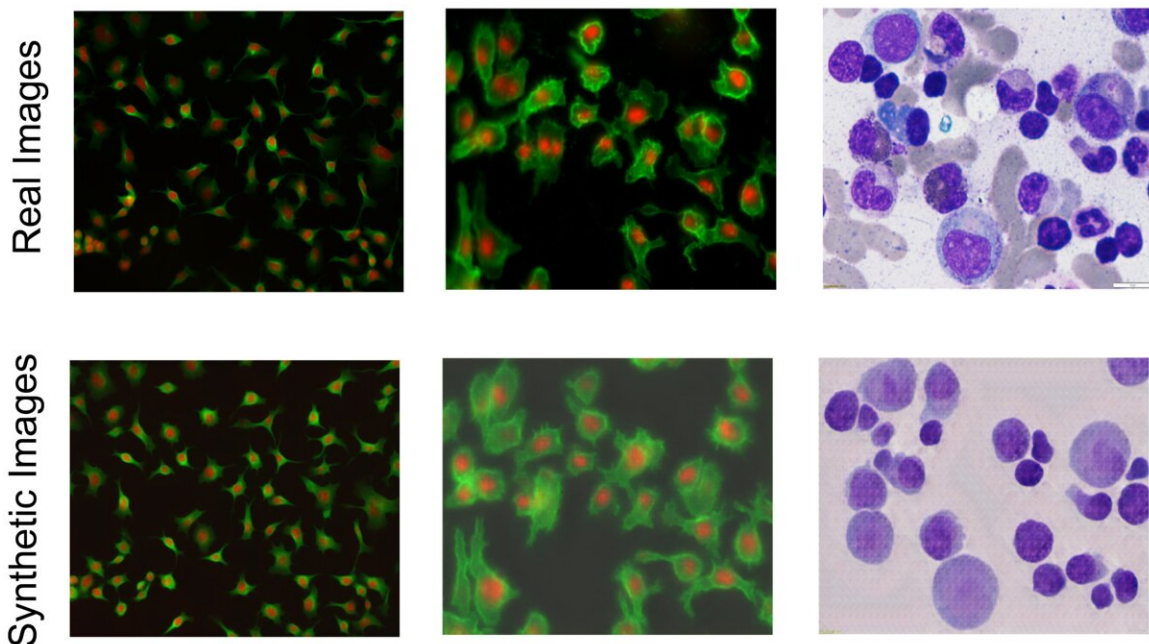# AI tool creates 'synthetic' images of cells for enhanced microscopy analysis

April 22 2024

Examples of real images of various cell types versus synthetic images produced by the researchers' generative model. Credit: Ali Shariati et al.

Observing individual cells through microscopes can reveal a range of important cell biological phenomena that frequently play a role in human diseases, but the process of distinguishing single cells from each other and their background is extremely time-consuming—and a task that is well-suited for AI assistance.

AI models learn how to carry out such tasks by using a set of data that are annotated by humans, but the process of distinguishing cells from their background, called "single-cell segmentation," is both time-consuming and laborious. As a result, there is a limited amount of annotated data to use in AI training sets. UC Santa Cruz researchers have developed a method to solve this by building a microscopy image generation AI model to create realistic images of single cells, which are then used as "synthetic data" to train an AI model to better carry out single cell-segmentation.

The new software is described in a new paper published in the journal *iScience*. The project was led by Assistant Professor of Biomolecular Engineering Ali Shariati and his graduate student Abolfazl Zargari. The model, called cGAN-Seg, is freely available on GitHub.

"The images that come out of our model are ready to be used to train segmentation models," Shariati said. "In a sense we are doing microscopy without a microscope, in that we are able to generate images that are very close to real images of cells in terms of the morphological details of the single cell. The beauty of it is that when they come out of the model, they are already annotated and labeled. The images show a ton of similarities to real images, which then allows us to generate new scenarios that have not been seen by our model during the training."

Images of individual cells seen through a microscope can help scientists learn about cell behavior and dynamics over time, improve disease detection, and find new medicines. Subcellular details such as texture can help researchers answer important questions, like whether a cell is cancerous or not.

Manually finding and labeling the boundaries of cells from their background is extremely difficult, however, especially in tissue samples where there are many cells in an image. It could take researchers several

days to manually perform cell segmentation on just 100 microscopy images.

Deep learning can speed up this process, but an initial data set of annotated images is needed to train the models—at least thousands of images are needed as a baseline to train an accurate deep learning model. Even if the researchers can find and annotate 1,000 images, those images may not contain the variation of features that appear across different experimental conditions.
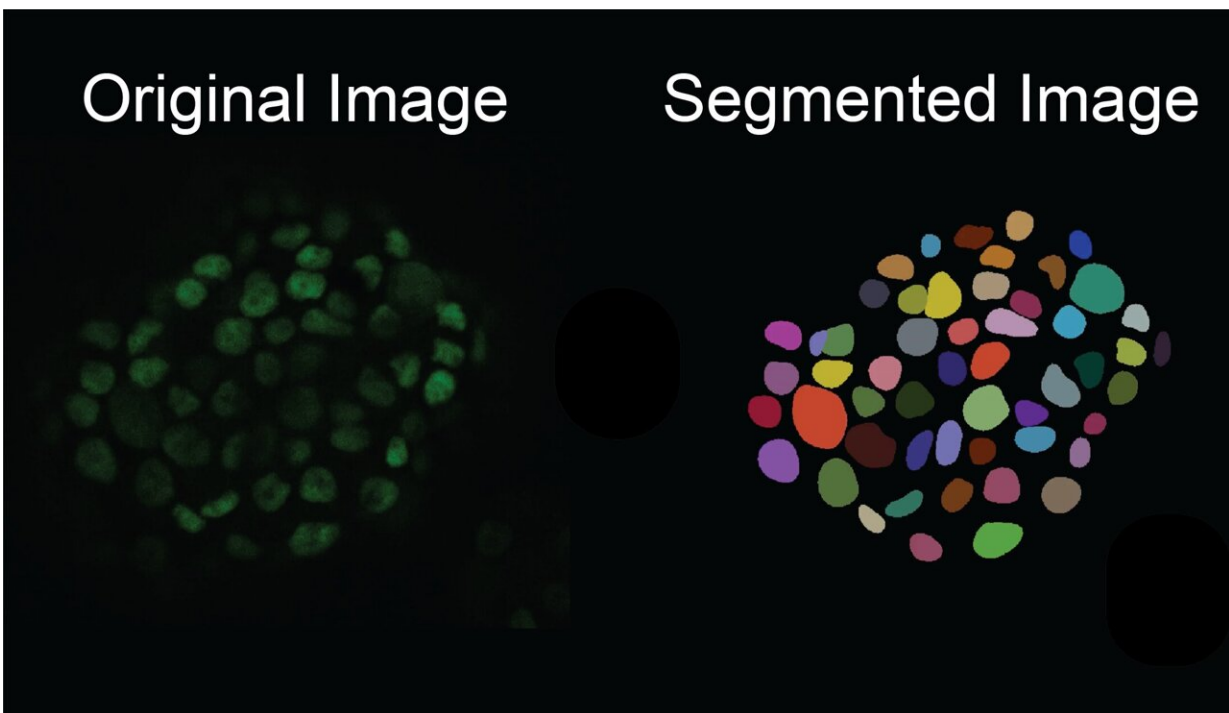
An example of a cell image before and after segmentation, a process which allows researchers to distinguish single cells from each other and their background. Credit: Ali Shariati et al

"You want to show your deep learning model works across different

samples with different cell types and different image qualities," Zargari said. "For example, if you train your model with high-quality images, it's not going to be able to segment the low-quality cell images. We can rarely find such a good data set in the microscopy field."

To address this issue, the researchers created an image-to-image generative AI model that takes a limited set of annotated, labeled cell images and generates more, introducing more intricate and varied subcellular features and structures to create a diverse set of "synthetic" images. Notably, they can generate annotated images with a high density of cells, which are especially difficult to annotate by hand and are especially relevant for studying tissues. This technique works to process and generate images of different cell types as well as different imaging modalities, such as those taken using fluorescence or histological staining.

Zargari, who led the development of the generative model, employed a commonly used AI algorithm called a "cycle generative adversarial network" for creating realistic images. The generative model is enhanced with so-called "augmentation functions" and a "style injecting network," which helps the generator to create a wide variety of high-quality synthetic images that show different possibilities for what the cells could look like. To the researchers' knowledge, this is the first time style injecting techniques have been used in this context.

Then, this diverse set of synthetic images created by the generator is used to train a model to accurately carry out cell segmentation on new, real images taken during experiments.

"Using a limited data set, we can train a good generative model. Using that generative model, we are able to generate a more diverse and larger set of annotated, synthetic images. Using the generated synthetic images we can train a good segmentation model—that is the main idea," Zagari

said.

The researchers compared the results of their model using synthetic training data to more traditional methods of training AI to carry out cell segmentation across different types of cells. They found that their model produces significantly improved segmentation compared to models trained with conventional, limited training data. This confirms to the researchers that providing a more diverse dataset during training of the segmentation model improves performance.

Through these enhanced segmentation capabilities, the researchers will be able to better detect cells and study variability between individual cells, especially among stem cells. In the future, the researchers hope to use the technology they have developed to move beyond still images to generate videos, which can help them pinpoint which factors influence the fate of a cell early in its life and predict their future.

"We are generating synthetic images that can also be turned into a time lapse movie, where we can generate the unseen future of cells," Shariati said. "With that, we want to see if we are able to predict the future states of a cell, like if the cell is going to grow, migrate, differentiate or divide."

Provided by University of California - Santa Cruz