Rapid tissue oxygenation mapping from snapshot structured-light images with adversarial deep learning

Mason T. Chen and **Nicholas J. Durr***
Johns Hopkins University, Department of Biomedical Engineering, Baltimore,
Maryland, United States

Abstract

Significance: Spatial frequency-domain imaging (SFDI) is a powerful technique for mapping tissue oxygen saturation over a wide field of view. However, current SFDI methods either require a sequence of several images with different illumination patterns or, in the case of single-snapshot optical properties (SSOP), introduce artifacts and sacrifice accuracy.

Aim: We introduce OxyGAN, a data-driven, content-aware method to estimate tissue oxygenation directly from single structured-light images.

Approach: OxyGAN is an end-to-end approach that uses supervised generative adversarial networks. Conventional SFDI is used to obtain ground truth tissue oxygenation maps for *ex vivo* human esophagi, *in vivo* hands and feet, and an *in vivo* pig colon sample under 659- and 851-nm sinusoidal illumination. We benchmark OxyGAN by comparing it with SSOP and a two-step hybrid technique that uses a previously developed deep learning model to predict optical properties followed by a physical model to calculate tissue oxygenation.

Results: When tested on human feet, cross-validated OxyGAN maps tissue oxygenation with an accuracy of 96.5%. When applied to sample types not included in the training set, such as human hands and pig colon, OxyGAN achieves a 93% accuracy, demonstrating robustness to various tissue types. On average, OxyGAN outperforms SSOP and a hybrid model in estimating tissue oxygenation by 24.9% and 24.7%, respectively. Finally, we optimize OxyGAN inference so that oxygenation maps are computed ~10 times faster than previous work, enabling video-rate, 25-Hz imaging.

Conclusions: Due to its rapid acquisition and processing speed, OxyGAN has the potential to enable real-time, high-fidelity tissue oxygenation mapping that may be useful for many clinical applications.

1 Introduction

Tissue oxygenation (StO_2) is a measure of the amount of oxygen in biological tissue and is often estimated by computing the fraction of oxygenated hemoglobin over total hemoglobin. StO_2 is a useful clinical biomarker for tissue viability, the continuous monitoring of which is valuable for surgical guidance and patient management.[1,2] Abnormal levels of StO_2 are indicative of many pathological conditions, such as sepsis, diabetes, and chronic obstructive pulmonary disease.[3–5]

One of the most widely used techniques for measuring physiological oxygen levels is pulse oximetry. Despite its ubiquity, robustness, and low cost, pulse oximetry requires a pulsatile arterial signal and only provides a systemic measure of oxygenation.[6,7] The majority of existing devices for local assessment of StO_2 are based on near-infrared (NIR) spectroscopy. NIR spectroscopy

---

*Address all correspondence to Nicholas J. Durr, ndurr@jhu.edu

quantifies the concentrations of oxygenated and de-oxygenated hemoglobin by characterizing tissue absorption of light at wavelengths typically between 650 and 1000 nm.[8] Similar to pulse oximetry, spectroscopic probes require direct contact with tissue. These measurements can be noisy as they are sensitive to pressure and sample movement.[9–11] Compared with tissue probes, spectroscopic imaging techniques are advantageous as they provide non-contact readings of oxygen saturation at a high spatial resolution and large field of view.[12] Nevertheless, continuous-wave NIR spectroscopy assumes constant scattering, which could be a source of error as scattering coefficients are often spatially non-uniform. Therefore, to accurately determine oxygen saturation, it is imperative to separate the effects of optical properties, including absorption (μ_a) and reduced scattering coefficients (μ_s'). Spectrally constrained reconstructions have been shown to be useful in measuring chromophore concentrations and μ_s', but this technique tends to require complex instrumentation and suffers from limited fields of view.[13]

In recent years, spatial frequency-domain imaging (SFDI) has emerged as a promising technique for measuring tissue optical properties. SFDI quantifies optical properties by projecting structured-light and characterizing the modulation transfer function of tissues in the spatial frequency domain.[14] Oxygenation can subsequently be determined by fitting chromophore concentrations to the measured absorption coefficients using the Beer–Lambert law. In addition to isolating the effect of tissue scattering, SFDI is a wide-field, non-contact technique that can be implemented using a simple setup that includes a camera and a projector. These advantages make SFDI suitable for many clinical applications that necessitate accurate StO_2 measurements, such as burn wound assessment,[15,16] pressure ulcer staging and risk stratification,[17] image-guided surgery,[7,11] and cancer therapy evaluation.[18]

Despite its growing use in various applications, there are several factors that limit the clinical translation of SFDI. First, compared with probe-based oximetry, SFDI components are costly. For example, digital micromirror devices or spatial light modulators are often used to produce programmable structured illumination. Second, SFDI requires carefully controlled imaging geometries, which can be difficult to achieve in clinical settings. Moreover, conventional SFDI requires six images per wavelength (0, $\frac{2}{3}\pi$, and $\frac{4}{3}\pi$ phase offsets at two spatial frequencies) and a pixel-wise lookup table (LUT) search to generate a single optical property map. For robust oxygenation estimates, absorption coefficients at a minimum of two wavelengths are needed, and an additional least square fitting step is performed [Fig. 1(a)].[19] Previous work has shown that real-time imaging can be achieved with single-snapshot acquisition[20] and either an optimized LUT[21] or a machine learning inversion method.[22] However, single image acquisition and frequency filtering often result in image artifacts and high per-pixel error.[23] Therefore, wide-field, rapid, and accurate StO_2 measurement still remains a challenge.

In recent years, convolutional neural networks (CNNs) have emerged as a powerful tool in many medical imaging-related tasks.[24,25] By employing convolutional filters followed by dimension reduction and rectification, CNNs are capable of extracting high-level features and interpreting spatial structures of input images.[26] For image translation tasks, generative adversarial networks (GANs) improve upon conventional CNNs by utilizing both a generator and a discriminator[27] to effectively model a complex loss function. The two components are trained simultaneously, with the generator learning to produce realistic output and the discriminator to classify the generator output as real or fake. Recently, GANs have been employed to predict optical properties from single structured-illumination images (GANPOP).[28] As a content-aware network, this technique significantly improves upon the accuracy of model-based single-snapshot techniques in estimating optical properties. However, to compute StO_2 with the GANPOP approach, multiple wavelength-specific networks must be run to first estimate absorption coefficients, followed by chromophore fitting, which compounds errors and increases computational demand [Fig. 1(b)]. In this study, we present an end-to-end technique for computing StO_2 directly from structured-illumination images using GANs (OxyGAN). OxyGAN maps StO_2 from single-snapshot images from 659- and 851-nm sinusoidal illumination. We train generative networks to estimate both uncorrected and profile-corrected StO_2 and compare the performance of the end-to-end architecture versus intermediately calculating optical properties. We accelerate OxyGAN model inference by importing the framework into NVIDIA TensorRT for efficient deployment. Finally, we demonstrate real-time OxyGAN by recording its estimation over the course of a 3-min occlusion experiment.
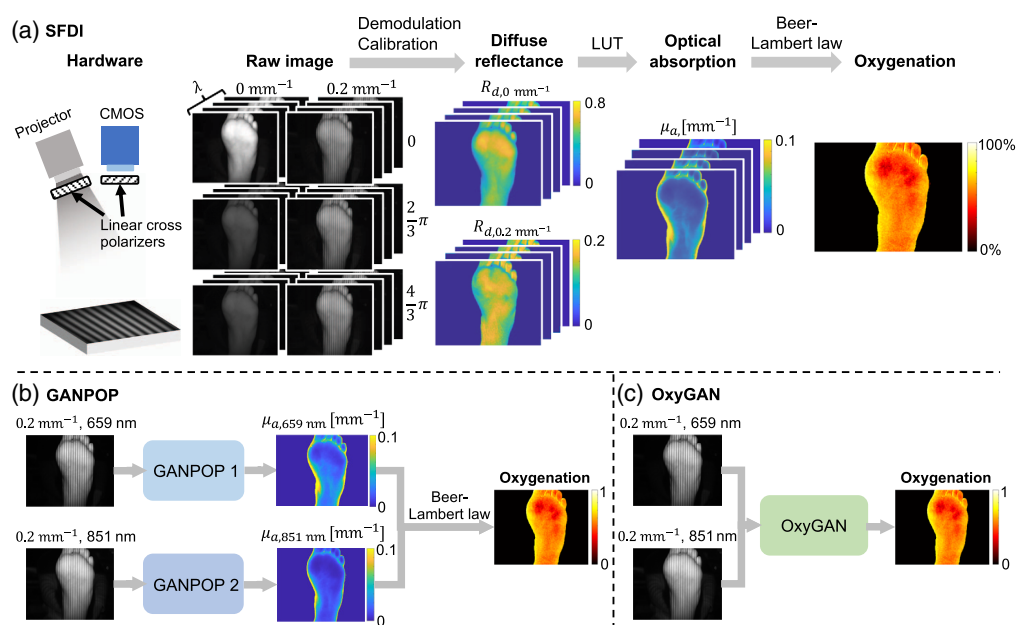
Fig. 1 Comparison of (a) SFDI, (b) GANPOP, and (c) OxyGAN StO$_2$ techniques. Ground truth SFDI uses 24 input images (2 spatial frequencies, 3 phases, and 4 wavelengths), while GANPOP and OxyGAN use 2 input images (1 spatial frequency, 1 phase, and 2 wavelengths). SFDI and GANPOP absorption maps are subsequently fit to basis chromophores to estimate StO$_2$. OxyGAN directly calculates oxygen saturation with a single network, reducing compounding errors and processing time.

2 Methods

For training and testing of OxyGAN, single structured-illumination images were acquired at two different wavelengths (659 and 851 nm) and paired with registered oxygenation maps. Ground truth oxygenation is obtained using the absorption coefficients measured by conventional SFDI at four wavelengths (659, 691, 731, and 851 nm). Experiments were conducted using both profile-corrected[29,30] and uncorrected ground truth. The training set included human *ex vivo* esophagus samples and *in vivo* feet. OxyGAN was evaluated using unseen tissues of the same type as the training data (human *in vivo* feet) and in different tissue types (human *in vivo* hands and a pig *in vivo* colon). Its performance was additionally compared with single-snapshot optical properties (SSOP)[23,31] as a model-based benchmark that utilizes a single structured-light image.

2.1 Ground Truth Tissue Oxygenation

In this study, conventional SFDI was used to obtain ground truth StO$_2$ maps. At each wavelength, structured-illumination images were captured using a commercial SFDI system (Reflect RS, Modulim Inc.) at two spatial frequencies (0 and 0.2 mm^{-1}) and three phase offsets (0, $\frac{2}{3}\pi$, and $\frac{4}{3}\pi$). The process was implemented for both the sample of interest and a reference phantom. The acquired images were then demodulated and calibrated against the response of the reference phantom at each frequency. The DC (0 mm^{-1}) and AC (0.2 mm^{-1}) diffuse reflectance of the sample were fit to an LUT generated by White Monte Carlo simulations.[32] This pixel-wise LUT search resulted in an optical property map of the sample, which consisted of scattering corrected absorption (μ_a) and reduced scattering (μ_s') coefficients. In experiments in which profile-corrected ground truth was used, we also implemented height and surface normal angle correction.[29,30] With μ_a measured at four different wavelengths (659, 691, 731, and 851 nm), we subsequently estimated chromophore concentrations using the Beer–Lambert law:

$$\mu_a(\lambda_i) = \sum_{n=1}^{N} \epsilon_n(\lambda_i) c_n, \tag{1}$$

where $\epsilon_n(\lambda_i)$ stands for the extinction coefficient of chromophore n at wavelength λ_i, c_n is its concentration, and N is the total number of chromophores. Oxygen saturation was then calculated as

$$\mathrm{stO_2} = \frac{c_{\mathrm{O_2Hb}}}{c_{\mathrm{O_2Hb}} + c_{\mathrm{HHb}}}, \tag{2}$$

where $c_{\mathrm{O_2Hb}}$ and c_{HHb} represent the concentrations of oxygenated and de-oxygenated hemoglobin, respectively. We estimated ground truth oxygenation maps using absorption coefficients at all of the NIR wavelengths available in the system (659, 691, 731, and 851 nm).

2.2 SSOP Benchmark

In this study, we implemented SSOP as a model-based benchmark. Briefly, this method calculates tissue optical properties from single structured-illumination images by 2-D filtering in the frequency domain.[23,31] We applied anisotropic low-pass filtering using a sine window and high-pass filtering using a Blackman window.[31] The absorption coefficients measured by SSOP at 659 and 851 nm were substituted into Eqs. (1) and (2) to estimate StO$_2$.

2.3 OxyGAN Framework

In this study, we pose StO$_2$ mapping as an image-to-image translation task. OxyGAN uses an adversarial training framework to accomplish this task (Fig. 2). Specifically, OxyGAN is a conditional generative adversarial network (cGAN) that consists of two CNNs—a generator and a discriminator. Both networks are conditioned on the same input data, which are single structured-light images in our case. First proposed in Ref. 33, the cGAN structure has been shown to be an effective solution to a wide range of image-to-image translation problems.[34] While conventional single-network CNNs require simple, hand-crafted loss functions, cGANs can be more generalizable because the discriminator can effectively learn a complex loss function.

For the OxyGAN generator, we implement a fusion network that combines the properties of a U-Net and a ResNet (Fig. 2).[35,36] Similar to a U-Net, the OxyGAN generator is an encoder–decoder setup with long skip connections between the two branches on the same level. However,
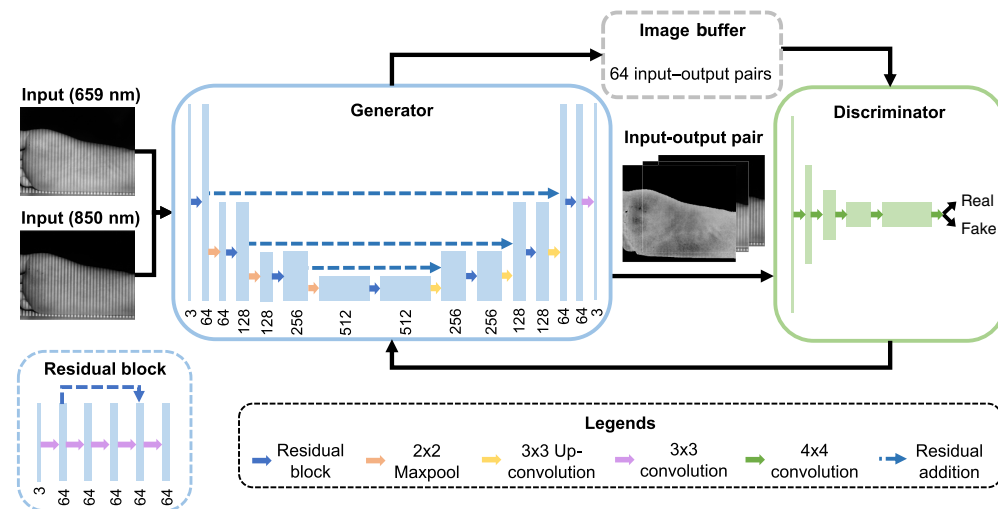

Fig. 2 The OxyGAN framework. OxyGAN produces StO$_2$ maps directly from single-phase SFDI images with 659- and 851-nm illumination. The generator is a fusion network that combines the properties of a U-Net and a ResNet. The number under each block describes the number of channels. The discriminator is a PatchGAN classifier with a receptive field of 70×70 pixels. The discriminator trains to classify the current image pair versus an input–output pair randomly sampled from a pool of 64 previously generated images.

OxyGAN also includes short skip connections within each level and replaces U-Net concatenation with additions, making the network fully residual.[28,37] The residual blocks on each level consist of five 3×3 convolutional layers, with residual additions between the outputs of the first and the fourth layers. Dimension reduction on the encoder side and expansion on the decoder side is achieved with 2×2 maxpooling and 3×3 up-convolutions, respectively. We use regular ReLUs for the encoder and leaky ReLUs with a slope of 0.2 for the decoder. A final 3×3 convolution followed by a $Tanh$ activation function is applied to generate the output. The discriminator is a three-layer PatchGAN with leaky ReLUs (slope $= 0.2$),[34] which results in a receptive field of 70×70 pixels. To stabilize the training process, we incorporated spectral normalization after each convolution layer in both the generator and the discriminator.[38] We use an adversarial loss of

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y}[\log(D(x, y))] + \mathbb{E}_x[\log(1 - D(x, G(x)))], \qquad (3)$$

where G is the generator ($G: X \rightarrow Y$) and D is the discriminator.[34] During training, G tries to minimize this objective while its adversary, D, tries to maximize it. The discriminator is trained to classify both the current input–ground truth pair and an image pair randomly sampled from a buffer of 64 previously generated images. Determine if a given pair of images forms a correct reconstruction for a given input. This classification is made from data that include the current input–ground truth pair and an image pair randomly sampled from a buffer of 64 previously generated pairs. Additionally, an L_1 loss is included to improve the generator performance and training stability:

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y}[\|y - G(x)\|_1]. \qquad (4)$$

The full objective function is expressed as

$$\min_G \max_D \mathcal{L}(G, D) = \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G), \qquad (5)$$

where λ is a hyperparameter that controls the weight of the L_1 loss term and was set to 60. OxyGAN models solved this objective using an "Adam" solver with a batch size of 1.[39] G and D weights were both initialized from a Gaussian distribution with a mean and standard deviation of 0 and 0.02, respectively. These models were trained for 200 epochs, and a constant learning rate of 0.0002 was used for the first 100 epochs. The learning rate was linearly decreased to 0 for the second half of the training process. The full algorithm was implemented using Pytorch 1.0 on Ubuntu 16.04 with a single NVIDIA Tesla P100 GPU on Google Cloud.[28]

2.4 Data Split and Augmentation

In this study, we conducted separate experiments to estimate both uncorrected (StO_2) and profile-corrected oxygenation ($StO_{2,corr}$) from the same single-snapshot structured-light image input. These networks were trained and tested on 256×256-pixel patches paired with registered oxygenation maps. To generate training datasets, each 520×696 image was segmented at a random stride size, which resulted in ~30 image pairs per sample. The input data were arranged in a way so that the three image channels normally used for color were efficiently utilized (Fig. 3). The first and second channels are flat-field corrected, single-phase illumination images at 659 and 851 nm, respectively. To account for system drift over time, we included the ratio between demodulated AC (M_{AC}) and DC magnitude (M_{DC}) of the reference phantom in the third channel. Reference measurements were taken on the same day as the tissue measurements, in the same way as conventional SFDI workflows. As shown in Fig. 3, the ratios at 659 and 851 nm alternate in a checkerboard pattern to account for any spatial variations.

To prevent overfitting of the models, we augmented the training data by flipping the images horizontally or vertically. During each epoch, both flipping operations occurred with a 50% chance and were independent of each other. Data augmentation was important for this study because of the small size of the training set and because the testing set includes new object types never seen in training. Additionally, since the classification task of the discriminator was
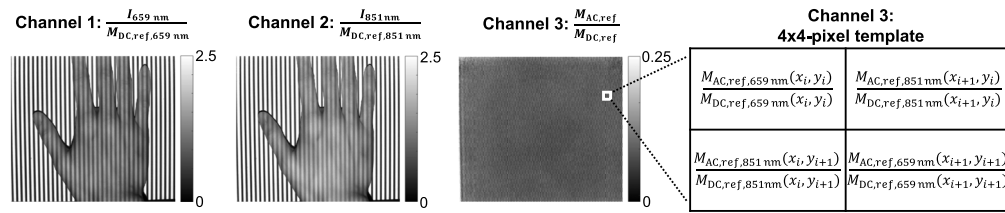
Fig. 3 Images of the three input channels. Channels 1 and 2 are flat-field corrected single-snapshot SFDI images at 659 and 851 nm. Channel 3 gives ratios between demodulated AC ($M_{AC, ref}$) and DC magnitude ($M_{DC, ref}$) of the reference phantom. The ratios at 659 and 851 nm are alternated to form a checkerboard pattern, as shown in the 4×4-pixel template on the right.

easier than the generator, we applied the one-sided label smoothing technique when training the discriminator. In short, the positive (real) targets with a value of 1 were replaced with a smoothed value (0.9 in our case). This was implemented to prevent the discriminator from becoming over-confident and using only a small set of features when classifying output.[40]

2.5 Samples

The training set of OxyGAN models included eight *ex vivo* human esophagectomy samples[41] and four *in vivo* human feet, which resulted in ~1200 image pairs after augmentation. The testing set consisted of two *in vivo* human hands and feet and an *in vivo* pig colon. All models were cross-validated by training on four of the six feet and testing on the remaining two each time. All summary results indicate the average performance of these three sets of trained networks. OxyGAN models never see data from hands or *in vivo* pig colon in training.

We additionally recorded a 400-s video of a healthy volunteer's hand during an occlusion study. We first applied a household pressure cuff (Walgreens Manual Inflate Blood Pressure Kit) to the upper arm of the volunteer and imaged the hand at baseline for a minute. Then, the cuff pressure was increased to 200 mmHg to occlude the arm for ~3 min. The pressure was then released, and the hand was imaged for another 2.5 min. Single-phase sinusoidal illumination was used; it was alternated between 659 and 851 nm so that oxygenation could be measured at each time point ($\Delta t = 0.73$ s). To obtain ground truth oxygenation, conventional six-image SFDI was implemented every 25 s, resulting in 15 measurements in total.

In this study, the protocols for *in vivo* imaging of human hands and feet (IRB00214149) and *ex vivo* imaging of esophagetomy samples (IRB00144178) were approved by Johns Hopkins Institutional Review Board. The *in vivo* imaging of the pig colon (SW18A164) was approved by Johns Hopkins Animal Care and Use Committee.

2.6 Performance Evaluation

In this study, we benchmarked OxyGAN by comparing it with SSOP. We additionally compared OxyGAN with an approach using GANPOP networks to first predict optical properties at 659 and 851 nm and to subsequently fit the concentrations of oxygenated and de-oxygenated hemoglobin using the Beer–Lambert law. These GANPOP networks were trained on the same dataset as OxyGAN with cross validation. The performance of all three methods was evaluated using normalized mean absolute error (NMAE), which is equivalent to absolute percentage error:

$$\text{NMAE} = \frac{\sum_{i=1}^{N} |StO_{2_i} - StO_{2_{i,GT}}|}{\sum_{i=1}^{N} StO_{2_{i,GT}}}. \tag{6}$$

StO_{2_i} and $StO_{2_{i,GT}}$ are predicted and SFDI ground-truth oxygen saturation, respectively. N is the total number of pixels. All testing datasets were manually masked to include only pixels that sampled the object.

3 Results

The average NMAEs are reported in Table 1 for SSOP, GANPOP, and OxyGAN tested on human feet, hands, and *in vivo* pig colon. The hands and feet are from different healthy volunteers with a wide range of pigmentation levels (Fitzpatrick skin types 1 to 5). It is worth emphasizing that the *in vivo* hands and pig colon were completely new tissue types that were not represented in the training set. On average, OxyGAN outperforms SSOP and GANPOP in accuracy by 24.04% and 6.88%, respectively, compared with uncorrected SFDI ground truth. Compared with profile-corrected ground truth, the improvement of OxyGAN over SSOP and GANPOP becomes 24.89% and 24.76%, respectively.

Figure 4 compares the results of profile-corrected SFDI, SSOP, and OxyGAN applied to a sample of each testing tissue type. Lower errors and fewer image artifacts are observed in OxyGAN results. Error plots highlight the fringe artifacts commonly observed parallel to the illumination patterns in SSOP. As expected, both SSOP and OxyGAN exhibited higher errors in the pig colon, which had more complex surface topography and made single-snapshot predictions more difficult.

We additionally implemented OxyGAN on a video of a volunteer's hand during an occlusion study (Fig. 5). The average oxygen saturation was calculated for a region of interest highlighted by the red box in Fig. 5(c) and was compared with the SFDI ground truth in Fig. 5(d). OxyGAN accurately measures a large range of oxygenation values and shows strong and stable agreement with conventional SFDI.

4 Discussion

In this study, we have described a fast and accurate technique for estimating wide-field tissue oxygenation from single-snapshot structured-illumination images using GANs. As shown in Table 1, OxyGAN accurately measures oxygenation not only for sample types represented

Table 1 NMAE of StO_2 predicted by SSOP, GANPOP, and OxyGAN on both uncorrected and profile-corrected SFDI ground truth.

	Uncorrected SFDI				Profile-corrected SFDI			
	Feet	Hands	Pig	Overall	Feet	Hands	Pig	Overall
SSOP	0.0396	0.0430	0.1508	*0.0778*	0.0601	0.0672	0.1404	*0.0892*
GANPOP	0.0438	0.0381	0.1085	*0.0635*	0.0795	0.0655	0.1220	*0.0890*
OxyGAN	**0.0358**	**0.0335**	**0.1080**	***0.0591***	**0.0536**	**0.0466**	**0.1007**	***0.0670***

Note: Bold values indicate best performance, and italic values indicate overall performance as the weighted average of the hands, feet, and pig samples.
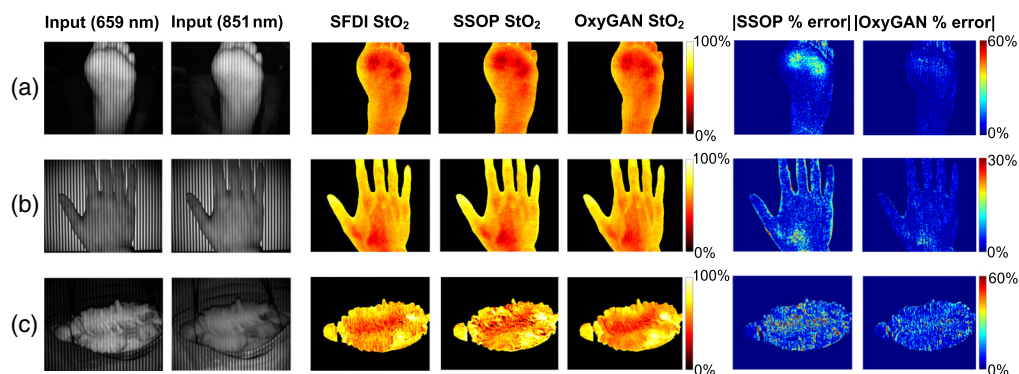

Fig. 4 Comparison of profile-corrected SFDI, SSOP, and OxyGAN StO_2 results: (a) *in vivo* human foot, (b) *in vivo* human hand, and (c) *in vivo* pig colon.
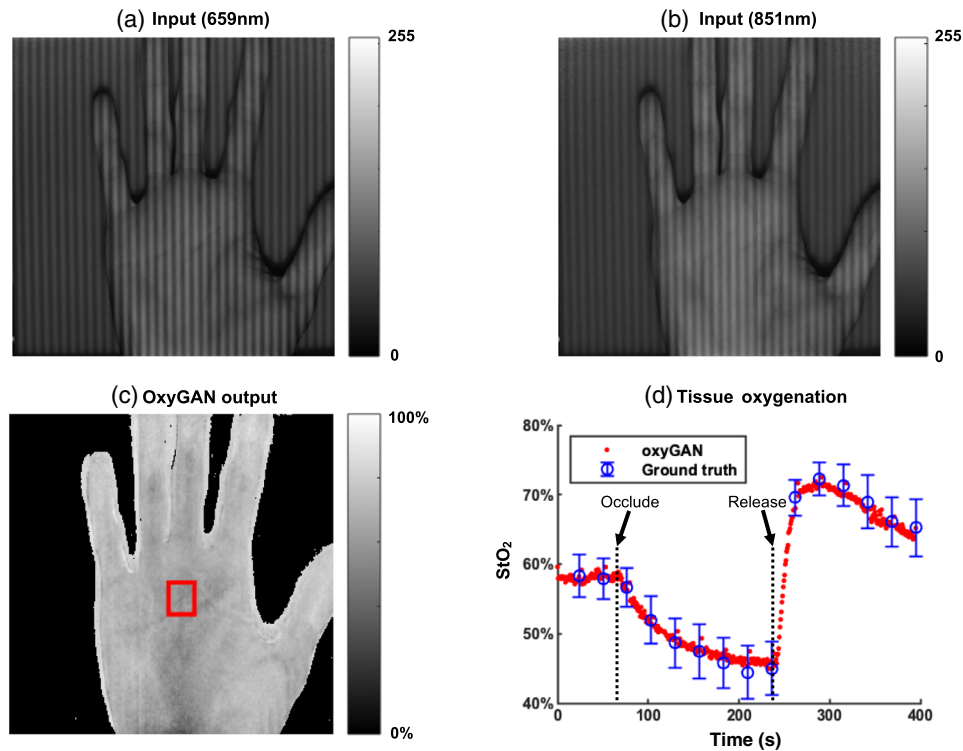
Fig. 5 Video of the occlusion-release experiment: (a) input at 659 nm, (b) input at 851 nm, (c) StO$_2$ map predicted by OxyGAN, and (d) StO$_2$ trend measured by OxyGAN and ground truth SFDI over time (Video 1, MP4, 3.7 MB [URL: https://doi.org/10.1117/1.JBO.25.11.112907.1]).

in the training set (human feet) but also for unseen sample types (human hands and pig colon). This supports the possibility that OxyGAN can be robust and generalizable. The occlusion video (Fig. 5) further demonstrates the ability of OxyGAN to accurately measure a wide range of tissue oxygenation levels and to detect changes over time.

Compared with training separate GANPOP networks to first estimate absorption coefficients, OxyGAN produces an average improved accuracy of 15.8%. Moreover, a greater improvement is observed in profile-corrected experiments. One potential explanation for this is that the errors in absorption coefficients due to uncertainties in profilometry estimation propagate and result in a larger error in oxygenation measurements. Additionally, compared with separate GANPOP models, the end-to-end OxyGAN approach requires only one network and bypasses the Beer–Lambert fitting step, thus greatly reducing the computational cost for training and inference. For example, training a network on 350 patches took ∼2.2 h or 40 s per epoch on an NVIDIA Tesla P100 GPU. Training separate GANPOP networks would take double the amount of time and memory. To achieve real-time StO$_2$ mapping, we first converted the trained model to the Open Neural Network Exchange (ONNX) format. We then imported the ONNX model into NVIDIA TensorRT 7 for reduced latency and optimized inference. For testing, OxyGAN inference on a Tesla P100 took ∼0.04 s to generate a 512×512 oxygenation map. This is 8 times faster than computing optical properties with two GANPOP networks and ∼10 times faster than two GANPOP inferences followed by a Beer–Lambert fitting step. We expect OxyGAN to process 1024×1024 images at a similar framerate (25 Hz) on a quad-GPU workstation.

To evaluate model performance, we benchmarked OxyGAN by comparing it with a single-snapshot technique based on a physical model (SSOP). Table 1 shows that, in estimating both uncorrected and profile-corrected oxygenation, OxyGAN achieves higher accuracy than SSOP in all tissue categories. In addition to improved average accuracy, OxyGAN results also contain fewer subjective image artifacts (Fig. 4). These benefits are more pronounced for samples with complex surface topography, such as the pig gastrointestinal sample. Unlike SSOP, which relies on Fourier domain filtering, OxyGAN utilizes both local and high-level features. As a

content-aware, data-driven approach, OxyGAN has the potential to learn the underlying distribution of the data and accurately infer oxygenation in regions with low signal or non-uniform surface structures.

In Table 1, we observe that GANPOP achieves similar accuracy to SSOP for profile-corrected ground truth. This is expected for several reasons. First, the training set used in this study is smaller than in the original GANPOP paper,[28] excluding *in vivo* hands and tissue-mimicking phantoms. Second, for physical model-based techniques, such as SSOP, the optical property errors due to surface topography variation are correlated across wavelengths and can later be reduced by chromophore fitting. For instance, for surface normal vectors pointing further away from the detector, the predicted absorption coefficients will be overestimated for both 659 and 851 nm. However, the fitting of hemoglobin concentrations oxygen saturation, which relies on the ratios of absorption coefficients, may mask the intermediate optical property errors. Because the GANPOP networks are trained independently for 659 and 851 nm, the loss function does not learn these correlations, resulting in smaller improvements in accuracy over SSOP for StO_2 measurements than for optical property measurements. This observation also provides some intuition for why the OxyGAN network might improve accuracy over GANPOP. Because OxyGAN is trained on multi-wavelength input and the loss function is computed from the StO_2 estimate, it is capable of modeling correlations between absorption at different wavelengths and learning to reduce the effects of varying surface topography. Furthermore, a higher error rate is observed for the pig colon sample for both GANPOP and OxyGAN, likely due to this tissue type not being included in the training set and the complex topography of the colon specimen compared with other training samples.

The architecture of OxyGAN is based on the GANPOP framework.[28] The generator combines the features of both the U-Net and the ResNet, in that it incorporates both short and long skip connections and is fully residual. As discussed in Ref. 28, this fusion generator has advantages over most other existing architectures because it allows information flow both within and between levels, which is important for the task of optical property prediction. In this study, we empirically trained a model with a standard U-Net generator. The model performed well on sample types included in the training set; however, it collapsed and was unable to produce accurate results when tested on unseen sample types, such as human hands. Compared with GANPOP, OxyGAN employs data augmentation in the form of horizontal and vertical flipping, which is important for preventing overfitting of models trained on small datasets. OxyGAN also utilizes label smoothing in training the discriminator, which further improves model performance and overall training stability. Finally, we found that adding a channel of checkerboard reference phantom measurements to the 2-wavelength structured-light inputs improves accuracy for measurements taken on different days, allowing OxyGAN to take system drift into account similarly to conventional SFDI.

In the future, more work could be done to optimize the algorithm of OxyGAN to further improve the data processing speed. The model could be trained and tested on larger datasets that span a wider range of tissue types or scenarios that might be encountered clinically. To develop a robust and generalizable model, future work should train on data with a range of spatial frequencies acquired on several different instruments. Domain adaptation techniques could also be implemented on the trained models to improve robustness to different imaging geometries. In addition, similar to other single-snapshot techniques, one limit of OxyGAN is its expected sensitivity to ambient light. Moreover, oxygenation mapping using SFDI structured illumination is currently limited because it has a shallow depth of field and requires precisely controlled imaging geometry, making its clinical adoption particularly challenging. One alternative is to use random laser speckle patterns as structured illumination, which could be less costly than SFDI projection systems and more easily incorporated into endoscopic applications and may avoid fringe artifacts due to sinusoidal illumination.[42] Monocular depth estimation could also be incorporated for profile-correction without requiring a projector and profilometry.[43,44] Furthermore, a more sophisticated LUT could be developed to directly estimate StO_2 from SSOP data, which models the correlations between reflectance measurements at different wavelengths and the underlying tissue oxygenation. This pixel-wise estimation may provide a more accurate baseline that will help quantify the benefit of the content-aware aspect of OxyGAN. Finally, data-driven methods may be useful for taking higher-order optical property effects, such as the scattering phase function, into account.

25

5 Conclusion

In this study, we have presented an end-to-end approach for wide-field tissue oxygenation mapping from single structured-illumination images using cGANs (OxyGAN). Compared with both uncorrected and profile-corrected SFDI ground truth, OxyGAN achieves a higher accuracy than model-based SSOP. It also demonstrates improved accuracy and faster computation than two GANPOP networks that first estimate optical absorption. This technique has the potential to be incorporated into many clinical applications for real-time, accurate tissue oxygenation measurements over a large field of view.

Disclosures

There are no conflicts of interest related to this article.

Acknowledgments

References

1. H. L. Edmonds, Jr., B. L. Ganzel, and E. H. Austin, "Cerebral oximetry for cardiac and vascular surgery," *Semin. Cardiothorac. Vasc. Anesth.* **8**(2), 147–166 (2004).
2. H. M. Swartz and J. F. Dunn, "Measurements of oxygen in tissues: overview and perspectives on methods," *Adv. Exp. Med. Biol.* **530**, 1–12 (2003).
3. M. Sair et al., "Tissue oxygenation and perfusion in patients with systemic sepsis," *Crit. Care Med.* **29**(7), 1343–1349 (2001).
4. J. Ditzel and E. Standl, "The problem of tissue oxygenation in diabetes mellitus: its relation to the early functional changes in the microcirculation of diabetic subjects," *Acta Med. Scand.* **197**(S578), 49–58 (1975).
5. G. Pitsiou et al., "Tumor necrosis factor—alpha serum levels, weight loss and tissue oxygenation in chronic obstructive pulmonary disease," *Respir. Med.* **96**(8), 594–598 (2002).
6. Y. Mendelson, "Pulse oximetry: theory and applications for noninvasive monitoring," *Clin. Chem.* **38**(9), 1601–1607 (1992).
7. S. Gioux et al., "First-in-human pilot study of a spatial frequency domain oxygenation imaging system," *J. Biomed. Opt.* **16**(8), 086015 (2011).
8. H. W. Siesler et al., *Near-Infrared Spectroscopy: Principles, Instruments, Applications*, John Wiley & Sons (2008).
9. D. A. Benaron et al., "Continuous, noninvasive, and localized microvascular tissue oximetry using visible light spectroscopy," *Anesthesiol.: J. Am. Soc. Anesthesiol.* **100**(6), 1469–1475 (2004).
10. J. M. Murkin and M. Arango, "Near-infrared spectroscopy as an index of brain and tissue oxygenation," *Br. J. Anaesth.* **103**(Suppl. 1), i3–i13 (2009).
11. J. T. Nguyen et al., "A novel pilot study using spatial frequency domain imaging to assess oxygenation of perforator flaps during reconstructive breast surgery," *Ann. Plastic Surg.* **71**(3), 308 (2013).
12. K. J. Zuzak et al., "Visible reflectance hyperspectral imaging: characterization of a noninvasive, in vivo system for determining tissue perfusion," *Anal. Chem.* **74**(9), 2021–2028 (2002).
13. S. Srinivasan et al., "Spectrally constrained chromophore and scattering near-infrared tomography provides quantitative and robust reconstruction," *Appl. Opt.* **44**(10), 1858–1869 (2005).
14. D. J. Cuccia et al., "Quantitation and mapping of tissue optical properties using modulated imaging," *J. Biomed. Opt.* **14**(2), 024012 (2009).

15. J. Q. M. Nguyen et al., "Spatial frequency domain imaging of burn wounds in a preclinical model of graded burn severity," *J. Biomed. Opt.* **18**(6), 066010 (2013).

16. D. M. Burmeister et al., "Utility of spatial frequency domain imaging (SFDI) and laser speckle imaging (LSI) to non-invasively diagnose burn depth in a porcine model," *Burns* **41**(6), 1242–1252 (2015).

17. A. Yafi et al., "Quantitative skin assessment using spatial frequency domain imaging (SFDI) in patients with or at high risk for pressure ulcers," *Lasers Surg. Med.* **49**(9), 827–834 (2017).

18. S. Tabassum et al., "Feasibility of spatial frequency domain imaging (SFDI) for optically characterizing a preclinical oncology model," *Biomed. Opt. Express* **7**(10), 4154–4170 (2016).

19. A. Mazhar et al., "Wavelength optimization for rapid chromophore mapping using spatial frequency domain imaging," *J. Biomed. Opt.* **15**(6), 061716 (2010).

20. M. Schmidt et al., "Real-time, wide-field, and quantitative oxygenation imaging using spatiotemporal modulation of light," *J. Biomed. Opt.* **24**(7), 071610 (2019).

21. J. Angelo et al., "Ultrafast optical property map generation using lookup tables," *J. Biomed. Opt.* **21**(11), 110501 (2016).

22. S. Panigrahi and S. Gioux, "Machine learning approach for rapid and accurate estimation of optical properties using spatial frequency domain imaging," *J. Biomed. Opt.* **24**(7), 071606 (2018).

23. J. Vervandier and S. Gioux, "Single snapshot imaging of optical properties," *Biomed. Opt. Express* **4**(12), 2938–2944 (2013).

24. H.-C. Shin et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imaging* **35**(5), 1285–1298 (2016).

25. D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annu. Rev. Biomed. Eng.* **19**, 221–248 (2017).

26. D. Raví et al., "Deep learning for health informatics," *IEEE J. Biomed. Health Inf.* **21**(1), 4–21 (2016).

27. I. Goodfellow et al., "Generative adversarial nets," in *Adv. Neural Inf. Process. Syst.*, pp. 2672–2680 (2014).

28. M. T. Chen et al., "Ganpop: Generative adversarial network prediction of optical properties from single snapshot wide-field images," *IEEE Trans. Med. Imaging* **39**(6), 1988–1999 (2020).

29. S. Gioux et al., "Three-dimensional surface profile intensity correction for spatially-modulated imaging," *J. Biomed. Opt.* **14**(3), 034045 (2009).

30. Y. Zhao et al., "Angle correction for small animal tumor imaging with spatial frequency domain imaging (SFDI)," *Biomed. Opt. Express* **7**(6), 2373–2384 (2016).

31. E. Aguénounon et al., "Single snapshot of optical properties image quality improvement using anisotropic two-dimensional windows filtering," *J. Biomed. Opt.* **24**(7), 071611 (2019).

32. A. Kienle and M. S. Patterson, "Determination of the optical properties of turbid media from a single Monte Carlo simulation," *Phys. Med. Biol.* **41**(10), 2221 (1996).

33. M. Mirza and S. Osindero, "Conditional generative adversarial nets," arXiv:1411.1784 (2014).

34. P. Isola et al., "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 1125–1134 (2017).

35. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).

36. K. He et al., "Deep residual learning for image recognition," in *IEEE Conf. Comput. Vision and Pattern Recognit.* (2016).

37. T. M. Quan, D. G. C. Hildebrand, and W.-K. Jeong, "FusionNet: a deep fully residual convolutional neural network for image segmentation in connectomics," arXiv:1612.05360 (2016).

38. T. Miyato et al., "Spectral normalization for generative adversarial networks," arXiv:1802.05957 (2018).

39. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," arXiv:1412.6980 (2014).
40. T. Salimans et al., "Improved techniques for training gans," in *Adv. Neural Inf. Process. Syst.*, pp. 2234–2242 (2016).
41. J. A. Sweer et al., "Wide-field optical property mapping and structured light imaging of the esophagus with spatial frequency domain imaging," *J. Biophotonics* **12**(9), e201900005 (2019).
42. M. T. Chen, M. Papadakis, and N. J. Durr, "Speckle illumination spatial frequency domain imaging for projector-free optical property mapping," arXiv:2006.03661 (2020).
43. F. Mahmood and N. J. Durr, "Deep learning and conditional random fields-based depth estimation and topographical reconstruction from conventional endoscopy," *Med. Image Anal.* **48**, 230–243 (2018).
44. R. Chen et al., "Rethinking monocular depth estimation with adversarial training," arXiv: 1808.07528 (2018).

Mason T. Chen is a PhD student of biomedical engineering at Johns Hopkins School of Medicine. He received his BE degree in biomedical engineering from Vanderbilt University in 2017. His current research interests include optical imaging, machine learning, and biophotonics.

Nicholas J. Durr is an assistant professor of biomedical engineering at Johns Hopkins University. He received his BS degree in electrical engineering and computer science from UC Berkeley in 2003 and MS and PhD degrees in biomedical engineering from UT Austin in 2007 and 2010, respectively. His current research interests include medical devices, machine learning, and biophotonics. He is a member of SPIE.