Contents lists available at ScienceDirect

Synthetic and Systems Biotechnology

Original Research Article

A generative benchmark for evaluating the performance of fluorescent cell image segmentation

Jun Tang [a,b], Wei Du [b,**], Zhanpeng Shu [c], Zhixing Cao [a,*]

[a] *State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, Shanghai, 200237, China*
[b] *MOE Key Laboratory of Smart Manufacturing in Energy Chemical Process, East China University of Science and Technology, Shanghai, 200237, China*
[c] *College of Electrical Engineering, Shanghai Dianji University, Shanghai, 201306, China*

ARTICLE INFO

ABSTRACT

Fluorescent cell imaging technology is fundamental in life science research, offering a rich source of image data crucial for understanding cell spatial positioning, differentiation, and decision-making mechanisms. As the volume of this data expands, precise image analysis becomes increasingly critical. Cell segmentation, a key analysis step, significantly influences quantitative analysis outcomes. However, selecting the most effective segmentation method is challenging, hindered by existing evaluation methods' inaccuracies, lack of graded evaluation, and narrow assessment scope. Addressing this, we developed a novel framework with two modules: StyleGAN2-based contour generation and Pix2PixHD-based image rendering, producing diverse, graded-density cell images. Using this dataset, we evaluated three leading cell segmentation methods: DeepCell, CellProfiler, and CellPose. Our comprehensive comparison revealed CellProfiler's superior accuracy in segmenting cytoplasm and nuclei. Our framework diversifies cell image data generation and systematically addresses evaluation challenges in cell segmentation technologies, establishing a solid foundation for advancing research and applications in cell image analysis.

1. Introduction

Fluorescent cell imaging, a non-invasive imaging tool, utilizes photons emitted by fluorescent probes to create images of cells, molecules, tissues, and various organs within the body [1]. The underlying principle of this technique is the phenomenon of fluorescence, also known as photoluminescence. When fluorescent molecules absorb energy from an external light source, they transition from a ground state to a higher energy excited state. Subsequently, these molecules return to the ground state by emitting photons, thereby producing fluorescence. This process typically involves a shift from high-energy (short-wave) incident light to lower-energy (long-wave) fluorescence, known as the Stokes shift [2,3]. This shift significantly influences the efficacy of fluorescence imaging. Fluorescence imaging systems achieve high-resolution imaging by capturing fluorescence within a specific wavelength range from the target area and integrating it into a display.

In the last several decades, owing to technological advancements,

fluorescence imaging technology has evolved to enable high-throughput generation of image data [4], markedly propelling forward research in the life sciences. Specifically, the advent of Single Molecule Fluorescence In Situ Hybridization (smFISH) technology, offering spatial localization data of RNA within cells [5], has facilitated the production of comprehensive fluorescence image datasets. This dataset empowers researchers to identify and enumerate individual RNA molecules within cells, thereby enabling a deeper investigation into the regulatory mechanism of gene expression noise [6–11]. Besides, live cell imaging technology employs high-throughput microscopy apparatus to incessantly document the dynamic alterations of RNA, encompassing synthesis, processing, transport, and degradation [12], thus permitting researchers to monitor the entire RNA lifecycle [13]. The image data garnered via these technologies, encompassing extensive temporal scales, high-throughput, multi-channel dimensions, and fluorescence intensity metrics [14], yield fundamental insights for the quantitative analysis of cellular differentiation and decision-making processes,

imperative for comprehending the intricate dynamics and interplays within cellular systems.

Within the domain of cell imaging and analysis, the burgeoning volume of image data necessitates progressively more precise analysis. The analytical pipeline encompasses multiple critical steps: image pre-processing, cell segmentation, data quantification, and a variety of downstream quantitative analysis. Among these steps, cell segmentation is crucial, as it strives to accurately define cell contours from image data for precise measurement of cellular attributes. Subsequently, researchers can conduct advanced analyses, such as inferring kinetic parameters, thereby not only uncovering dynamic characteristics of gene expression but also yielding further insights into cellular collective behaviors and tissue development [15,16]. Traditionally, cell segmentation was manually executed, entailing domain experts' hand-selecting Regions of Interest (ROIs) [17]. However, with the escalating volume of data, the efficiency of manual cell segmentation markedly diminishes. To surmount these challenges, res earchers have recently innovated automated cell segmentation techniques grounded in deep learning [18]. These methodologies expedite the segmentation process and enhance repeatability, thus alleviating the workload associated with manual segmentation. For instance, technologies such as GeneSegNet, MARS-Net, DeLTA leverage diverse deep learning instruments to augment the precision and efficiency of segmentation [19–21]. Among these technologies, GeneSegNet converts RNA location data into continuous two-dimensional probability maps [19]. Subsequently, these distributions are integrated with imaging data and inputted into a neural network, thereby enabling the precise delineation of cellular boundaries. Furthermore, this approach utilizes a recursive training strategy to mitigate noise in training labels, consequently augmenting both the performance and precision of cell segmentation. MARS-Net [20], conceived by Junbong Jang et al., integrates the VGG-19 encoder and U-Net decoder, applying transfer learning and datasets from diverse microscopes to accomplish exact segmentation of live cell images. DeLTA employs two sequential U-Net deep learning models [21], enhancing the precision of intercellular boundary identification via training with thousands of image-segmentation mask pairs. Beyond the previously mentioned neural networks necessitating training, a spectrum of more user-friendly, readily deployable cell segmentation technologies exists, including CellPose, DeepCell, and CellProfiler [22–24]. A notable advantage of these technologies is their pre-trained nature, enabling direct application to diverse cell imaging data types without necessitating further training by researchers. CellPose [22], proposed by Marius Pachitariu et al., utilizes a simulated diffusion-based method to create topological maps, deploying neural networks to forecast these maps' gradients for cell segmentation. DeepCell amalgamates the Mesmer algorithm and marker point segmentation technology [23], utilizing the devised TissueNet dataset and manually annotated data to realize precise demarcation of inter-cellular boundaries. CellProfiler implements the Sauvola adaptive thresholding algorithm to delineate cells from the background in an image [24]. Subsequent to thresholding, CellProfiler leverages the capabilities of the scikit-image library to accurately delineate and quantify the contours and area of individual cells [25].

Despite substantial advancements in cell segmentation technology, accurately determining the most effective method for practical applications continues to be challenging due to the absence of systematic evaluations. For instance, Caicedo et al. proposed an evaluation framework that includes a dataset of 23,165 manually annotated cell nuclei [26], conceptualizing cell segmentation as an instance segmentation problem and reducing it to a tripartite pixel classification issue: background, interior of the cell nucleus, and cell boundary. This methodology discerns accurately segmented cells by computing the Intersection over Union (IoU) metric, which quantifies the congruence between the true cell contours and their algorithmically predicted counterparts [27]. However, the reliability of conclusions derived from this method is compromised, because the dataset may contain inaccurate annotations and lack image diversity [28]. Firstly, due to limited involvement of experienced professionals, cell contour annotation in existing datasets may not be accurate. Subsequently, current cell image datasets may not be representative of a diverse range of cellular characteristics, including a wide variety of cellular morphologies and densities. Consequently, even if an algorithm demonstrates high performance on a particular dataset, this does not invariably ensure equivalent generalizability in actual biological specimens.

To remedy the shortfall in existing evaluation methods concerning accurate and comprehensive datasets, we used generative networks to devise a data generation framework for evaluating the performance of fluorescent cell image segmentation. This framework primarily comprises two components: a contour generation module and a cell rendering module. Within the contour generation module, we employed the StyleGAN2 network, equipped with style control capabilities, enabling the generation of cell contours of diverse sizes, shapes, and densities, thus facilitating density grading of the image data. The cell rendering module then utilizes Pix2PixHD for image rendering, converting these contours into realistic cell images. This approach not only guarantees the high quality and realism of the images but also yields a comprehensive cell image dataset. Utilizing this dataset enables a more systematic evaluation of diverse cell segmentation methodologies, thereby providing critical insights for the development of increasingly precise segmentation tools.

The structure of the subsequent article is as follows: Section 2 delves into the technical nuances of our proposed approach, detailing the utilization of the StyleGAN2 network in the contour generation module for creating varied styles of cell contours, and the application of Pix2PixHD in rendering these contours into realistic cell images. Section 3 primarily concentrates on affirming the high fidelity of our generated data and contrasting the evaluation metrics of extant cell segmentation methodologies. Section 4 presents the experimental results, encompassing thorough analyses of the contour generation module, image rendering module, and segmentation outcomes. Finally, Section 5 engages in an exhaustive discussion, exploring the merits of the framework and its potential implications and value in practical applications.

2. Methods

This study endeavors to generate a more diverse and density-controllable dataset of cell images, and to construct a comprehensive fluorescent cell dataset that characterizes the diverse characteristics of cells. To accomplish this objective, a detailed processing workflow (illustrated in Fig. 1) was developed, consisting mainly of two principal modules: contour generation and image rendering. The fundamental task of the contour generation module is to harness the sophisticated style control functions of the StyleGAN2 algorithm to produce a variety of cell contours, encompassing an extensive array of cell densities and morphologies. These contour images form the basis for the ensuing steps in image rendering. In the image rendering module, the Pix2PixHD generative adversarial network technology is utilized to convert these contour images into lifelike fluorescent cell images, thereby finalizing the cell image rendering process.

2.1. Contour generation module

In our study, we utilized the StyleGAN2 generative adversarial network, as proposed by Karras et al. [29], to facilitate density-graded generation of cell contours. The StyleGAN2 network's capacity to precisely manipulate style information is instrumental in our process. It allows us to accurately generate cell contours, simulating a diverse range of variations of cellular biological conditions and density variations.

2.1.1. Network architecture

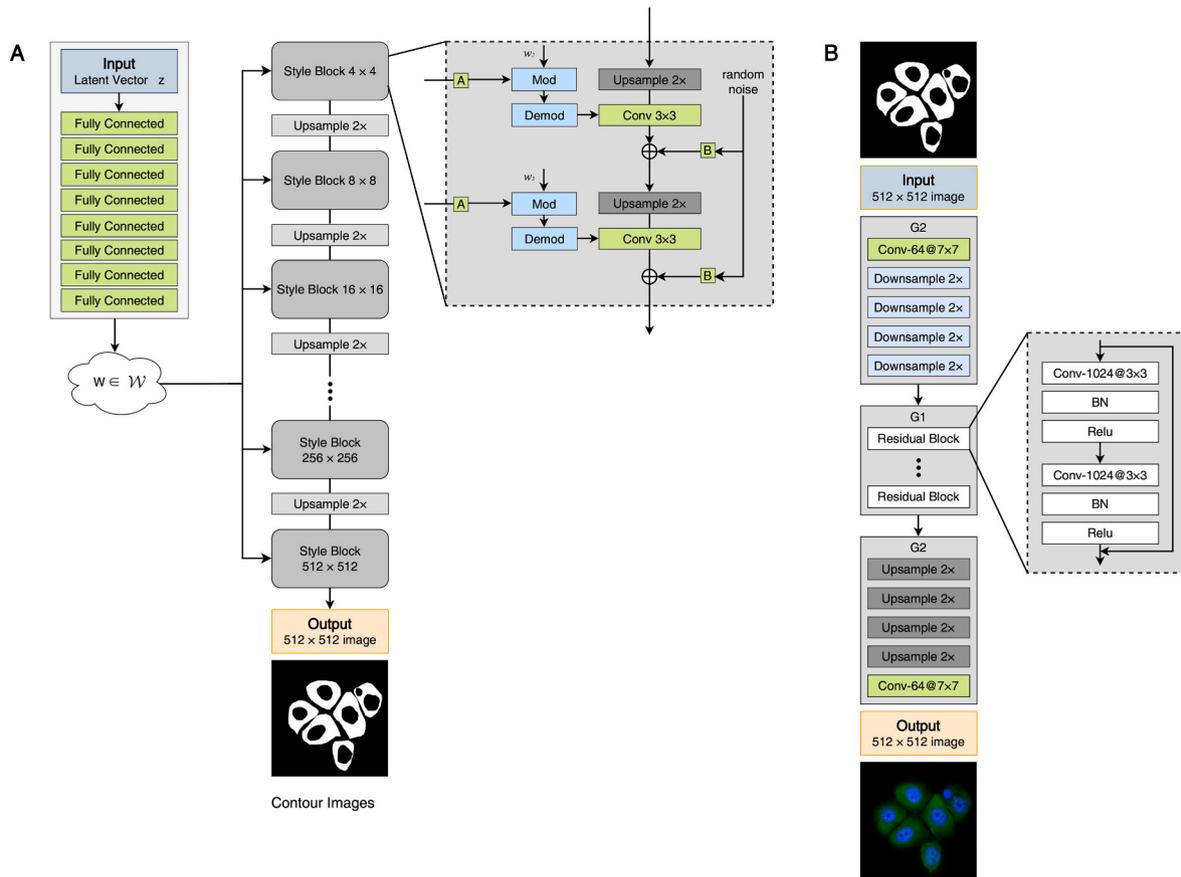In the initial training phase of the StyleGAN2 network, the generator

Fig. 1. Synthesizing cellular images with varied densities using generative adversarial networks. (A) Schematic of the StyleGAN2 Network Structure. The StyleGAN2 network receives a latent vector z as input, which traverses numerous fully connected layers and sequential style blocks, culminating in the production of 512×512 pixel cell contour images. Each style block is composed of upsampling and convolutional layers. (B) Schematic of the Pix2PixHD Network Structure. The input contour image undergoes a succession of convolutional and residual block processes, succeeded by an upsampling layer, to yield the ultimate fluorescent cell images.

network initially focuses on producing low-resolution images. This strategy allows the network to concentrate on learning the basic structure and features of images in the early stages. As training progresses, the network gradually increases the resolution of the images, thereby refining and enhancing the high-level details of the images. This strategy of progressively increasing resolution, drawing on the principles of Progressive GAN [30], effectively balances detail capture and structure establishment during the learning process. Using this method, Style-GAN2 can avoid the interference of complex details in the early stages of training, focusing on constructing the overall architecture of the image, then gradually moving to higher resolutions to precisely capture details, thus generating high-quality and highly realistic images.

The StyleGAN2 network receives a d-dimensional Gaussian distributed random vector z as input, forming the Z space [31]. As illustrated in Fig. 1A, the Z space is transposed into an intermediate latent space W following processing through a mapping network comprising eight fully connected layers. This transformation process effectively alters the input from a random space to a style space. In this style space, distinct image attributes are encoded in dimensions that are nearly orthogonal, thereby allowing for smooth interpolation of specific image attributes within the W space. Subsequently, as depicted in the Style Block of Fig. 1A, the output from the W space is conveyed to each layer of the generator network via a style modulation process, which introduces style characteristics by modulating the weights of the convolutional layers [32,33]. The modulation process is mathematically represented as follows.

$$\mathbf{w}'_{ijk} = \mathbf{s}_i \cdot \mathbf{w}_{ijk} \tag{1}$$

where i is the index of the input feature map, j and k represent the indices

of the output feature map and the spatial dimensions of the convolution.

Following the style modulation, a demodulation process is introduced to ensure the consistency of the feature maps and enhance model stability. This process entails calculating the L2 norm of the modulation weights and utilizing this norm to normalize the output feature maps, thereby adjusting their standard deviation to a unitary value. The formula for calculating the demodulation weights is as follows.

$$\mathbf{w}''_{ijk} = \frac{\mathbf{w}'_{ijk}}{\sqrt{\sum_{ij}\left(\mathbf{w}'_{ijk}\right)^2 + \epsilon}} \tag{2}$$

where $\epsilon = 10^{-8}$ represents a constant introduced to prevent the issue of the denominator becoming zero.

Through these two phases, the feature maps not only retain the style information but also have their dimensions and statistical properties suitably adjusted. Additionally, StyleGAN2 integrates noise injection at each resolution layer, infusing random details into the images. In the contour generation phase, solely the generator component of the network is utilized. It synthesizes fluorescent cell mask images possessing diverse style attributes by amalgamating the latent vector z with a fixed constant input, thus guaranteeing the diversity of the generated data.

2.1.2. Contour generation

During the contour generation phase, as delineated in Section 2.1.1, our study employs the generator network of StyleGAN2. Initially, a latent vector z is inputted and converted into an intermediate vector w

through a mapping network. Subsequently, the synthesis network utilizes this intermediate vector w, processed via the style modulation and demodulation procedures, as delineated in Equations (1) and (2). This process, in conjunction with a fixed constant input (const 4 × 4 × 512 vector), facilitates the creation of fluorescent cell mask images. The strength of this approach is its flexibility in producing cell contours with diverse style attributes by manipulating the input latent vector z, thereby ensuring substantial diversity in the data. To exercise precise control over the cell density in the generated images, we have incorporated the Truncation Trick in our study [34,35]. This technique modulates the position of the intermediate vector w relative to its average value $\overline{w}$ in the style space W, effectively modulating cell density. This is achieved by varying the distance between the latent vector and its mean. Consequently, this enables the images to display varying levels of cell density while maintaining their stylistic features. The modulation process is formulated as follows.

$$\mathrm{w}^{'} = \overline{\mathrm{w}} + \psi(\mathrm{w} - \overline{\mathrm{w}}) \tag{3}$$

where w represents the mean of the latent space; ψ is a truncation threshold, used to balance the quality and diversity of the generated images; w is the original random vector drawn from the latent space.

By modulating the ψ value, we can effectively control the resemblance between the generated image and the mean image, thereby achieving precise regulation of cell density. A smaller ψ value leads to the image being more akin to the mean image, hence reducing the cell density; conversely, a larger ψ value permits greater deviation, resulting in images with higher cell density. This method effectively regulates the variation in cell density in the generated images while maintaining image quality. Following the contour generation module in Section 2.1, we acquired three cell contour datasets with varying density gradients. These datasets will serve as inputs in the ensuing image rendering phase, establishing the groundwork for further image rendering endeavors.

2.2. Image rendering module

To accomplish the translation from fluorescent cell contour images to realistic cell images, we utilized the Pix2PixHD model in the image rendering module [36]. The Pix2PixHD model is tailored for high-resolution and multi-scale image translation tasks, guaranteeing that the output images exhibit high consistency with the input images in structure, detail, and semantics. In the generation phase, integrating the fluorescent cell contour images from the contour generation module, as outlined in Section 2.1, ensures that the resultant cell images are both precise and detailed.

2.2.1. Network architecture

The Pix2PixHD framework is distinguished by its distinctive coarse-to-fine generator architecture, comprising two integral components: the Global Generator Network (GGN) [37], referred to as G1, and the Local Enhancer Network (LEN), known as G2. As demonstrated in Fig. 1(B), the GGN (G1) primarily focuses on capturing the overall structure and major features of an image. In contrast, the LEN (G2) specializes in refining and enhancing finer details. As illustrated in Fig. 1B, the process begins with the LEN (G2), which extracts features from high-resolution 1024 × 1024 real label images. These high-resolution images are then downsampled to 512 × 512 label maps and fed into the GGN (G1). Here, G1 primarily processes the image's outer contours and overall geometric structure [38]. Subsequently, the output from G1, now upscaled back to 1024 × 1024, is merged with the high-resolution feature extraction results from LEN (G2). This fusion not only maintains consistency across various scales but also significantly improves the local detail richness and precision, such as accurately reproducing cellular textures.

2.2.2. Image rendering

During the generation phase, only the generator component of the

image rendering module, as described in Section 2.2.1, is employed. With a cell contour label map input to G2 (LEN), G2 captures local details and texture information in the label map through its advanced feature extraction capability at this stage. These features are then fed into G1 (GGN), where G1 is tasked with processing and assimilating the global information of the image, including overall shape and geometric structure. Additionally, G1's output, a 512 × 512 feature map, is recombined with features in G2's downsampling module, augmenting the quality of local details in the generated image while preserving the overall structural consistency. After the fusion, G2's upsampling module further processes these features, ultimately yielding an output image with a resolution of 1024 × 1024. The generated cell images not only exhibit high visual similarity to actual cell images but also maintain semantic consistency with the input label images. This semantic consistency guarantees that the details and general structure of the generated images correspond with the semantic information in the label maps, thus enhancing their reliability and applicability in cytological applications while preserving realism. Through the application of this methodology, this study successfully converted the images generated in Section 2.1 contour generation module into fluorescent cell images, creating a comprehensive cell image dataset encompassing multi-level density information.

2.3. Training procedures

In our study, we implemented a two-stage training strategy, tailored to the specific goals of each module. The training of the contour generation module was conducted on a Linux server equipped with two NVIDIA P40 GPUs. During this phase, StyleGAN2 was employed to receive random latent vectors z and generate two-dimensional cell contour images, focusing on creating diverse and density-controllable contours. The optimizer was set to the momentum-based Adam algorithm for training, with momentum parameters $\beta_1 = 0$ and $\beta_2 = 0.99$ [39]. Following the method in Ref. [30], the weights were subjected to an exponential moving average with a decay rate of 0.999. The initial learning rates for both the generator and discriminator were set at 0.001 and 0.0001 for the mapping network transforming latent vectors from Z space to W space. The network updates for the generator and discriminator involved two types of loss functions: the Wasserstein loss with gradient penalty (WGAN-GP) [40] and the non-saturating loss with R1 regularization [41,42]. We progressively reduced the batch sizes from 256 to 8 to address the learning challenges associated with increasing resolutions.

The training of image rendering module was centered on converting the generated cell contours into realistic images. Conducted on the same Linux server, this phase utilized the Pix2PixHD model. We maintained a consistent batch size of 8 and capped the training at 200 epochs, each comprising 200 iterations. The Adam algorithm was chosen as the optimizer, with a stable initial learning rate for the first 100 epochs, followed by a gradual decay to zero. For the generator and discriminator network updates, we employed CGAN loss and feature-matching loss functions. The successful outcome of this training phase was the production of images that closely align visually and semantically with the input contour label maps, achieving high-quality image replication.

2.4. Implementation details of segmentation techniques

Within the realm of automated cell segmentation technology, CellPose, DeepCell, and CellProfiler are broadly recognized as the foremost and most prevalent methodologies. Consequently, we chose these three technologies for the systematic evaluation of our comprehensive and accurately assembled image dataset. This section will furnish detailed insights into the user application specifics of the three automated cell segmentation technologies utilized in this study: CellPose, DeepCell, and CellProfiler.

2.4.1. CellPose

CellPose features an intuitive interface that facilitates users in easily uploading and processing images. Subsequently, based on the image content, a pre-trained model is selected, such as the cyto model for segmenting cytoplasm-stained cells, and the nuclei model for segmenting cell nuclei. Users then need to select the appropriate channel for the RGB or multi-channel image to be segmented. Additionally, the training of the cytoplasm model in CellPose is based on dual-channel images, wherein the first channel is the one targeted for segmentation, and the second is an optional nuclear channel. Finally, to ensure precise segmentation, CellPose requires users to input information about the cell diameter, which can be manually set or automatically estimated by the system. In using CellPose, the cell nuclei are designated to the nuclei model, and the cytoplasm is assigned to the cyto model for prediction.

2.4.2. DeepCell

In our study, we accessed DeepCell's Mesmer model through a Python program and prepared the density-graded image data for analysis. The image data comprised two channels: one for the nucleus (blue channel) and another for the cytoplasm (green channel). Subsequently, we verified that the image resolution matched the resolution utilized during the model's training. During prediction, the model's default settings can be modified to enhance segmentation accuracy. For example, adjusting the interior_threshold parameter can regulate the size of the cells, while the maxima_threshold parameter influences the number of independent cells predicted.

2.4.3. CellProfiler

The workflow in CellProfiler is principally predicated on its modular design, enabling users to devise their own image analysis pipeline by selecting and configuring a series of modules. In this study, the IdentifyPrimaryObject module is initially employed to delineate the primary areas of cells, such as the nucleus. This module is capable of effectively detecting bright spots in darker background areas. Subsequently, the IdentifySecondaryObjects module utilizes the objects identified in the previous step to further define and delineate the boundaries of the cells. Once segmentation is completed, the MeasureObjectIntensity module quantifies the intensity of each cell. Finally, all measurement data can be exported as image files using the SaveImages module, thereby facilitating subsequent data analysis.

3. Evaluation metrics

This study focuses on generating more diversified and density-manageable cell image data, and constructing an exhaustive fluorescent cell dataset that accurately reflects the diversity in cell morphology. To ascertain that the quality of the generated images aligns with that of real data, this chapter will delve into the evaluation metrics and their respective analysis methods. These metrics are designed to quantitatively measure the performance of the generated images and systematically evaluate their resemblance to the original dataset. Concurrently, this collection of generated image datasets will be employed for a systematic evaluation of various cell segmentation methods. Different methods will be compared in terms of their effectiveness in image data processing, utilizing a series of cell segmentation metrics. In subsequent sections, these evaluation metrics will be explored in greater detail, further elucidating their significance and application in practical scenarios.

3.1. Performance assessment for contour generation

To efficaciously appraise the similarity between the synthesized cell contour images and the original dataset, the Fréchet Inception Distance (FID) was employed as the evaluation metric [43,44]. FID is a methodology to quantify the differences in statistical distribution between synthesized and authentic images within the latent feature space of a designated deep learning model. This is accomplished by computing the Fréchet distance between the multivariate Gaussian distributions of two image sets. A low FID value signifies substantial closeness between the distributions of synthesized and real images, implying a high degree of similarity. The formula for computing FID is delineated as follows.

$$\text{FID} = \|\mu_{\text{real}} - \mu_{\text{fake}}\|^2 + \text{Tr}\left(C_{\text{real}} + C_{\text{fake}} - 2\sqrt{C_{\text{real}}C_{\text{fake}}}\right) \tag{4}$$

where μ_{real} represents the mean of real images in the latent space of the deep model, μ_{fake} denotes the mean of generated images in the same latent space, C_{real} is the covariance matrix of real images within that latent space, C_{fake} is the covariance matrix of generated images in said latent space, Tr signifies the trace operation on a matrix, and $|\cdot|$ denotes the Euclidean norm of a vector.

3.2. Effectiveness evaluation for image rendering

To thoroughly assess the effectiveness of our image rendering module, we utilized evaluation methods aligned with those established in previous image rendering research. The primary objective of image rendering is to produce images that realistically mirror the original label maps. Consequently, evaluating the quality of these generated images is a critical component of determining the overall model performance. For this purpose, we conducted semantic segmentation on the synthesized images and compared the segmented outcomes with the original input labels [45]. This methodology is predicated on the assumption that if the model successfully generates realistic images corresponding to the input label maps, then a competent semantic segmentation model (such as the FCN-UNet employed in our study) should accurately predict the original labels. In line with this, we employed several pivotal evaluation metrics: average accuracy (aAcc), mean Intersection over Union (mIoU) [46], mean accuracy (mAcc), mean Dice coefficient (mDice) [47], mean precision (mPrecision), and mean recall (mRecall) [48]. These metrics are integral not only to the image rendering task but also essential for semantic segmentation. To further validate the realism and visual quality of the generated images, we also employed the FID metric, as discussed in Section 3.1, to measure the consistency between the generated images and real images. The subsequent sections will offer a comprehensive overview and detailed explanations of each of these evaluation metrics.

Average accuracy (aAcc): This metric quantifies the model's overall performance across all categories. It is the mean of accuracy for each individual class, and it can be expressed as follows.

$$\text{aAcc} = \frac{\sum_{i=1}^{n} \text{TP}_i}{\sum_{i=1}^{n} (\text{TP}_i + \text{FN}_i)} \tag{5}$$

where i refers to each class in semantic segmentation. Each class corresponds to a specific pixel value. For instance, 0 represents the background, 128 represents the nucleus, and 255 represents the cytoplasm. TP_i and FN_i are the true positives and false negatives for the ith class, and n is the total number of classes.

Mean intersection over union (mIoU): This metric is employed for segmentation tasks, and it measures the overlap between model predictions and actual labels. It indicates the mean IoU over all classes.

$$\text{mIoU} = \frac{1}{n} \sum_{i=1}^{n} \frac{\text{TP}_i}{(\text{TP}_i + \text{FP}_i + \text{FN}_i)} \tag{6}$$

Mean accuracy (mAcc): This parameter provides a more granular view of accuracy, which is calculated within each class and then averaged.

$$\text{mAcc} = \frac{1}{n} \sum_{i=1}^{n} \frac{\text{TP}_i}{(\text{TP}_i + \text{FN}_i)} \tag{7}$$

Mean dice coefficient (mDice): This metric gauges the similarity between model predictions and the actual labels.

$$mDice = \frac{1}{n} \sum_{i=1}^{n} \frac{2 \times TP_i}{(2 \times TP_i + FP_i + FN_i)} \tag{8}$$

Mean precision (mPrecision): Precision is the ratio of true positives to the sum of true and false positives.

$$mPrecision = \frac{1}{n} \sum_{i=1}^{n} \frac{TP_i}{(TP_i + FP_i)} \tag{9}$$

Mean recall (mRecall): Recall is the ratio of true positives to the sum of true positives and false negatives.

$$mRecall = \frac{1}{n} \sum_{i=1}^{n} \frac{TP_i}{(TP_i + FN_i)} \tag{10}$$

The comprehensive evaluation of these metrics provides an in-depth, multi-dimensional understanding of image rendering module's performance and accuracy.

3.3. Comprehensive analysis of cell segmentation methods

To conduct a systematic evaluation of the generated dataset, we elected to evaluate segmentation performance across varying cell density levels utilizing four fundamental metrics: mIoU (mean Intersection over Union), mDice (mean Dice coefficient), mFscore (mean F-score), and mPrecision (mean Precision). These metrics are indispensable for quantifying the precision and efficacy of segmentation. For comprehensive formulas and elucidations of these evaluation metrics, please consult section 3.2 of this study. This structured methodology permits a detailed evaluation of each segmentation framework, underscoring their strengths and limitations in managing diverse cell densities in the generated dataset.

4. Results

4.1. Performance evaluation of contour generation

This section details the performance outcomes of the contour generation module. Employing the previously outlined contour generation method and the Fréchet Inception Distance (FID) as an evaluation metric, a thorough assessment was conducted on the generated cell contour images. The results, as depicted in Table 1, show that our contour generation module achieved an FID score of 3.04. This score not only surpasses those of generation models like DCGAN and PGGAN but also expresses the efficacy of the StyleGAN2 model in producing high-quality and diverse fluorescent cell images. Consequently, these findings proved the practical relevance of our study in the realm of cell image generation.

As depicted in Fig. 2, which illustrates the qualitative evolution of images generated at various stages of model training, we observed a progressive enhancement in the quality of these images. Initially, during the early phases of training, the images were often blurry, overly uniform, or structurally implausible. However, with an increase in the number of training iterations (measured in 'Kimg', where 'Kimg' signifies thousands of images processed by the network), there was a marked improvement in image clarity and realism. This improvement was demonstrated in our experimental results. Additionally, the model's ability to adapt to various styles and content became increasingly refined, enabling the generation of a wider array of high-quality images. Following our implementation of the truncation trick, we generated cell

Table 1
Comparative evaluation of cell contour image generation models based on FID scores.

Methods	DCGAN [49]	PGGAN [30]	StyleGAN2 (Ours)
FID	9.34	6.07	3.04

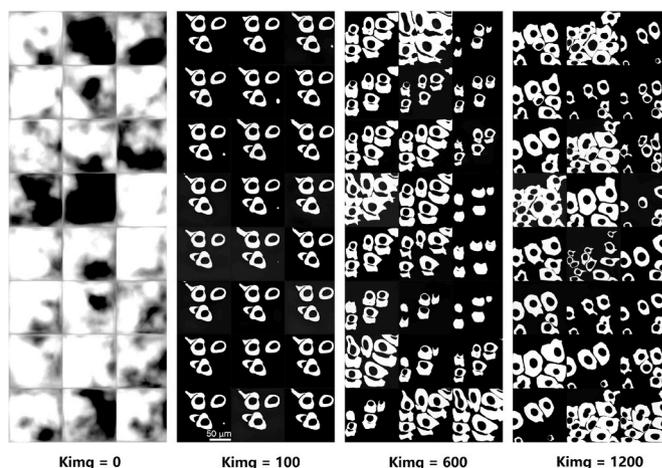| Kimg = 0 | Kimg = 100 | Kimg = 600 | Kimg = 1200 |

Fig. 2. Enhancement in image quality with increasing Kimg training iterations. At Kimg = 0, the generated contour images are markedly blurry and irregular, rendering cell structure identification almost unfeasible; at Kimg = 100, the contour images exhibit a lack of detailed information, appearing as homogeneous areas; upon reaching Kimg = 600, the generated contour images become more detailed, yet continue to display areas with structurally unrealistic features; ultimately, at Kimg = 1200, the generated contour images become nearly indistinguishable from authentic images, showcasing exceptional clarity and richness in detail. Scale Bar: 50 μm.

contours using 300 different random seeds. For each seed, we set a truncation value ψ at -1, -0.5, 0, 0.5, and 1, leading to the production of a total of 1500 cell contour images. Some of these images are presented in Fig. 3. We also categorized the generated images based on their density into three distinct levels of difficulty. Employing this methodology, we regulated the density of the generated cell images and classified them into three distinct density levels. This approach offers a robust data basis for future assessments of cell segmentation models, guaranteeing the efficiency and methodical nature of the evaluation process.

4.2. Performance evaluation of image rendering

In this experiment, the performance of Pix2PixHD, specifically in image rendering, was analysed along with CycleGAN and Pix2Pix. For the evaluation of the model's performance, assessment metrics previously delineated, including aAcc, mIoU, mAcc, mDice, mPrecision, and mRecall, were utilized, with detailed experimental outcomes exhibited in Table 3. The data reveals that the Pix2PixHD model achieved an aAcc of 95.16 %, outperforming CycleGAN's 83.39 % and Pix2Pix's 83.84 %.
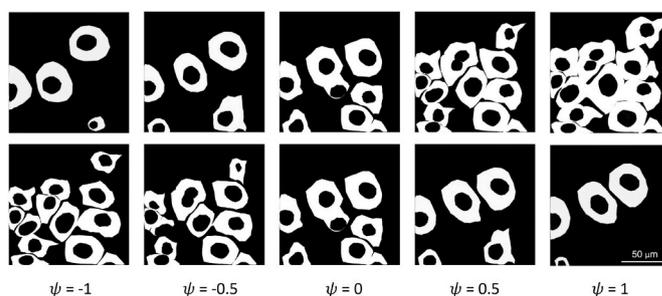

| $\psi = -1$ | $\psi = -0.5$ | $\psi = 0$ | $\psi = 0.5$ | $\psi = 1$ |

Fig. 3. Variation in cell contour image density across truncation values from -1 to 1. This observation reveals that irrespective of the gradual increase or decrease in truncation values, the resultant images persistently exhibit a monotonic trend of alteration concerning density and complexity. It is significant to note that the precise direction of this change trend (namely, whether incrementally increasing or decreasing) is contingent upon the selected random seed.

Table 2

Evaluation of cell render image generation models based on FID scores.

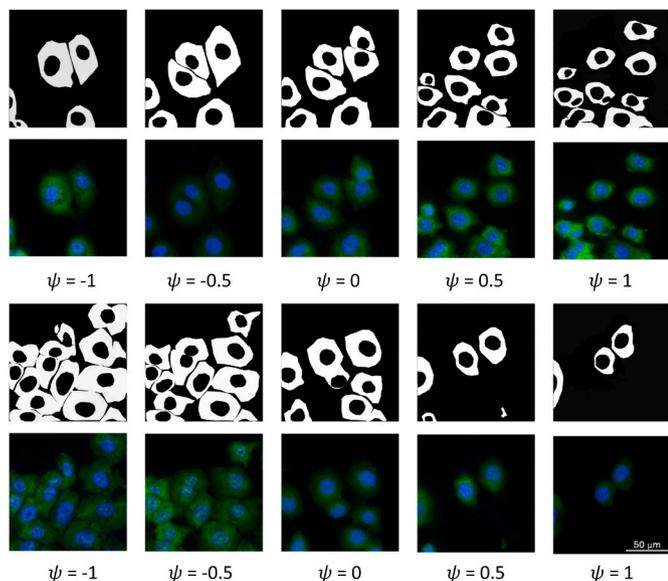| Density	CycleGAN [50]	Pix2Pix [51]	Pix2PixHD(Ours)
Low	23.51	27.29	1.61
Medium			2.79
High			6.62

Fig. 4. Synthesis of fluorescent cell images via Pix2PixHD from contour inputs. It showcases the stylized contours generated using varying truncation values (ranging from −1 to 1) and the corresponding fluorescent cell images produced by the Pix2PixHD network.

This result highlights Pix2PixHD's ability to distinguish foreground from background. Moreover, Pix2PixHD scored 88.27 % in mIoU, compared to 64 % and 63 % for CycleGAN and Pix2Pix respectively, indicating that images generated by Pix2PixHD have a higher pixel-level consistency with actual labels. Pix2PixHD also demonstrated good performance in other key metrics like mAcc, mDice, and mPrecision. These aggregate metrics attest to the comprehensive superiority of Pix2PixHD in image rending tasks.

Furthermore, this study also incorporated the FID to evaluate the similarity between the rendered cell images and actual images, comparing the results with those from Pix2Pix and CycleGAN. As shown in Table 2, for the FID scores, CycleGAN and Pix2Pix exhibited higher scores of 23.51 and 27.29, indicating a lower visual similarity to real images. In contrast, Pix2PixHD achieved favorable results on the FID metric, recording a score of 1.61 for low density images, and scores of 2.79 and 6.62 for medium and high densities. This result indicates that Pix2PixHD maintains a strong visual similarity to actual images across a range of densities, underscoring its effectiveness in image rendering. Combined with the earlier contour generation module, we successfully completed the rendering from contours to images, forming a complete dataset with density gradation. Fig. 4 showcases ten fluorescent cell images generated from contour pictures in the 3.1 Contour Generation Module. This provides a rich data foundation for subsequent

applications and lays a solid foundation for future research in related fields.

4.3. Performance of segmentation results

To systematically assess the performance of different segmentation methods on our generated dataset, we first applied CellPose, DeepCell, and CellProfiler to the real dataset categorized into various density levels. This step not only established a realistic performance baseline but also ensured the effectiveness of our generated data when subjected to segmentation methods. We then assessed the performance of these three methods on our generated image data, aligning with similar density levels as those of the real data. Additionally, considering the noise issues commonly encountered in fluorescence microscopy imaging, we introduced Gaussian and Poisson noise into our generated data for another set of experiments following the aforementioned density configurations. This approach allowed us to evaluate the segmentation methods' performance under ideal conditions and test their robustness in challenging scenarios, such as low signal-to-noise ratio (SNR) conditions.

Fig. 5 presents the segmentation results of three methods on real data, spanning different densities. In the dataset with the lowest density, CellProfiler's segmentation of cytoplasm achieved an mDice of 88.34 % and an mIoU of 79.63 %, displaying effective accuracy in segmenting low-density images. Its nuclear segmentation results were also strong, with an mDice of 86.37 % and an mIoU of 76.89 %. CellPose exhibited commendable performance in this density range as well, with cytoplasmic mDice and mIoU scores of 82.27 % and 71.9 %, and nuclear mDice and mIoU scores of 86.19 % and 77.06 %.

Moving to the medium density dataset, CellProfiler's reliable segmentation performance is reflected in cytoplasmic mDice and mIoU scores of 88.27 % and 79.59 %, and nuclear mDice and mIoU scores of 87.14 % and 77.77 %. CellPose also showed consistent performance, particularly in nuclear segmentation, with mDice and mIoU scores of 84.7 % and 75.59 %. DeepCell's results for medium-density cytoplasm segmentation, with mDice and mIoU scores of 71.01 % and 58.94 %, suggest areas for potential enhancement.

In the high density dataset, the steady performance of CellProfiler is again observed with cytoplasmic mDice and mIoU scores of 87.68 % and 79.12 %, and nuclear mDice and mIoU scores of 90.39 % and 82.69 %. CellPose maintained its strong performance in nuclear segmentation with mDice and mIoU scores of 85.37 % and 76.56 %. DeepCell's performance in the high-density setting, particularly for nuclear segmentation with mDice and mIoU scores of 47.39 % and 32.61 %, highlights areas where further improvements could be beneficial.

Fig. 6 showcases performance trends of segmentation methods on our generated image dataset as cell density increases from lowest to highest. In the dataset with the lowest cell density, CellProfiler's mDice and mIoU for cytoplasm were impressive at 92.01 % and 85.41 %, respectively, indicating its robust performance. For cytoplasm, CellPose achieved mDice and mIoU scores of 71.81 % and 65.56 %, and for nuclei, its performance was solid with mDice and mIoU scores of 76.19 % and 67.46 %.

In the medium density dataset, CellProfiler continued to perform well, with mDice and mIoU for cytoplasm at 92.44 % and 86.03 %, respectively. Meanwhile, CellPose demonstrated its capabilities in nuclear segmentation with mDice and mIoU scores of 82.03 % and 71.68 %. DeepCell showed less optimal performance, particularly in nuclear
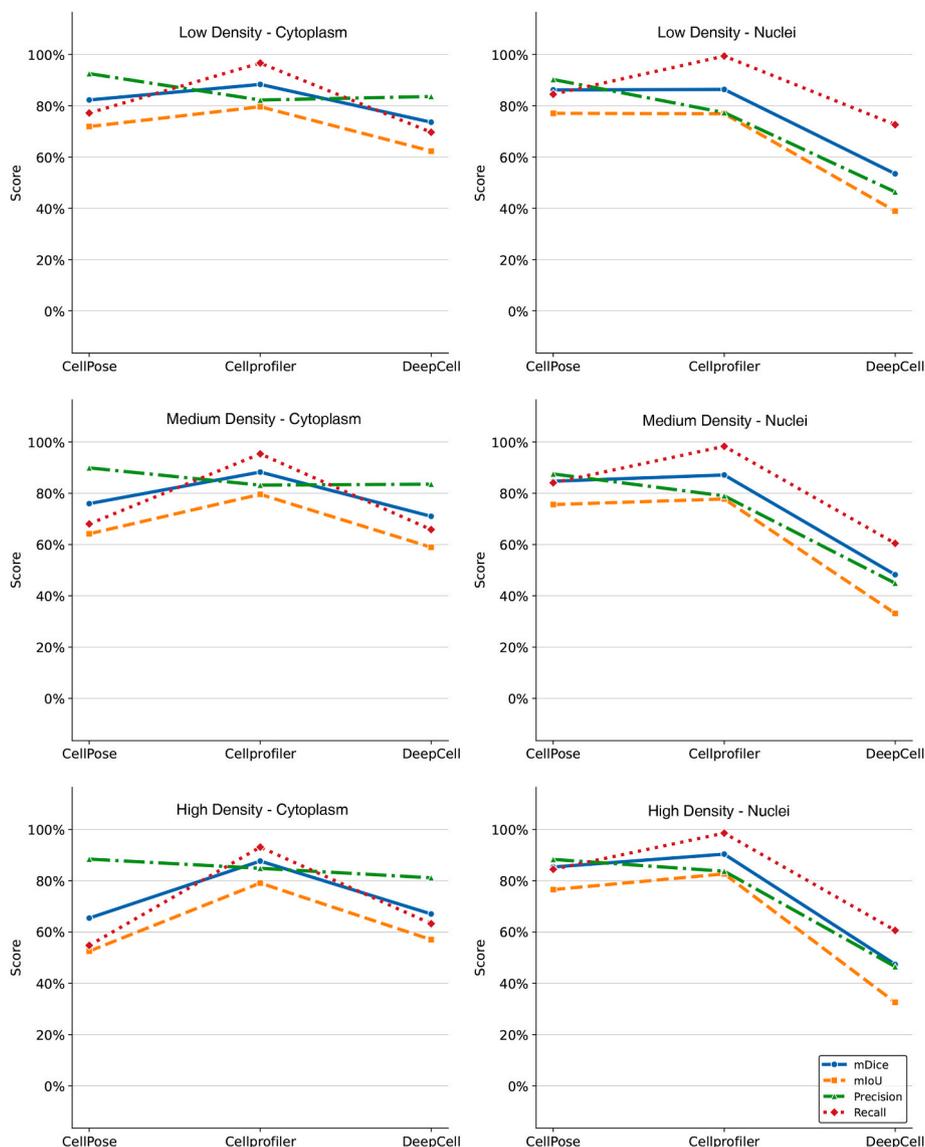
Table 3

Performance comparison of image rendering models on fluorescent cell dataset.

Methods	aAcc	mIoU	mAcc	mDice	mPrecision	mRecall
CycleGAN [50]	83.39	64.650	74.84	77.10	85.03	74.84
Pix2Pix [51]	83.84	63.14	73.75	76.39	81.24	73.75
Pix2PixHD(Ours)	95.16	88.27	94.44	93.72	93.04	94.44

Fig. 5. Segmentation outcomes for cytoplasm and nucleus across varied densities on real images.

segmentation, where mDice and mIoU were recorded at 53.89 % and 37.88 %.

Within the highest density dataset, CellProfiler's performance remained consistent, achieving 92.23 % in mDice and 85.65 % in mIoU for cytoplasm. It also exhibited excellent performance in nuclear segmentation with mDice and mIoU scores of 90.39 % and 82.55 %. On the other hand, DeepCell's performance in nuclear segmentation was less satisfactory, with mDice and mIoU scores of 54.54 % and 38.53 %.

Fig. 7 details the segmentation outcomes of methods applied to generated image datasets subjected to Gaussian and Poisson noise, across varying densities. In the dataset with the lowest density, Cell-Profiler showed strong performance in cytoplasm segmentation, achieving an mDice of 93.27 % and an mIoU of 87.56 %, demonstrating its effectiveness even in the presence of noise. Its nuclear segmentation results were solid, with an mDice of 80.65 % and an mIoU of 70.08 %. CellPose, in this density, registered cytoplasmic mDice and mIoU scores of 57.46 % and 56.43 %, and for nuclei, mDice and mIoU scores of 44.3 % and 32.66 %, indicating a decrease in performance attributable to noise.

In the medium density dataset, CellProfiler continued to show high accuracy in segmentation, with cytoplasmic mDice and mIoU scores of 93.22 % and 87.38 %, and nuclear mDice and mIoU scores of 91.17 %

and 84.11 %. CellPose, however, experienced a drop in performance, with cytoplasmic mDice and mIoU scores of 48.49 % and 36.74 %, and nuclear mDice and mIoU scores of 49.42 % and 36.48 %.

In the high density dataset, CellProfiler's performance remained good, with cytoplasmic mDice and mIoU scores of 92.2 % and 85.6 %, and nuclear mDice and mIoU scores of 92.05 % and 85.48 %. CellPose showed further reduction in effectiveness, with cytoplasmic mDice and mIoU scores of 34.45 % and 23.41 %, and nuclear mDice and mIoU scores of 44.71 % and 32.62 %. DeepCell's results across densities indicated a good performance in cytoplasm segmentation but high-lighted challenges in nuclear segmentation, especially in high-density settings with mDice and mIoU scores of 43.04 % and 28.82 %.

In summary, when compared to segmentation on real data, the generated datasets in our study yielded results that are comparable to those obtained from actual datasets. This underscores the effectiveness and applicability of our synthetic images for use in fluorescence imaging research. In scenarios without added noise, both CellProfiler and DeepCell showed good performance in segmenting the cytoplasm, with CellProfiler edging ahead in the segmentation of nuclei. However, when noise was introduced, CellProfiler stood out by maintaining consistent performance in both cytoplasm and nuclear regions. DeepCell saw a modest decrease in its ability to segment the cytoplasm and a more
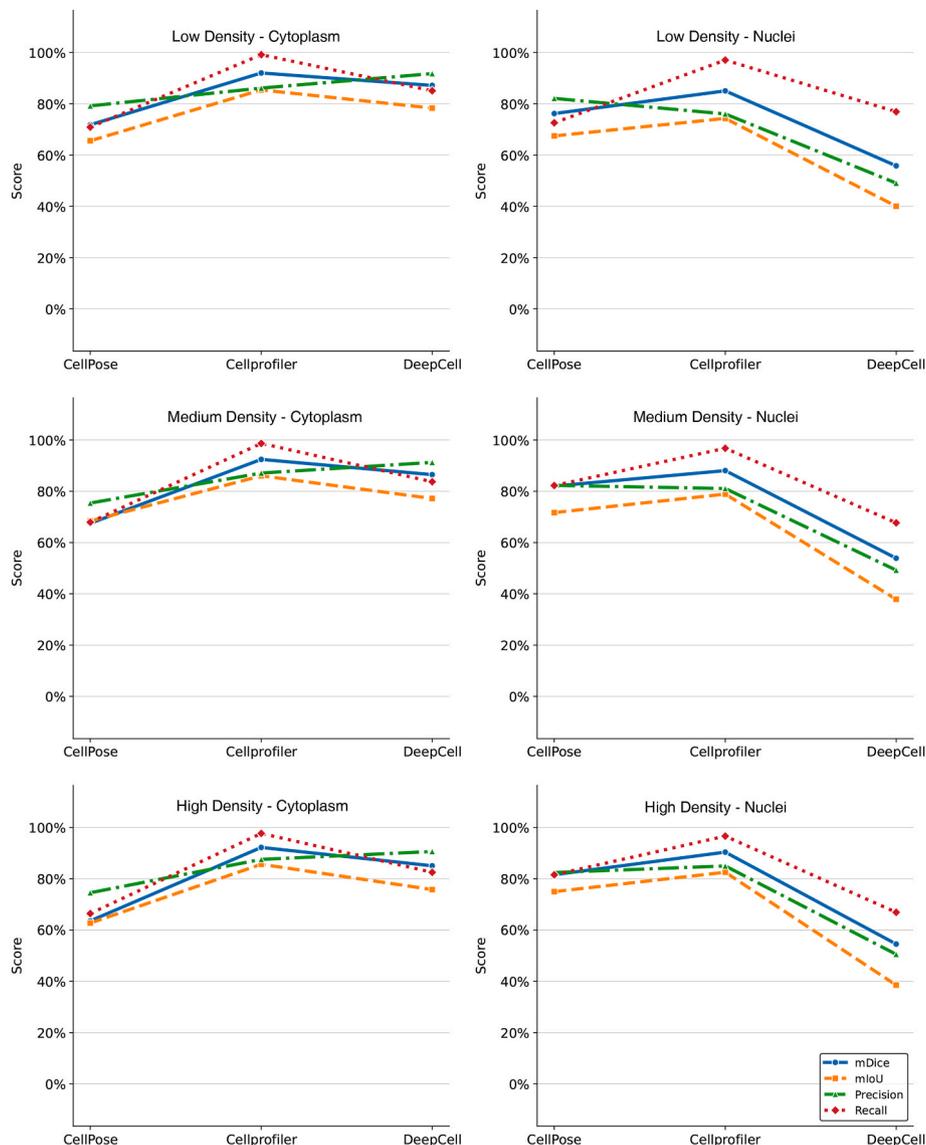
Fig. 6. Segmentation outcomes for cytoplasm and nucleus across varied densities on generated images.

noticeable drop in nuclear segmentation, while CellPose experienced substantial reductions in performance for both.

This observation indicates CellProfiler's resilience to noise, suggesting it as a dependable choice for cell segmentation under varied imaging conditions. The decline in performance for DeepCell and Cell-Pose, especially under noisy conditions, signals opportunities for further improvement in these models.

5. Discussion

In recent decades, technological advancements have facilitated the generation of high-throughput image data via fluorescence imaging, advancing research in the life sciences. Cell segmentation, a fundamental step in data analysis, encounters challenges in systematic evaluation due to the absence of standardized, comprehensive, and graded image datasets, complicating the determination of superior segmentation methods. This study represents the inaugural endeavor to employ Generative Adversarial Networks (GANs) in addressing these challenges, developing a fluorescence dataset with graded density difficulty standards. This not only addresses a void in fluorescence imaging research but also furnishes valuable data resources for future endeavors in image generation and segmentation.

Firstly, the study conducts an exploration of the StyleGAN2 network's performance, as utilized in the contour generation module for producing contour images. Comparative evaluations with DCGAN and PGGAN indicate that StyleGAN2 has an advantage in the FID metric. This supports StyleGAN2's ability to generate quality and diverse images, balancing complexity and training efficiency. With progressive training iterations, both the quality and the generative density of images have markedly improved, indicating StyleGAN2's potential in generating fluorescent cell contours.

Secondly, the study highlights the performance of the Pix2PixHD model in generating fluorescent cell images within the image rendering module. Through analysis, it was found that Pix2PixHD outperforms CycleGAN and Pix2Pix in metrics like aAcc and mIoU. This not only demonstrates Pix2PixHD's capability in image segmentation accuracy but also evidences the consistency between its generated images and the actual dataset. Consequently, Pix2PixHD possesses practical value in image analysis within the life sciences, particularly in the generation and segmentation of fluorescent cell images.

Additionally, the study utilized the generated dataset to evaluate three widely recognized cell segmentation methods: DeepCell, CellProfiler, and CellPose. The findings revealed a general trend: as cell density increased, the performance of the segmentation algorithms generally
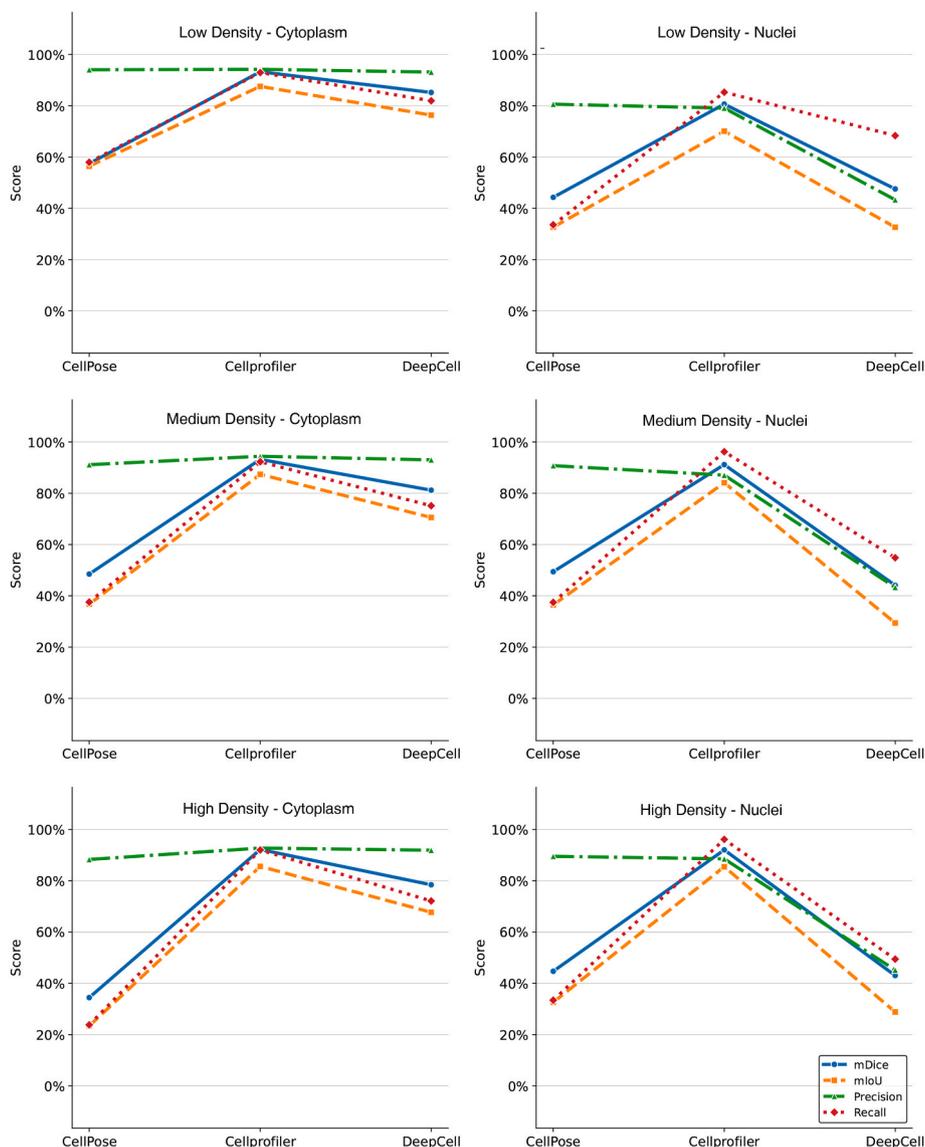
Fig. 7. Segmentation performance on generated datasets with added noise across different densities.

decreased. Notably, in low-density settings, DeepCell excelled in cytoplasm segmentation, while CellPose showed some limitations in nuclear segmentation metrics. In contrast, at higher cell densities, CellProfiler demonstrated good performance in cytoplasmic mDice scores. The results indicated that CellProfiler showed good performance across various evaluation metrics. This discovery offers insights for the domain of life science image analysis and provides guidance to researchers.

Regarding future research trajectories, our intent is to substantially enhance and broaden this study across multiple critical domains. Initially, we intend to investigate the applicability of this data generation framework to an expanded spectrum of biomedical image types, including images of cells in tissue sections or within microfluidic devices. Subsequently, we propose to refine and adapt the algorithms to amplify their control over cellular features. Finally, we will delve into the utilization of algorithms in multi-channel imaging to enhance analytical capabilities for complex biological specimens. Moreover, while our study has successfully demonstrated the potential of GANs in generating fluorescent cell images with varied densities, a critical aspect that warrants future exploration is the quantitative analysis of cell densities and shape complexities. Until now, our endeavors have largely focused on evaluating the overarching capabilities of our generative models, with less emphasis placed on the meticulous control of

particular attributes, including monotonic density variations or the intricacies of the cell shapes produced. Acknowledging the significance of these attributes in fluorescence imaging and their prospective influence on the precision and efficacy of segmentation methodologies, forthcoming studies will adopt a more stringent and quantitative methodology for the examination of these traits.

Data and code availability

All datasets [52] used for training the contour generation and image rendering modules are available on Mendeley data: https://doi.org/10.17632/6hsf4fyhsn.2; https://doi.org/10.17632/9s9m4wytfw.1; https://doi.org/10.17632/cv7n2bbcb4.1.

The source code and documentation of this study are available at https://github.com/edwardcao3026/SegBenchmark. The raw images datasets, along with their segmentation annotations are available for download at https://github.com/edwardcao3026/SegBenchmark/blob/main/datasets/datasets.tar.

CRediT author statement

Jun Tang performed research, analysed the data and wrote the

manuscript. Wei Du supervised research and wrote the manuscript. Zhanpeng Shu analysed the data. Zhixing Cao designed research, supervised research, acquired funding and wrote the manuscript with input from the others.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

References

[1] Samantara AK, Acharya C, Satpathy D, Panda CR, Bhaskara PK, Sasmal A. Functionalized graphene: an unique platform for biomedical application. In: Fullerens, graphenes and nanotubes. Elsevier; 2018. p. 545–84.

[2] Lakowicz JR. Principles of fluorescence spectroscopy. Springer; 2006.

[3] Guilbault GG. Practical fluorescence. CRC Press; 2020.

[4] Zhao W, Zhao S, Han Z, Ding X, Hu G, Qu L, Huang Y, Wang X, Mao H, Jiu Y, et al. Enhanced detection of fluorescence fluctuations for high-throughput super-resolution imaging. Nat Photonics 2023:1–8.

[5] Haimovich G, Gerst JE. Single-molecule fluorescence in situ hybridization (smFISH) for RNA detection in adherent animal cells. Bio-protocol 2018;8(21): e3070. e3070.

[6] Cao Z, Grima R. Analytical distributions for detailed models of stochastic gene expression in eukaryotic cells. Proc Natl Acad Sci USA 2020;117(9):4682–92.

[7] Fu X, Patel HP, Coppola S, Xu L, Cao Z, Lenstra TL, Grima R. Quantifying how post-transcriptional noise and gene copy number variation bias transcriptional parameter inference from mRNA distributions. Elife 2022;11:e82493.

[8] Cao Z, Filatova T, Oyarzún DA, Grima R. A stochastic model of gene expression with polymerase recruitment and pause release. Biophys J 2020;119(5):1002–14.

[9] Cao Z, Grima R. Linear mapping approximation of gene regulatory networks with stochastic dynamics. Nat Commun 2018;9(1):3305.

[10] Jiang Q, Fu X, Yan S, Li R, Du W, Cao Z, Qian F, Grima R. Neural network aided approximation and parameter inference of non-markovian models of gene expression. Nat Commun 2021;12(1):2618.

[11] Wang Y, Yu Z, Grima R, Cao Z. Exact solution of a three-stage model of stochastic gene expression including cell-cycle dynamics. J Chem Phys 2023;159(22).

[12] Ettinger A, Wittmann T. Fluorescence live cell imaging. Methods Cell Biol 2014; 123:77–94.

[13] Stephens DJ, Allan VJ. Light microscopy techniques for live cell imaging. Science 2003;300(5616):82–6.

[14] Laine RF, Heil HS, Coelho S, Nixon-Abell J, Jimenez A, Wiesner T, Martínez D, Galgani T, Régnier L, Stubb A, et al. High-fidelity 3D live-cell nanoscopy through data-driven enhanced super-resolution radial fluctuation. Nat Methods 2023:1–8.

[15] Xu H, Sepúlveda LA, Figard L, Sokac AM, Golding I. Combining protein and mRNA quantification to decipher transcriptional regulation. Nat Methods 2015;12(8): 739–42.

[16] Eng CL, Lawson M, Zhu Q, Dries R, Koulena N, Takei Y, Yun J, Cronin C, Karp C, Yuan G, et al. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. Nature 2019;568(7751):235–9.

[17] Sità L, Brondi M, Lagomarsino de Leon Roig P, Curreli S, Panniello M, Vecchia D, Fellin T. A deep-learning approach for online cell identification and trace extraction in functional two-photon calcium imaging. Nat Commun 2022;13(1): 1529.

[18] Aydın M, Kiraz B, Eren F, Uysallı Y, Morova B, Ozcan SC, Acilan C, Kiraz A. A deep learning model for automated segmentation of fluorescence cell images. In: Journal of physics: conference series, vol. 2191. IOP Publishing; 2022, 012003.

[19] Wang Y, Wang W, Liu D, Hou W, Zhou T, Ji Z. Genesegnet: a deep learning framework for cell segmentation by integrating gene expression and imaging. Genome Biol 2023;24(1):235.

[20] Jang J, Wang C, Zhang X, Choi HJ, Pan X, Lin B, Yu Y, Whittle C, Ryan M, Chen Y, et al. A deep learning-based segmentation pipeline for profiling cellular morphodynamics using multiple types of live cell microscopy. Cell Reports Methods 2021;1(7).

[21] Lugagne J, Lin H, Dunlop MJ. DeLTA: automated cell segmentation, tracking, and lineage reconstruction using deep learning. PLoS Comput Biol 2020;16(4): e1007673.

[22] Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. Nat Methods 2021;18(1):100–6.

[23] Greenwald NF, Miller G, Moen E, Kong A, Kagel A, Dougherty T, Fullaway CC, McIntosh BJ, Leow KX, Schwartz MS, et al. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. Nat Biotechnol 2022;40(4):555–65.

[24] Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA, Moffat J, et al. Cellprofiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol 2006;7:1–11.

[25] Van der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, Gouillart E, Yu T. scikit-image: image processing in Python. PeerJ 2014;2:e453.

[26] Caicedo JC, Roth J, Goodman A, Becker T, Karhohs KW, Broisin M, Molnar C, McQuin C, Singh S, Theis FJ, et al. Evaluation of deep learning strategies for nucleus segmentation in fluorescence images. Cytometry 2019;95(9):952–65.

[27] Wang N, Hu L, Walsh AJ. POSEA: a novel algorithm to evaluate the performance of multi-object instance image segmentation. PLoS One 2023;18(3):e0283692.

[28] Chen H, Murphy RF. Evaluation of cell segmentation methods without reference segmentations. Mol Biol Cell 2023;34(6):ar50.

[29] Karras T, Laine S, Aittala M, Hellsten J, Lehtinen J, Aila T. Analyzing and improving the image quality of stylegan. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2020. p. 8110–9.

[30] Karras T, Aila T, Laine S, Lehtinen J. Progressive growing of gans for improved quality, stability, and variation. 2017. arXiv preprint arXiv:1710.10196.

[31] Karras T, Laine S, Aila T. A style-based generator architecture for generative adversarial networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2019. p. 4401–10.

[32] Salimans T, Kingma DP. Weight normalization: a simple reparameterization to accelerate training of deep neural networks. Adv Neural Inf Process Syst 2016;29.

[33] Xiang S, Li H. On the effects of batch and weight normalization in generative adversarial networks. arXiv preprint arXiv:1704.03971 2017.

[34] Brock A, Donahue J, Simonyan K. Large scale GAN training for high fidelity natural image synthesis. 2018. arXiv preprint arXiv:1809.11096.

[35] Marchesi M. Megapixel size image creation using generative adversarial networks. arXiv preprint arXiv:1706.00082 2017.

[36] Wang T, Liu M, Zhu J, Tao A, Kautz J, Catanzaro B. High-resolution image synthesis and semantic manipulation with conditional gans. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2018. p. 8798–807.

[37] Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. In: Computer vision–ECCV 2016: 14th European conference, Amsterdam, The Netherlands, october 11-14, 2016, proceedings, Part II 14. Springer; 2016. p. 694–711.

[38] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 770–8.

[39] Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 2014.

[40] Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville AC. Improved training of wasserstein gans. Adv Neural Inf Process Syst 2017;30.

[41] Mescheder L, Geiger A, Nowozin S. Which training methods for GANs do actually converge. In: International conference on machine learning. PMLR; 2018. p. 3481–90.

[42] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. Adv Neural Inf Process Syst 2014;27.

[43] Kynkäänniemi T, Karras T, Laine S, Lehtinen J, Aila T. Improved precision and recall metric for assessing generative models. Adv Neural Inf Process Syst 2019;32.

[44] Heusel M, Ramsauer H, Unterthiner T, Nessler B, Hochreiter S. Gans trained by a two time-scale update rule converge to a local nash equilibrium. Adv Neural Inf Process Syst 2017;30.

[45] Gupta D. Image segmentation keras: implementation of segnet, fcn, unet, pspnet and other models in keras. 2023. arXiv preprint arXiv:2307.13215.

[46] Everingham M, Van Gool L, Williams CK, Winn J, Zisserman A. The pascal visual object classes (voc) challenge. Int J Comput Vis 2010;88:303–38.

[47] Dice LR. Measures of the amount of ecologic association between species. Ecology 1945;26(3):297–302.

[48] Dex S. The reliability of recall data: a literature review. Bulletin of Sociological Methodology/Bulletin de Methodologie Sociologique 1995;49(1):58–89.

[49] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434 2015.

[50] Zhu J, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision; 2017. p. 2223–32.

[51] Isola P, Zhu J, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 1125–34.

[52] Zuckerman B, Ron M, Mikl M, Segal E, Ulitsky I. Gene architecture and sequence composition underpin selective dependency of nuclear export of long RNAs on NXF1 and the TREX complex. Mol Cell 2020;79(2):251–67.