Video-Rate Acquisition Fluorescence Microscopy via Generative Adversarial Networks

Tahir Bachar Issa, Claudio Vinegoni, Andrew Shaw,
Paolo Fumene Feruglio, Ralph Weissleder and David Uminsky

Video-rate acquisition fluorescence microscopy via generative adversarial networks

1ˢᵗ T. B. Issa
Data Institute
University of San Francisco
San Francisco, CA, USA
tissa2@usfca.edu

1ˢᵗ C. Vinegoni
Center for Systems Biology
Harvard University
Boston, MA, USA
cvinegoni@mgh.harvard.edu

3ʳᵈ A. Shaw
Data Institute
University of San Francisco
San Francisco, CA, USA
ashaw3@usfca.edu

4ᵗʰ P. F. Feruglio
Center for Systems Biology
Harvard University
Boston, MA, USA
paolo.fumeneferuglio@univr.it

5ᵗʰ R. Weissleder
Department of Systems Biology
Harvard University
Boston, MA, USA
rweissleder@mgh.harvard.edu

6ᵗʰ D. Uminsky
Data Institute
University of San Francisco
San Francisco, CA, USA
duminsky@usfca.edu

Abstract—Laser scanning microscopy is a powerful imaging modality ideal for monitoring spatial and temporal dynamics in both *in vitro* and *in vivo* models. To accurately resolve dynamic changes, particular to the neuroimaging field, fast acquisition rates are in great need. Unfortunately, the video-rate acquisition required to capture these changes comes with a trade-off between resolution, high spatial distortion, and low signal-to-noise ratio due to the electronics and Poisson noise. By combining microscopy fast acquisition methods with a Generative Adversarial Network (GAN), we show here, for the first time, that video-rate image acquisition, up to 20x the speed of equivalent standard high resolution acquisition, can be obtained without significant reduction in image quality. Specifically, we present a GAN based training approach that is able to simultaneously 1) super-resolve, 2) denoise and 3) correct distortion on fast scanning acquisition microscopy images. In addition, we show that our method generalizes on unseen data, requires minimal ground truth images for training and can easily be fine-tuned on different biological samples.

Index Terms—Video-rate microscopy, GAN, Deep learning, Computer vision, Minimal data

I. Introduction

Optical imaging modalities are ideal for *in vitro* and *in vivo* single-cell phenotypic screening [1]. In particular, confocal and multiphoton laser scanning fluorescence microscopy have become fundamental imaging tools, offering high spatial and temporal resolution [2]–[4], extended tissue penetration depth [5]. When used in combination with fluorescent proteins, these methods offer the capability to increase our knowledge and understanding of both cellular and molecular phenomena [6]–[9].

Both these techniques are typically based on single point axial scanning and allow for optical sectioning and 3D imaging [3]. While many scanning mechanisms are available such as acousto-optical deflectors, Nipkow discs, or digital mirror devices, galvanometric scanning is by far the most straightforward and popular scheme. In a standard laser scanning mode acquisition, the position of the raster scan is controlled by two galvanometric-based scanning mirrors oriented at 90 degrees with each other [10]–[12]. One of the two mirrors operates in raster mode in a unidirectional way controlling the horizontal scanning, while the other controls the vertical deflection [13]. The laser beam is then focused to a diffraction-limited point and moved over time in space (Fig. 1a-c) where it describes a flat-plane shaped support covering the field of view to be imaged [14]. Images are finally reconstructed from the illuminated sample on a pixel-by-pixel basis by collecting fluorescence-emitted photons created along the raster scan path and by converting them to electric signals by means of a photomultiplier tube (PMT). In this way virtual optical sectioning through the sample [3], [15] can be performed and three-dimensional reconstructions can be obtained.

While galvanometric scanning provides accurate high axial resolution and increased SNR due to the reduced background scattering signal, it is also very slow with strong constraints on the scanning speed. Image acquisitions at these frame rates (0.1-0.5 frames per second) are affected by strong motion artifacts due to physiological motion [16]–[19] and can not help in resolving fast dynamic biological changes.

A. Improving acquisition rates

Acquisition rates can be slightly improved by increasing the unidirectional speed of the galvanometer and reducing the number of acquired pixels. This problem has been previously approached and the loss of image resolution mitigated using super resolution neural network approaches [20]. In addition, it is important to bear in mind that increasing the acquisition speed always translates in a reduced amount of light collected by the detector, which results in very noisy images. A trade-off is therefore always present between the necessity to work at high speed acquisition and the inevitable reduced amount of light which accompanies this modality. In microscopy, this is known as the "eternal triangle of compromise" between speed, sensitivity, and resolution.
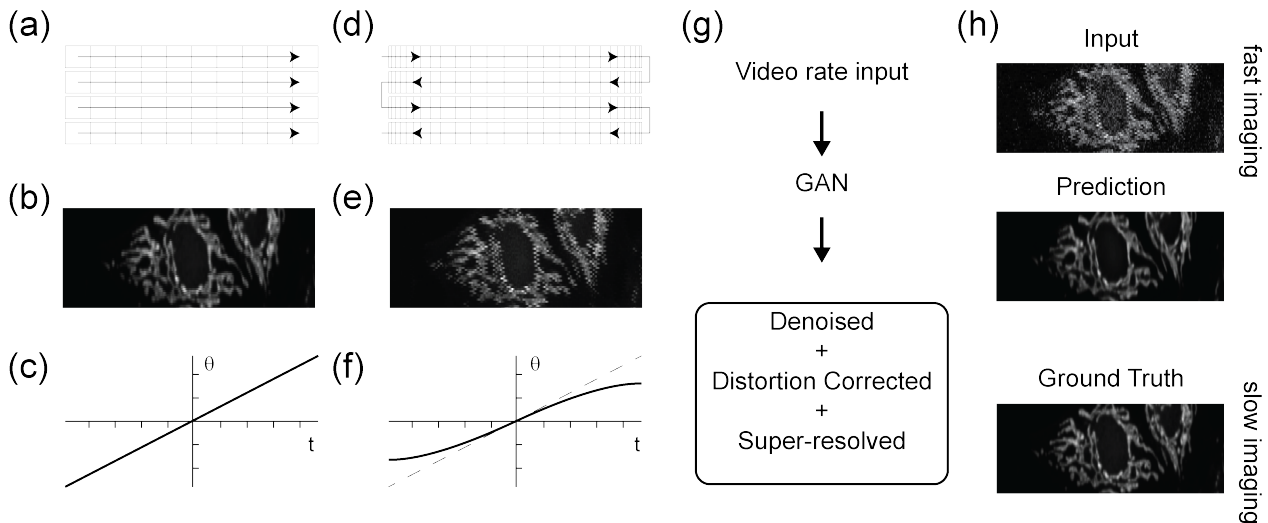
Fig. 1. Acquisition imaging scheme and processing pipeline. (a) Unidirectional scanning microscopy raster scan path. (b) Resulting undistorted image obtained on a pixel-by-pixel basis. (c) Position of the scanning mirror as a function of time, during line acquisition. (d) Bidirectional video-rate scanning microscopy raster scan path. (e) Resulting distorted image. (f) Distortion along the edges of the images are caused by the nonlinearity of the scanning galvo. (g) Proposed processing pipeline. (h) Starting from a video-rate image (input), image distortions are corrected, the noise mitigated and a new resulting upscaled image (prediction) is obtained. The obtained predictions are then compared to the ground truth HR images.

B. Video-rate acquisition

A possible way to increase acquisition rates further consists in using a bidirectional scanning configurations [21], [22]. Instead of scanning the laser beam exclusively in the same direction, the beam is moved back and forth multiple times. Consequently, within an image, odd lines are scanned from left-to-right while even lines are imaged from right-to-left, effectively doubling the scan rate of the acquisition.

These faster acquisition rates in the video-rate range can be reached when operating in resonant scanning mode [21], [23]–[26] but at the price of significant image distortion, on top of the increased accompanying noise. In this configuration, a resonant scanning mirror is used in place of the slow servo-controlled one for the horizontal linear scanning, leading to imaging frame rates on the order of 30 fps at several fields of view. Unfortunately, one inherent disadvantage of bidirectional imaging is that it introduces spatial distortions in the acquired images due to the fact that the mirror oscillates back-and-forth with a speed that is a sinusoidal function of time (Fig. 1d-f).

Despite these challenges, this imaging modality is extremely valuable in a number of settings. For example, at this increased frame rate, one can effectively study blood cell trafficking in capillary networks [19], [27], [28], record *in vivo* neuronal activity [29]–[33], obtain acquisitions of multigigabyte-sized whole slides accelerating histopathological data acquisition, and also facilitate whole organ mapping at the microscopic level [34], [35]. Thus, it is critical to address the challenges presented by video-rate acquisition modalities as we do so here with a deep learning approach, in order to improve overall image quality and fidelity.

C. A deep learning framework for fast acquisition.

The problem we intend to address in this work is how to better reconstruct image data obtained in fast (video-rate) acquisition modalities and mitigate the accompanying noise and distortion effects intrinsically present (Fig. 1g-h).

Deep learning super-resolution methods have shown significant promise in recovering high resolution images from noisy low resolution images acquired at faster scanning rates. These deep learning methods often achieve state-of-the art results on super-resolution tasks [36] and range from early Convolutional Neural Networks [37], [38] to the recent promising approaches using Generative Adversarial Networks (GANs) [39], [40]. GANs have demonstrated to be capable of enhancing visual details with remarkable resolution. For a survey on deep learning methods in computer vision, we refer the reader to [41]. Recent works in deep learning microscopy range from bright-field microscopy [42], [43] to fluorescence microscopy [43], [44]. Manifold et al. [45] have demonstrated that deep learning is effective at super-resolving and denoising microscopy images. Wang et al. [46] have showed that deep learning enables super-resolution in microscopy, while Ouyang et al. [43] have demonstrated that deep learning can accelerate super-resolution localization microscopy. However, these approaches typically focus on images acquired *without distortion effects*, which are typically present in images collected in galvo-based bidirectional video-rate acquisition mode as described in Section I-B.

Here, we specifically introduce a framework to acquire microscopy images at video-rate and examine the efficacy of deep learning Generative Adversarial Networks (GANs) on mitigating the trade-offs introduced through video-rate acquisition (Fig. 1g-h) . Our main contributions are the following:

- We train a GAN model that is successfully able to simultaneously 1) super-resolve, 2) denoise and 3) correct distortions on fast scanning acquisition microscopy images.
- We demonstrate that very few high resolution images ("ground truth" HR images) are required to successfully train an accurate deep learning super resolution model.
- We show that our GAN approach generalizes well on unseen, unrelated microscopy data with minimal fine tune training, a crucial feature for a real-life practical implementation on generic biological specimens.

II. DATA COLLECTION

A. Distortion, noise and fast acquisition

As noted in Section I, acquiring images at video-rate scanning speed introduces two main effects: distortion and noise. The distortion effects introduced in video-rate acquisition are due to the motion of one of the two oscillating mirrors in bidirectional imaging. The mirror gradually accelerates and decelerates as it moves through its cycles (Fig. 1f), slowing at the edges where it changes in direction, and reaching a maximal speed in the middle of the scanned field [48]. Such a pattern leads to multiple implications. Pixels at different locations map different specimen areas (bigger at the center, smaller at the edges) causing a space-dependent blur and giving rise to an image which appears visually stretched toward the edges [13]. In addition, image rows display a horizontal shift between adjacent lines as microscopes are not designed to accurately compensate for this uneven motion (Fig. 1e).

Apart from the geometrical distortions, the fluorescence signal acquisition process is inevitably affected by several sources of noise ultimately affecting the system performances and the images signal-to-noise ratio (SNR). In standard imaging condition, acquisitions can be optimized by choosing appropriate setup parameters, such as increasing the integration time or averaging multiple times, leading to high image quality collection which is crucial to allow quantitative information extraction and analysis of complex dynamic cellular events [14]. Unfortunately, high acquisition speeds inherently reduce the amount of light collected by the PMT, which results in very noisy images. A trade-off is therefore always present between the necessity to work at high speed acquisition and the inevitable reduced amount of light which accompanies this modality. Some noise contributions are due to the imaged sample with the presence of scattering and/or autofluorescence signal components, while others can be traced to the system optical components (e.g. stray light detection, filter bleed-through, etc). But above all, the more relevant sources of noise components are shot noise and dark noise. Shot noise, also known as photon- or Poisson noise, is independent from the detecting system and represents the fundamental limit on noise performances. This heterostocastic noise component is signal dependent and inherent to the statistical uncertainty in the arrival of the fluorescence photons at the sensing detector [14], [49]. It is particularly cumbersome at low photon

fluxes in cases such as bidirectional or resonant scanning. Dark noise is another source of noise which is not related to the photon flux incident on the detector. This noise is dependent on the detector's temperature and the exposure time and tends to become predominant at very low photon fluxes [50]. As a result, fast acquisition microscopy images generally suffer from 1) severe distortion artifacts which can not be straightforwardly compensated, 2) they present low SNR due to the Poisson noise and the reduced photon numbers acquired during each voxel effectively reducing the detector's bit depth, and 3) frequently present a reduced number of sampling pixels.

All these factors are very difficult to correct separately, and such an approach generally leads to a sub-optimal solution. For instance, it is possible to correct for the misalignment and the blur, by applying a complex non-uniform registration. However, further processing with a denoise algorithm should be designed to work on a transformed version of the noise, which is no longer well approximated by a Poisson-Gaussian random variable. In addition, deriving for such a noise a variance stabilization transformation along with its unbiased inversion is also an intricate task. Alternatively, also swapping the registration and denoise steps is unsuitable. In fact, due to the un-matching of the image lines, the state of the art denoising algorithms based on patches-matching, e.g. BM3D [51], [52], would lose efficacy in finding similarities among and within image patches. This leads to a less sparse representation of patches, resulting in a loss of performance. All these considerations indicate how it is better to avoid the use of a cascade of process driven models in favor of an all-in-one data-driven approach based on deep learning as illustrated in the introduction.

TABLE I
COMPARISON OF IMAGE ACQUISITION

Acquisition Type	Frame Rate	Scanning Mode	Pixel Integration Time (μ sec)	Image Size (pixels)	Pixel Size (μ m)
HR (averaged)	0.08	Unidirectional	10	512x512	0.2
(LR + Noise + Distortion)	20	Bidirectional	0.5	256x256	0.4

B. Our Data

To demonstrate the capability of our image processing approach we used two different samples showcasing biological features at different scales: a microscope slide (FluoCells, Invitrogen) containing BPAE cells with mitochondria stained with MitoTracker Red CMXRos (sample A), and a separate H&E stained histological section of the small and large mouse intestine (sample B). We chose to image these two samples because one (sample A) consists of 1) very small features extending few pixels in size within the acquired images and 2) present high image contrast, while the other (sample B) tends to showcase distinct features over large areas of distributed signal. High resolution HR representative images
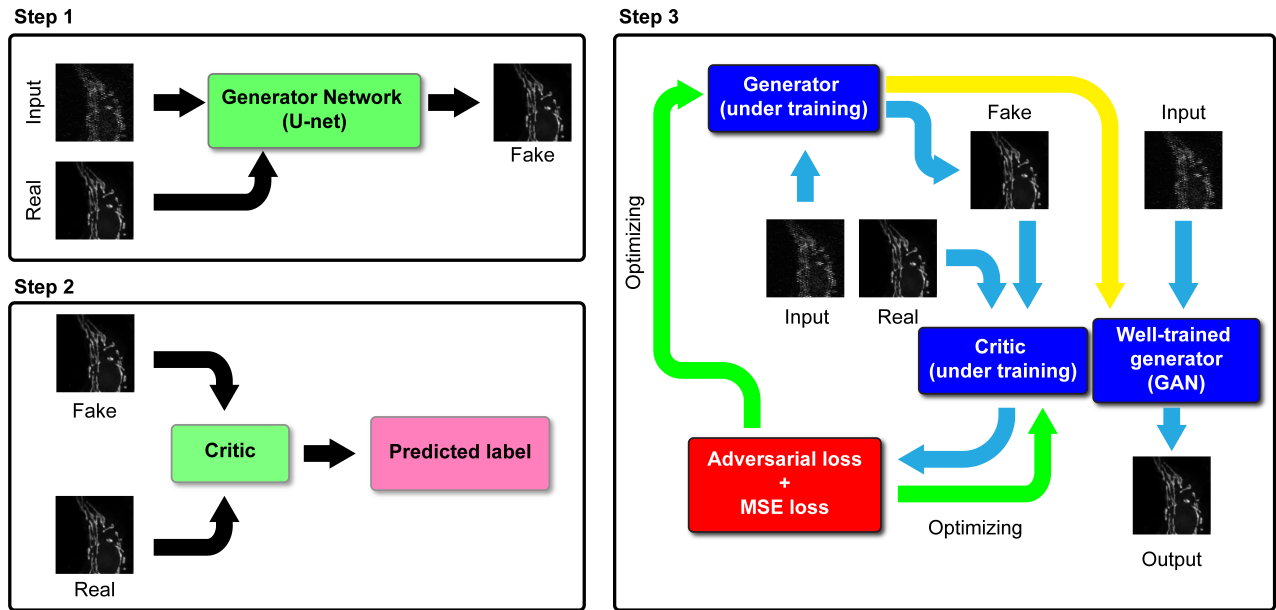
Fig. 2. Scheme of principle for the training of our GAN. The training is divided into three distinct steps. In step 1, the U-net generator is trained. In step 2, the critic is trained to distinguish between real high resolution images from the fake images generated by the trained generator. In step 3, the GAN is trained by simultaneously training the generator and the critic.

of the two samples are given in 3 together with an histogram representation of their fluorescence signal distribution within the dynamic range of the detector.
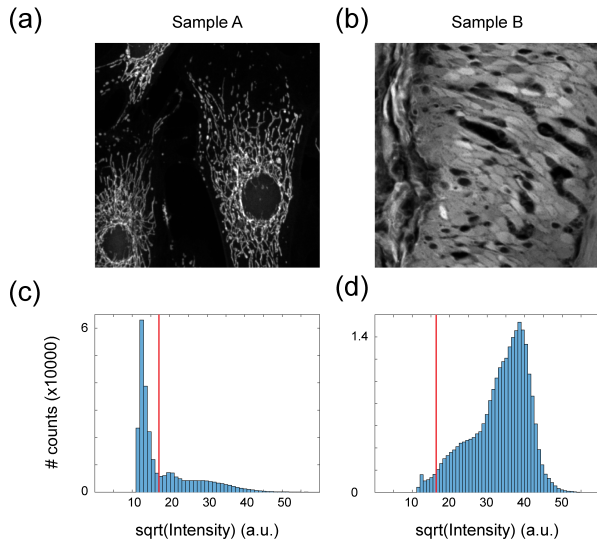

Fig. 3. Representative images of sample A (a) and sample B (b). (c,d) Pixel-by-pixel fluorescence intensity histograms for sample A (c) and B (d) respectively. The red bar indicates the image background level.

All images were acquired in confocal mode using a customized Olympus FV1000 system based on a BX61-WI confocal microscope (Olympus America). For data collection, we used a XLUMPLFLN 20x water immersion (NA 1.0, Olympus America, 2mm WD) imaging objectives. MitoTracker has

a maximum excitation at 554nm and a 576 nm maximum emission and was excited using a 559nm diode laser. Fluorescence was collected using an appropriate combination of beam splitters (SDM560) and emission filters (BA575–620).

High resolution, ground truth images (HR, Table I) were acquired at a planar resolution of 0.2 microns/pixel (512x512, zoom factor x6, 106x106 microns field of view). To decrease the overall noise within the image, the integration time was kept high (10 microsecond/pixel), the PMT voltage was kept low and images were averaged 4 times.

To maximize the frame-rate acquisition speed, low resolution (LR) images acquired at fast speed in bidirectional mode in video-rate mode ((LR+Noise+Distortion)=VideoRate, Table I) were collected over the same field of view as for the ground truth images HR but with 1) a reduced number of pixels (256x256) corresponding to a planar resolution of 0.4 microns/pixel, 2) a lower integration time (0.5 microseconds/pixel) and 3) higher PMT voltage to increase the signal. Series of 100 VideoRate images were acquired for each ground truth image HR over the same field of view, in order to increase the number of noisy images for the training.

III. METHOD AND RESULTS

A. Neural Network Algorithm

Our neural network algorithm for training is a fast.ai version of the generative adversarial network [53]. GANs were introduced by Goodfellow et al. in [39] and have been used to achieve state-of-the-art performances on various super-resolution tasks [43], [44]. GANs simultaneously train two competing neural network models: a generator and a critic, in order to generate new, synthetic instances of data that

pass for real data. The generator will try to capture the data distribution by making new images similar to the images in the data set, while the critic will try to distinguish real images from the fake ones the generator produces. Our generator is a ResNet34-based U-net [54] in the form of encoder-decoder with skip connections. The encoder downsamples an input image and the decoder upsamples the image back to its original size. The U-net generator uses a ResNet34 pre-trained on ImageNet as a backbone for the encoder. Furthermore, we use MSE loss, progressive resizing, and cyclic learning rates with momentum. See [55] for more details on the specific U-net architecture. In addition, we also apply data augmentation techniques such as random zoom, lightning, and contrast changes to increase the size and variety of our training data. We avoided heavy distortion augmentations such as flip, shift, and warp in order to avoid collision with the spatial distortions introduced in video-rate acquisition process. We found these heavy distortions actually hindered model performance.

We divided the training of our GAN into three steps as visually illustrated in Fig. 2 in order to facilitate the task of training a stable GAN model. In step 1, we train the U-net generator using MSE loss to achieve fast microscopy super-resolution on our training data. Next, in step 2, we train a critic to distinguish real high resolution images from the fake images generated by our trained generator using corresponding lower resolution images. Finally, in step 3, we train the full GAN by simultaneously training the generator and critic models that were pre-trained in steps 1 and 2. For more details on fast.ai GAN and its training method, we refer the reader to [53].

B. Metrics and evaluations.

To assess the quality of the GAN-based image reconstructions we have selected two metrics widely used in image processing and appropriate for our task: the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [56], [57]. PSNR is commonly used to measure the reconstruction quality of lossy transformations (e.g., image compression, image inpainting) and is currently the most widely used evaluation criteria for super-resolution models. The PSNR is typically expressed in terms of the logarithmic decibel scale and is calculated between the original and the reconstructed image: the higher the PSNR, the better is the quality of the reconstructed image. SSIM instead is proposed for measuring the structural similarity between images, based on three relatively independent comparisons, namely luminance, contrast, and structure. Since SSIM evaluates the reconstruction quality from the perspective of the human visual system (HVS), it better meets the requirements of perceptual assessment [56], [58] and thus widely used by super-resolution models. The SSIM metric-range extends from -1 to +1, and equals +1 when the original and the reconstructed image are equal. For a more comprehensive review and a precise mathematical definition of both PSNR and SSIM, the reader is referred to [56].

It is also clear that any finite set of metrics may not capture the quality of the super resolution for the microscopy community. We therefore also provide here representative sample images to demonstrate the capability of our GAN-based image reconstructions to pick up both the coarse and fine grain structures present within the images.

C. Simultaneous denoising, super-resolving, and distortion removal.

We have trained the GAN model on 6,000 paired images of mitochondria (sample A) acquired at video-rate speed (VideoRate images, Table I). We used 5,600 paired images for training and 400 for validation. In Table II we present the average structural similarity (SSIM) and the peak signal-to-noise ratio (PSNR) metrics calculated on this validation set, for both our GAN model as well as a simple baseline bilinear upsample model for comparison purposes. The results we have obtained demonstrate the very strong capability of our GAN model to simultaneously super-resolve, denoise and remove distortions from the images in the validation set. To better evaluate the generalization of our deep learning model, we tested our model on a new unseen test set of 192 paired images that were never used during the training. The box-plot in Fig. 5 shows the distribution of the newly obtained SSIM and PSNR metrics considering this test set. We have found a mean value of 0.87 and 29.73 for the SSIM and PSNR respectively (Table II). In addition to a vast improvement over the baseline bilinear model, these results fall within the range (25-32 for PSNR and 0.7-0.9 for SSIM) of other successful and recent deep learning microscopy results see [45] for PSNR and [36] for SSIM and PSNR. This is the first important piece of evidence that our GAN model is capable of simultaneously super-resolve, denoise and remove distortion from images obtained at video-rate acquisition speed. We also present in Fig. 4, a representative visual example of the outstanding improvement of the GAN image (prediction) over the VideoRate image (input), and how closely the prediction matches the ground truth high resolution image HR. The exceptional recovery of the mitochondrial features within the prediction image is also better emphasized in greater details in the zoomed areas of Fig. 4d-f.

TABLE II
PERFORMANCE ON VALIDATION

Model	SSIM	PSNR
Bilinear	0.521 ± 0.071	22.65 ± 2.42
GAN	0.873 ± 0.038	29.73 ± 1.635

D. Optimal number of paired images

In fluorescence microscopy, gathering a large amount of paired images is not always feasible, recommended, and/or practical. It would therefore be useful to see if good predictions could be obtained with a reduced number of images in the training dataset, with the intent of limiting the training process during normal acquisition sessions. We have seen that our GAN model, introduced in section III-C and trained on 5600 paired images, was able to effectively simultaneously super-resolve, denoise and correct distortion on images obtained at
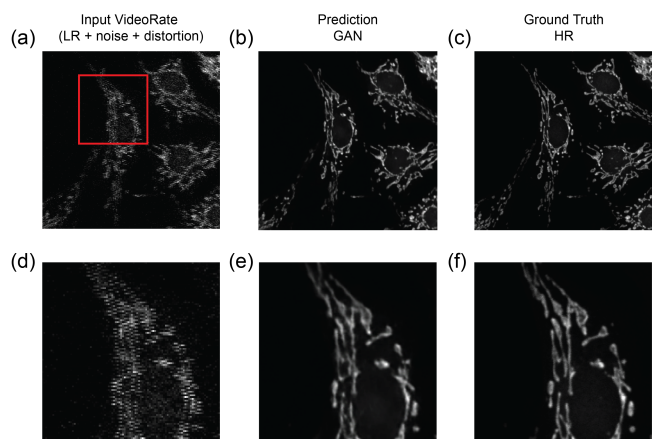
Fig. 4. GAN predictions on typical acquired images. (a) Representative input image acquired at video-rate acquisition speed and heavily affected by distortion and noise. (b) GAN prediction. (c) Corresponding target (ground truth) HR image. (d,e,f) Analogue magnified images corresponding to the red box area indicated in (a).
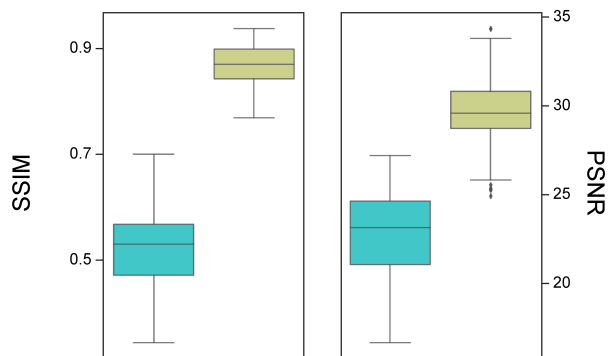

Fig. 5. Box plot of the SSIM (Left) and PSNR (Right) metrics on a new unseen test set of 192 paired images obtained on Sample A. For comparison a simple baseline bilinear upsample (cyan), is shown in addition to our GAN model (yellow).

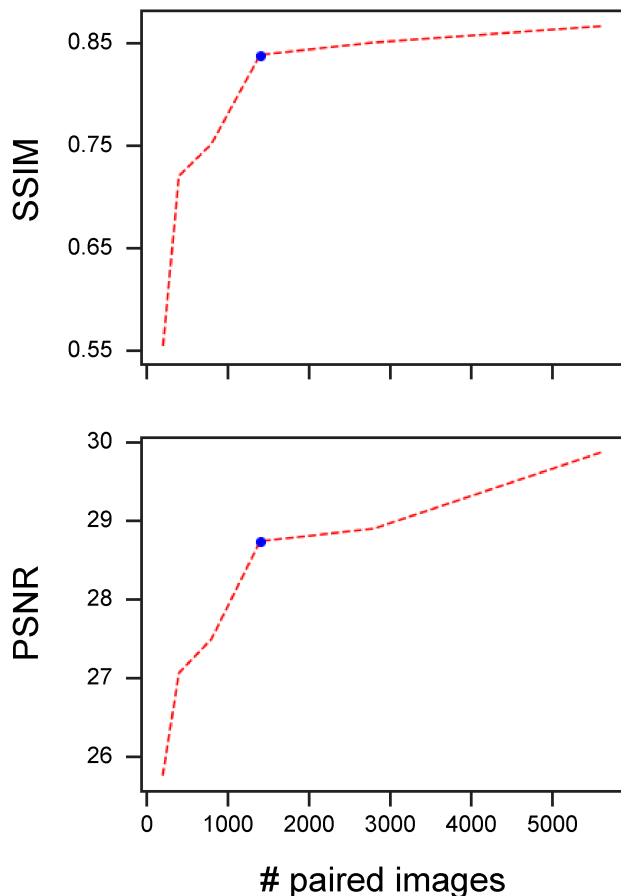| Number of paired images	SSIM	PSNR
5600	0.866	29.862
2800	0.85	28.894
1400	0.838	28.734
800	0.751	27.492
400	0.72	27.062
200	0.551	25.762

Fig. 6. SSIM and PSNR metrics calculated on the predicted images as a function of total number of paired images. Blue dot indicates the minimum number of images necessary for obtaining adequate fast microscopy image reconstructions, at 1400 paired images. In this example, 1400 paired images is a reasonable minimum number of images necessary for obtaining adequate fast microscopy image reconstructions (blue dot in Fig. 6).

E. Generalization to unseen and unrelated biological structures

It is also imperative to understand how GANs generalize to unseen and unrelated biological structures (e.g. nuclei, cytoplasm, organelles, membranes, etc). In fact, biological specimens tend not to be uniform across a large field of view, with heterogeneity in cell populations, organization, structure, and signal. To measure the generalizability of our model, we acquired a completely new dataset of 848 paired video-

video-rate acquisition speed. We now explore the sensitivity of our performance metrics as the number of paired images used in the training set decreases. The results are given in Fig. 6 and Table III and clearly show how the metrics on the predicted images change as we decrease the number of paired images from 5600 to 200. From these results, it is clear that is therefore possible to strike a balance trade between efforts spent on acquiring enough images to train on and image accuracy. In this example, 14 ground truth images each one with 100 temporal images to train on the noise(1400) is a reasonable minimum number of images necessary for obtaining adequate fast microscopy image reconstructions (blue dot in Fig. 6).

rate images taken on an histological section of the mouse intestines (Sample B, Fig. 7). As noted in section II-B, this sample is vastly different in physiology and morphology from the original mitochondria dataset (Sample A). To evaluate performance, we used a GAN model, previously trained on Sample A, and directly applied it to the new dataset. Separately, we evaluate the benefits of model fine-tuning by evaluating image reconstruction performance on a model fine-tuned with both 400 and 800 paired images from sample B itself (GAN Finetuned4, and GAN Finetuned8 respectively). In Fig. 7, we present representative visual results obtained by the direct application of the model with no fine-tuning (Fig. 7c) and one fine-tuned on 400 paired images (Fig. 7d). As clearly evident, image artifacts present in the predictions obtained without fine tuning (Fig. 7c) are absent in the one fine tuned with 400 paired images (Fig. 7d). This provides promising visual evidence that with minimal fine tuning we can achieve robust image reconstructions on vastly different biological samples and within large imaging field of views. We also note the promising metrics given in Table IV. Here, the mean and the standard deviation on the SSIM and the PSNR are calculated for all three models as well as a baseline using bilinear upsampling on an unseen test set of 48 paired images from Sample B. In addition to the clear improvement over bilinear upsampling, we see that light fine tuning can result in SSIM scores in the .7-.9 ranges and PSNR scores above 25 as well.

TABLE IV
PERFORMANCE ON TEST SET

Model	SSIM	PSNR
Bilinear	0.52 ± 0.059	22.03 ± 3.277
GAN	0.69 ± 0.097	19.75 ± 5.066
GAN Finetuned4	0.76 ± 0.057	25.36 ± 3.461
GAN Finetuned8	0.7847 ± 0.049	26.65 ± 2.794

IV. CONCLUSION

We have presented here a new approach for combining deep learning super resolution networks with fast acquisition microscopy to achieve fast microscopy deep learning super-resolution. Specifically, we demonstrate that bidirectional image scanning provides a significant speed boost over previous acquisition methods without any significant loss in image quality when combined with GAN methods. This is confirmed both by image quality assessment metrics and more importantly also by showcasing typical visual examples of the outstanding improvements of the GAN model predictions and the exceptional recovery of cellular features. Additionally, we show that super resolution microscopy may not need a large volume of data to train. In our study, we have observed diminishing gains in performance when using only a quarter of the number of paired training images. Lastly, We provide evidence that our GAN approach has effective generalizability (with light fine tuning) to new, unseen microscopy images.

To conclude, our deep learning approach is starting to chip away at previous limitations of fast imaging microscopy
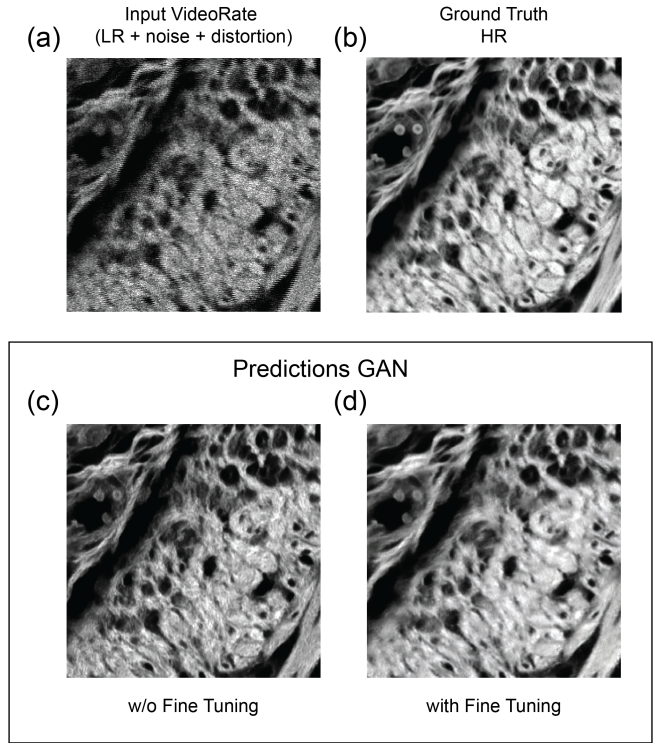

Fig. 7. Prediction on novel dissimilar unseen image datasets with minimal or no fine tuning of the GAN model. (a) Representative input VideoRate image obtained on sample B. (b) Corresponding ground truth image. (c,d) Predictions obtained with the previously obtained GAN model, without fine tuning (c) and with fine tuning (d).

including the fundamental triangle of compromise in microscopy. Our work demonstrates that it's possible to now acquire images at a greatly increased frame-rate without sacrificing image quality. In our opinion, the proposed methodology will allow microscopy researchers a cheap and easy new, open-source tool to approach challenging imaging problems such as the aforementioned examples (e.g. blood cell trafficking in capillary network, capturing neuronal activity, etc.) where video-rate acquisition is necessary to resolve fast dynamics events, as well as to accelerate multigigabyte-sized whole slides histopathological data acquisitions and whole cleared organ mapping.

REFERENCES

[1] C. Vinegoni, P.F. Feruglio, and R. Weissleder, "High dynamic range fluorescence imaging," IEEE Journal of Selected Topics in Quantum Electronics, vol. 25, pp. 1–7, 2018.
[2] C. Vinegoni, et al., "Real-time high dynamic range laser scanning microscopy," Nature communications, vol. 7, pp. 1–13, 2016.
[3] J.A. Conchello, and J.W. Lichtman, "Optical sectioning microscopy," Nature methods, vol. 2, pp. 920–931, 2005.
[4] F. Helmchen, and W. Denk, "Deep tissue two-photon microscopy," Nature Methods, vol. 2, pp. 932–940, 2005.
[5] J. Condeelis, and R. Weissleder, "In vivo imaging in cancer," Cold Spring Harbor perspectives in biology, vol. 2, a003848, 2010.
[6] E. Gallo, and J.W. Jarvik, "Breaking the color barrier–a multi-selective antibody reporter offers innovative strategies of fluorescence detection," J. Cell. Sci., vol. 130, pp. 2644–2653, 2017.

[7] M.J. Pittet, and R. Weissleder, "Intravital imaging," Cell, vol. 147, pp. 983–991, 2011.

[8] A.D. Aguirre, C. Vinegoni, W. Sebas, and R. Weissleder, "Intravital imaging of cardiac function at the single-cell level," Proceedings of the National Academy of Sciences, vol. 111, pp. 11257–11262, 2014.

[9] R, Yuste, "Fluorescence microscopy today," Nature methods, vol. 2, pp. 902–904, 2005.

[10] R.P. Aylward, "Advanced galvanometer-based optical scanner design," Sensor Review, 2003

[11] J. Pawley, "Handbook of biological confocal microscopy," Springer Science & Business Media, 2010.

[12] P. Saggau, "New methods and uses for fast optical scanning," Current opinion in neurobiology, vol. 16, pp.543–550, 2006.

[13] M. Sanderson, "Acquisition of multiple real-time images for laser scanning microscopy mirrors," Microsc. Anal., vol. 18, pp. 2–6, 2004.

[14] C. Vinegoni, P.F. Feruglio, I. Gryczynski, R. Mazitschek, and R. Weissleder, "Fluorescence anisotropy imaging in drug discovery," Advanced drug delivery reviews, vol. 151, pp. 262–288, 2019.

[15] T. Wilson, "Resolution and optical sectioning in the confocal microscope," Journal of microscopy, vol. 244, pp. 113–121, 2011.

[16] C. Vinegoni, A.D. Aguirre, S. Lee, R. Weissleder, "Imaging the beating heart in the mouse using intravital microscopy techniques," Nature protocols vol. 10, pp. 1802, 2015.

[17] S. Lee, et al., "Real-time in vivo imaging of the beating mouse heart at microscopic resolution," Nature communications, vol. 3, pp. 1–8, 2012.

[18] A.Y. Chan, "Biomedical device technology: principles and design," Charles C Thomas Publisher, 2016.

[19] Y. Chen, et al., "Video-rate in vivo fluorescence imaging with a line-scanned dual-axis confocal microscope," Journal of biomedical optics, vol. 20, 106011, 2015.

[20] L. Fang, et al., "Deep learning-based point-scanning super-resolution imaging," bioRxiv p. 740548, 2019.

[21] V.F. Duma, "Optimal scanning function of a galvanometer scanner for an increased duty cycle," Optical Engineering, vol. 49, 103001, 2010.

[22] G. E. Fantner, et al., "Data acquisition system for high speed atomic force microscopy," Review of Scientific Instruments, vol 76, 026118, 2005.

[23] H.W. Yoo, S. Ito, and G. Schitter, "High speed laser scanning microscopy by iterative learning control of a galvanometer scanner," Control Engineering Practice, vol. 50, pp. 12–21, 2016.

[24] V.F. Duma, K.S. Lee, P. Meemon, and J.P. Rolland, "Experimental investigations of the scanning functions of galvanometer-based scanners with applications in oct," Applied Optics, vol. 50, pp. 5735–5749, 2011.

[25] N. Ji, H. Shroff, H. Zhong, and E. Betzig, "Advances in the speed and resolution of light microscopy," Current opinion in neurobiology, vol. 18, pp. 605–616.

[26] K. Leang, and S. Devasia, "Design of hysteresis-compensating iterative learning control for piezo-positioners: Application to atomic force microscopes," Mechatronics, vol. 16, pp. 141–158, 2006.

[27] I. Veilleux, J.A. Spencer, D.P. Biss, D. Cote, and C.P. Lin, "In vivo cell tracking with video rate multimodality laser scanning microscopy," IEEE Journal of selected topics in quantum electronics, vol. 14, pp. 10–18, 2008.

[28] C. Halin, J. Rodrigo Mora, C. Sumen, and U.H. von Andrian, "In vivo imaging of lymphocyte trafficking," Annu. Rev. Cell Dev. Biol., vol.21, pp. 581–603, 2005.

[29] B.N. Ozbay, et al., "Three dimensional two-photon brain imaging in freely moving mice using a miniature fiber coupled microscope with active axial-scanning," Scientific reports, vol. 8, pp. 1–14, 2018.

[30] K.N.S. Nadella, et al., "Random-access scanning microscopy for 3d imaging in awake behaving animals," Nature methods vol. 13, pp. 1001, 2016.

[31] W. Göbel, B.M. Kampa, and F. Helmchen, "Imaging cellular network dynamics in three dimensions using fast 3d laser scanning," Nature Methods, vol. 4, pp. 73–79, 2007.

[32] B.F. Grewe, F.F. Voigt, M. van't Hoff, and F. Helmchen, "Fast two-layer two-photon imaging of neuronal cell populations using an electrically tunable lens," Biomedical optics express, vol. 2, pp. 2035–2046, 2011.

[33] W. Yang, et al., "Simultaneous multi-plane imaging of neural circuits," Neuron, vol. 89, pp. 269–284, 2016.

[34] M.N. Economo, et al., "A platform for brain-wide imaging and reconstruction of individual neurons," Elife, vol. 5, e10566, 2016.

[35] N. Tanaka, et al., "Whole-tissue biopsy phenotyping of three-dimensional tumours reveals patterns of cancer heterogeneity," Nature Biomedical Engineering, vol. 1, pp. 796–806, 2017.

[36] Y. Zhang, et al., "A poisson-gaussian denoising dataset with real fluorescence microscopy images," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 11710–11718, 2019.

[37] C. Dong, C.C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," European Conference on Computer Vision (ECCV), vol. 8692, pp. 184–199, Springer 2014.

[38] C. Dong, C.C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, pp. 295–307, 2016.

[39] I. Goodfellow, et al., "Generative adversarial nets," Proceedings of the 27th International Conference on Neural Information Processing Systems, vol. 2, pp. 2672–2680, 2017.

[40] C. Ledig et al., "Photorealistic single image super-resolution using a generative adversarial network," Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 105–114, 2017.

[41] Z. Wang, J. Chen, S.C. Hoi, "Deep learning for image super-resolution: A survey," IEEE transactions on pattern analysis and machine intelligence, 2020.

[42] Y. Rivenson, et al., "Deep learning enhanced mobile-phone microscopy," ACS Photonics, vol. 5, pp. 2354–2364, 2018.

[43] W. Ouyang, et al., "Deep learning massively accelerates super-resolution localization microscopy," Nat. Biotechnol., vol. 36, pp. 460–468, 2018.

[44] H. Wang et al. "Cross-modality deep learning achieves super-resolution in fluorescence microscopy," Conference on Lasers and Electro-Optics (CLEO), pp. 1–2, 2019.

[45] B. Manifold, E. Thomas, A.T. Francis, A.H. Hill, and D. Fu, "Denoising of stimulated raman scattering microscopy images via deep learning," Biomedical optics express, vol. 10, pp. 3860–3874, 2019.

[46] H. Wang, et al., "Deep learning enables cross-modality super-resolution in fluorescence microscopy," Nature Methods, vol. 16, pp. 103–110, 2019.

[47] C. Belthangady, and L.A. Royer, "Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction," Nat. Methods, vol. 16, pp. 1215–1225, 2019.

[48] L. Leybaert, A. de Meyer, C. Mabilde, and M. Sanderson, "A simple and practical method to acquire geometrically correct images with resonant scanning-based line scanning in a custom-built video-rate laser scanning microscope," Journal of microscopy, vol. 219, pp. 133–140, 2005.

[49] A.R. Hibbs, G. MacDonald, K. Garsha, "Practical confocal microscopy," pp. 650–671, Springer, 2006.

[50] R.K.P. Benninger, W.J. Ashby, E.A. Ring, and D.W. Piston, "Single-photon-counting detector for increased sensitivity in two-photon laser scanning microscopy," Optics letters, vol. 33, pp. 2895–2897, 2008.

[51] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," IEEE Transactions on image processing, vol. 16, pp. 2080–2095, 2007.

[52] M. Makitalo, and A. Foi, "Optimal inversion of the anscombe transformation in low-count poisson image denoising," IEEE transactions on Image Processing, vol. 20, pp. 99–109, 2010.

[53] J. Howard, R. Thomas, and S. Gugger, "FastaiGitHub Homepage," https://docs.fast.ai/vision.gan.html.

[54] O. Ronneberger, P. Fischer, T. Brox, "Convolutional networks for biomedical image segmentation," Medical Image Computing and Computer-Assisted Intervention, vol. 9351, Springer, 2015.

[55] J. Howard, R. Thomas, and S. Gugger, "FastaiGitHub Homepage," https://docs.fast.ai/vision.models.unet.html.

[56] A. Horé, and D. Ziou, "Image quality metrics: Psnr vs. ssim," 20th International Conference on Pattern Recognition, pp. 2366–2369, 2010.

[57] Z. Wang, and A. Bovik, "Mean squared error: Love it or leave it? a new look at signal fidelity measures," IEEE signal processing magazine, vol. 26, pp. 98–117, 2009.

[58] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE transactions on image processing, vol. 13, pp. 600–612, 2004.