them to the groundtruth. This process is performed by generating anchor boxes on the input image which are then matched to the groundtruth by taking Intersection over Union (IoU) between anchors and groundtruth. If IoU is larger than the defined threshold of 0.7, the anchor is linked to one of the groundtruth boxes and assigned to the foreground. If the IoU is greater than 0.3, it is considered background and otherwise ignored. The anchor strides and aspect ratio parameter used to detect and segment objects in MS-COCO [6] dataset overlooks most of the small cell instances when transferred to this task. Unlike MS-COCO [6] and other commonly used image datasets, the area of some cells especially BV-2 cell culture in the LIVECell [3] dataset is exceedingly small. After extensive experimentation, the anchor sizes and anchor aspect ratios were selected that fit the task. The details about the anchor parameters are given in Section V.

C. Prediction Head

The job of the prediction head is to predict the class, bounding box, and binary mask for each region of interest. We are using Cascade Mask R-CNN [10] as the prediction head, which is an extension of Cascade R-CNN by adding a mask branch to the cascade. Cascade Mask R-CNN addresses the problem of making predictions that are more accurate on a pixel level. Cascade Mask R-CNN [10] is a multi-stage network with the IoU threshold increasing for each stage to refine the final output. In the proposed methodology, the segmentation branch is added at the last stage of the Cascade R-CNN. The box head classifies the object within the ROI and fine-tunes the shape and position of the box. The mask head is composed of a small Fully Convolutional Network (FCN) applied to each ROI, which predicts a segmentation mask in a pixel-to-pixel manner to achieve the task of instance segmentation.

IV. Dataset

In the cell segmentation domain, all datasets either have images for one cell culture or do not differentiate between different cell cultures. Only two datasets have multiple cell lines in the monoculture microscopic images i.e., EVICAN60

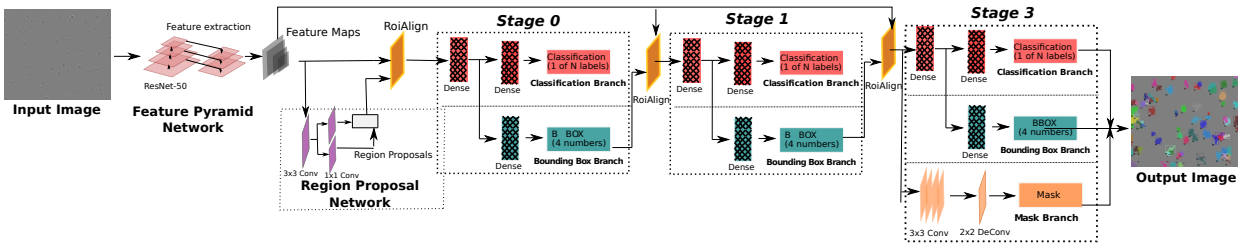

Fig. 2: System overview of DeepMuCS-Segmentation. Input image is passed to the proposed pipeline and the output image with cell-type aware detection and segmentation is produced.

[5] and the LIVECell [3]. EVICAN60 dataset is partially annotated and the instances per class are exceptionally low [4]. LIVECell dataset, on the other hand, consists of 5,239 fully annotated, expert-validated, phase contrast microscopy images with a total of 1,686,352 individual cells annotated from eight different cell cultures. That is the reason we opted for the LIVECell dataset for this study. Synthetic co-culture images are generated using the original LIVECell dataset as described in Section II. We have generated three different subsets for training namely DeepMuCS800, DeepMuCS1600, and DeepMuCS4000, each containing 800, 1600, and 4,000 images with 10,137, 19,826, and 49,613 cell instances, respectively. The validation and test sets are composed of 570 and 1,564 images, containing 7,120 and 19,408 instances, respectively.

V. Experimental Setup

Different subsets of synthetic co-culture images are used for training. The same validation and test set is used for the evaluation of all the models trained on different training subsets. Training for all the experiments is performed with a base learning rate of 0.02 and momentum of 0.9. Anchor sizes and aspect ratios were set to 8, 16, 32, 64, 128, and 0.5, 1, 2, 3, 4 for all the settings, respectively.

Evaluation Metrics: To evaluate the performance of the proposed pipeline we are following the standard COCO evaluation protocol [6] with some modifications as reported in [3] for the area ranges. Mean average precision (mAP) is reported at different IoU thresholds i.e., mAP50 and mAP75, and on three different area ranges i.e., mAPs, mAPm, and mAPl.

Results: Table I gives the overall detection and segmentation mAP scores averaged over all the eight-cell classes for the different training subsets. For the model trained on DeepMuCS800, mAP scores of 70.17% and 69.31% are achieved for detection and segmentation, respectively. By increasing the images and cell instances in DeepMuCS1600, the model performs approximately 7% better in terms of detection and segmentation mAP. Further increasing the data for training in DeepMuCS4000, the detection and segmentation scores improve by around 1%.

Per class evaluation results for the models trained on the three different synthetic co-culture training subsets show that as the training data increases, the performance across all the eight-cell culture classes improves. For the DeepMuCS800, the best segmentation mAP score is achieved for the cell culture SkBr3 (77.86), followed by A172 (76.15). The worse performance in terms of segmentation mAP is seen for BV-2 (60.11) followed by SH-SY5Y (62.81). The best performance in terms of segmentation mAP for the DeepMuCS1600 trained model is seen across the cell culture A172 followed by SkBr3 and the worse performance is recorded against the cell culture SH-SY5Y. Similar to DeepMuCS1600, the best performance is seen for the cell culture A172 (83.24) in terms of segmentation mAP, and the worst performance is detected for SH-SY5Y (70.63).

VI. Analysis and Discussion

This section discusses the results of the culture-dependent cell segmentation in co-culture using the DeepMuCS framework. It can be seen from Table I that by increasing the images and cell instances for each cell culture, the segmentation performance increases. These results manifest the potential of the proposed framework for cell type-aware segmentation in a co-culture environment in the microscopic images.

Fig. 3 shows the inference results on some samples where our proposed pipeline performed adequately and inadequately. The left column shows adequate results, and the right column represents the inadequate results. The red, blue, and green background around each image depict the results for DeepMuCS800, DeepMuCS1600, and DeepMuCS4000 trained models, respectively. The AP50 on top of every prediction sub-image is the segmentation average precision score at IoU threshold of 0.5.

First row in the adequate column represents the results for the model trained on DeepMuCS800. The groundtruth of the image contains a total of 18 cell instances from four different cell cultures i.e., MCF-7, SkBr3, BT-474, and SH-SY5Y. All the cell instances are accurately detected, segmented, and classified according to their cell culture, hence the AP50 score of 100%. In the second row, we have adequate results for the DeepMuCS4000. It can be observed that all instances are correctly predicted. In the inadequate column, we have the results for the DeepMuCS800 trained model in the first row. The groundtruth for the image contains 16 cell instances from 4 different cell cultures. The model on a few occasions confused the cell cultures SH-SY5Y and SkBr3, hence the AP50 score of 61.0%. For the DeepMuCS1600 results, all the cell instances are correctly segmented by the model, but on one occurrence it confused the SH-SY5Y cell instance for MCF7, hence the AP50 score of 66.7%.

TABLE I: Overall detection and segmentation results on different Intersection over union threshold and area ranges.

Train dataset	mAP		mAP50		mAP75		mAPs		mAPm		mAPl	
	Det.	Seg.	Det.	Seg.	Det.	Seg.	Det.	Seg.	Det.	Seg.	Det.	Seg.
DeepMuCS800	70.17	69.31	84.38	87.53	79.07	83.92	64.72	59.37	75.04	75.00	61.32	81.44
DeepMuCS1600	77.22	76.14	90.18	93.45	85.54	90.94	76.97	70.63	80.82	79.97	64.63	84.12
DeepMuCS4000	**78.84**	**77.04**	**90.65**	**94.09**	**86.62**	**91.63**	**77.00**	**71.63**	**82.33**	**80.93**	**66.58**	**84.67**

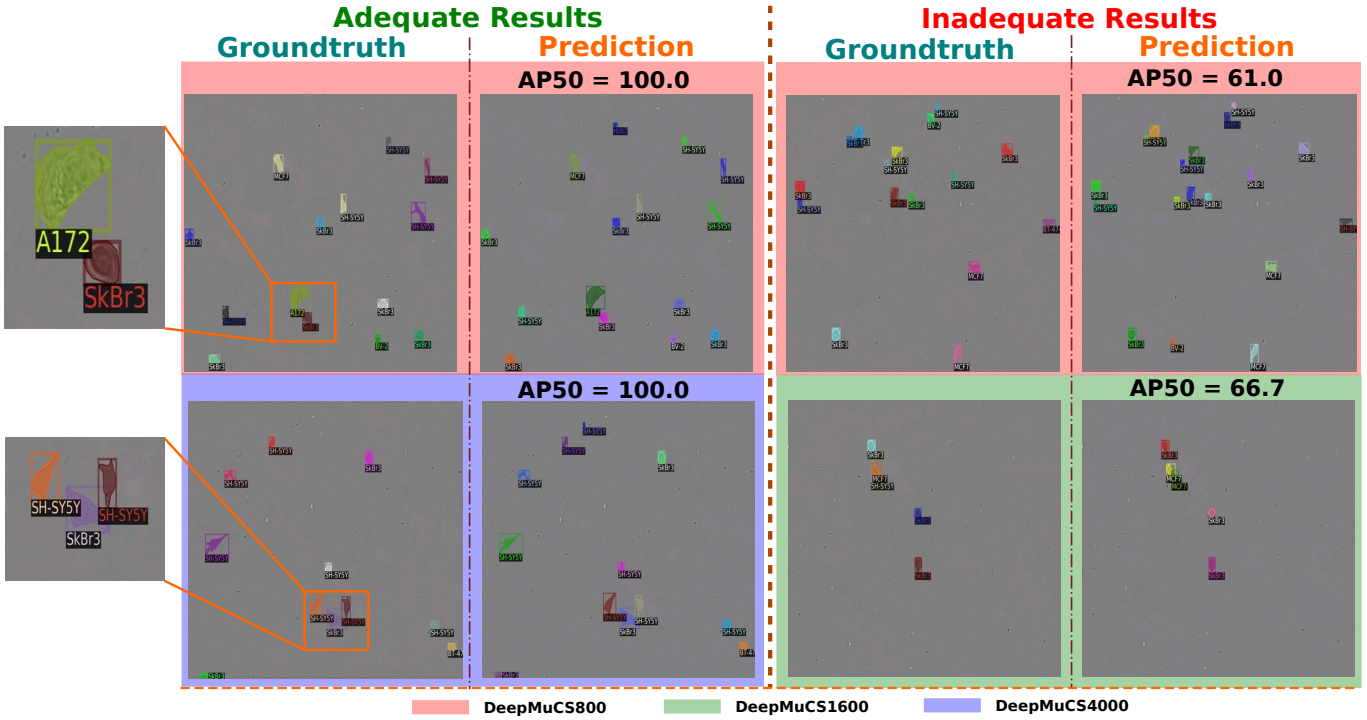

Fig. 3: Inference results of some samples where DeepMuCS-Segmentation trained models performed adequately and inadequately. The red, blue, and green background around each image depict the results for different trained models.

VII. Conclusion

DeepMuCS provides a method to detect, segment, and classify each cell instance in the co-culture microscopic images, which helps better quantification of individual cell appearance and behavior of different cell cultures. We have validated with the help of the proposed pipelines that it is possible to distinguish different cell cultures in the co-culture microscopic images and achieve good performance. This study can help biologists to study the interaction between different cell populations and assist in drug research. The LIVECell dataset used in this study is composed of cells from various kinds of cancer. The future direction for this work would be to validate the proposed workflow on a more clinically relevant co-culture model, like cancer cells mixed with stromal cells.

Acknowledgment

This work was supported by SAIL (Sartorius Artificial Intelligence Lab) project. We thank all members of the Deep Learning Competence Center at the DFKI for their comments and support.

References

[1] L. Axanova, D. J. Morré, and D. M. Morré, "Growth of lncap cells in monoculture and coculture with osteoblasts and response to tnox inhibitors," *Cancer letters*, 2005.

[2] N. Khalid, M. Munir, C. Edlund, T. R. Jackson, J. Trygg, R. Sjögren, A. Dengel, and S. Ahmed, "Deepcens: An end-to-end pipeline for cell and nucleus segmentation in microscopic images," in *2021 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2021.

[3] 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*, 2021.

[4] N. Khalid, M. Munir, C. Edlund, T. R. Jackson, J. Trygg, R. Sjögren, A. Dengel, and S. Ahmed, "Deepcis: An end-to-end pipeline for cell-type aware instance segmentation in microscopic images," in *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*. IEEE, 2021.

[5] M. Schwendy, R. E. Unger, and S. H. Parekh, "Evican–a balanced dataset for algorithm development in cell and nucleus segmentation," *Bioinformatics*, 2020.

[6] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *European conference on computer vision*. Springer, 2014.

[7] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017.

[8] H. Zhang, C. Wu, Z. Zhang, Y. Zhu, Z. Zhang, H. Lin, Y. Sun, T. He, J. Mueller, R. Manmatha *et al.*, "Resnest: Split-attention networks," *arXiv preprint arXiv:2004.08955*, 2020.

[9] J. Dai, H. Qi, Y. Xiong, Y. Li, G. Zhang, H. Hu, and Y. Wei, "Deformable convolutional networks," in *Proceedings of the IEEE international conference on computer vision*, 2017.

[10] Z. Cai and N. Vasconcelos, "Cascade r-cnn: Delving into high quality object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018.