Reconstruction of 3D Cellular Behavior

Chanhong Min(cmin11), Hyunji Park(hpark111), Anubhav De(ade11),
Emily Guan(eguan3)

1 Background & Motivation

1.1 Live Cell Imaging

Prior to live-cell imaging, we were only able to study cells that had been frozen in time with their structures preserved. Although we were able to identify cell structures and steps of certain biological functions for the first time in high resolution, measuring and visualizing dynamic processes was extremely difficult, as we were only able to estimate an average in population movement. Time-lapsed imaging has allowed us to visualize cell healing processes, learn more about cell development and mitosis, as well as visualizing our body's immune response to foreign threats.

1.2 Spatial/Temporal Trade Off in Live Cellular Imaging

Within the world of 3D cellular imaging, one of the largest tradeoffs a researcher must make is between time and space. The higher resolution in z direction, the higher quality of spatial dynamics captured, but takes more time to scan through z-axis, resulting in long gaps in timing between each image slice. Given we are trying to image a live cell, information is lost in between each image taken. In contrast, 2D imaging is advantageous in terms of experimental simplicity and retaining high temporal resolution. Therefore, if we can leverage the advantages of both 2D and 3D imaging by reconstructing the 3D images from 2D, we can obtain high resolution in both Z-axis and temporal axis.

1.3 Objective

Our objective is to use 2D image slices to reconstruct the 3D shape and behaviors in single-cell by utilizing deep neural network.

In a 2D slice, as a cell moves up and down in the z-direction, it appears to move in and out of focus.

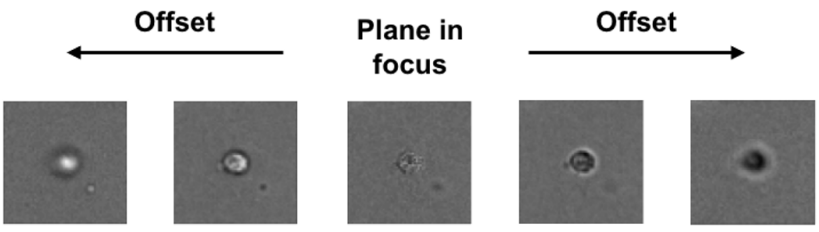

Figure 1

Leveraging this property, we trained two different neural networks for the reconstruction of 3D shape:

Attention 3D U-Net: U-Net is a well-established NN architecture used for image segmentation, generally in the field of medical imaging. Our 3D U-Net implements the U-Net architecture in 3D space as opposed to traditional 2D, and includes attention mechanisms to improve segmentation of important features. The model inputs multiple slices of the brightfield to infer the 3D shape, getting more information about the z displacement.

2D 3D Reconstruction U-Net: The model inputs the brightfield and segmented slice of image, where the brightfield is used to infer the z positions of the cell and segmented slice is used to infer the 3D reconstructed shape.

For the reconstruction of the 3D trajectory, we used following model to extract latent features that describes the vector representation of the trajectory:

Temporal 1D & 2D CNN Autoencoder: : The 1D causal convolution captures the temporal information and 2D convolution captures the spatial association between x, y and z positions. The model learns to reconstruct the same 3D trajectory and learns one-dimensional vector representation from the latent layer to describe the trajectories.

2 Dataset

Our data, sourced from the Phillip Lab at JHU, consisting of 3D time-lapse image with two channels, which first is brightfield image and second is fluorescent image. The brightfield image is used for testing the prediction ability of the model, and fluorescent image is used to create ground truth shape of the cells.

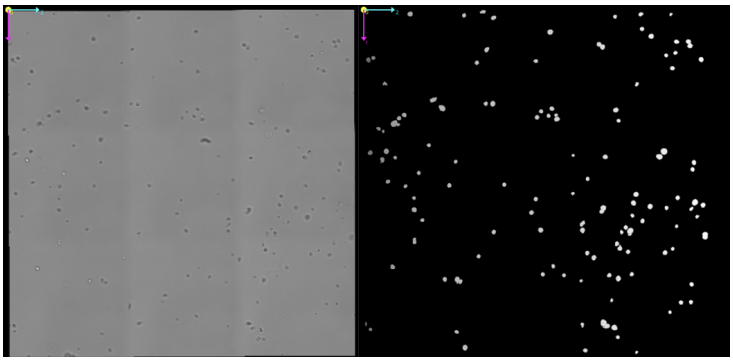

Figure 2: *Example of brighfield image (left) and segmented image (right) in dataset.*

This was split into ∼10,000 images of training data, and ∼1000 images of testing data.

3 Methodology

3.1 Relevant Tools, Libraries

Name	Usage
SciPy	Image Pre-Processing (Interpolation), Feature Extraction
scikit-image	Image Pre-Processing (Segmentation)
TensorFlow	Model Definitions, Training
pyclesperanto-prototype	3D Instance Segmentation
trackpy	Tracking of 3D trajectory
cv2	Image Pre-Processing (Resizing)
umap-learn	UMAP dimensionality reduction of trajectory features

3.2 Pre-Processing

`prepare_training.py utils.py img_processing.py`

Data was normalized, instance-segmented (small objects removed as well as objects with touching borders), interpolated to increase z resolution, and resized to create a stack of planes with uniform dimensions. Fluorescent channel data was used to generate 3D ground truth shape via performing 3D segmentation using Voronoi-otsu labeling algorithm.

3.3 Feature Extraction

`anisotropic_mottility.py motility.py extract_feature.py`

Features of each cell are extracted, such as speed, angle, direction, and mean squared displacement over a given time based on trajectory to compare ground truth cell movement with predicted cell movement. Displacement distribution properties are also calculated.

3.4 Model Definitions

`models.py layers.py loss_func.py`

Attention 3D U-Net

i. Input: 3D Data Tensor, Grayscale

ii. Encoder: Use res_conv3d_block to extract features, downsample via pooling. # of filters doubles at each level.

iii. Bottle Neck: Deepest layer, reduced dimensionality

iv. Decoder: gating_signal and attention_block used to focus model on relevant features, upsample via size doubling

v. Output: Model is compiled using Jaccard loss

$$\textbf{Jaccard Loss} = 1 - \frac{A \cap B}{A \cup B} \tag{1}$$

2D 3D Reconstruction U-Net

i. Input: Two channels of 2D Image slices consisting of brightfield and segmented.

ii. Encoder: Use res_conv3d_block to extract features, downsample via pooling. # of filters doubles at each level.

iii. Bottle Neck: Deepest layer, reduced dimensionality

iv. Decoder: gating_signal and attention_block used to focus model on relevant features, upsample via size doubling

v. Output: Model is compiled using Jaccard loss

Temporal 1D & 2D CNN Autoencoder

i. Input: Sequence of 3D coordinates

ii. Encoding:

   i. res_conv1d_block extracts time dependent patterns

   ii. res_conv2d_block extracts spatial association patterns

iii. Output: Model is compiled using MSE loss

3.5 Training

`train2D3Drecons.py train_3dattention.py`

Data loaded and trained according to each model.

4 Results

The 2D_3D ReconstructionNet proved to be capable of reconstructing 3D cell shapes from 2D inputs, with considerably high fidelity on the XY-projections, while struggling with the Z-axis positions. At the same time, the 3D AttentionNet did a comparatively better job of predicting the accurate 3D position and displacement.

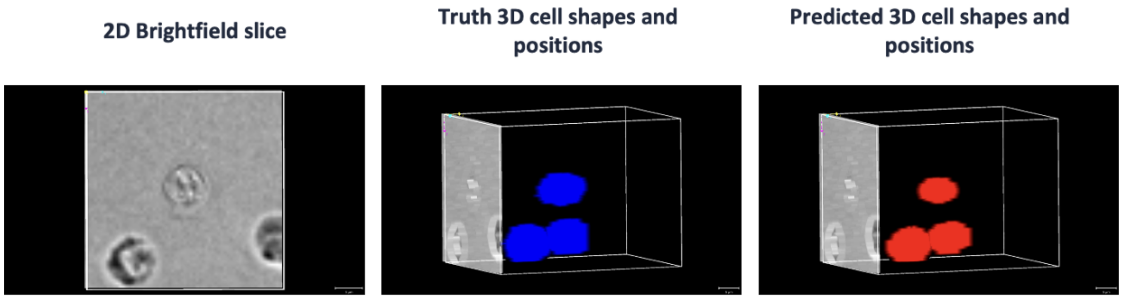

Figure 3: *Example 2D Brighfield image (left), 3D true shape (center) and 3D predicted shape (right).*

The first big challenge for us was the inference of accurate Z-axis information. Though 2D_3D ReconstructionNet gave accurate results for reconstructing the 3D shape and locating cells in XY planes, their lack of accuracy locating cells in Z-axis degraded the quality of reconstructing 3D cellular behavior. We utilized 3D Attention Net to infer 3D cell behaviors for further analysis.

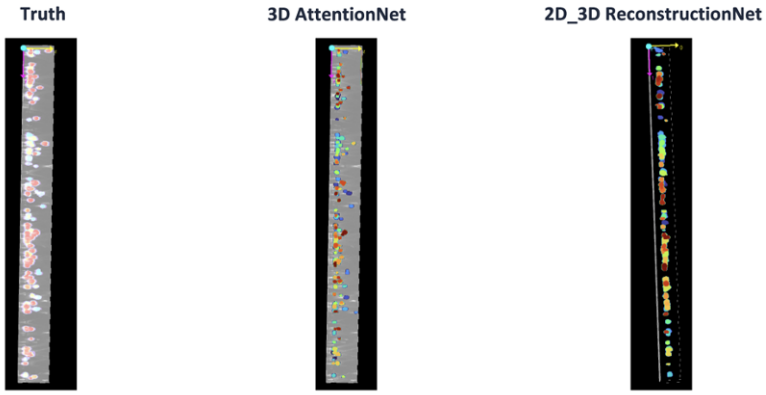

Figure 4: *YZ plane projection of the whole ROI for ground truth (left), prediction from 3D AttentionNet (center) and prediction from 2D_3D ReconstructionNet (right)*

After obtaining 3D positions of each cell in time, we stitched into movement trajectories for each cell by applying a tracking algorithm called Crocker-Greier algorithm. Briefly, it finds the nearest cell for the next frame assuming a Brownian motion, reindexing the cell labels at each frame until the algorithm reaches the final frame, or misses due to high uncertainty. These 3D movement trajectories are then fed into the temporal 1D & 2D CNN autoencoder to extract latent features. The model successfully reconstructed the 3D cell behaviors using latent features, capturing the overall trajectory and dynamics.

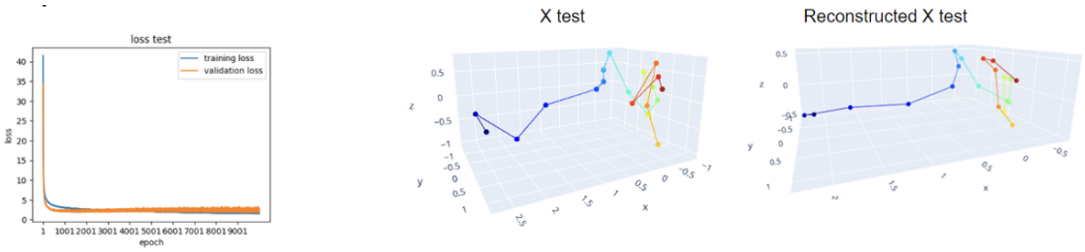

Figure 5: *Training and validation loss for each epochs (left), ground truth 3D trajectory (center), and reconstructed 3D trajectory (right)*

While the models could predict 3D behaviors, there is scope for improvement. This is highlighted by the discrepancies between the predicted and actual trajectories in the overall behavior UMAP space. We questioned which feature of trajectory leads to this discrepancy, and compared set of motility features like average turning angle, alpha (coefficient for the mode of migration) and instantaneous speed coefficient of variance. While the average turning angle and alpha range has no significant difference between ground truth and prediction, there was high discrepancy in the speed coefficient of variance. This implies that there was some fluctuations in the prediction including the missed frames and subtle inconsistencies and affected the prediction robustness.

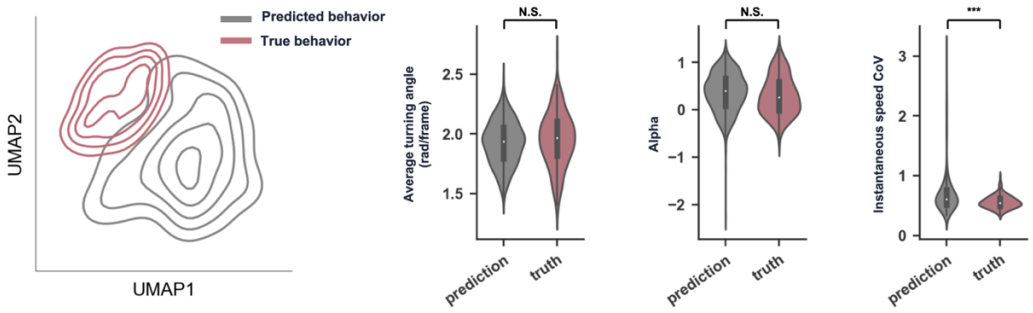

Figure 6: *Overall UMAP behavior space (left), motility features comparison between ground truth and prediction (right)*

5 Conclusion

First of all, our work involves quite a few important biomedical applications besides novel imaging technology. Methods developed by us enable 3D reconstruction from 2D time-lapse imaging-an alternative approach for those so-called direct 3D methods of imaging, which are substantially inexpensive and computationally undemanding. Such progress should significantly contribute to studies on rather complex processes, like tumor metastasis, immune reactions, and wound healing. Furthermore, the insights derived from this research provide a sound basis for future modeling efforts to increase the accuracy and applicability of 3D cellular behavior reconstruction. These approaches could also have an impact beyond biology, inspiring new developments in other fields, such as robotics or autonomous systems, where trajectory prediction and 3D reconstruction are essential.

Our work has several scopes for further improvement and research, such as the improvement of the Z-position inference in the 2D_3D ReconstructionNet. This could be done by subdividing the prediction into a two-step process, prediction of z slice segmentation and then merging with brightfield slice to reconstruct 3D shape. Here, we used Cellpose3 model to segment z slice, which is merely trained to segment 2D shape, not incorporating z deflections. This might have led the model to confusion in locating the cells in z direction, choosing center as a safe guess. Also, investigations into the causes of behavioral discrepancies may result in better feature extraction methods or alternative loss functions that decrease variance. Furthermore, increasing the model's training dataset with cells of more irregular morphologies will also enhance the model's generalization. Lastly, this developed framework could go well beyond biomedical engineering and prove particularly valuable for applications where highly accurate 3D modeling and trajectory prediction are necessary.