

# TIE-GANs: single-shot quantitative phase imaging using transport of intensity equation with integration of GANs

Vikas Thapa, Ashwini Subhash Galande<sup>ORCID</sup>, Gurram Hanu Phani Ram,  
and Renu John<sup>ORCID\*</sup>

Indian Institute of Technology Hyderabad, Medical Optics and Sensors Laboratory, Department of Biomedical Engineering, Hyderabad, Telangana, India

**ABSTRACT.** **Significance:** Artificial intelligence (AI) has become a prominent technology in computational imaging over the past decade. The expeditious and label-free characteristics of quantitative phase imaging (QPI) render it a promising contender for AI investigation. Though interferometric methodologies exhibit potential efficacy, their implementation involves complex experimental platforms and computationally intensive reconstruction procedures. Hence, non-interferometric methods, such as transport of intensity equation (TIE), are preferred over interferometric methods.

**Aim:** TIE method, despite its effectiveness, is tedious as it requires the acquisition of many images at varying defocus planes. The proposed methodology holds the ability to generate a phase image utilizing a single intensity image using generative adversarial networks (GANs). We present a method called TIE-GANs to overcome the multi-shot scheme of conventional TIE.

**Approach:** The present investigation employs the TIE as a QPI methodology, which necessitates reduced experimental and computational efforts. TIE is being used for the dataset preparation as well. The proposed method captures images from different defocus planes for training. Our approach uses an image-to-image translation technique to produce phase maps and is based on GANs. The main contribution of this work is the introduction of GANs with TIE (TIE:GANs) that can give better phase reconstruction results with shorter computation times. This is the first time the GANs is proposed for TIE phase retrieval.

**Results:** The characterization of the system was carried out with microbeads of 4  $\mu\text{m}$  size and structural similarity index (SSIM) for microbeads was found to be 0.98. We demonstrated the application of the proposed method with oral cells, which yielded a maximum SSIM value of 0.95. The key characteristics include mean squared error and peak-signal-to-noise ratio values of 140 and 26.42 dB for oral cells and 100 and 28.10 dB for microbeads.

**Conclusions:** The proposed methodology holds the ability to generate a phase image utilizing a single intensity image. Our method is feasible for digital cytology because of its reported high value of SSIM. Our approach can handle defocused images in such a way that it can take intensity image from any defocus plane within the provided range and able to generate phase map.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.29.1.016010](https://doi.org/10.1117/1.JBO.29.1.016010)]

**Keywords:** quantitative phase imaging; transport of intensity equation; generative adversarial networks; phase map; deep learning

Paper 230268GR received Sep. 12, 2023; revised Nov. 18, 2023; accepted Jan. 9, 2024; published Jan. 30, 2024.

\*Address all correspondence to Renu John, [renujohn@bme.iith.ac.in](mailto:renujohn@bme.iith.ac.in)

## 1 Introduction

Cellular analysis of biological cells is one of the most popular tools for early diagnosis of human diseases. Cytological investigations unveil the latent topographical characteristics of cells, thereby aiding clinicians in making crucial diagnostic decisions concerning diseases.<sup>1</sup> Cytology became more informative and useful as biology and microscopic methods advanced in conjunction. The invention of the microscope<sup>2</sup> in 1600 had a profound impact on the scientific community. Conventional brightfield microscopy techniques have the capability to visualize amplitude objects, as they solely exhibit intensity. Transparent objects, such as biological cells and microorganisms, present a challenge for visualization under a brightfield microscope due to their transparent nature.

Staining the cells is considered one of the potential solutions to address this issue. But staining is not a recommended option because it can cause morphological or chemical alterations in the cells. The examination of cells utilizing qualitative feature maps<sup>3</sup> does not disclose the morphological characteristics of the specimen. Consequently, the utilization of solely intensity-based information does not allow for accurate characterization based on morphology. The label-free high contrast optical imaging of living cells was made possible by Zernike's invention of the phase contrast microscope (PCM) in 1930.<sup>4</sup> The Normarski differential interference contrast (DIC) microscope produces phase maps in a similar manner.<sup>5,6</sup> Phase contrast techniques, such as PCM and DIC, produce good contrast images but do not reveal thickness information. Quantitative phase imaging (QPI) is a label-free methodology that measures the phase delay caused by the sample,<sup>7-9</sup> thereby enhancing the precision of disease diagnosis. The utilization of QPI live cell imaging has allowed researchers to prevent the expenses associated with laborious staining procedures. With subwavelength accuracies, it provides the optical thickness profile of a transparent object, which is vital information for cell biology and meteorology. Although there are several benefits associated with QPI, it continues to be a laborious commitment.

The integration of QPI and deep learning (DL) has demonstrated significant potential in providing diverse applications in recent times. While QPI has the potential to address issues pertaining to precise thickness mapping of biological samples DL is better suited for managing extensive datasets and automating processes.<sup>10,11</sup> DL techniques aim to replicate the functionality of biological neurons in the human brain.<sup>12</sup>

Fast phase retrieval using only a single intensity image was initially reported by Rivenson et al.,<sup>13</sup> demonstrating the power of DL. This method not only solved the problem of reconstruction, but it can generate phase and amplitude images, using only single hologram intensity. However, one needs to be specifically trained to reconstruct the phase and amplitude to eliminate subjective. Uncertainty maps describe shortcomings, such as noise, model error, missing training data, and out-of-sample testing data, that are common in practice but rarely accounted for. DL predictions are not checked for reliability and hence mistakes are often identified only after the results are achieved. Xue et al.<sup>14</sup> overcome this problem. Lensless imaging is novel in terms of its low cost, portability, and fast retrieval times. A lensless imaging system to recover phase from a single intensity diffraction pattern was proposed by Sinha et al.<sup>15</sup> as a solution to the inverse computational imaging issue known as "phase retrieval."<sup>16,17</sup> Some of the techniques employ holography with convolutional neural network (CNN) and generative adversarial networks (GANs) to quickly retrieve the phase and amplitude from a single hologram image.<sup>18-21</sup> Ptychography<sup>22-24</sup> is one of the popular reconstruction techniques for phase recovery. Kaiqiang et al. discussed the TIE with deep learning (termed as dTIE).<sup>25</sup> The method shows DL ability to produce phase map from single intensity image. The reported average structural similarity index (SSIM) for the proposed method is 0.95. Wu et al. discussed the fundamental principles of phase imaging with the DL-based technique to create a model-based phase retrieval framework.<sup>26</sup> The low-rank total variation based regularization approach reports a peak-signal-to-noise ratio (PSNR) of 11.63 and an SSIM of 0.92, whereas the model-based network regularization method reports PSNRs of 23.70 and SSIMs of 0.90. Xiaofeng et al. created a cascaded deep neural network using the forward and inverse physics models to create a physics-informed neural network (PINN).<sup>27</sup> PINN method reports the SSIM value of 0.91 for human buccal epithelial cells 25.23 dB in PSNR value. The extended use of CNN and U-Net to achieve large space

bandwidth and high-resolution imaging is possible with ptychography method. The integration of CNN with lensless imaging shows promising results and has clinical applications in real world.<sup>28,29</sup> One of the modalities for phase imaging is tomography, which shows high resolution and better performance for voxel reconstruction<sup>30,31</sup> using deep convolutional neural network. Transport of intensity equation (TIE) is a non-interferometric QPI technique that enables the reconstruction of phase information from defocused intensity measurements in a regular microscope.<sup>32</sup> Though the TIE is elegant and simple, the need for multiple defocus measurements and a precise phase characterization prior to use made it unpopular in practical real time use.

In this work, we introduce an efficient method known as TIE-GANs to overcome multi-shot scheme of conventional TIE. The proposed approach is advantageous as it can be implemented in real time as a single shot technique. Our approach is based on the mapping relationship between intensity and phase image. The use of GANs<sup>33</sup> in the field of computer vision is a groundbreaking work with applications in various operations, such as image-to-image translation,<sup>34</sup> text-to-image translation,<sup>35</sup> and semantic image to photo translation.<sup>36</sup> Image-to-image translation application is exploited in demonstrating virtual phase staining of the images.<sup>37</sup> This approach makes use of image-to-image translation using GANs modality.<sup>38</sup> The image-to-image translation method finds the relationship between input and output images. It is unique in terms of loss learning because this approach not just learns image translation from input to output map but also learns a loss function simultaneously. Due to the time and effort required for execution, the network no longer needs to rely on hand-engineered loss functions. For dataset preparation, TIE is being used for creating phase and intensity image pair.<sup>39</sup> In optical and electron microscopy, the TIE is a computational method for reconstructing the phase of a complex wave.<sup>40</sup> One of the main bottlenecks for DL algorithms is that it requires a large dataset, i.e., images for proper training. Hence, during the experiments our focus was to generate an appropriate number of images for training the neural network. In this study, we used oral buccal cells for training the GANs-based model. There are two sets of image pairs generated using TIE microscope: the first set consists of 600 and the second one consists of 700 images consecutively. Out of these image stacks, 100 images are reserved for testing from each set while the remaining images are used for training the network. We use SSIM, mean squared error (MSE), PSNR, and universal image quality index (UIQI) as the main image quality assessment parameters to evaluate the translated and reconstructed images.<sup>41</sup>

## 2 Principle

### 2.1 Transport of Intensity Equation

In the recent past, several QPI imaging techniques have been developed for biomedical applications.<sup>42</sup> In QPI techniques based on interferometric principles, one of the main challenges is to address the coherence induces disturbances<sup>43</sup> that can hinder the quality and accuracy of phase reconstructions. Traditional intensity-based imaging methods in QPI capture only the object's intensity information, whereas the phase information, which is related to the object's refractive index variations, is lost.<sup>44</sup> The TIE method, which is a non-interferometric QPI technique, enables the phase information to be recovered from intensity measurements.<sup>45-47</sup>

TIE is based on Poisson equation, which takes into consideration the Laplacian, of the phase of wave and its variation along the optical axis.<sup>48</sup> The TIE makes use of a well-focused image and two defocused images to extract the complex wave information of the object. A plane wave in terms of amplitude and phase term can be represented as

$$u(x, y) = \sqrt{I(x, y)} \exp(i\varphi(x, y)), \quad (1)$$

where  $I$  is the captured intensity and  $\varphi(x, y)$  is the spatial distribution of the phase. Paraxial approximation of the propagation gives the TIE,<sup>49</sup> as

$$\frac{\partial I(x, y)}{\partial z} = -\frac{\lambda}{2\pi} \nabla \cdot (I \nabla \varphi(x, y)), \quad (2)$$

where  $\lambda$  represents the spectrally weighted mean wavelength. By slightly defocusing the image in both the positive and negative  $z$ -direction  $z$  derivative can be obtained. Phase retrieval can be calculated,<sup>50</sup> using Eq. (2) to compute the inverse Laplacian.

The  $\varphi(x, y)$  can be obtained as

$$\varphi(x, y) = \nabla^{-2} \left[ \nabla \cdot \left[ \frac{1}{I(x, y)} \nabla \left[ \nabla^{-2} \left( -\frac{2\pi}{\lambda} \frac{\partial I(x, y)}{\partial z} \right) \right] \right] \right], \quad (3)$$

where  $I(x, y)$  represents the infocus image intensity and the  $\nabla^{-2}$  is the inverse Laplacian operator.

Finally, to obtain the quantitative measurement of the phase in terms of the thickness of the object  $h(x, y)$  can be obtained as

$$h(x, y) = \frac{\lambda}{2\pi\Delta n} \Delta\varphi, \quad (4)$$

where  $\Delta n$  represents the refractive index difference.

Our approach utilizes the non-interferometric property of TIE to record multiple image stack for training our model.<sup>51</sup> TIE methodology is based on multiple recordings at different defocus planes, hence while reconstructing the phase image one must consider one infocus and two defocus images.<sup>52</sup> Finding infocus image plane could be another difficult task if one must decide manually from a stack of images. We considered Tamura coefficient to calculate the best infocus image so that the reconstructed image has the highest possible SNR.<sup>53</sup> Training with the wide spectrum of defocus images leads to a robust network and therefore, this modality could reconstruct the object information at larger defocus planes. Our method incorporates about  $\pm 60 \mu\text{m}$  defocus distance hence, a well defocused image can also produce a reliable reconstructed phase image.

## 2.2 Generative Adversarial Networks

DL can extract features from multimodal dataset and recognizes the patterns.<sup>54</sup> This property makes it suitable for various tasks, such as classification,<sup>55</sup> cancer detection,<sup>56</sup> and segmentation.<sup>57</sup> A DL model is composed of multiple hidden layers, at every layer input is the linear combination of previous layer output, which then passes through a nonlinear function on weighted sums.<sup>58</sup>

Output of  $k$ 'th node in layer  $n + 1$  is denoted as

$$Z_k^{(n+1)} = h(a_k^{(n+1)}) = h\left(\sum_{i=0}^{N_n} w_{ki}^{(n)} x_i^{(n)}\right) \quad k = 0, \dots, K; \quad n = 1, \dots, N_{n+1}, \quad (5)$$

which is derived from the input  $x_1^n, x_2^n, x_3^n, \dots, x_N^n$ .  $a_k^{(n+1)}$  is the linear combination of inputs, and  $w_{ki}^{(n)}$  are weights associated with these inputs these are also known as convolutional maps.<sup>59</sup> Number of nodes in the layer  $l$  to  $L$  represents the number of hidden layers; usually a single hidden layer represents a shallow and more than three layers is known as the deep neural network.<sup>60</sup> The nonlinear activation function choices are sigmoid function  $h(a) = \frac{1}{(1+\exp(-a))}$ , RELU  $h(a) = \max(0, a)$ , and hyperbolic tangent  $\tanh(a)$  function.<sup>61</sup>

GANs work on a principle where a model can generate data in an adversarial training setup.<sup>62</sup> GANs are different from conventional discriminative model where an image is given as  $X$  and model try to predict the label  $Y \rightarrow P(Y|X)$ . The main disadvantage associated with the discriminative model is that they cannot model  $P(X)$ , i.e., they are not capable of modeling input images. Hence, these models cannot decide the probability map of input images therefore unable to generate new images. GANs are unique in terms of learning loss, which classify the output images as real or fake and simultaneously train a generative model to minimize this loss. Our approach is focused on a modality know as conditional generative adversarial networks (cGANs). This method is capable of converting an image from one domain to another using conventional GANs in a conditional setting. The main advantage of this technique is that it is best suited for image-to-image translation task in which an input image is conditioned and generates corresponding output image.<sup>63,64</sup> The training procedure associated with the GANs is shown in Fig. 1.

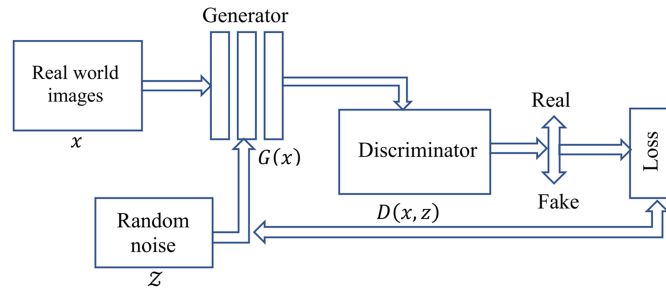


Fig. 1 GANs training map.

Mainly a GAN consists of a generator  $G$  and a discriminator  $D$ , GAN works as a minmax game theory where discriminator tries to maximize its reward and generator tries to maximize discriminator loss.

GANs learn to convert a random noise vector  $z$  to output image map  $y$ ,  $G : z \rightarrow y$ . In the case of cGANs, it learns mapping from observed images  $x$  and random noise vector  $z$ ,  $z \rightarrow y$ ,  $G : \{x, z\} \rightarrow y$ .

The objective function of conditional GANs is given as

$$\mathcal{L}_{\text{cGAN}}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x, z)))], \quad (6)$$

$G$  minimizes the objective function against an adversarial  $D$ , which is trying to maximize the cost function.

It can be represented as

$$G^* = \arg \min_G \max_D \mathcal{L}_{\text{cGAN}}(G, D). \quad (7)$$

The detailed description of this network and method is shown in Figs. 5(a) and 5(b).

### 3 Materials and Methods

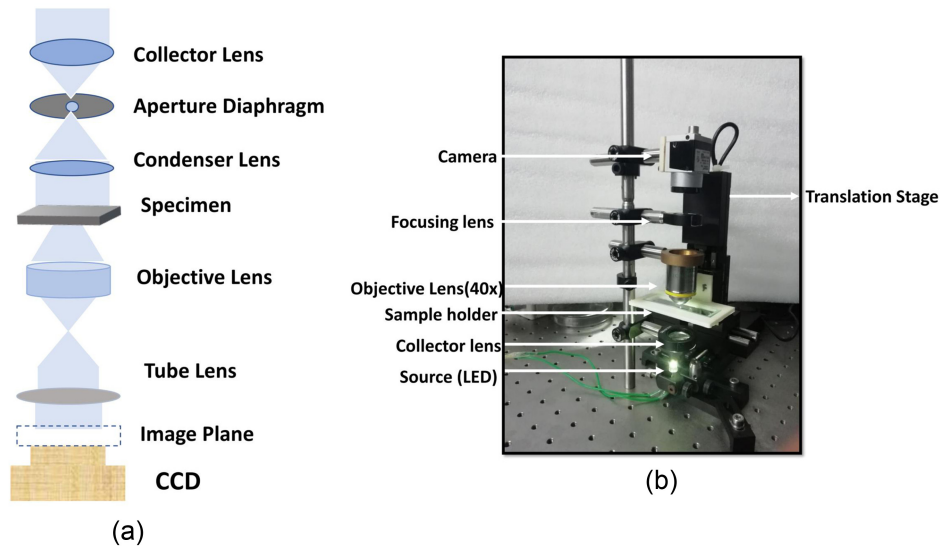
#### 3.1 Experimental Set-Up

Oral smears were produced from the buccal mucosa of six healthy 25- to 30-year-old volunteers using a sterile wooden spatula on a clean sterile glass slide. Before taking samples, all of the respondents were asked to gargle with standard saline. After fixing the cell with alcohol, samples were put right on a microscope slide so that images can be recorded. No staining procedures were followed in this process. The oral cells were illuminated using a partially coherent LED light source having wavelength of 627 nm. The camera used for the experiment is (Basler acA2440-75um USB 3.0) with (Sony IMX250 CMOS) sensor of 5 MP resolution operating at a frame rate of 75 fps, with  $3.45 \mu\text{m} \times 3.45 \mu\text{m}$  pixel size. A microscope objective with 40 $\times$  magnification and a numerical aperture (NA) of 0.65 is used in the microscope. An automated translation stage (Thorlabs MTS50-Z8) having a travel range of 50 mm with 0.05  $\mu\text{m}$  minimum achievable incremental movement or step size is used for scanning the focus in  $z$  direction. The schematic of the experimental set-up for TIE is shown in Fig. 2(a), and the optical setup for the experiment is shown in Fig. 2(b).

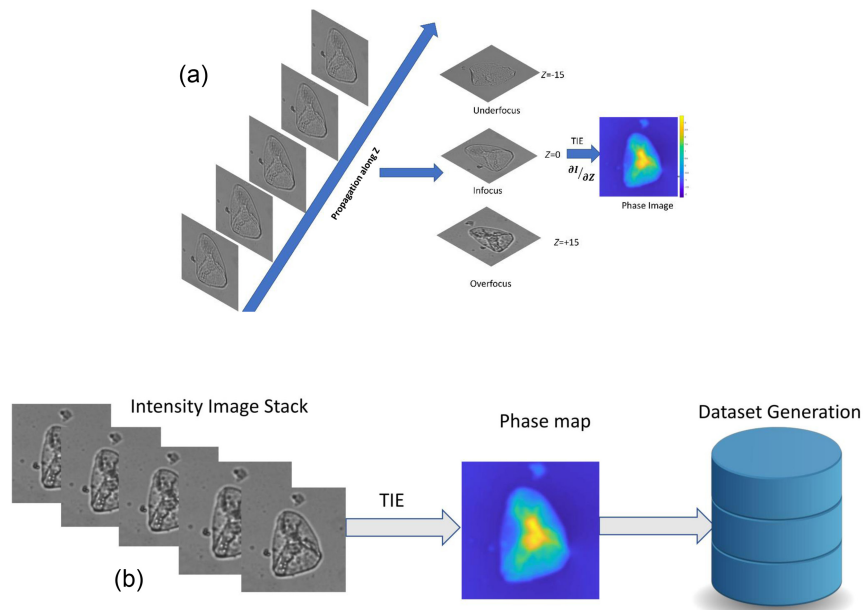
The stack of images was recorded with a defocus distance up to  $\pm 60 \mu\text{m}$  and a step size of 15  $\mu\text{m}$  along the  $Z$  axis, for creating the training dataset images. This image stack data are further used to train GANs. The captured intensity image is of size  $1388 \times 1040$  pixels from which a region of interest (ROI) containing each individual cell size of  $256 \times 256$  pixels was extracted out. Reconstructions were performed on a computer with an Intel Core i9 9820X (10Cores 20Threads up to 4.2 GHz) CPU, 64 GB of RAM, and an NVIDIA RTX 2080Ti Dual GPU setup. The training took between 1 and 1.5 h, based on how many sets of images were used.

#### 3.2 Reconstruction Algorithm of TIE

First the imaging stack was created using the TIE microscope as TIE establishes the connection between intensity variation and the phase gradient.<sup>65</sup> We use the TIE for extraction of the phase from the intensity variation.<sup>66</sup> Figure 3(a) shows the TIE phase retrieval block diagram.



**Fig. 2** (a) Schematic block diagram of TIE experiment. (b) The optical setup for the experiment.



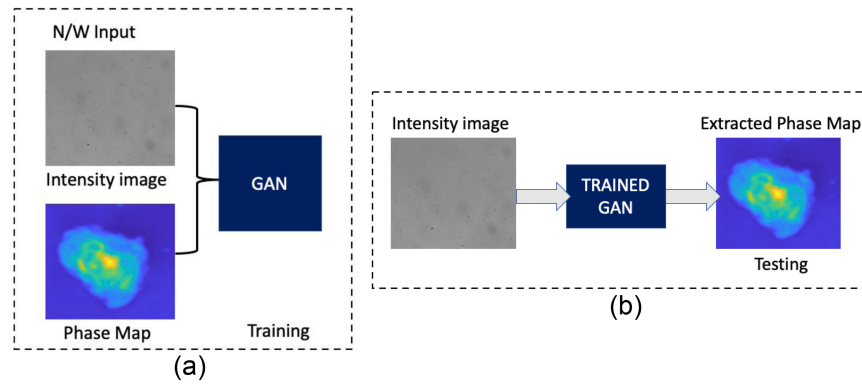
**Fig. 3** (a) Block diagram of phase extraction using TIE. (b) Dataset generation of phase map using TIE data.

After phase extraction, the dataset is created using pairs of intensity and phase images, which is given for the training. The methodology diagram for the dataset generation is presented in Fig. 3(b).

### 3.3 TIE-GANs

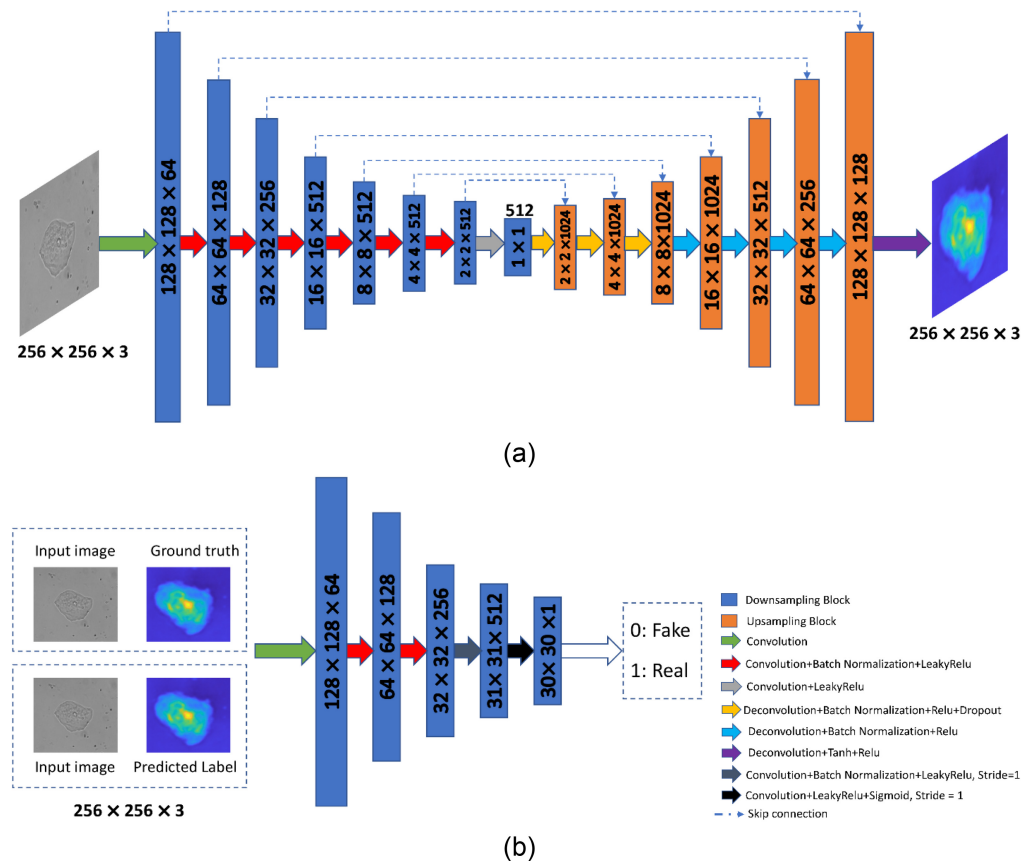
Following dataset development, GANs are used to train image pairs consisting of an intensity and phase map. The network's output is a phase map that corresponds to the input image's intensity after proper training. Our approach incorporates data from a variety of focal lengths and positions along the optical axis in  $z$  direction to strengthen the resulting model. The training and testing procedure for the neural network is shown in Figs. 4(a) and 4(b).

In our approach, the TIE-GANs are based on cGANs architecture consisting of two networks: generator and discriminator. The generator is inspired from U-Net architecture, which



**Fig. 4** (a) Training methodology of TIE-GANs, (b) the testing methodology and network output as phase image.

is widely accepted in biomedical image segmentation.<sup>67</sup> The network proposed here uses the encoder–decoder scheme at the generator side with skip connections.<sup>64</sup> The input image is down-sampled using eight convolutional layers in the encoder part followed by eight upsampling layers in the decoder part.<sup>68</sup> The activation function used after every encoder node is Leaky ReLU.<sup>69</sup> In order to avoid the overfitting during training of the neural network, we have implemented the dropout at the decoder part. Batch normalization has been used after each layer in order to make the input to each network standardized. The schematic of the network architecture used for generator is shown in Fig. 5(a). Discriminator network follows the PatchGAN structure where the input to the network is a pair of real set and generated set of images. The primary goal of this network is to determine whether or not the produced image is genuine. The discriminator takes



**Fig. 5** (a) Generator network to generate images from random noise vectors. (b) Discriminator network.

on a new function by instructing the generating network on how similar the data should be to real-world images. The picture pair is encoded into a feature vector by the discriminator's five convolutional layers. After sigmoid activation functions are applied to the last layer of the network, a binary output of 0 or 1 is produced, representing an actual picture pair. The schematic of the discriminator network is shown in Fig. 5(b).

### 3.4 Accuracy Assessment of TIE versus TIE-GANs

The SSIM, MSE, PSNR, and the UIQI are used to measure the quality of an image with respect to ground truth images.

The MSE is the first parameter to be considered because it is widely used and its calculation is easy and inexpensive. One of the primary characteristics of the MSE is its independence, which makes it memoryless and therefore capable of being evaluated on each sample. The MSE's physical significance is that it is a measure of signal fidelity in terms of the error signal's energy.

The MSE was calculated between translated, i.e., GANs output image and ground truth TIE image, if  $x = \{x_i|1,2, \dots, N\}$  and  $y = \{y_i|1,2, \dots, N\}$  two signals i.e., images the MSE can be calculated as<sup>70</sup>

$$\text{MSE}(x, y) = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2, \quad (8)$$

where  $N$  represents the number of image samples.

The PSNR is one of the most used objective methods for quality assessment metric for translated and ground truth image.<sup>71</sup> High PSNR value implies a high-quality translated image as the MSE between translated and ground truth image becomes minimum.

The further extension to the MSE leads to the PSNR calculation, which can be represented as

$$\text{PSNR} = 10 \log_{10} \frac{L^2}{\text{MSE}}, \quad (9)$$

where  $L^2$  is the dynamic range of the image.

The UIQI is applied since the performance of MSE and PSNR does not incorporate luminance as well as contrast factor of the images. UIQI is a unique parameter that gives the image distortion in terms of three factors: loss of correlation, luminance distortion, and contrast distortion.<sup>72</sup>

The UIQI is the measure of three parameters so it can be expressed as

$$\text{UIQI} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\mu_x \mu_y}{(\mu_x)^2 + (\mu_y)^2} \cdot \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}. \quad (10)$$

The first term represents the correlation coefficient between image  $X$  and  $Y$ , the dynamic range of this factor is  $[-1,1]$ .

The second term represents how close the luminance is in between both the images, the value of this factor ranges between  $[0,1]$ . The third term represents the contrast and gives the similarity factor between both input and translated images, the value of this factor ranges between  $[0,1]$ . Hence by combining all these parameters, the UIQI can be expressed as

$$\text{UIQI} = \frac{4\sigma_{xy}\mu_x\mu_y}{(\sigma_x^2 + \sigma_y^2)[(\mu_x)^2 + (\mu_y)^2]}. \quad (11)$$

Above all the quality assessment metrics performs well but do not consider the human visual system (HVS). The HVS perception is based on the structural information of the scene, and evaluation based on structural loss is vital in terms of quality assessment. The SSIM index is a well-known parameter to compare the structural similarity between two reconstructed and ground truth images. It gives the quality measurement in terms of luminance, contrast, and structure, SSIM is a widely used and accepted parameter for quality assessment.<sup>73,74</sup>



The SSIM between two images  $X, Y$  can be calculated as

$$\text{SSIM}(X, Y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}. \quad (12)$$

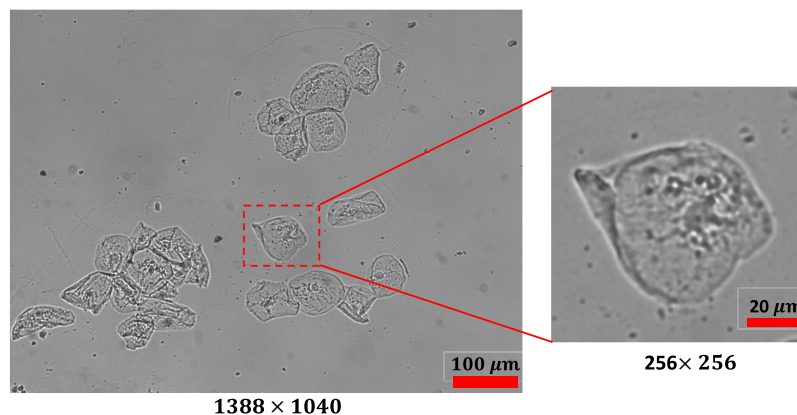
For UIQI and SSIM,  $\mu_x$  and  $\mu_y$  the average image intensity of images  $X$  and  $Y$ , respectively,  $\sigma_x^2$  and  $\sigma_y^2$  are the variance of both the images.  $\sigma_{xy}$  is the cross-covariance of images,  $c_1 = (K_1L)^2$ ,  $c_2 = (K_2L)^2$  are the stabilization constant, where  $K_1$  and  $K_2$  is smaller than 1 and  $L$  is the dynamic range of the image.

## 4 Results and Discussions

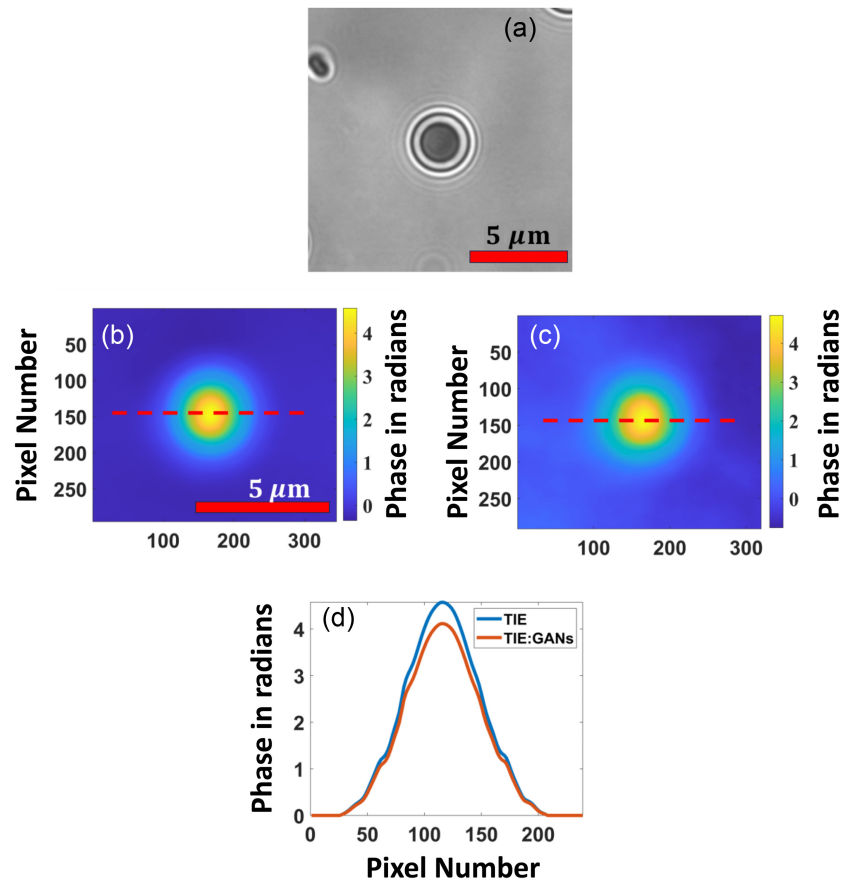
One of the findings during our testing is that if we increase the ROI bounding box having multiple cells for training in the region, it increases the time of training, but the SSIM and other parameter do not show any significant improvements. The captured image and extracted ROI images are shown in Fig. 6. Hence, the experiment was conducted on images having single oral cell, with a size of  $256 \times 256$  pixels. The dataset contains 700 images for training and 100 for testing and 100 images for validation. Similarly for the characterization, the microbeads are used with known refractive index and thickness. The proposed TIE-GANs were used to reconstruct the phase image of microbeads measuring  $4 \mu\text{m}$  in size (Sigma-Aldrich) with a known refractive index of 1.68 for characterization. The specimens were immersed in distilled water with a refractive index of 1.33, which has a linear attenuation coefficient of  $0.843 \mu\text{m}^{-1}$ . The results in Fig. 7(a) show input intensity image, Fig. 7(b) depicts the phase reconstruction using TIE, Fig. 7(c) represents the proposed TIE-GANs reconstruction, and Fig. 7(d) shows phase profile for both methods. The characterization data showed 0.98 SSIM value, which clearly verifies the validity of our proposed method with microbeads.

The initial learning rate and batch size were set to be at 0.0001 and 4, respectively, and the optimizer used was Adam, which is based on the stochastic gradient descent.<sup>75</sup> Our method performed two sets of iterations, one with 600 images and 200 epochs and the second with 700 images and 300 epochs. After 300 epochs, the network seems to be saturated and does not show any significant increase in accuracy to produce high SSIM reports. Figure 10(a) shows the results with 600 images and 200 epochs where we observed the higher number of SSIM were reported to that of the ground truth images. While in Fig. 10(b), the SSIM shows the data lying toward the 0.75 range, which also depicts after certain epochs despite the increase in the number of images the network becomes saturated, i.e., similarity to the ground truth data becomes less. The results from both the experiments are shown in Fig. 8. The maximum SSIM observed during our experiment was found to be 0.95 with 600 images and 200 epochs.

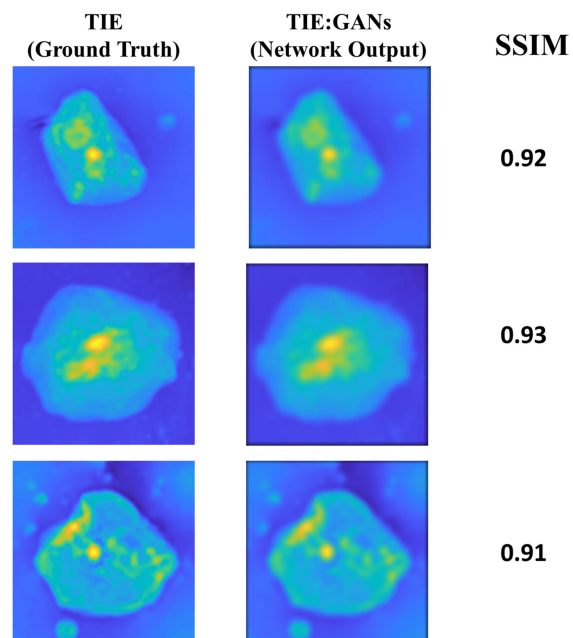
The detailed results of SSIM index are presented in Table 1 for both the experiments. We observed that after 200 epochs the network started losing its capability to generate new samples near to actual ground truth images gradually. The experimental parameters of both the experiments have been shown in Table 1.



**Fig. 6** Captured intensity image and extracted ROI for training.



**Fig. 7** (a) Input intensity image, (b) TIE phase reconstruction, (c) proposed TIE-GANs reconstruction, (d) phase profile for TIE and proposed TIE-GANs methods.



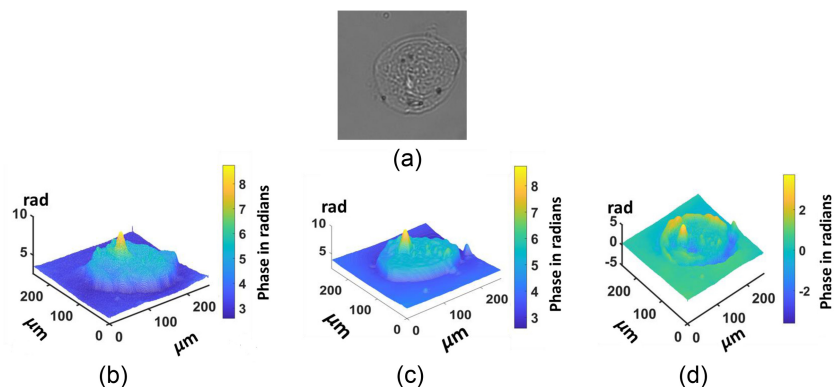
**Fig. 8** Conventional TIE and TIE:GANs reconstruction with SSIM values.

**Table 1** Experimental parameters.

Parameter	Experiment 1	Experiment 2
Number of images	600	700
Epochs	200	350
Max SSIM	0.95	0.94
Min SSIM	0.78	0.76
Max MSE	1364.98	1856.61
Min MSE	150.14	148.20
Max PSNR	26.36 dB	26.42 dB
Min PSNR	16.78 dB	15.45 dB
Max UIQI	0.89	0.90
Min UIQI	0.33	0.42
Time taken	1 h	1.5 h

We calculated the various parameters based on these observations and found that the experiment works best with 200 epochs and around 600 images. Figure 9(a) shows the intensity image, which is the input to the TIE: GAN's network, Fig. 9(b) shows the ground truth image. We can observe TIE-GANs output in Fig. 9(c), which clearly depicts the quantitative phase map and Fig. 9(d) shows the error map between TIE-GANs and ground truth image.

One more factor that is visible during training process is that along with the high number of images we need to introduce robustness to the network as well to avoid overfitting.<sup>76</sup> To do that, a skip connection was introduced in the network. However, it is observed that it is highly uncertain to determine the optimized parameter to get the desired results. The parameter needs to be tested through several iterations until the likely output is achieved. Overfitting of the network is one of the problems that arise during training. Our method utilizes skip connection and dropout in the deconvolution part of the network to avoid overfitting. TIE-GANs are found to be effective, and their execution time after training is as low as  $\sim 0.010$  s for generating phase map image. We do expect that the training cycles, the volume of training data (number of phase images) that is needed for practically reconstructing objects with different phases and thicknesses will be much high compared to what is demonstrated as a proof of concept in this paper. We have used beads of thicknesses  $\sim 4 \mu\text{m}$  with a microscope objective magnification of (63 $\times$ , 1.4 NA) in one set of images. We found that a defocus distance of  $1 \mu\text{m}$  was optimal for this NA and magnification. A



**Fig. 9** Reconstruction results of TIE-GANs. (a) Input defocused intensity image, (b) ground Truth data conventional TIE, (c) TIE-GANs output network generated output, and (d) error map between ground truth and TIE-GANs.

microscope objective (40 $\times$ , NA: 0.65) was used for imaging the exfoliate buccal cells. Light source exposure of the camera settings has been adjusted according to the sample for good quality images.

To introduce randomness to the network during training process, some of the images given to the network is unpaired, so the network does not get saturated over time. Although, unpaired image dataset is a well-known method that is known as cycle GANs.<sup>77</sup> During our training process, we observed that introducing false image pair increases the accurate image pair generated.

Our method is quite simple, and it works on the defocused images and the defocus distance that is trained during training of the samples. During testing, a single intensity defocused image is input to the algorithm and without any decision on the best focus plane or numerical iterative wave propagations, we retrieve the best phase image. This saves time and computational labor. The one to one mapping of pixels from the input defocused object to the output phase object can be elegantly achieved through training using GANs for a larger defocused distance, which is an advantage of TIE:GANs. Another important advantage of our technique is that the proposed algorithm works on DL derived GANs networks. Hence unlike other iterative phase retrieval algorithms, we retrieve the phase image instantly from the trained network. This gives us a remarkable edge in terms of real time operations compared to most other iterative and numerical propagation techniques.

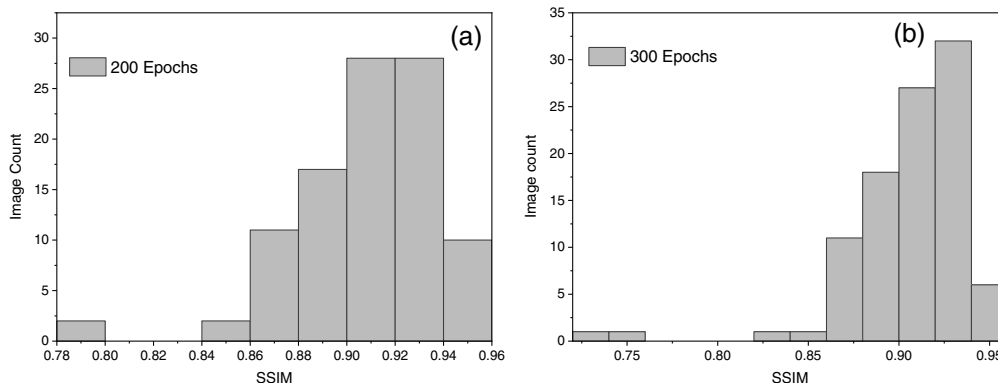
Figure 10(a) shows the SSIM histogram for both experiments, which reveals that the first iteration with 200 epochs has a greater percentage of SSIM values in the range of 0.92 to 0.94. Figure 10(b) for iteration 2 with 300 epochs also displays values near the 0.90 range, but it is evident that it has a low SSIM index compared to iteration 1. Maximum SSIM reported with 200 epochs and 600 image sample size was 0.95. Therefore, TIE-GANs-based reconstruction can reconstruct images with 95% similarity to the conventional TIE scheme, i.e., ground truth images.

The second parameter for quality assessment for reconstruction is MSE, which is calculated using Eq. (8) the histogram and the MSE distribution is shown in Figs. 11(a) and 11(b).

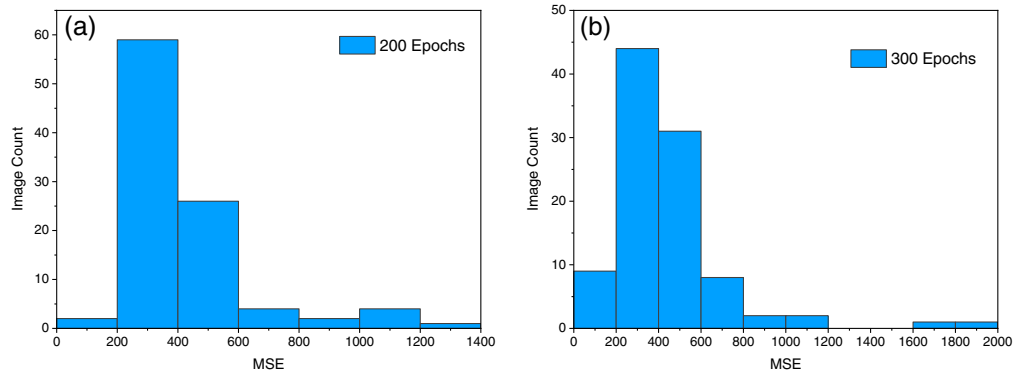
MSE computes the average squared error between each pixel in the two images being compared. From Fig. 11(a) with 200 epochs, we can observe that the maximum MSE is coming around the 1400 range. While in the second experiment it is exceeding the 2000 range means that the value of MSE is quite high as compared to the first iteration. Since we are looking for lower number of MSE hence first experiment makes the optimum sense regarding this experiment. Because the maximum number of values lies between 200 and 400 range, the result shows that the error is low as expected and it shows closeness to the ground truth images.

The PSNR value calculated using Eq. (9) and the maximum value for 200 epochs was found to be 26.36 dB and for 300 epochs with 700 training image pair the maximum value found to be 26.42 dB. The experimental parameter for PSNR is shown in Table 1.

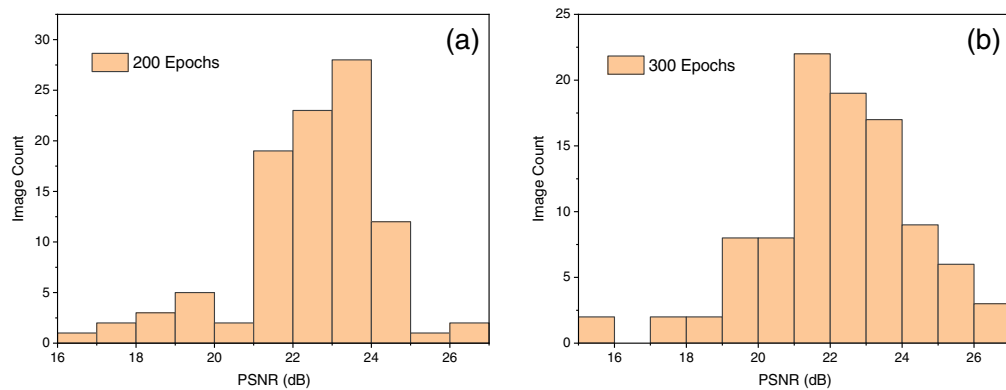
There was no such significant increase observed from 200 to 300 epochs hence it verified our observation with SSIM that after 200 epochs network start saturating. The histogram and the distribution of the PSNR value on test images are shown in Figs. 12(a) and 12(b).



**Fig. 10** (a) Figure with 200 epoch and 600 image sample size. (b) SSIM histogram with 300 epoch and 700 image sample size.

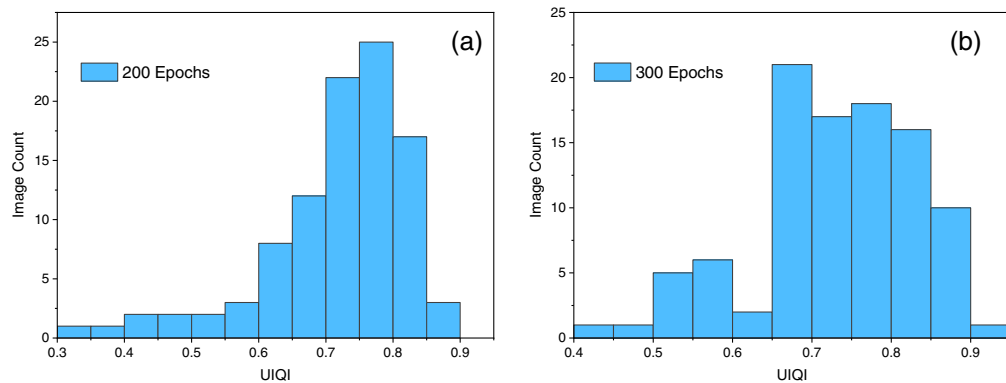


**Fig. 11** (a) MSE value for 200 epochs and 600 sample. (b) MSE for 300 epochs and 700 image sample.

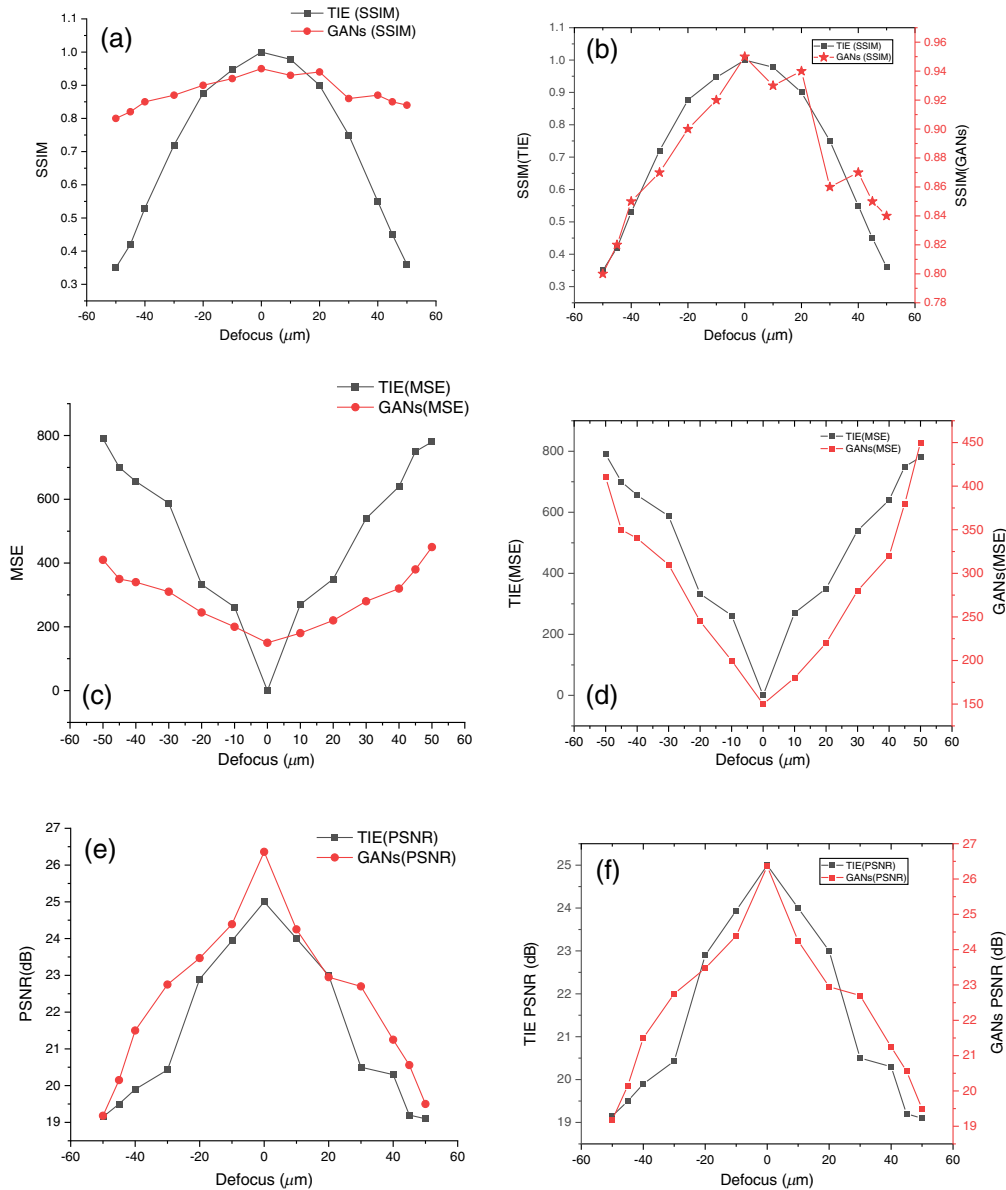


**Fig. 12** (a) PSNR value for 200 epochs and 600 image sample. (b) PSNR for 300 epochs and 700 image sample.

The final parameter for consideration is UIQI, which is calculated using Eq. (11). The UIQI is the combined measure of correlation, luminance, and contrast, and the histogram for epochs 200 and 300 and distribution of values is shown in Figs. 13(a) and 13(b). The UIQI ranges between 0 and 1, 0 shows no similarity between two images and 1 shows identical images to the ground truth images. The maximum UIQI observed is 0.90 for experiment 2, although the experiment 1 also gave significant results and reached about 0.89 UIQI value. From these four parameters, we came to the conclusion that despite of increment in the number of images in the second experiment the first one with 600 images worked well for optimum results.



**Fig. 13** (a) UIQI value for 200 epochs and 600 image sample. (b) UIQI for 300 epochs and 700 image sample.



**Fig. 14** (a), (b) SSIM with conventional TIE versus TIE-GANs; (c), (d) MSE TIE versus TIE-GANs; (e), (f) PSNR TIE versus TIE-GANs.

Since the conventional TIE is not a single shot scheme, it requires multiple images to be recorded at different defocus planes provided that the sample is non-homogeneous. Since we cannot have a single shot scheme for non-homogeneous samples, one of the advantages of our method is that it can work on any sample despite its nature. For non-homogeneous samples, the conventional TIE works in such a way we must find the best infocus image from intensity image stack using tamura coefficient as mentioned in the methodology section. Afterward it demands to compute SNR value by putting two defocus images one from overfocus and one from underfocus plane for phase reconstruction. This process always takes tedious parameter calibration. However, in the TIE-GANs dataset has been prepared using the different defocus plane images with oral cells so that it learns the features from different defocus distances during training itself. Figures 14(a) and 14(b) show the SSIM reconstructed at different defocus planes, which shows the SSIM reconstructed using conventional TIE and TIE-GANs. It is observed that if we propagate through the different defocus planes the TIE-GANs are able to produce phase maps closer to ground truth. As we shift away from the infocus plane and reconstruct phase map using three defocused images, the SSIM decreases gradually in case of conventional TIE shown in

Fig. 14(a). In the case of conventional TIE, it ranges from [0.35, 1] while in case of TIE-GANs it never reaches the value of 1 but it never drops below 0.80 range. It shows the consistency of the TIE-GANs to reconstruct phase map over long defocus distances as shown in Figs. 14(a) and 14(b).

Similar results can be observed for the MSE and PSNR values as well. In Figs. 14(c) and 14(d), it is clearly observed that the conventional TIE MSE can be ranged between [0,800], whereas TIE-GANs range between [150,450]. So, we can conclude that the TIE-GANs results have consistency over reconstruction in terms of accuracy. It produces minimal noise as low as possible even if we propagate along the optical axis in the Z direction. The final parameter for consideration is the PSNR value we can observe that in Figs. 14(e) and 14(f) the maximum PSNR value in terms of the TIE-GANs reaches up to 26.42 dB while in case of conventional TIE it reaches the value of the 24.5 dB, the values here even exceed the conventional TIE value. The minimum value that we observed during our experiment was 19 dB in the conventional TIE while in the case of the TIE-GANs it never drops below 19.5 dB. From comparative analysis of all these parameters, we can conclude that although conventional TIE performs well for non-homogeneous samples, TIE-GANs can perform well as a single-shot scheme. The TIE-GANs can reconstruct phase for long defocus distance in optical axis. Hence, TIE-GANs and other related schemes for the phase reconstruction based on DL hold the future for phase contrast microscopy. The challenge that one has to face during implementation is that we need to train the network with the same sample then only we are able to get the desirable results. In the case of conventional TIE, one does not have to hold the knowledge of the sample beforehand the mathematical implementation of the algorithm does the job for phase reconstruction. In case of TIE-GANs, the network must learn the feature map of the sample through training. After training, only TIE-GANs can predict feature map for the phase image and works as a single shot technique. Although unpaired image translation techniques have potential for phase reconstruction, we have not seen better statistical improvements as of now hence did not consider them for reconstruction in this study.

## 5 Conclusion

The incorporation of QPI with DL methodologies is a noteworthy achievement within the domain of microscopy and data processing. The convergence of these two innovative disciplines has the potential to fundamentally transform our understanding of biological systems, materials, and beyond. The label-free feature of QPI and DL algorithms enables automatic extraction of complicated information, enabling more accurate and efficient analysis. This study has emphasized the novelty and benefits of using DL algorithms, such as TIE-GANs for quantitative phase microscopy. The qualities of TIE such as label-free nature and less complex experimental set-up works as an add-on to QPI analysis with DL. This study is able to employ a method based on GANs to reduce the necessity of manual labor. Our method solved complexity in capturing a large number of images at different defocus planes by enabling single shot technique. The utilization of GANs as a methodology has been found to be proficient in generating phase images, provided that the dataset has been suitably prepared. TIE-GANs are able to convert single-intensity images to phase maps in rapid manner, and the results are encouraging for oral cells. The characterization is done with microbeads, which produces 98% SSIM and improved PSNR value. When trained on a structured dataset, as demonstrated by the microbead validation results, the proposed method can be used to characterize nanometric samples as well. The SSIM for oral cell was reported to be 95% and there is room for further improvement when the network is fine-tuned. There are some demerits of DL in terms of computation as it requires a heavy hardware setup. High configuration requirements make these algorithms and experiments data-, and computation-hungry. However, the future of artificial intelligence holds great potential along with QPI and other sectors of biomedical imaging. One of the key findings during our experiment was that it works for images on any defocus distance up to  $\pm 60 \mu\text{m}$  in  $z$  direction along the optical axis. The novelty of our approach showed that even if we provide intensity image at any distance in optical axis, and it is able to reconstruct the phase map of these intensity images under provided range. Future work for this research includes the accuracy assessment as well as tuning of hyperparameters that will improve our current parameters. Another dimension of this research is

to extend this similar capability for lensless holographic system as well for the virtual staining of the samples based on thickness or other properties. The future work for this research consists of the unpaired image translation for phase reconstruction. The future work of our method includes the nanometric resolution characterization for RBC and sperm cells.

---

## Disclosures

The authors declare no conflicts of interest.

## Code and Data Availability

All relevant data, materials, and software code used in this research are available upon request from the corresponding author.

## References

1. R. Mehrotra et al., "Application of cytology and molecular biology in diagnosing premalignant or malignant oral lesions," *Mol. Cancer* **5**(1), 11 (2006).
2. W. J. Croft, *Under the Microscope*, World Scientific (2006).
3. P. M. Zavos, J. R. Correa, and P. N. Zarmakoupis-Zavos, "Measurement of the sperm motility index via the sperm quality analyzer and its relationship to other qualitative sperm parameters," *Theriogenology* **46**(3), 421–427 (1996).
4. F. Zernike, "How I discovered phase contrast," *Science (1979)* **121**(3141), 345–349 (1955).
5. X. Wang et al., "Single-shot isotropic differential interference contrast microscopy," *Nat. Commun.* **14**(1), 2063 (2023).
6. R. D. Allen and G. B. David, "The Zeiss-Nomarski differential interference equipment for transmitted-light microscopy," *Z. Wiss. Mikrosk.* **69**(4), 193–221 (1969).
7. K. Lee et al., "Quantitative phase imaging techniques for the study of cell pathophysiology: from principles to applications," *Sensors* **13**(4), 4170–4191 (2013).
8. G. Popescu, *Quantitative Phase Imaging of Cells and Tissues*, McGraw-Hill Education (2011).
9. Y. K. Park, C. Depeursinge, and G. Popescu, "Quantitative phase imaging in biomedicine," *Nat. Photonics* **12**(10), 578–589 (2018).
10. J. Cai et al., "Automated Gleason grading of prostate cancers via deep learning in label-free multiphoton microscopic images," *Proc. SPIE* **11244**, 112442A (2020).
11. H. S. Park et al., "Automated detection of *P. falciparum* using machine learning algorithms with quantitative phase images of unstained cells," *PLoS One* **11**(9), e0163045 (2016).
12. Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**(7553), 436–444 (2015).
13. Y. Rivenson et al., "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light Sci. Appl.* **7**(2), 17141 (2018).
14. Y. Xue et al., "Reliable deep-learning-based phase imaging with uncertainty quantification," *Optica* **6**(5), 618–629 (2019).
15. A. Sinha et al., "Lensless computational imaging through deep learning," *Optica* **4**(9), 1117–1125 (2017).
16. J. C. Ye, Y. Han, and E. Cha, "Deep convolutional framelets: a general deep learning framework for inverse problems," *SIAM J. Imaging Sci.* **11**(2), 991–1048 (2018).
17. K. H. Jin et al., "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.* **26**(9), 4509–4522 (2017).
18. Y. Wu et al., "Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery," *Optica* **5**, 704–710 (2018).
19. W. Jeon et al., "Speckle noise reduction for digital holographic images using multi-scale convolutional neural networks," *Opt. Lett.* **43**(17), 4240 (2018).
20. Y. Wu et al., "Bright-field holography: cross-modality deep learning enables snapshot 3D imaging with bright-field contrast using a single hologram," *Light Sci. Appl.* **8**(1), 25 (2019).
21. E. Nehme et al., "DeepSTORM3D: dense 3D localization microscopy and PSF design by deep learning," *Nat. Methods* **17**(7), 734–740 (2020).
22. T. Nguyen et al., "Deep learning approach for Fourier ptychography microscopy," *Opt. Express* **26**(20), 26470 (2018).
23. S. Jiang et al., "Solving Fourier ptychographic imaging problems via neural network modeling and tensor-flow," *Biomed. Opt. Express* **5**(7), 7873–7892 (2018).
24. L. Boominathan et al., "Phase retrieval for Fourier ptychography under varying amount of measurements," in *British Machine Vision Conf.* (2018).



25. K. Wang et al., "Transport of intensity equation from a single intensity image via deep learning," *Opt. Lasers Eng.* **134**, 106233 (2020).
26. X. Wu, Z. Wu, and Y. Zhu, "Learning a model-based neural network for quantitative phase imaging based on the transport of intensity," *Proc. SPIE* **11396**, 1139604 (2020).
27. X. Wu et al., "Physics-informed neural network for phase imaging based on transport of intensity equation," *Opt. Express* **30**(24), 43398–43416 (2022).
28. A. S. Galande et al., "Untrained neural network with explicit denoiser for lensless inline holographic microscopy," in *Comput. Opt. Sens. and Imaging*, Optica Publishing Group, p. CTh7A-7 (2021).
29. P. Panta et al., "Label-free analysis of oral cytology specimens through digital holographic microscopy and deep-neural networks," in *Digital Hologr. and Three-Dimens. Imaging*, Optical Society of America, p. Th3A-11 (2019).
30. U. S. Kamilov et al., "Learning approach to optical tomography," *Optica* **2**(6), 517 (2015).
31. H. Qiao and J. Wu, "GPU-based deep convolutional neural network for tomographic phase microscopy with  $\ell_1$  fitting and regularization," *J. Biomed. Opt.* **23**(6), 066003 (2018).
32. A. Barty et al., "Quantitative optical phase microscopy," *Opt. Lett.* **23**(11), 817–819 (1998).
33. I. J. Goodfellow et al., "Generative adversarial networks," *Neural Inf. Process. Syst.*, arXiv:1406.2661 (2014).
34. P. Isola et al., "Image-to-image translation with conditional adversarial networks," in *IEEE Conf. Comput. Vision and Pattern Recognit. (CVPR)*, Honolulu, Hawaii, pp. 5967–5976 (2017).
35. S. K. Gorti and J. Ma, "Text-to-image-to-text translation using cycle consistent adversarial networks," arXiv:1808.04538 (2018).
36. G. Couairon et al., "Flexit: towards flexible semantic image translation," in *Proc. IEEE/CVF Conf. Comput. Vision and Pattern Recognit.*, pp. 18270–18279 (2022).
37. Y. Rivenson et al., "PhaseStain: the digital staining of label-free quantitative phase microscopy images using deep learning," *Light Sci. Appl.* **8**(1), 2047–7538 (2019).
38. J. Cao et al., "Remix: towards image-to-image translation with limited data," in *Proc. IEEE/CVF Conf. Comput. Vision and Pattern Recognit.*, pp. 15018–15027 (2021).
39. P. K. Poola and R. John, "Label-free nanoscale characterization of red blood cell structure and dynamics using single-shot transport of intensity equation," *J. Biomed. Opt.* **22**(10), 106001 (2017).
40. P. K. Poola et al., "Quantitative label-free sperm imaging by means of transport of intensity," *Proc. SPIE* **9718**, 971800 (2016).
41. U. Sara, M. Akter, and M. S. Uddin, "Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study," *J. Comput. Commun.* **7**(3), 8–18 (2019).
42. N. Yoneda et al., "Single-shot higher-order transport-of-intensity quantitative phase imaging based on computer-generated holography," *Opt. Express* **29**(4), 4783–4801 (2021).
43. T. Cacace, V. Bianco, and P. Ferraro, "Quantitative phase imaging trends in biomedical applications," *Opt. Lasers Eng.* **135**, 106188 (2020).
44. J. Li et al., "Transport of intensity diffraction tomography with non-interferometric synthetic aperture for three-dimensional label-free microscopy," *Light Sci. Appl.* **11**(1), 154 (2022).
45. N. Streibl, "Phase imaging by the transport equation of intensity," *Opt. Commun.* **49**(1), 6–10 (1984).
46. Y. Li et al., "Quantitative phase microscopy for cellular dynamics based on transport of intensity equation," *Opt. Express* **26**(1), 586–593 (2018).
47. I. Ghori, V. Thapa, and R. John, "X-ray phase contrast imaging of coronary arteries," in *Opt. Mol. Probes, Imaging and Drug Deliv.*, Optical Society of America, p. JTu4A-13 (2021).
48. C. Zuo et al., "Transport of intensity equation: a tutorial," *Opt. Lasers Eng.* **135**, 106187 (2020).
49. C. Zuo, Q. Chen, and A. Asundi, "Boundary-artifact-free phase retrieval with the transport of intensity equation: fast solution with use of discrete cosine transform," *Opt. Express* **22**(8), 9220 (2014).
50. R. Ali et al., "On the use of low-pass filters for image processing with inverse Laplacian models," *J. Math. Imaging Vision* **43**(2), 156–165 (2012).
51. C. Zuo et al., "Deep learning in optical metrology: a review," *Light Sci. Appl.* **11**(1), 39 (2022).
52. M. Beleggia et al., "On the transport of intensity technique for phase retrieval," *Ultramicroscopy* **102**(1), 37–49 (2004).
53. M. Tamamitsu et al., "Comparison of Gini index and Tamura coefficient for holographic autofocusing based on the edge sparsity of the complex optical wavefront," arXiv:1708.08055 (2017).
54. A. Mahjoubfar, C. L. Chen, and B. Jalali, *Artificial Intelligence in Label-Free Microscopy: Biological Cell Classification by Time Stretch*, Springer International Publishing (2017).
55. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM* **60**(6), 84–90 (2017).
56. S. Albarqouni et al., "AggNet: deep learning from crowds for mitosis detection in breast cancer histology images," *IEEE Trans. Med. Imaging* **35**(5), 1313–1321 (2016).
57. A. Garcia-Garcia et al., "A review on deep learning techniques applied to semantic segmentation," arXiv:1704.06857 (2017).

58. I. Goodfellow, Y. Bengio, and A. Courville, “Deep learning,” MIT Press (2016).
59. K. Simonyan, A. Vedaldi, and A. Zisserman, “Deep inside convolutional networks: visualising image classification models and saliency maps,” arXiv:1312.6034 (2013).
60. A. Koutsoukas et al., “Deep-learning: investigating deep neural networks hyper-parameters and comparison of performance to shallow methods for modeling bioactivity data,” *J. Cheminform* **9**(1), 42 (2017).
61. M. Varshney and P. Singh, “Optimizing nonlinear activation function for convolutional neural networks,” *Signal Image Video Process.* **15**(6), 1323–1330 (2021).
62. I. Goodfellow, “NIPS 2016 tutorial: generative adversarial networks,” arXiv:1701.00160 (2016).
63. T.-C. Wang et al., “High-resolution image synthesis and semantic manipulation with conditional GANs,” in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 8798–8807 (2018).
64. M. Mirza and S. Osindero, “Conditional generative adversarial nets,” arXiv:1411.1784 (2014).
65. M. Mitome, “Transport of intensity equation method and its applications,” *Microscopy* **70**(1), 69–74 (2021).
66. R. Close et al., “Towards quantitative, atomic-resolution reconstruction of the electrostatic potential via differential phase contrast using electrons,” *Ultramicroscopy* **159**, 124–137 (2015).
67. O. Ronneberger, P. Fischer, and T. Brox, “U-Net: convolutional networks for biomedical image segmentation,” *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).
68. S. M. K. Hasan and C. A. Linte, “U-NetPlus: a modified encoder-decoder U-Net architecture for semantic and instance segmentation of surgical instruments from laparoscopic images,” in *41st Annu. Int. Conf. of the IEEE Eng. in Med. and Biol. Soc. (EMBC)*, IEEE, pp. 7205–7211 (2019).
69. S. Mastromichalakis, “ALReLU: a different approach on leaky ReLU activation function to improve neural networks performance,” arXiv:2012.07564 (2020).
70. Z. Wang and A. C. Bovik, “Mean squared error: love it or leave it? A new look at signal fidelity measures,” *IEEE Signal Process. Mag.* **26**(1), 98–117 (2009).
71. A. Hore and D. Ziou, “Image quality metrics: PSNR vs. SSIM,” in *20th Int. Conf. Pattern Recognit.*, IEEE, pp. 2366–2369 (2010).
72. Z. Wang and A. C. Bovik, “A universal image quality index,” *IEEE Signal Process. Lett.* **9**(3), 81–84 (2002).
73. Z. Wang et al., “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. Image Process.* **13**(4), 600–612 (2004).
74. H. Cheong et al., “OCT-GAN: single step shadow and noise removal from optical coherence tomography images of the human optic nerve head,” *Biomed. Opt. Express* **12**(3), 1482–1498 (2021).
75. N. S. Huda and M. S. Mubarak, “A multi-label classification on topics of quranic verses (English translation) using backpropagation neural network with stochastic gradient descent and Adam optimizer,” in *7th Int. Conf. Inf. and Commun. Technol. (ICoICT)*, IEEE, pp. 1–5 (2019).
76. X. Ying, “An overview of overfitting and its solutions,” *J. Phys.: Conf. Ser.* **1168**, 022022 (2019).
77. J.-Y. Zhu et al., “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *Proc. IEEE Int. Conf. Comput. Vision*, pp. 2223–2232 (2017).

**Vikas Thapa** is a PhD research scholar in the Department of Biomedical Engineering at Indian Institute of Technology (IIT), Hyderabad, India. He received his master’s degree from UPES, Dehradun, in 2018. His PhD research work is dedicated to deep learning based microscopic image analysis. He has hands-on experience with various biomedical devices and instruments. He is also actively participating in international and national conferences.

**Ashwini Subhash Galande** is pursuing a PhD in IIT, Hyderabad, India. She is working on the physics aware deep networks for holographic imaging. Her main area of research is deep learning methods to solve inverse problems in holography.

**G. Hanu Phani Ram** holds a PhD in biomedical engineering from the IIT, Hyderabad, which was completed in 2020. His areas of research are biophotonics specifically aligned toward developing low cost microscopy. Prior to his PhD, he received his master’s degree (MTech) in biomedical engineering from the same institution in 2014. His educational journey began with a bachelor’s degree (BTech) in electronics and communication engineering from BSIT, JNTUH, Hyderabad, which he completed in 2012. Presently, he is working as the assistant professor in the Department of Biomedical Engineering, B V Raju Institute of Technology, Narsapur.

**Renu John** is a professor in the Department of Biomedical Engineering at IIT, Hyderabad, India. He received his PhD from IIT Delhi in 2006 and received his postdoc from Duke University, Durham, North Carolina, United States, and the University of Illinois Urbana-Champaign, Champaign, Illinois, United States, from 2006 to 2010. His research covers several key areas in biomedical optics, such as quantitative phase imaging and optical coherence imaging.