
Progressive GAN for High-Resolution Live Cell Fluorescence Microscopy Images

Tianyao Wu, Tianchi Liu
{twu31, kitliu5}@stanford.edu

1 Introduction

The code for this project can be downloaded at

[https://drive.google.com/drive/folders/1MU287Zcy2kuTLxYC4suwx1DP1sdw--S3?
usp=sharing](https://drive.google.com/drive/folders/1MU287Zcy2kuTLxYC4suwx1DP1sdw--S3?usp=sharing)

One of the intriguing application of Generative Adversarial Networks (GANs) is image super-resolution. It could be a useful technique to enhance various type of microscopy images, thereby providing additional scientific value in different research fields. In the field of cell biology, it is very important for us to resolve small cellular structures in order to study a wide variety of biological problems (1). However, the resolution of optical microscopes is often limited by the diffraction limit. Diffraction happens when a beam of light passes through a narrow aperture. When a circular aperture is used, a concentric geometrical pattern is created through combination of constructive and destructive interference (2) (Figure 1). When there are two neighboring patterns generated from two point sources, the diffraction limit sets the minimum distance that these two points can be distinguished (3). There exists other microscopy methods such as electron microscopy that can resolve even the smallest cellular structures, but all of them require fixed cells, severely limiting their use in the studies of live cells. Among the optical techniques, fluorescence microscopy is the single most implemented technique due to its specificity and being noninvasive to cells (4). To study a specific structure in a live cell, fluorophores are bond to this specific structure. In a fluorescence microscopy image, only the emission of these fluorophores is observed. Figure 2 show two examples of fluorescence microscopy images. Some structures have filament shapes while others have distinctive dots. In this study, we focus on the fluorescence images with distinctive dots as they are more widely used for quantitative analysis.

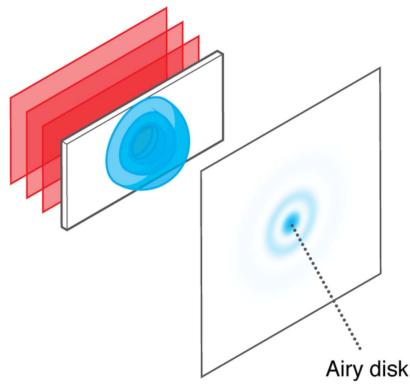


Figure 1: A circular aperture yielding an Airy (concentric geometrical) pattern.(2)

There have been a few studies of GANs on fluorescence images (5) (6) (7). However, it suffers from biodiversity and lack of training data and reference data, same as many other biomedical data.

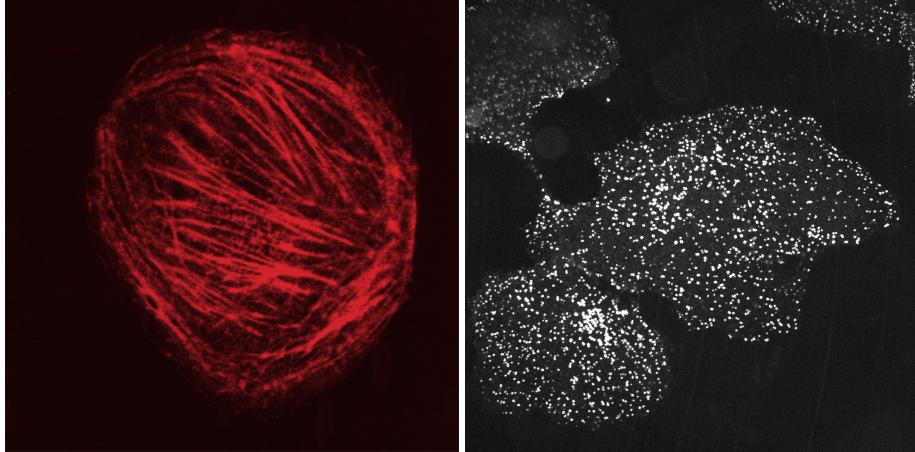


Figure 2: Examples of Fluorescence Microscopy Images. Left: Actin, Right: AP2

The common strategy was to implement example-based framework in which low-resolution images are obtained in a same or very similar condition as the high-resolution images. In the preceding studies of medical images, high-resolution images are downsampled to create low-resolution images, forming pairs for the network training (8). It is a common practice for everyday images such as face recognition data, but it presents a problem for biomedical data where it is not realistic to obtain enough data with diverse conditions. As a result, the heavy reliance on SRGAN-style network (9) and paired content loss make the model very inversatile, often requiring the model to be retrained when the environment or imaging system change slightly. This is expected because GAN is known for its unstable training process and difficulty balancing between generator and discriminator (10). A new technique of growing GAN progressively has been emerged recently to address these issues and performed well on high-dimension images. We aim to explore this progressively growing GAN method and build a versatile generative model for fluorescence microscopy data.

2 Problem Statement

We acquired the fluorescence microscopy image dataset from the imaging systems in Kural lab at the Ohio State University. The imaging system consists of an Eclipse TI-E microscope (Nikon) equipped with a temperature controlled chamber, a CSU-W1 spinning disk confocal unit (Yokogawa Electric Corporation), a $100\times$ objective lens (Nikon CFI Plan-Apochromat Lambda, NA 1.45) and an EMCCD camera (iXon DU897 Ultra, Andor Technology). We acquired 3 types of fluorescence labels, including RedNile Beads, AP2 labelled SUM159 cells, and single molecules. Each movie was taken in a way that two exposure times were used and a low exposure time acquisition is followed by a high exposure time acquisition, and vice versa. Using this method, even though the fluorescence labels in the movie are constantly changing, the adjacent pairs can be considered as matched image pairs. An example of these pairs is shown in Fig.3.

The high exposure and low exposure images are treated as high-resolution and low-resolution images respectively. We aim to build a progressively growing GAN using high exposure images as input. The input has size of 256×256 , the original 1200×1200 size. No labels of fluorescence labels are used. The image has type of 16-bits ranging from integer value 0 to 65535 that differs from the commonly used 8-bits data in other progressively growing GANs. The trained generator is expected to generate fluorescence microscopy images that retain the properties of high-resolution fluorescence microscopy images. The paired low exposure images are used along with high exposure images to build a U-Net for us to have a reference point during evaluation due to lack of similar studies and evaluation methods.

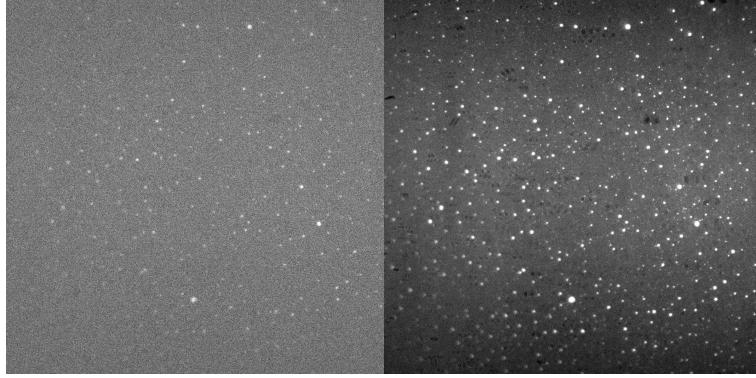


Figure 3: Sample image taken at low exposure (left) and high exposure (right)

3 Approach

3.1 U-Net as baseline

In need of a reference for our final GAN model to compare to, we used U-Net, the most widely used network used in the study of biomedical images. The architecture is shown in the end (Fig.14). Pairs of low exposure (LR) and high exposure (HR) images are cropped from 1200x1200 original size to 256x256 size and are then fed into the model. Two versions of U-Net are used, one same as the original U-Net intended for segmentation tasks that uses binary crossentropy as loss function, the other with mae as loss function and relu in the last activation layer instead of sigmoid.

3.2 Progressively growing GAN

We adopted the concept of progressively growing GAN from *Karras et al* (10) where the training starts from a regular GAN using 4x4 downsampled input. Then, layers are added to generator and discriminator pairs gradually so that large-scale structure of images could be discovered first and increasingly finer details of image distribution could be gradually learned. Furthermore, The added layers will also be smoothly faded with residual-block nature to avoid sudden shock to existing layers with lower resolution, with a coefficient α that ranges from 0 to 1 controlling the weight of newly added higher-resolution layer. This procedure is illustrated by the simplified model architecture in Fig.4 . The detailed architecture is included in the end (Fig.15).

For better stabilization, the Wasserstein GAN with gradient penalty (WGAN-GP) shown in Fig.5 is used as the loss function. In addition, we implemented equalized learning rate as *Karras et al* (10) suggests that $\mathcal{N}(0, 1)$ initialization and explicit pre-layer weight scaling at runtime ensures stable learning speed for different weights . We did not implement pixel normalization as it is not applicable to fluorescence microscopy images. The downscaling and upscaling are important factors used in this network and more details about them will be discussed in the next section.

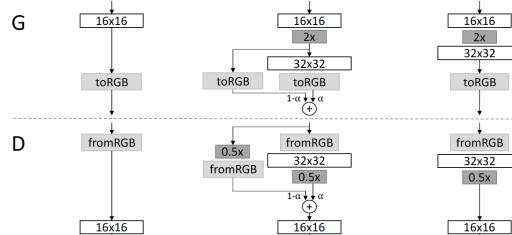


Figure 4: Progressive GAN Architecture

$$L = \underbrace{\mathbb{E}_{\hat{x} \sim \mathbb{P}_g} [D(\hat{x})] - \mathbb{E}_{x \sim \mathbb{P}_r} [D(x)]}_{\text{Original critic loss}} + \lambda \underbrace{\mathbb{E}_{\hat{x} \sim \mathbb{P}_{\hat{\phi}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2]}_{\text{Our gradient penalty}}.$$

Figure 5: WGAN-GP loss

3.3 Another look at fluorescence microscopy images

Our first attempt on the progressively growing GAN, only visualization modification is applied to the proposed method by *Karras et al.* It did not manage to achieve anything close to a fluorescence microscopy image (Fig.6). After the initial failure, we took another look at our fluorescence microscopy data. It is clear that the fluorescence microscopy dataset, and in a broader view all biomedical data, differ greatly from everyday images that existing deep learning approaches cannot be easily re-applied to these imaging tasks.

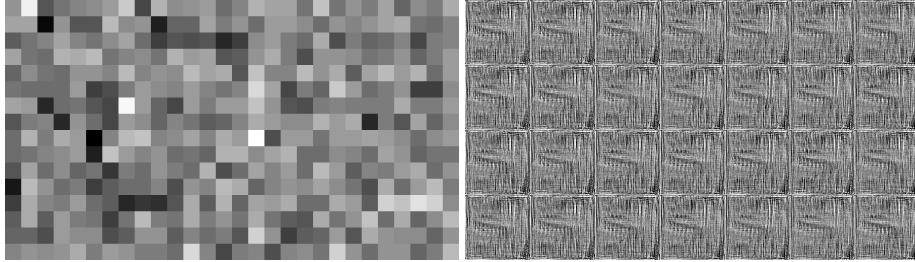


Figure 6: 4x7 grid of predicted sample images. Left: early stages of training with size 4x4; Right: 2 days of training with size 128x128

To investigate the special property of fluorescence microscopy images, we first took a look at the distribution of intensities (Fig.7). While the range of intensities go up to 65535 and there exists pixels that reach close to this maximum value, they account for less than 1 percent of all the signals. Among most analyses on fluorescence microscopy images, the absolute intensity value holds limited information and little importance. On the low end of intensities, the lowest signals have such low intensity that sometimes they have only 2-3 intensity unit away from background. This makes the distribution of background and signals overlapping and makes separation between background and signal non-trivial. We decided to adopt a similar method used by *Wu et al.* (7) to preprocess our data. First, since the distribution of intensities are close to triangular distribution, we applied a triangular thresholding(11) is applied to randomly sampled data to determine a single threshold value as breakpoint between background and signals. Then, the mean of background and 99 percentile of signals are treated as new minimum and maximum respectively and the entire training dataset is clipped to this new minimum and maximum value. This method improves the dynamic drastically while retaining important information of a fluorescence microscopy image.

We then investigate the validity of applying upsampling and downsampling methods that are the key components in progressively growing GAN. For a common facial image, we can easily identify its perceptual identity even it is downsampled to size of 8x8. However, for a fluorescence microscopy image, the details on the signal are critical for identification and further analysis. This phenomenon is universal in many other biomedical images that small structures are on the top priority of diagnostic information of an image (12). It has become clear that the original progressively growing GAN is not suited for this type of image data as most of burden are placed on the last few growing layers. To find a way to gradually growing the diagnostic information of a fluorescence microscopy image, we turned our eye on the frequency space. In frequency space, details of small structures are presented in high frequency region while the broad structure is presented in low frequency region. Structured illumination microscopy is a novel super-resolution microscopy technique that combines multiple images in different phases and orientations in frequency space to surpass the diffraction limit of fluorescence microscopy (13). We drew inspiration from this technique and modified the upscaling and downscaling methods. A fast Fourier transform is applied upon entering upscaling and downscaling layer. The real components then go through either tiling or meanpooling in the downscaling or upscaling layer respectively. Afterwards, an inverse Fourier transform is applied and the real components are treated as output. To demonstrate this idea, we tested this downscaling

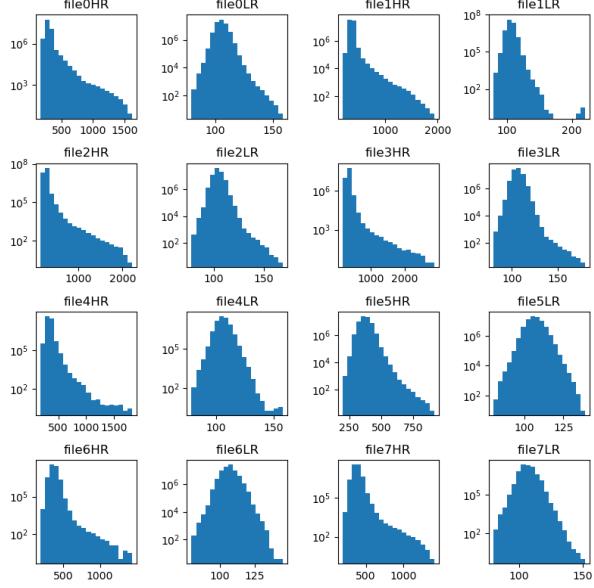


Figure 7: Histogram of intensities in sample fluorescence microscopy images. Each individual figure represents a separate movie session.

method on fluorescence microscopy images (Figure 8). While not perfect, the downsampled images showed significantly more high frequency details compared to the original downscaling method (Fig. 6 left image).

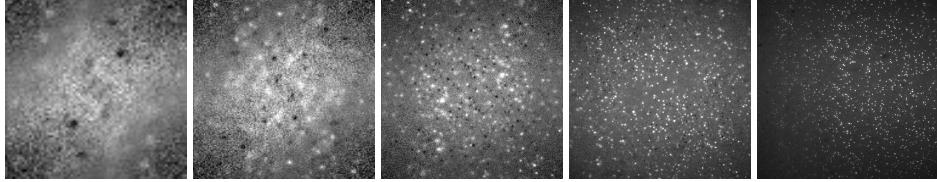


Figure 8: Fluorescence image after meanpooling in frequency space. The rightmost image is the original image. Downscaling $\times 2$ is applied on frequency space step by step towards the leftmost image.

4 Results

4.1 U-Net

8500 256x256 16-bit live cell fluorescence image pairs (with low exposure time and high exposure time) are used as training data and 1500 image pairs of same type are used as test data. The special preprocessing procedure is covered in earlier section and is applied to the training data. Overall, U-Net provides reasonable high resolution images with good accuracy of fluorescence label identification. Examples of predicted results using two different loss functions (binary crossentropy or mae) are shown in Fig.9 and Fig.10. On surface, the trained results resemble high resolution images quite well. However, it lacks the shape of structures and precision of intensities once we enhance the brightness and contrast to see signals with lower intensities (Fig.11). The radius of each dots are universally bigger in U-Net predicted samples, especially the samples with binary crossentropy as loss. This is expected because U-Net was introduced as a semantic segmentation tool. While it has been proved to be able to perform image super-resolution tasks on biomedical data, it lacks the sharpness we need in order to identify the small structure. All of the predicted labels in U-Net trained results are close to a 2D Gaussian distribution, making it impossible to do further analysis on studies that rely on small

structure changes. Overall, the U-Net results provided us a solid reference point for us to evaluate progressively growing GAN results.

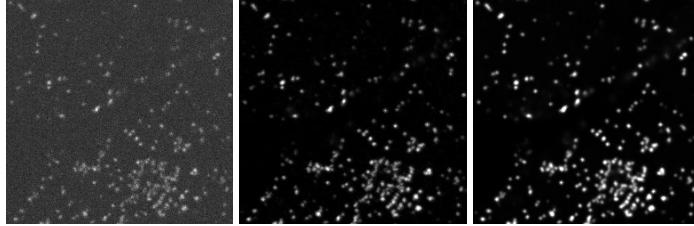


Figure 9: U-Net training using binary crossentropy loss. From left to right: LR, HR, predicted

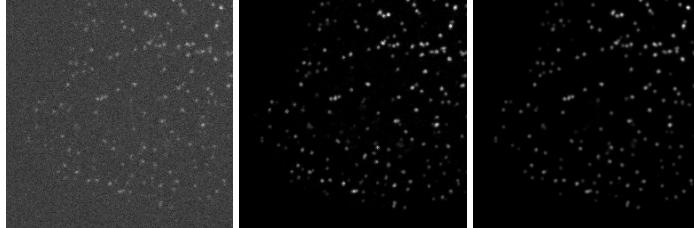


Figure 10: U-Net training using mae loss. From left to right: LR, HR, predicted

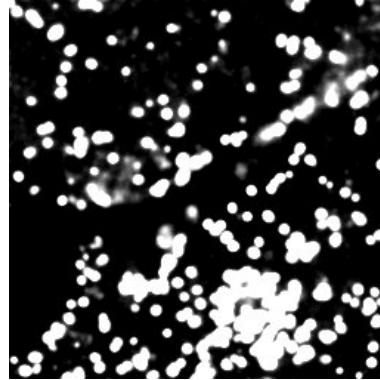


Figure 11: High contrast predicted image from U-Net training using binary crossentropy loss.

4.2 Progressively growing GAN

The same 8500 256x256 16-bit live cell fluorescence images are used for Progressively growing GAN training. However, only the high exposure time images are used. Due to computational power and time constraints, a total of 6000k images are processed through the model during training, half of the epochs used by *Karras et al.* in their original progressively growing GAN paper (10). The method of upscaling and downscaling in frequency space is performed in every upscaling and downscaling incidents throughout the training. Randomly generated samples from trained generator are shown in Fig.12.

4.3 Analysis

It is tricky to evaluate the quality of generated live-cell microscopy images even with U-Net results as our baseline. We first looked through zoomed in results in high contrast mode and compared them with the U-Net predicted results (Fig.13). It is clear that only the progressive GAN generated results manage to generate labels with non-Gaussian shape. While some of them look distorted compared to the truth samples, we see it as a success towards generating fluorescence labels with

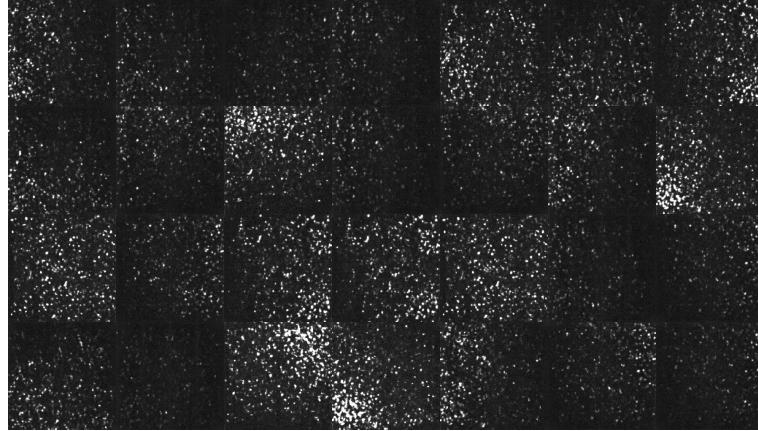


Figure 12: Randomly generated samples from trained generator. 4x7 total samples with each having size of 256x256.

actual shapes rather than 2D Gaussian approximation. Next, we analyzed the intensity distribution of 28 randomly selected results (Table 1). Progressively growing GAN outperforms two U-Net results in every categories except the mean background value, which is much less important than signal values. Overall, progressively growing GAN showed improvement over U-Net in both small structure shape and intensity precision. We planned to quantitatively analyze the improvement on the shapes of small structures, but there is no easy way to classify different fluorescence labels with small structure difference other than hand labeling enough structures to build a classifier, so only qualitative assessment on small structures is performed for now.

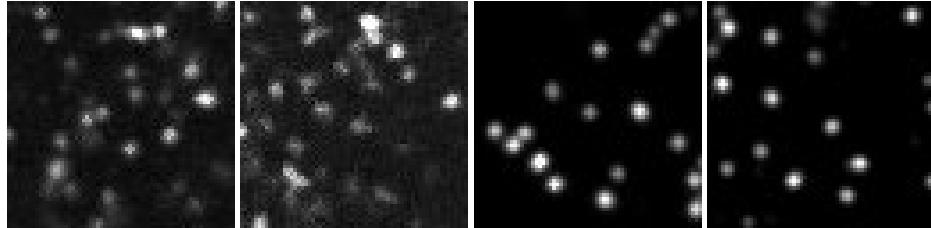


Figure 13: Zoomed in high contrast 64x64 sample results. From left to right: Truth; GAN generated result; U-net predicted result with binary crossentropy loss; U-net predicted result with mae loss.

	Triangular threshold	Mean Background	99 percentile Signal	90 percentile Signal	50 percentile Signal
Truth	151	119	424	257	180
progressive GAN	147	116	376	263	173
U-Net (Binary Crossentropy)	144	125	690	459	211
U-Net (mae)	128	119	414	279	161

Table 1: Intensity statistics of 28 randomly selected high exposure images (as truth), generated progressive GAN images, and U-net predicted images. Triangular thresholds are calculated separately for each image and the mean values are shown above.

5 Conclusion

In this work, we used GAN to perform super-resolution on fluorescence microscopy images. An U-Net baseline model is first established due to lack of preceding studies in this direction. It is followed by implementation of progressive GAN with modifications outlined in section 3.2 and 3.3. The unsatisfactory results lead to more careful investigation on properties of fluorescence microscopy images, especially on their pixel intensity distributions and the validity of applying upsampling and downsampling on this type of microscopy images.

The core novelty of this work lies in the tailored processing of fluorescence microscopy images for progressive growth of GAN layers to work. Due to the discrepancies between high and low frequency components and the importance of high frequency components, scaling in frequency space makes it possible to run progressively growing GAN on fluorescence microscopy images. Additional preprocessing such as triangular thresholding also proved to be crucial in improving the dynamic range of the input. While we still lack quantitative analysis on the small structures of generated result, this project is a progressive step towards generating high-resolution fluorescence microscopy images for studies of small cellular structures.

References

- [1] R. Heintzmann and T. Huser, “Super-resolution structured illumination microscopy,” *Chemical Reviews*, vol. 117, pp. 13890–13908, 12 2017.
- [2] J. Vangindertael, R. Camacho, W. Sempels, H. Mizuno, P. Dedecker, and K. P. F. Janssen, “An introduction to optical super-resolution microscopy for the adventurous biologist,” *Methods and Applications in Fluorescence*, vol. 6, no. 2, p. 022003, 2018.
- [3] S. G. Lipson, H. Lipson, and D. S. Tannhauser, *Optical Physics*. Aug. 1995.
- [4] B. Huang, M. Bates, and X. Zhuang, “Super-resolution fluorescence microscopy,” *Annual Review of Biochemistry*, vol. 78, no. 1, pp. 993–1016, 2009. PMID: 19489737.
- [5] H. Zhang, C. Fang, X. Xie, Y. Yang, W. Mei, D. Jin, and P. Fei, “High-throughput, high-resolution deep learning microscopy based on registration-free generative adversarial network,” *Biomedical optics express*, vol. 10, pp. 1044–1063, Feb 2019. 30891329[pmid].
- [6] H. Wang, Y. Rivenson, Y. Jin, Z. Wei, R. Gao, H. Günaydin, L. A. Bentolila, C. Kural, and A. Ozcan, “Deep learning enables cross-modality super-resolution in fluorescence microscopy,” *Nature Methods*, vol. 16, no. 1, pp. 103–110, 2019.
- [7] Y. Wu, Y. Rivenson, H. Wang, Y. Luo, E. Ben-David, L. A. Bentolila, C. Pritz, and A. Ozcan, “Three-dimensional virtual refocusing of fluorescence microscopy images using deep learning,” *Nature Methods*, vol. 16, no. 12, pp. 1323–1331, 2019.
- [8] D. Mahapatra, B. Bozorgtabar, and R. Garnavi, “Image super-resolution using progressive generative adversarial networks for medical image analysis,” *Computerized Medical Imaging and Graphics*, vol. 71, pp. 30–39, 2019.
- [9] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. P. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, “Photo-realistic single image super-resolution using a generative adversarial network,” *CoRR*, vol. abs/1609.04802, 2016.
- [10] T. Karras, T. Aila, S. Laine, and J. Lehtinen, “Progressive Growing of GANs for Improved Quality, Stability, and Variation,” *arXiv e-prints*, p. arXiv:1710.10196, Oct 2017.
- [11] G. W. Zack, W. E. Rogers, and S. A. Latt, “Automatic measurement of sister chromatid exchange frequency.,” *Journal of Histochemistry & Cytochemistry*, vol. 25, no. 7, pp. 741–753, 1977. PMID: 70454.
- [12] K. Armanious, C. Jiang, M. Fischer, T. Küstner, T. Hepp, K. Nikolaou, S. Gatidis, and B. Yang, “Medgan: Medical image translation using gans,” *Computerized Medical Imaging and Graphics*, vol. 79, p. 101684, 2020.
- [13] M. G. L. Gustafsson, “Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy,” *Journal of Microscopy*, vol. 198, no. 2, pp. 82–87.

Appendix - Additional Figures

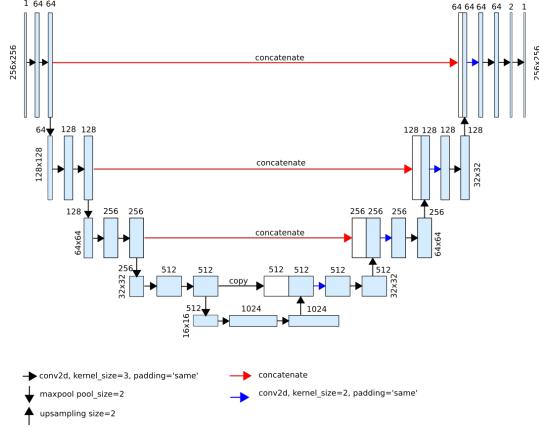


Figure 14: U-Net Architecture

D	Params	OutputShape	WeightShape	G	Params	OutputShape	WeightShape
---	---	---	---	---	---	---	---
images_in	-	(?, 1, 256, 256)	-	latents_in	-	(?, 512)	-
lod	-	()	-	labels_in	-	(?, 0)	-
FromRGB_lod0	128	(?, 64, 256, 256)	(1, 1, 1, 64)	lod	-	()	-
256x256/Conv0	36928	(?, 64, 256, 256)	(3, 3, 64, 64)	4x4/PixelNorm	-	(?, 512)	-
256x256/Conv1_down	73856	(?, 128, 128, 128)	(3, 3, 64, 128)	4x4/Dense	4194816	(?, 512, 4, 4)	(512, 8192)
Downscale2D	-	(?, 1, 128, 128)	-	4x4/Conv	2359808	(?, 512, 4, 4)	(3, 3, 512, 512)
FromRGB_lod1	256	(?, 128, 128, 128)	(1, 1, 1, 128)	ToRGB_lod6	513	(?, 1, 4, 4)	(1, 1, 512, 1)
Grow_lod0	-	(?, 128, 128, 128)	-	8x8/Conv8_up	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)
128x128/Conv0	147584	(?, 128, 128, 128)	(3, 3, 128, 128)	8x8/Conv1	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)
128x128/Conv1_down	295168	(?, 256, 64, 64)	(3, 3, 128, 256)	ToRGB_lod5	513	(?, 1, 8, 8)	(1, 1, 512, 1)
Downscale2D_1	-	(?, 1, 64, 64)	-	Upscale2D	-	(?, 1, 8, 8)	-
FromRGB_lod2	512	(?, 256, 64, 64)	(1, 1, 1, 256)	Grow_lod5	-	(?, 1, 8, 8)	-
Grow_lod1	-	(?, 256, 64, 64)	-	16x16/Conv8_up	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)
64x64/Conv0	590080	(?, 256, 64, 64)	(3, 3, 256, 256)	16x16/Conv1	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)
64x64/Conv1_down	1180160	(?, 512, 32, 32)	(3, 3, 256, 512)	ToRGB_lod4	513	(?, 1, 16, 16)	(1, 1, 512, 1)
Downscale2D_2	-	(?, 1, 32, 32)	-	Upscale2D_1	-	(?, 1, 16, 16)	-
FromRGB_lod3	1024	(?, 512, 32, 32)	(1, 1, 1, 512)	Grow_lod4	-	(?, 1, 16, 16)	-
Grow_lod2	-	(?, 512, 32, 32)	-	32x32/Conv8_up	2359808	(?, 512, 32, 32)	(3, 3, 512, 512)
32x32/Conv0	2359808	(?, 512, 32, 32)	(3, 3, 512, 512)	32x32/Conv1	2359808	(?, 512, 32, 32)	(3, 3, 512, 512)
32x32/Conv1_down	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)	ToRGB_lod3	513	(?, 1, 32, 32)	(1, 1, 512, 1)
Downscale2D_3	-	(?, 1, 16, 16)	-	Upscale2D_2	-	(?, 1, 32, 32)	-
FromRGB_lod4	1024	(?, 512, 16, 16)	(1, 1, 1, 512)	Grow_lod3	-	(?, 1, 32, 32)	-
Grow_lod3	-	(?, 512, 16, 16)	-	64x64/Conv8_up	1179904	(?, 256, 64, 64)	(3, 3, 256, 512)
16x16/Conv0	2359808	(?, 512, 16, 16)	(3, 3, 512, 512)	64x64/Conv1	590080	(?, 256, 64, 64)	(3, 3, 256, 256)
16x16/Conv1_down	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)	ToRGB_lod2	257	(?, 1, 64, 64)	(1, 1, 256, 1)
Downscale2D_4	-	(?, 1, 8, 8)	-	Upscale2D_3	-	(?, 1, 64, 64)	-
FromRGB_lod5	1024	(?, 512, 8, 8)	(1, 1, 1, 512)	Grow_lod2	-	(?, 1, 64, 64)	-
Grow_lod4	-	(?, 512, 8, 8)	-	128x128/Conv8_up	295040	(?, 128, 128, 128)	(3, 3, 128, 256)
8x8/Conv0	2359808	(?, 512, 8, 8)	(3, 3, 512, 512)	128x128/Conv1	147584	(?, 128, 128, 128)	(3, 3, 128, 128)
8x8/Conv1_down	2359808	(?, 512, 4, 4)	(3, 3, 512, 512)	ToRGB_lod1	129	(?, 1, 128, 128)	(1, 1, 128, 1)
Downscale2D_5	-	(?, 1, 4, 4)	-	Upscale2D_4	-	(?, 1, 128, 128)	-
FromRGB_lod6	1024	(?, 512, 4, 4)	(1, 1, 1, 512)	Grow_lod1	-	(?, 1, 128, 128)	-
Grow_lod5	-	(?, 512, 4, 4)	-	256x256/Conv0_up	73792	(?, 64, 256, 256)	(3, 3, 64, 128)
4x4/MinibatchStddev	-	(?, 1, 4, 4)	-	256x256/Conv1	36928	(?, 64, 256, 256)	(3, 3, 64, 64)
4x4/Conv	2364416	(?, 512, 4, 4)	(3, 3, 513, 512)	ToRGB_lod0	65	(?, 1, 256, 256)	(1, 1, 64, 1)
4x4/Dense0	4194816	(?, 512)	(8192, 512)	Upscale2D_5	-	(?, 1, 256, 256)	-
4x4/Dense1	513	(?, 1)	(512, 1)	Grow_lod0	-	(?, 1, 256, 256)	-
scores_out	-	(?, 1)	-	images_out	-	(?, 1, 256, 256)	-
labels_out	-	(?, 0)	-	---	---	---	---
---	---	---	---	Total	23039303	---	---
Total	23047361						

Figure 15: Fully Grown Discriminator and Generator of Progressive GAN