**AUTOMATED AND CURATED ANNOTATION OF
CELL-CYCLE STAGE & MITOTIC CLASSIFICATION**

To enrich our data, we are providing mitotic classification for the cells. This is produced by algorithmically annotating cell images with a draft estimate of their cell-cycle stage and a manual curation of the results.

TRAINING DATA ANNOTATION:

Mitotic classification e.g. "non-mitotic vs mitotic"

Approximately 7500 cell images were first manually annotated as belonging to one of two classes: either non-mitotic or mitotic. For this process, composite images were generated for every cell to facilitate manual annotation in a feasible time frame. For each cell these images contained its DNA image and membrane image at the center-z-slice cropped to a region slightly larger than the cell to provide information on neighboring cells, and projections of the DNA image in x, y, and z directions. These composite images enabled the biologist to see sufficient detail relevant to determining the mitotic state in one 2D view without the need to open every 3D multi-channel microscopy z-stack image.

Mitotic classification e.g. subdividing "mitotic" cells into 4 stages of mitosis

The ~2500 cell images that were annotated as mitotic were further annotated as belonging to 4 stages of mitosis in the same manner as described above. These stages are further defined in https://www.allencell.org/hips-cells-during-mitosis.html and include prophase, early prometaphase, prometaphase/metaphase, anaphase/telophase/cytokinesis.

MODEL TRAINING:

From this hand-annotated data, we trained a neural network to predict the mitotic state and the mitotic stage of any unannotated cell, directly from the image.

While deep convolutional neural networks are the state of the art in classifying images, training a neural network from scratch on a relatively small set of data does not usually generalize well or produce viable models. To overcome this issue, we adopted a transfer-learning strategy and employed a model that was pre-trained on a large set of unrelated auxiliary image data. Specifically, we used the Resnet18[1] model implemented in PyTorch[2], pre-trained on the ImageNet data set[3]. Initializing our model from a pre-trained network and altering only the last layer allowed us to leverage the generalizable features learned on a large set of highly heterogeneous images. The subsequent refinement of the network was achieved by training it on the hand-annotated cell images labeled as *mitotic* or *non-mitotic*, tailoring it to our specific task (to predict *mitotic* or *non-mitotic*, given a completely new cell image). This same approach was taken when further classifying the mitotic cells into 4 stages of mitosis.

EXPANDED METHODS AND METRICS

Our image data is three-dimensional and multi-channel. Each image has two channels in common with all other images, and we used those two channels (the cell membrane and DNA stains) for all our predictions. Since our images are three-dimensional and not suitable for use with networks that are pre-trained on 2D images, we instead trained three independent models on x-, y-, and z-axis maximum-intensity 2D projections of our three-dimensional two-channel cell image data. Once trained, these models were used in concert, and the final *mitotic*/*non-mitotic* annotation for each cell was taken to be the most confident of the three predictions. Specifically, the output of each model is a real number between zero (*non-mitotic*) and one (*mitotic*), and the overall prediction was taken to be the output of the model whose prediction was the lowest in entropy, i.e. closest to zero or one.

The results of this predictive model are summarized in the table below. The model weights were updated using the training set, with hyper-parameter optimization dictated by performance on a validation split (to which the model was otherwise blind). Finally, generalizability was assessed using the performance of the model on a held-out, out-of-sample test set.

Split	Num. Images	Mitotic Frac.	Accuracy	Precision	Recall
Train	6039	0.077	0.996	0.934	1.000
Validation	1509	0.082	0.977	0.850	0.871
Test	1134	0.149	0.980	0.919	0.941

MANUAL CURATION OF CELL STATE PREDICTED BY THE MODEL

The trained predictive model was applied to the entirety of the released production data available on the website. The results of the model were examined primarily based on cell and DNA volume. Two obvious qualitative data clusters were seen when comparing cells predicted to be mitotic or non-mitotic. A group of cells with the opposite cell cycle state label was seen within each of these clusters. In addition, a group of cells was identified between these two clusters. Approximately 1200 potentially mislabeled cells (including extreme outliers) were identified and manually examined for their cell cycle state using the same procedure as for training data annotations. This manual curation corrected 106 total cells out of the entire population (~2500 mitotic and ~32,000 non-mitotic cells) and improved the likelihood that cells within obvious cell and nuclear volume clusters correspond to their proper cell cycle state label. Both the original model predictions and the manually-corrected results are available for all applicable cells.

MANUAL CURATION OF CELL STATE PREDICTED BY THE MODEL

From this initial trained and curated single mitotic cell population (~2500 mitotic cells), further manual curation was performed to classify cells into four mitotic stages: prophase, early prometaphase, prometaphase/metaphase, and anaphase/telophase/cytokinesis using x-, y-, and z-axis maximum-intensity 2D projections of our three-dimensional two-channel cell image data. Our manual curation identified 140 prophase, 238 early prometaphase, 858 prometaphase/metaphase, and 1255 anaphase / telophase /

cytokinesis cells. In our examination, we identified a few anaphase-telophase cells whereby segmentation failed to capture the faint connection of the separating cells, and only contained one part of the separating cells. Thus, we searched all anaphase/telophase/cytokinesis cells and matched 177 pairs of segmented cells that comprised a complete cell when combined. These combined cells replaced the incomplete cells in the data set.

1. *Deep Residual Learning for Image Recognition*, He *et al.*, arXiv.org 2015-12-10
   https://arxiv.org/pdf/1512.03385.pdf

2. Torchvision.models, PyTorch Docs, https://pytorch.org/docs/stable/torchvision/models.html

3. *ImageNet Large Scale Visual Recognition Challenge*, Russakovsky *et al.*, IJCV 2015
   doi:10.1007/s11263-015-0816-y