

3D-Organoid-SwinNet: High-content profiling of 3D organoids

Muhammad Sohaib¹, Siyavash Shabani¹, Sahar A Mohammed¹, Qingsu Cheng², Garrett Winkelmaier¹, Bahram Parvin^{1,3}

¹Department of Electrical and Biomedical Engineering, University of Nevada, Reno, USA

²University of Wisconsin – Milwaukee, USA

³Pennington Cancer Institute
[msohaib, bparvin@unr.edu]

Abstract—Profiling of Patient-Derived organoids is necessary for drug screening and precision medicine. This step requires accurate segmentation of three-dimensional cellular structures followed by protein readouts. While fully Convolutional Neural Networks are widely used in medical image segmentation, they struggle to capture long-range dependencies necessary for accurate segmentation. On the other hand, transformer models have shown promise in capturing long-range information across domain boundaries. Motivated by this, we present 3D-Organoid-SwinNet, a unique segmentation model explicitly designed for organoid semantic segmentation. We evaluated the performance of our technique using an Organoid dataset from four breast cancer subtypes. We demonstrated consistent top-tier performance in both the validation and testing phases, achieving a Dice score of 94.91 while reducing the number of parameters to 21 million. Our findings indicate that the proposed model offers a foundation for transformer-based models designed for high-content profiling of organoid models.

Keywords—3D organoid segmentation, Transformers, High content profiling

I. INTRODUCTION

Organoid models have emerged as a powerful tool for cancer therapeutics and investigating biological processes [1–4]. The scope of drug screening in cancer therapeutics is not limited to apoptotic assays but also differentiation to a phenotype with a better prognosis [5]. In 2D monocultures, cells are exposed only to neighboring cells, attached to a plastic substrate, and lack access to the extracellular matrix (ECM). However, cells in 3D organoid models are organized in a three-dimensional structure, there are more cell-cell contacts, and basal cells have access to the ECM. As a result, 3D cultures can better recapitulate *in vivo* models for investigating biological processes and drug responses. For example, studies have shown that 3D systems are closer to *in vivo* in terms of tumor characteristics such as hypoxia and drug resistance [6] and can potentially reduce animal studies and improve the predictive value of drug efficacy [7]. Organoid models are often imaged using a confocal or a microscope equipped with an apotome and deconvolution, resulting in a 3D volumetric image stack. Identifying each nucleus in a 3D colony is essential to compute colony organization and positional protein measurements on a cell-by-cell basis. The rationale is that for tumor cells, 3D colony organization can predict tumor subtypes [6] whereas 2D cultures reveal nothing about the subtype for the same cell lines. In addition, therapeutics can differentiate tumor subtypes, which can only be recognized in 3D [5]. These observations are the motivating factors in characterizing colony organization in terms of its nuclei and necessitating nuclei segmentation in 3D images.

The current study aims to develop a dictionary of known breast cancer subtypes based on well-characterized cell lines. The dictionary can be represented as a heatmap and a classifier in terms of breast cancer subtypes. The heatmap is a powerful visualization tool since it gives insights into the comparative analysis of Patient-Derived organoids with known cell lines. The representation for each colony/organoid is based on its morphometric properties (e.g., roundness, elongation), which are important in breast cancer subtypes. For example, normal colonies should form hollow spheres, whereas malignant colonies should form solid spheres, elongated, or non-convex colonies. However, the basis for a colony representation is 3D nuclear segmentation, which is the focus of this manuscript. Here, we introduce 3D-Organoid-SwinNet, as shown in Fig. 1, and it is motivated by recent advances such as Swin-UNET [8] and SegFormer [9]. Here, we are motivated by developing and testing a simple architecture with a reduced number of parameters. Several model architectures were investigated, but the system in Fig. 1 produced the best results. 3D-Organoid-SwinNet consists of an encoder and a decoder module. The encoder module consists of a Swin-transformer and multiscale convolutional layers. The decoder module consists of a multiscale MLP layer, upsampling, and a convolutional block. The latter convolutional block operates on a concatenation of scale-space upsampling blocks. The central insight is that a rich representation can bypass secondary convolutional layers and replace them with the MLP layers with much fewer parameters. One disadvantage of transformers is that more training data is typically required. However, taking advantage of global consistency outweighs the requirements for a more extensive training set. Finally, the efficacy of the proposed model is evaluated on four model systems of breast cancer subtypes.

The organization of this paper is as follows. Section II reviews prior literature. Section III outlines the approach. Section IV summarizes the results. Section V concludes the paper.

II. RELATED WORKS

Computational approaches for profiling organoid models are either at low resolution using wide-field microscopy [10], high resolution using a confocal microscope, or a combination of the two [11]. One of the main barriers to high-resolution microscopy is the sample thickness that exceeds the working distance of the microscope objectives. Some protocols include cryo-sectioning each spheroid for 2D imaging and nuclear segmentation [12]. Still, this approach requires accurate registration of adjacent sections and stacking the results together later. Because of the complexities associated with the imaging and analysis of organoid models, many protocols quantify organoids in terms of their gross size, morphometric features, and various

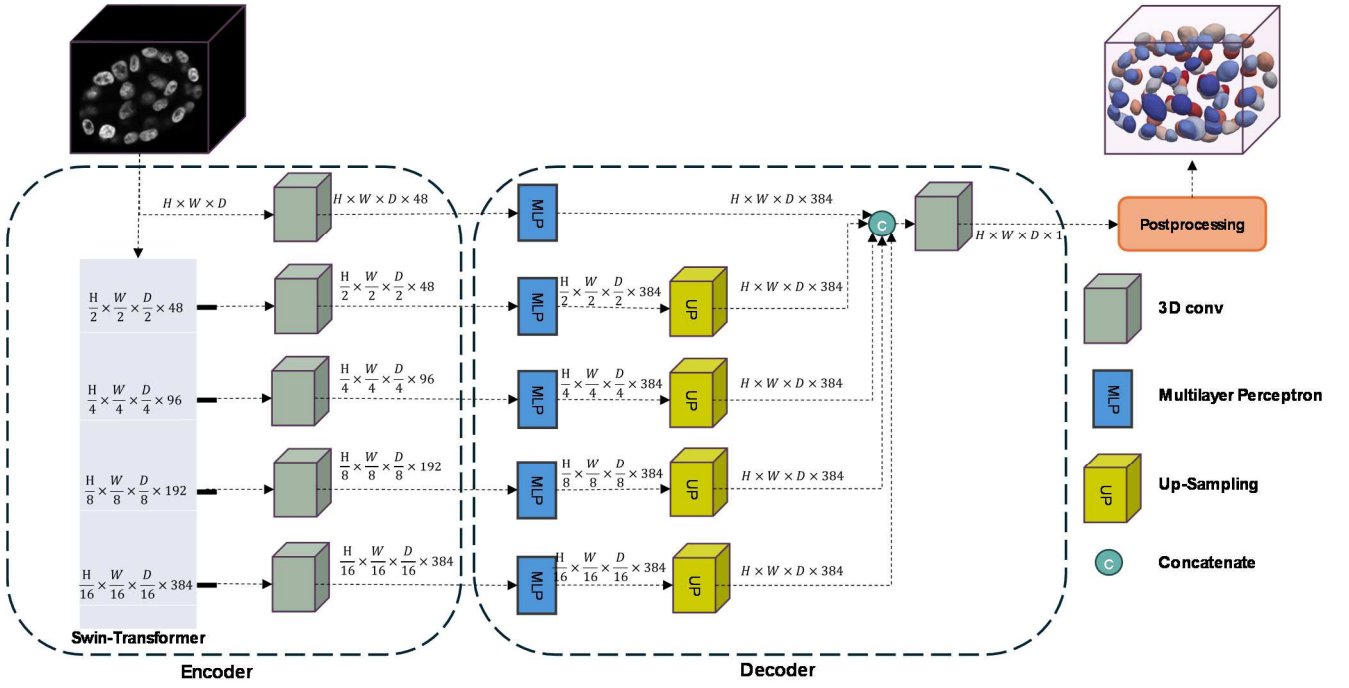


Fig. 1. The architecture of 3D-Organoid-SwinNet consists of an encoder and a decoder. The encoder consists of a Swin Transformer and multiscale convolutional layers. The decoder consists of multiscale MLP layers, upsampling, and one convolutional block for processing concatenated scale-space upsampled outputs.

fluorescent markers [13-16]; however, in some cases, these systems are also complemented with organoids cultured in microchips to allow for high throughput microscopy. Concerning high-resolution microscopy and nuclear segmentation, the main challenge is the variation in fixation and staining that leads to clumps of cells in 3D. One group of researchers used classical image analytic methods such as identifying points of maximum curvature or marker-based watershed for partitioning overlapping nuclei [17, 18]. During the last ten years, deep learning (DL) methods have significantly improved over classical techniques because of their intrinsic generalization associated with machine learning methods. Examples of these techniques include using shallow CNN for the detection of nuclei in 3D [19], UNET [20], and UNET with a modified loss function [21], among others. One of the rationales for incorporating a modified loss function is to enhance the partitioning of overlapping nuclei. DL provides a suitable generalization of nuclear signatures (e.g., vesicular phenotypes, a continuum landscape of apoptosis, or fluorescent signature as a cell cycle function). However, DL is sometimes challenged by a clump of nuclei with no apparent boundaries due to fixation and staining. However, this issue can be alleviated by modifying the loss function [21][22].

III. METHODS

In this section, we discuss the structure of our framework, illustrated in Fig. 1, which consists of Swin-Transformer blocks, 3D convolutional blocks, and MLP blocks.

A. 3D-Organoid-SwinNet

As illustrated in Fig. 1, the proposed framework comprises a Swin Transformer block, 3D convolutional blocks, and

MLP (Multilayer Perceptron) blocks. The Swin Transformer block outputs feature at four different resolution scales, with further details provided in the subsequent section. Following the extraction of global features by the Swin Transformer block, the architecture utilizes five 3D convolutional blocks designed for local feature extraction. Subsequently, MLP blocks are employed to reconstruct the output shape, drawing inspiration from [9]. In the final stage, an additional 3D convolutional layer, equipped with a 1×1 kernel size, serves as the model's terminal layer.

B. Swin-Transformer

Fig. 2 illustrates the Swin Transformer block, which comprises a 3D patch partition block, linear embedding, patch merging, and base Swin Transformer blocks, as presented in the Swin UNETR. In the initial step, the extracted patch, denoted as x and with dimensions represented by $H \times W \times D \times C$, is fed into the patch partition block, where H , W , and D denote the three dimensions of the 3D patch and C represents the channel.

In greater detail, x is divided into N non-overlap patches where $N = \frac{H \times W \times D}{p^3}$ and $x = \{x_p^i | i \in [1, N]\}$. In the subsequent layer, through an embedding layer, x is projected into a C -dimensional space. According to Fig. 2, the base Swin Transformer block comprises two layers. The first layer houses the W-MSA (Window-based Multi-head Self-Attention) block, which partitions the inputs into non-overlapping windows and calculates the local self-attentions for each window as follows [23]:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (1)$$

Where Q, K, V denote query, key, and value, respectively, the window size for the first layer is $M - M - M$. For the

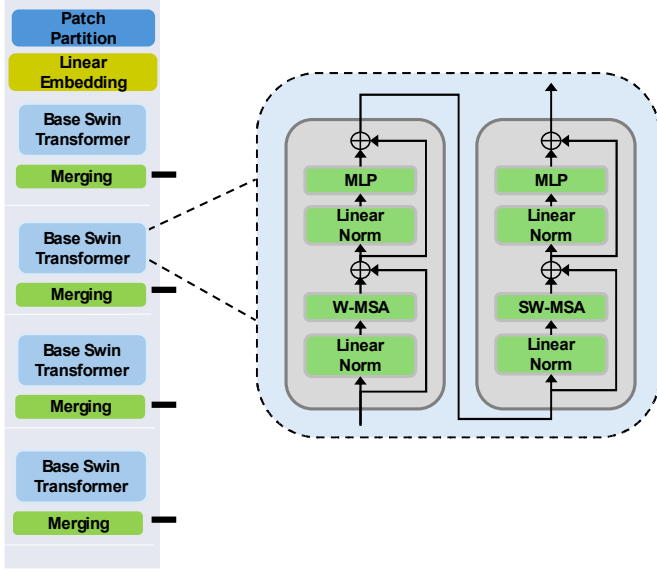


Fig. 2. The Swin Transformer block used in the 3D-Organoid-SwinNet architecture.

second layer, the window is shifted by $\frac{M}{2}, \frac{M}{2}, \frac{M}{2}$ voxels. Similar to the first layer, the attentions for each window are computed within the SW-MSA (Shifted Window Multi-Head Self-Attention) block. The final output is calculated as:

$$\begin{aligned} z_1 &= W\text{-MSA}(LN(x)) + x \\ \hat{z}_1 &= MLP(LN(z_1)) + z_1 \\ z_2 &= SW\text{-MSA}(LN(\hat{z}_1)) + \hat{z}_1 \\ \hat{z}_2 &= MLP(LN(z_2)) + z_2 \end{aligned} \quad (2)$$

Where LN and MLP stand for Linear Normalization and Multi-Layer Perceptron. Whereas $W\text{-MSA}$ and $SW\text{-MSA}$ are multi-head self-attention modules for regular and shifted partitioned windows, respectively.

Figure 1 illustrates how the Swin Transformer processes the input data through four hierarchical stages, each incrementally reducing spatial resolution while increasing feature depth, i.e., $\left[\frac{H}{2}, \frac{W}{2}, \frac{D}{2}, 48\right]$, $\left[\frac{H}{4}, \frac{W}{4}, \frac{D}{4}, 96\right]$, $\left[\frac{H}{8}, \frac{W}{8}, \frac{D}{8}, 192\right]$ and $\left[\frac{H}{16}, \frac{W}{16}, \frac{D}{16}, 384\right]$. These outputs then serve as inputs for 3D convolutions, which extract spatial features in a high dimensional space. This model captures long range dependencies and similarities for 3D organoid segmentation.

C. Implementation Details

The model was trained on a server with 8 Nvidia RTX 2080TI GPUs with a total of 80GB RAM using PyTorch and MONAI. The learning rate is set to 0.0008. Input images with non-zero voxels are modified to have a zero mean and unit standard deviation. Random patches of $96 \times 96 \times 96$ are cropped from 3D image volumes for training purposes. The model trained for 800 epochs with a batch size 1 per GPU. A

sliding window approach is used for inference, with a voxel overlap 0.8 between nearby voxels. We incorporate a weight decay of $1e-4$, dropout of 0.5 and utilize the Adam optimizer for model training. We implement a five-fold cross-validation strategy to evaluate the model's performance robustly. An 80:20 split ratio between the training and validation data is maintained during each cross-validation fold. In this study, we utilized a range of data augmentation techniques to enhance the diversity and robustness of our training dataset. Specifically, we implemented RandFlipd, RandRotate90d, and RandShiftIntensityd functions. RandFlipd introduced random flips along specified spatial axes, RandRotate90d performed random 90-degree rotations, and RandShiftIntensityd randomly adjusted the intensity values of images. These augmentation methods were applied to our data's 'image' and 'label' keys with carefully chosen probabilities.

D. Loss Function

Our loss function combines the widely adopted soft dice loss [24] with cross-entropy loss, leveraging the complementary strengths of both. This combined loss function is calculated on a voxel-wise basis. The resulting loss function is defined as:

$$L(Y, P) = (1 - \sum_{i=1}^I \frac{2 \times \sum_{v=1}^V Y_{v,i} P_{v,i}}{\sum_{v=1}^V Y_{v,i}^2 + \sum_{v=1}^V P_{v,i}^2}) + \sum_{v=1}^V Y_{v,i} \log P_{v,i} \quad (3)$$

The first term represents the Dice loss, while the second for binary cross-entropy loss. I denotes the number of classes, and Y and P represent the ground truth and the output probabilities of the model, respectively.

E. Postprocessing

The final step is thresholding the prediction, another hyperparameter used during the training, followed by a 3D connected component. Each colony is then represented in terms of nuclei organization and shape. Each nucleus within the colony is also represented in morphometric properties, protein expression, and location (e.g., basal versus luminal).

F. Dataset details

One of the attributes of our approach is that our validation data included organoid cultures that are fixed on different days, and cells cover a wide mutation landscape of breast cancer cell lines. These include MCF10A, MCF7, MDA-MB-231, and MDA-MB-468. MCF10A is non-malignant; MCF7 is estrogen- (ER) and progesterone- receptor (PR) positive and ERBB2 negative, and MDA-MB-231 MDA-MB-468 are triple-negative (TNBC), i.e., ER, PR, and ERBB2. In 3D cultures, these cells can form a hollow sphere, a solid sphere, a sheet-like, or a grape-like structure [25, 26]. Notice that although MDA-MB-231 and MDA-MB-468 are both TNBC, their phenotypes differ. This is an important facet of our future research since we aim to classify phenotypes of primary cells into one of the four categories: hollow sphere, solid sphere, sheet-like, or grape-like. Furthermore, morphometric properties of nuclei tend to be different. For example, nuclei in MCF7 are larger and

malignant nuclei tend to have a larger variation in shape and chromatin contents.

Most of the dataset includes public data collected from a Zeiss confocal microscope. This dataset was extended with six new image stacks using Zeiss Axiovert equipped with an apotome and a 40X objective with a working distance of 650 microns. This new dataset is continuously being extended, and we plan to include them in our ever-growing datasets. Furthermore, one significant feature of the dataset is the heterogeneity in image quality and microscopy, which promises a model with improved generalization.

IV. RESULTS

The proposed model was compared to state-of-the-art models for 3D organoid segmentation on our data.

A. Validation and comparison

Using our organoid dataset, we compared 3D-Organoid-SwinNet against the previous year's winning approach, Swin UNETR. Table I summarizes the evaluation results. Our proposed 3D-Organoid-SwinNet model outperforms Swin UNETR [8], Large Language Model (e.g. SAM) [27] and 3D U-Net [28]. The 3D-Organoid-SwinNet achieves the highest performance, attaining a Dice score of 94.91 ± 0.50 and showcasing the trade-off between model complexity and segmentation accuracy of HD95 of 04.03 ± 0.05 , with 21M parameters. This comparison highlights the advancements in segmentation techniques.

TABLE I
3D-ORGANOID-SWINNET IMPROVES SEGMENTATION PERFORMANCE OVER ALTERNATIVE MODELS.

Method	Dice	HD95	Parameters
OTSU Thresholding	64.67 ± 1.80	19.51 ± 0.13	N/A
Segment Anything 2D	77.73 ± 0.75	14.36 ± 0.12	91M
3D U-NET	80.65 ± 0.56	11.55 ± 0.09	95M
Swin UNETR	89.74 ± 0.44	07.42 ± 0.08	61M
3D-Organoid-SwinNet	94.91 ± 0.50	04.03 ± 0.05	21M

Fig. 3 visually compares the results obtained from our approach and those from other techniques. It demonstrates that our model, 3D-Organoid-SwinNet, outperforms the alternatives, particularly in the context of morphogenesis analysis. In morphogenesis analysis, the precise delineation of individual cells is crucial for understanding how tissues develop and organize themselves. The merging of cells, as observed in the outputs of other methods, can obscure essential details and hinder accurate analysis of morphogenetic processes. Therefore, 3D-Organoid-SwinNet's superior performance in preserving cell boundaries is advantageous for conducting thorough

morphogenesis studies and gaining deeper insights into tissue development.

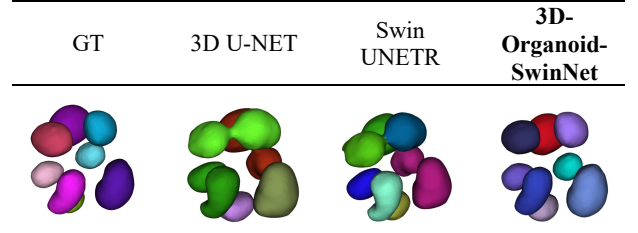


Fig. 3. Comparison of 3D-Organoid-SwinNet with prior art indicates superior performance in terms of segmentation quality since adjacent nuclei are not merged. As a result, a more robust proliferation index is computed.

B. Ablation study

Ablation studies include the effect of different numbers of nodes in MLP layers and the different loss functions.

Effect of the Number of Nodes in MLP Layers: We empirically analyzed our proposed framework by varying the number of MLP (Multilayer Perceptron) layers.

The metrics evaluated include the Dice coefficient, HD95, and the number of parameters for each configuration. The number of nodes considered in this study were 128, 192, 384, and 768. Table II visually represents the effectiveness of different numbers of nodes of MLP layers within our framework. This analysis helps understanding and understanding the trade-offs between model performance and complexity, guiding the selection of an optimal number of nodes, which was 384 for the proposed method.

TABLE II
SIMULATION STUDY FOR DETERMINING THE NUMBER OF NODES IN THE MLP LAYERS

No. of Nodes	Dice	HD95	Parameters
128	88.50	07.89	18.19 M
192	89.59	07.10	19.28 M
768	85.01	10.14	24.50 M
384	94.91	04.03	21.04 M

Effect of Loss Function: In our study, we also conducted an empirical analysis of our proposed framework using various loss functions, as detailed in Table III.

TABLE III
THE IMPACT OF EACH LOSS FUNCTION FOR THE PROPOSED FRAMEWORK

Loss Function	Dice	HD95
L_{dice}	93.07	04.87
L_{focal}	93.11	04.58
L_{CeDice}	94.91	04.03

When trained with the Dice loss function (L_{dice}), the model achieved a Dice score of 93.07 and HD95 of 04.87. Using the Focal loss function (L_{focal}), the score improved slightly to 93.11 with 04.58 of HD95. The L_{CeDice} loss function achieved the best performance with a Dice score of 94.91 and an HD95

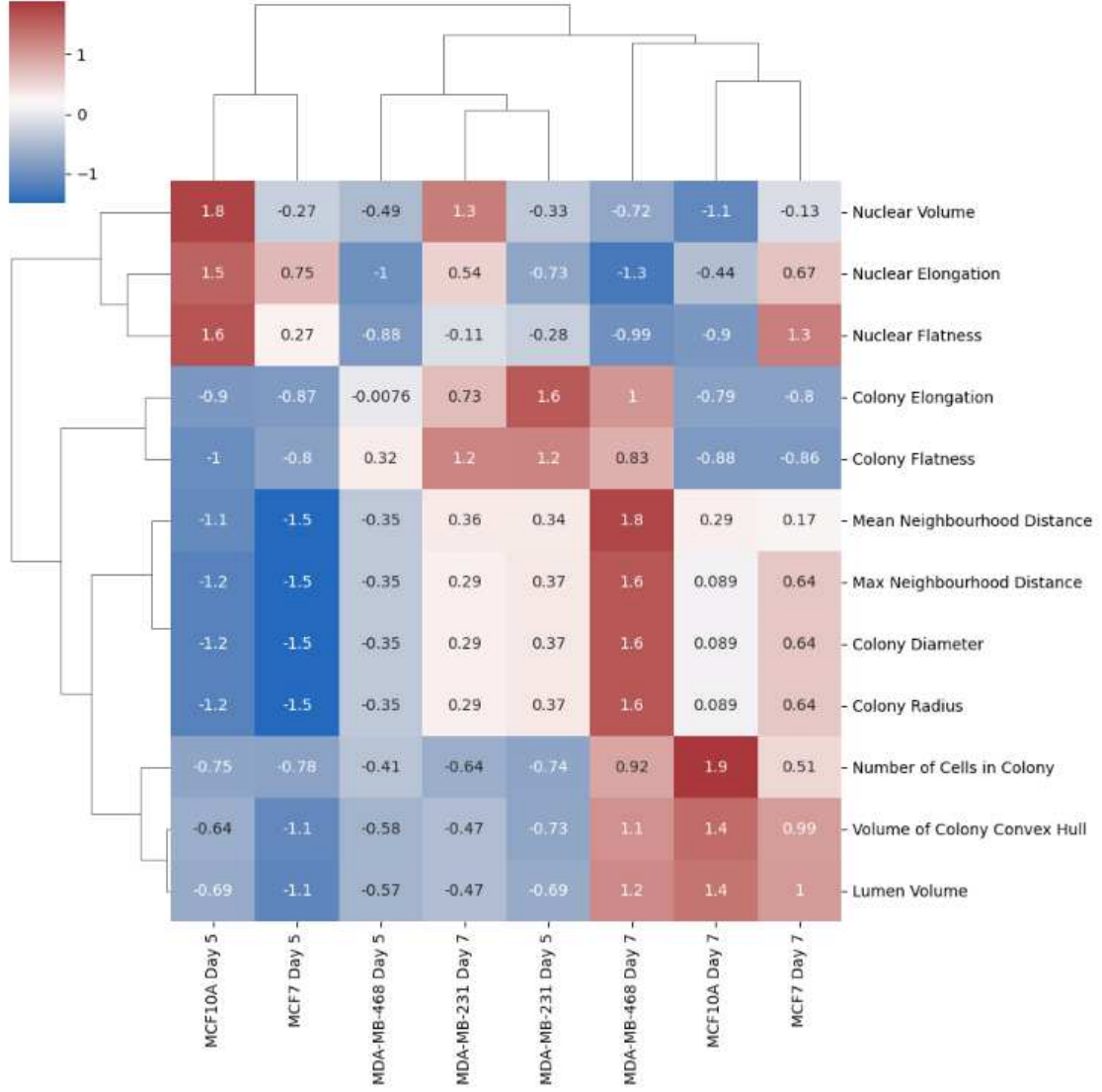


Fig. 4. The multiparametric analysis is visualized for colony organization across various cell lines. This analysis incorporates multiple parameters and offers insights into cellular dynamics and organization across different cell lines.

of 04.03, outperforming other loss functions in our segmentation framework.

C. An application of high content profiling in terms of morphogenesis

In this study, we conducted morphogenesis analysis across multiple model cell lines, where the organization of colonies was profiled using multiparametric techniques, notably employing Delaunay triangulation for neighborhood analysis. The mean computed indices are then visualized with a heatmap, as shown in Fig. 4. These four cell lines correspond to breast cancer subtypes with known molecular aberrations that lead to unique phenotypes. This is evident by MDA-MB-231, which grows in a planar geometry (cultured using the “ontop” method) and is more elongated than the non-malignant MCF10A. As a result, cell lines can serve as a dictionary or a model for predicting the molecular aberrations from Patient-Derived organoids with unknown characteristics. For example, indices computed from the Patient-Derived organoids at a specific date of fixation can

correlate with the relevant columns of the heatmap to find the closest match. Although similar analyses can be performed with transcriptomic or other molecular profiling, they are not functional assays and do not capture relevant 3D structures.

V. CONCLUSION

In this study, we proposed 3D-Organoid-SwinNet, a unique architecture designed for the semantic segmentation of organoids. Our model has a U-shaped network architecture, with a Swin transformer as the encoder and a decoder augmented with an MLP layer that is inspired by a SegFormer framework. We established the efficacy of our approach by validating it on our organoid dataset. The proposed method has only 21M parameters and achieves a new state-of-the-art in terms of the mean Dice and HD95 scores. Our model performs better than the Swin UNETR with fewer parameters. One of the attributes of our design is to validate organoid formation using multiple cell lines with diverse molecular aberrations. For example, MCF7 is ER- and PR-positive; MDA-MB-231 is ER-, PR-, and ERBB2-negative; and MDA-MB-468 is ERBB2-positive. As a result, the

computational method is stress-tested under a diversity of molecular aberrations. Furthermore, the 3D phenotypes of these cell lines serve as a *dictionary* for correlative studies of Patient-Derived organoids with unknown molecular characteristics. Our future research focuses on profiling Patient-Derived organoids and predicting their molecular signatures based on the model systems constructed in our current efforts.

Acknowledgment: This research is supported by a grant from NIH RO1-CA279408.

REFERENCES

- [1] G. Lee, P. Kenny, E. Lee, and M. NBissell, "Three-dimensional culture models of normal and malignant breast epithelial cells," *Nature Method*, vol. 4, no. 4, pp. 359-365, 2007.
- [2] C. C. Bilgin, G. Fontenay, Q. Cheng, H. Chang, J. Han, and B. Parvin, "BioSig3D: High content screening of three-Dimensional cell culture models," (in eng), *PLoS One*, vol. 11, no. 3, p. e0148379, 2016, doi: 10.1371/journal.pone.0148379.
- [3] Q. Cheng, M. Khoshdeli, B. S. Ferguson, K. Jabbari, C. Zang, and B. Parvin, "YY1 is a Cis-regulator in the organoid models of high mammographic density," *Bioinformatics*, vol. 36, no. 6, Nov 5 2020, doi: 10.1093/bioinformatics/btz812.
- [4] V. Srivastava, T. R. Huycke, K. T. Phong, and Z. J. Gartner, "Organoid models for mammary gland dynamics and breast cancer," *Curr Opin Cell Biol*, vol. 66, pp. 51-58, Oct 2020, doi: 10.1016/j.ceb.2020.05.003.
- [5] Q. Cheng *et al.*, "Overexpression of CD36 in mammary fibroblasts suppresses colony growth in breast cancer cell lines," *Biochem Biophys Res Commun*, vol. 526, no. 1, pp. 41-47, May 21 2020, doi: 10.1016/j.bbrc.2020.03.061.
- [6] Y. Imamura *et al.*, "Comparison of 2D- and 3D-culture models as drug-testing platforms in breast cancer," *Oncol Rep*, vol. 33, no. 4, pp. 1837-43, Apr 2015, doi: 10.3892/or.2015.3767.
- [7] M. A. G. Barbosa, C. P. R. Xavier, R. F. Pereira, V. Petrikaite, and M. H. Vasconcelos, "3D Cell Culture Models as Recapitulators of the Tumor Microenvironment for the Screening of Anti-Cancer Drugs," *Cancers (Basel)*, vol. 14, no. 1, Dec 31 2021, doi: 10.3390/cancers14010190.
- [8] N. A. Hatamizadeh, V., Tang, Y., Yang, D., Roth, H.R., Xu, D, "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images," presented at the Lecture Notes in Computer Science, 2022.
- [9] E. Z. Xie, W. H. Wang, Z. D. Yu, A. Anandkumar, J. M. Alvarez, and P. Luo, "SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers," (in English), *Adv Neur In*, vol. 34, pp. 12077-12090, 2021.
- [10] E. R. Spiller *et al.*, "Imaging-Based Machine Learning Analysis of Patient-Derived Tumor Organoid Drug Response," *Front Oncol*, vol. 11, p. 771173, 2021, doi: 10.3389/fonc.2021.771173.
- [11] K. Fei, J. Zhang, J. Yuan, and P. Xiao, "Present Application and Perspectives of Organoid Imaging Technology," *Bioengineering (Basel)*, vol. 9, no. 3, Mar 16 2022, doi: 10.3390/bioengineering9030121.
- [12] M. E. Boutin, T. C. Voss, S. A. Titus, K. Cruz-Gutierrez, S. Michael, and M. Ferrer, "A high-throughput imaging and nuclear segmentation analysis protocol for cleared 3D culture models," *Sci Rep*, vol. 8, no. 1, p. 11135, Jul 24 2018, doi: 10.1038/s41598-018-29169-0.
- [13] A. St-Georges-Robillard *et al.*, "Fluorescence hyperspectral imaging for live monitoring of multiple spheroids in microfluidic chips," *Analyst*, vol. 143, no. 16, pp. 3829-3840, Aug 6 2018, doi: 10.1039/c8an00536b.
- [14] S. Gopal, S. J. Kwon, B. Ku, D. W. Lee, J. Kim, and J. S. Dordick, "3D tumor spheroid microarray for high-throughput, high-content natural killer cell-mediated cytotoxicity," *Commun Biol*, vol. 4, no. 1, p. 893, Jul 21 2021, doi: 10.1038/s42003-021-02417-2.
- [15] J. M. Matthews *et al.*, "OrganoID: A versatile deep learning platform for tracking and analysis of single-organoid dynamics," *PLoS Comput Biol*, vol. 18, no. 11, p. e1010584, Nov 2022, doi: 10.1371/journal.pcbi.1010584.
- [16] X. Wang *et al.*, "A novel deep learning segmentation model for organoid-based drug screening," *Front Pharmacol*, vol. 13, p. 1080273, 2022, doi: 10.3389/fphar.2022.1080273.
- [17] C. C. Bilgin, Kim, S., Leung, E., Chang, H., Parvin, B., "Integrated profiling of three dimensional cell culture models and 3D microscopy," (in eng), *Bioinformatics*, Research Support, N.I.H., Extramural vol. 29, no. 23, pp. 3087-93, Dec 1 2013, doi: 10.1093/bioinformatics/btt535.
- [18] T. Kaseva *et al.*, "Marker-controlled watershed with deep edge emphasis and optimized H-minima transform for automatic segmentation of densely cultivated 3D cell nuclei," *BMC Bioinformatics*, vol. 23, no. 1, p. 289, Jul 21 2022, doi: 10.1186/s12859-022-04827-3.
- [19] M. Khoshdeli and B. Parvin, "Feature-Based Representation Improves Color Decomposition and Nuclear Detection Using a Convolutional Neural Network" *IEEE Trans Biomed Eng*, vol. 65, no. 3, pp. 625-634, Mar 2018, doi: 10.1109/TBME.2017.2711529.
- [20] F. Long, "Microscopy cell nuclei segmentation with enhanced U-Net," *BMC Bioinformatics*, vol. 21, no. 1, p. 8, Jan 8 2020, doi: 10.1186/s12859-019-3332-1.
- [21] G. Winkelmaier and B. Parvin, "An enhanced loss function simplifies the deep learning model for characterizing the 3D organoid models," *Bioinformatics*, Feb 23 2021, doi: 10.1093/bioinformatics/btab120.
- [22] R. O. e. al., "U-net: convolutional networks for biomedical image segmentation," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234-241, 2015.
- [23] A. Dosovitskiy, Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N., "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.
- [24] N. F. Milletari, N., Ahmadi, S.A, " V-net: Fully convolutional neural networks for volumetric medical image segmentation," presented at the fourth international conference on 3D vision (3DV), 2016.
- [25] P. A. K. e. al, "The morphologies of breast cancer cell lines in three-dimensional assays correlate with their profiles of gene expression," *Molecular Oncology*, vol. 1, pp. 84-96, 2007.
- [26] J. H. e. al, "Molecular Predictors of 3D morphogenesis by breast cancer cells in 3D culture," *PLoS Computational Biology*, 2010.
- [27] A. Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao et al, "Segment anything," presented at the Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023.
- [28] Ö. Çiçek, et al, "3D U-Net: learning dense volumetric segmentation from sparse annotation," *19th International Conference, Athens, Greece*, pp. 17-21, 2016.