Deep Computational Imaging Techniques for Microscopy

A Project Report

submitted by

MAYUG MANIPARAMBIL, EE13B088

in partial fulfilment of requirements
for the award of the dual degree of

BACHELOR OF TECHNOLOGY AND MASTER OF TECHNOLOGY

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY MADRAS

MAY 2018

THESIS CERTIFICATE

This is to certify that the thesis titled **Deep Computational Imaging Techniques for Microscopy**, submitted by **Mayug Maniparambil, EE13B088**, to the Indian Institute of Technology Madras, for the award of the dual degree of **Bachelor of Technology and Master of Technology**, is a bona fide record of the research work done by him under our supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. Kaushik Mitra
Research Guide
Assistant Professor
Dept. of Electrical Engineering
IIT Madras, 600036

Place: Chennai

Date: 10th May 2018

ABSTRACT

KEYWORDS: Computational Imaging,Deep Learning, Fourier Ptychography Microscopy, Image inpainting, Biomedical Images, Serial Two Photon Tomography

This project explores and solves two different problems in microscopy of biomedical images using deep learning based computational imaging techniques. The first problem is on phase retrieval in Fourier Ptychography Microscopy.Fourier Ptychography is a recently proposed imaging technique that yields high-resolution images by computationally transcending the diffraction blur of an optical system. At the crux of this method is the phase retrieval algorithm, which is used for computationally stitching together low-resolution images taken under varying illumination angles of a coherent light source. However, the traditional iterative phase retrieval technique relies heavily on the initialization and also need a good amount of overlap in the Fourier domain for the successively captured low-resolution images, thus increasing the acquisition time and data. We show that an auto-encoder based architecture can be adaptively trained for phase retrieval at 40% and 65% of overlap and outperform traditional methods like AP. For high overlap, we show that optimizing the generator for reducing the forward model error is an appropriate choice. Using simulations for the challenging case of uncorrelated phase and amplitude, we show that our method outperforms many of the previously proposed Fourier ptychography phase retrieval techniques. The second problem is 'Denoising High Density Gene Expression in Whole Mouse Brain Images scanned using Serial Two Photon Tomography Microscopy'. In particular this project was aimed to solve a particular de-noising problem in the automated ex vivo mouse brain imaging STP setup of Cold Spring Harbor Labs.It is a method that achieves high-throughput fluorescence imaging of mouse brains by integrating two-photon microscopy and tissue sectioning. STP tomography generates high-resolution datasets that are free of distortions and can be readily warped in three dimensions, for example, for comparing multiple anatomical tracings and opens the door to routine systematic studies of neuroanatomy in mouse

models of human brain disorders.In certain slice high density cellular/nuclear gene expression affects registration accuracy via the pixel intensity of the anatomical features in downsampled images.We propose to use a deep neural network with adversarial loss for denoising the gene expression to improve the registration accuracy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

IITM	Indian Institute of Technology Madras
FPM	Fourier Ptychography Microscopy
CGAN	Conditional Generative Adversarial Networks
DIP	Deep Image Prior
CDIP	Conditional Deep Image Prior
STP	Serial Two-Photon Tomography
CGAN-FP	Conditional Generative Adversarial Network-Fourier Ptychography
CSHL	Cold Spring Harbor Laboratory

CHAPTER 1

INTRODUCTION

1.1 Introduction to Fourier ptychography phase retrieval

Sub-optimal throughput of traditional imaging systems can be attributed to the fundamental limitation of its optics, known as the Space Bandwidth Product (SBP). The SBP of an optical system characterizes the total resolvable pixels, and imposes a trade-off between the field of view and resolution of captured images. Fourier Ptychography (FP) Zheng *et al.* (2013) is a powerful imaging technique that circumvents this physical limitation using computation, yielding gigapixel-scale intensity and phase images. It has applications in both microscopic bio-medical imaging Zheng *et al.* (2013), as well as long range imaging for surveillance and remote sensing Holloway *et al.* (2017)Holloway *et al.* (2016).

The technique of FP includes the acquisition of many SBP limited images using varying illumination angles of a coherent light source. Since conventional image sensors can capture only the intensity of light falling on them, there's a loss of phase information. Hence, phase retrieval algorithm is then applied to the captured set of images to reconstruct a high resolution, high field-of-view image. However, traditional phase retrieval algorithms suffer from a common artifact known as phase-amplitude leakage - an artifact caused due to phase information leaking into amplitude and vice versa. This artifact becomes more predominant when phase and amplitude are highly uncorrelated.

In this paper, we exploit the rapid progress of deep learning based techniques in solving inverse reconstruction problems Ulyanov *et al.* (2017); Chang *et al.* (2017); Dave *et al.* (2018, 2017) for the task of phase retrieval in FP. We propose a non data-driven technique that optimizes the weights of an auto encoder based generator network, using an objective function that reduces the loss between estimated and observed low resolution measurements. This method, inspired by Ulyanov *et al.*'s Deep Image PriorUlyanov *et al.* (2017), doesn't necessarily require any prior training and the optimization is done only for a given set of test measurements.

In summary, we propose an auto-encoder based non-data driven framework, where we optimize over the generator parameters by minimizing the forward measurement error of FP. While other traditional algorithms also optimize in a similar way, the proposed method additionally exploits the low-level image statistics captured by generator network's inherent structure Ulyanov *et al.* (2017), thereby making it more robust to phase-amplitude leakage. Using simulations for uncorrelated phase and amplitude in both low and high overlaps, we show that our algorithm outperforms previously proposed FP phase retrieval techniques.

1.1.1 Background on Fourier ptychography

Traditional imaging systems have a trade off between the field-of-view and resolution at which it can capture images, due to the physical limitation of its optics. The objective of Fourier ptychography is to circumvent this limitation by computation. For this, multiple low resolution images are captured using a high field of view objective lens, where each acquisition is done for different angles of illumination. Using concepts from Fourier optics, this can be understood as sampling different regions of object's high resolution Fourier domain, and capturing only its corresponding intensity. As the phase

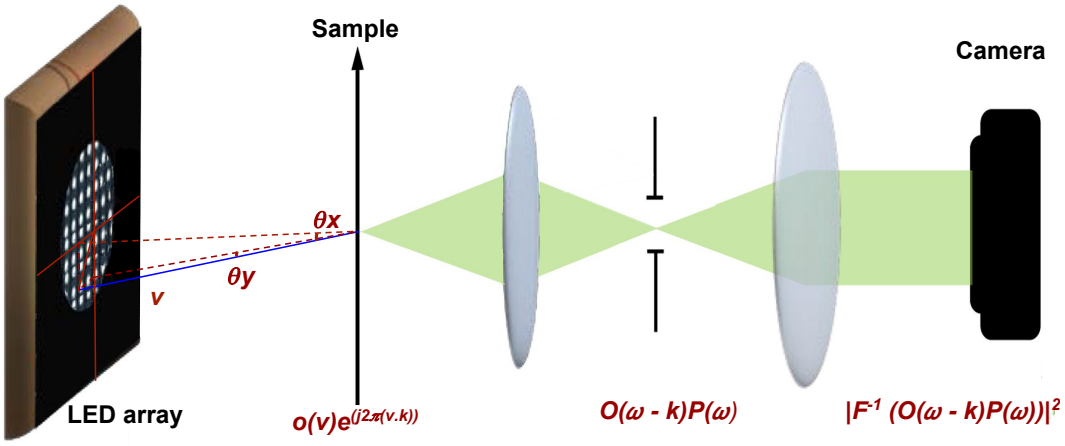

Figure 1.1: *Schematic illustration of :* Here we show the three parts of FP forward model. The first being the angular illumination from LED array to target sample. Light from target sample then passes through the objective lens, in the process the high frequency components gets removed due to the limited bandwidth of objective lens' pupil function. The camera then captures the intensity of light coming from the objective lens, during which phase information of light is lost.

information is lost in each of those acquisitions, reconstruction is no longer straightforward. There is a need for capturing a certain amount redundant information during each acquisition, so as to retrieve the lost phase. This comes in terms of high overlap between successive measurements in Fourier domain, making required number of acquisitions sub-optimal. The high resolution intensity and phase images are then reconstructed, by finding an estimate that satisfies both the spatial and Fourier domain constraints imposed by the observed measurements.

In the following subsections, we mathematically model the steps involved in FP technique, and briefly discuss some of the recent phase retrieve techniques used for solving FP.

FP forward model:

Fourier ptychography forward model involves the angular illumination of target using LEDs, interaction of light from target sample with the objective lens, and finally its acquisition using the camera. Below, we mathematically model each of these steps (See Figure 1.1).

- LED illumination: Let us consider the imaging of a thin target sample having transmission function given by $o(v) = A(v)\mathrm{e}^{j\phi(v)}$, where v is the spatial vector, and $A(v)$ and $\phi(v)$ are the amplitude and phase attenuation respectively. If the target sample is illuminated using an LED that emits light with spatial frequency k, then light from the target sample that reaches the objective lens is $o(v)\mathrm{e}^{j2\pi(v.k)}$. Therefore, if $O(\omega)$ is Fourier transform of the target sample under normal illumination, then the Fourier transform under angular illumination would be $O(\omega - k)$.
- Objective lens: Let $P(\omega)$ be the pupil function of the objective lens, then the light passing through it would be $O(\omega - k)P(\omega)$. Since under normal illumination $P(\omega)$ acts as a low pass filter with cut-off frequency given by its numerical aperture, under angular illumination it would behave as a band pass filter.
- Image acquisition: However, since camera sensors can capture only the intensity and not the phase of light coming from the objective lens, the captured image is given as $|\mathcal{F}^{-1}(O(\omega - k)P(\omega))|^2$, where $\mathcal{F}^{-1}$ is Fourier inverse. Multiple such images are taken using varying angles of illumination to capture different regions of the object's Fouier domain.

FP inverse model:

Phase retrieval can be posed as a non-convex optimization problem that estimates the high resolution complex object, subject to the constraints imposed by the observed low resolution measurements. In the context of this paper, we segregate such optimization algorithms as non-data driven approaches and data-driven deep-learning based approaches.

Non-data driven algorithms for FP: As discussed in Yeh *et al.* (2015), non-data driven algorithms can be broadly classified as sequential or global, based on the region of high resolution Fourier domain that is updated in each iteration. In each iteration, global techniques update the entire Fourier domain in order to minimize the cumulative loss with respect to all observed measurements, whereas sequential techniques update only the region corresponding to one observed measurement at a time. In terms of the order of derivative used for optimization, these algorithms can be further divided as first order sequential Zheng *et al.* (2013),Ou *et al.* (2014), second order sequential Tian *et al.* (2014), first order global Bian *et al.* (2015), and second order global Yeh *et al.* (2015). It has also been shown that the non-convex phase retrieval algorithm can be cast as a low-rank semi-definite programming problemRecht *et al.* (2010)Candes *et al.* (2015)Candes *et al.* (2013)Burer and Monteiro (2003)Horstmeyer *et al.* (2015), which has been solved in he context of FP in Horstmeyer *et al.* (2015). From studies shown in Yeh *et al.* (2015), it was inferred that among these, global Newton's method Yeh *et al.* (2015) gives the best reconstruction at the cost of complex computation, whereas sequential Gauss-Newton Tian *et al.* (2014) offers the best trade off between computational efficiency and robustness for large-scale applications.

Data-driven deep learning based approaches: However, due to the non-convex nature of the problem, all non-data driven algorithms tend to get stuck in local-minima when the number of constraints imposed by measurements are less, which is the case when overlap in the Fourier domain between successive captured low-resolution images are less. Under such an ill-posed setting, we obviously expect data-driven techniques that learns the prior structure of high resolution intensity and phase to perform well. PtychNet Kappeler *et al.* (2017), prDeep Metzler *et al.* (2018) and Rivenson *et al.* (2017) are examples of data-driven techniques that were used for different such ill posed phase retrieval settings.

While PtychNet have shown promising results for intensity reconstruction under reduced overlap, it was limited to the special case of using constant phase in the forward model. In most real-world scenarios, the assumption of light from an object having only constant phase does not always hold true, and hence the reconstruction might suffer from severe phase-amplitude leakage Yeh *et al.* (2015). In this paper, we show that the proposed algorithm can reconstruct both high resolution intensity and phase under reduced overlap, even when they are highly uncorrelated. Rivenson *et al.* (2017) have also proposed such an approach of reducing the number of needed observations by applying deep learning, but have shown its application only for holography images with correlated phase and amplitude.

Metzler *et al.* (2018) have shown that the robustness of phase retrieval algorithm to noise can be further improved by using a denoising based deep-learning prior. The formulation includes optimization of high resolution estimate subjected to constrains imposed by both the observed measurements, as well as a trained denoising convolutional neural network (CNN) which ensures that the estimate suffers from less noise. However, the results shown are limited to amplitude-only targets, and hence phase is not reconstructed. Although this work Metzler *et al.* (2018) is not in the context of reduced measurements or uncorrelated phase and amplitude, it demonstrates how deep learning based priors can be exploited for phase retrieval. It is also worth noting that, unlike Metzler *et al.* (2018), our algorithm is fully deep learning based where even the architecture weights are directly updated based on loss with respect to observed measurements.

Recently, *Boominathan & Mitra, 2018* proposed a conditional adversarial based network for Fourier Ptychography (cGAN-FP), which aims to reconstruct high resolution amplitude and phase, specially when the amount of overlap between successive measurements in the Fourier domain is less. Although this shows promising results for the low overlap case, cGAN-FP's performance is much inferior to traditional algorithms for the high overlap case. This is because cGAN-FP doesn't directly take into account measurement based loss, and makes prediction only based on the average statistics it learned over the entire dataset it was trained for. This motivates us to explore ways in which the estimate of the deep learning framework not only uses image prior, but also aims to satisfy the likelihood constraints imposed by the measurements.

1.2 Introduction to denoising of high density gene expression in whole mouse brain images

STP tomography, that enables automated high-throughput imaging of fluorescently labeled mouse brains. This method uses whole-mount two-photon microscopy and thus generates datasets of well-aligned, high-resolution serial optical sections.Using STP tomography cell types specific fluorescent protein expression can be studied and by systematic mapping of input and output connections of mouse, somatosensory cortex can be done easily and robustly. This can help to study mouse models of human

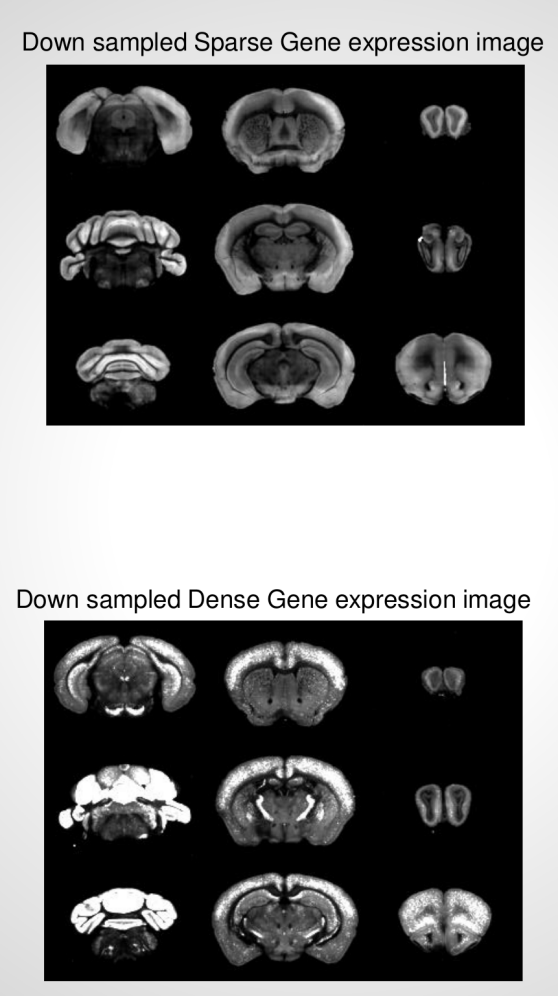

Figure 1.2: *Problem statement:* Here we show the down-sampled images of various brain slices. The top set of brain slices only have sparse gene expressions, whereas the bottom set has thick chumps of dense expressions. As a result of which, performing registration for the the former is much is easier than for latter. We aim to solve this problem of registration for slices with dense gene expression by applying deep learning based inpainitng.

brain disorders in standard laboratory settings and revolutionize drug discovery. Cold Spring Harbor Laboratory has been pioneering a method for automate ex vivo STP. Their pipeline invlolves collecting the slices at 8557*11377 resolution, downsampling them by 20 times and registering the imaged 3D volume to a reference model. The high expression in certain slices become white patches when we downsample by 20 times as shown in Figure **??**. This affects the registration accuracy badly. Our objective in this project is to remove the high expression data in the image, effectively a denoising problem. To solve this problem we decided to inpaint the white patches in the high resolution image and then downsample.For this inpainting we used the context-encoder architecture proposed by Pathak et al and then downsample the image by 20 times to be used in the pipeline. n contrast, our context encoder needs to solve a much harder task: to fill in large missing areas of the image, where it canâĂŹt get âĂIJhintsâĂİ from nearby pixels. This requires a much deeper semantic understanding of the scene, and the ability to synthesize high-level features over large spatial extents.

1.2.1 Related works

As suggested by its name, exemplar-based inpainting like patch-match uses an iterative solution to generate the unknown region based on the source region of the image. The pixel synthesis of the fill-region begins at the fill-front, or the edge between the known/unknown region, and gradually moves inwards to complete the missing zone. The intuition of this algorithm came from fluid dynamics and partial differential equations. The filling mechanism is derived from Dr.Bertalmo, Bertozzi and Sapiros concept of smooth continuation of information in the level-lines direction.[3] According to BertalmoâĂŹs 2001 CVPR paper, the filling rule is to extend the isophotes, or linear structures, while matching gradient vectors at the contiguous edge of the fill-region.[4] It is important to point out that our hole-filling task cannot be handled by classical inpainting [4, 32] or texture synthesis [2, 11] approaches,since the missing region is too large for local non-semantic methods to work well. In computer graphics, filling in large holes is typically done via scene completion [19], involving a cut-paste formulation using nearest neighbors from a dataset of millions of images. However, scene completion is meant for filling in holes left by removing whole objects,and it struggles to fill arbitrary holes, e.g. amodal completion of partially occluded objects. Further-

more, previous completion relies on a hand-crafted distance metric, such as Gist [31] for nearest-neighbor computation which is inferior to a learned distance metric. We show that our method is often able to inpaint semantically meaningful content in a parametric fashion, as well as provide a better feature for nearest neighbor-based inpainting methods.

CHAPTER 2

Methodology

2.1 Proposed method for FP phase retrieval

As in the above mentioned cGAN-FP framework (illustration of its training shown in subfigure (a) of Figure 2.1) uses only a conditional prior learned over a large dataset, the reconstruction lacks the sharpness and details that can otherwise be achieved in the

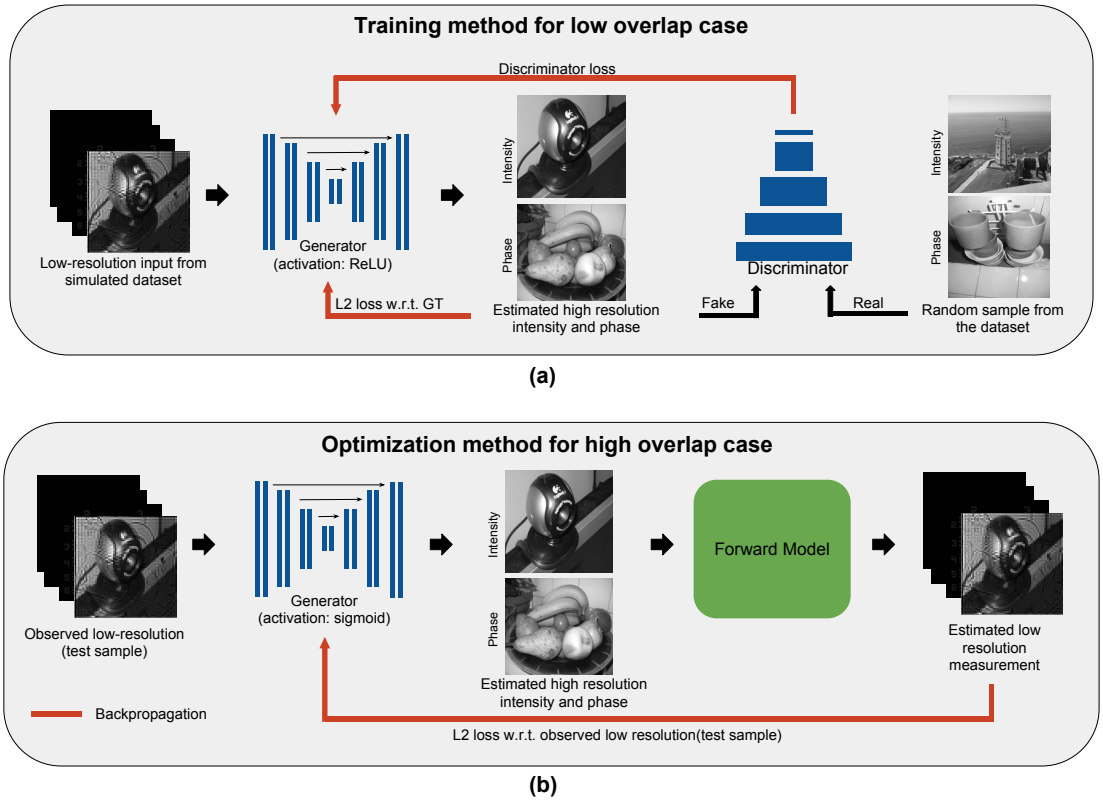

Figure 2.1: *Proposed approach for low and high overlap cases* For the low overlap case we train the network via supervised training approach, see subfigure (a). The input to the generator are the simulated low resolution images. The network weights are learned by minimizing L_2 loss between network's output and the ground truth high resolution intensity and phase images. Subfigure (b) shows our method for the high overlap case. We optimize the generator's weights by minimizing the loss between the input and the estimated low resolution images (obtained by passing the the generator's reconstructed intensity and phase images through the FP forward model.)

high overlap case by exploiting measurement based loss. This motivates us to directly optimize the generator network, such that the reconstruction when passed through the forward model yields values close to the observed measurements. This can be formulated as

$$\arg\min_{\theta} \; loss(x, M(G_\theta(x))) \; \forall \; x \; \in \; \{training \; dataset\} \tag{2.1}$$

where x is the low resolution measurements, G_θ is the generator network with learnable parameters θ, M is the FP forward model. The above optimization problem basically means that the optimized generator G_θ is an approximation for the inverse of M. This would be difficult to learn, more so when the number of measurements are not sufficient for solving the inverse problem. Thus, instead of finding the generator that works for all x, we propose to optimize the generator parameter θ for only a given test set of measurements x_0, i.e.

$$\arg\min_{\theta} \; loss(x_0, M(G_\theta(x_0))) \;\; for \; a \; given \; x_0, \tag{2.2}$$

which is a much easier optimization problem and also produces better results, see Figure 2.2. We observe that the formulation in 2.2 is related to that used in Deep Image Prior

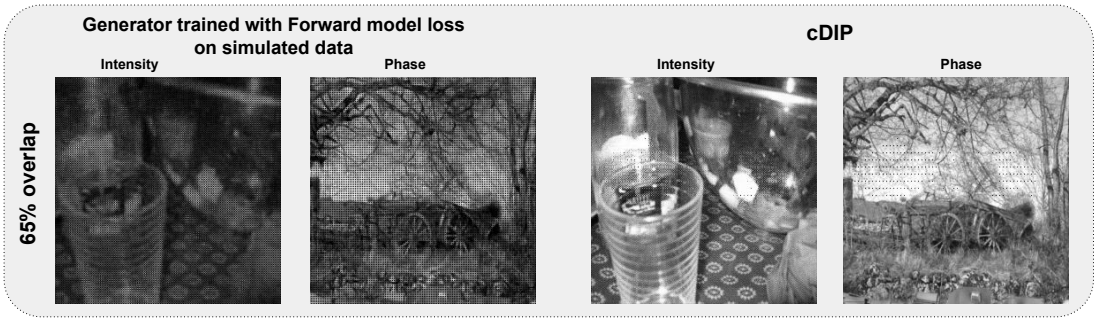

Figure 2.2: *Training for entire dataset vs optimizing for one test sample, for 65% overlap case:* On the left is the result obtained by training (on the simulated dataset) our auto-encoder based generator network using just forward model loss. On the right is the result obtained for the same generator network with forward model loss, but with weights optimized only for a given test sample. We observe that the generator finds it more difficult to reconstruct high resolution phase and amplitude, when trained for entire dataset, as compared to optimizing just for the one test sample.

Ulyanov *et al.* (2017), with the difference that we use the low resolution measurements as input instead of random noise. This is more suitable for our case, as it is easier to reconstruct from low resolution input as compared to noise, which is clearly evident

from Figure 3.1. Hence, we call this as conditional deep image prior (cDIP). For cDIP, we found sigmoid to be a better non linearity function after generator's last convolution layer. An illustrative example of its optimization is shown in subfigure (b) of Figure 2.1.

Just as mentioned in Ulyanov *et al.* (2017), the network structure's high impedance to noise makes it more robust to artifacts such as phase-amplitude leakage, resulting in atleast a good looking local optimum with minor variations depending on the network initialization. We note that initializing cDIP using learned cGAN-FP weights yields significantly better reconstruction with faster convergence.

2.2 Proposed method for denoising of high density gene expression in whole mouse brain images

Given an image with a missing region, we train a convolutional neural network to regress to the missing pixel values. We call our model context encoder, as it consists of an encoder capturing the context of an image into a compact latent feature representation and a decoder which uses that representation to produce the missing image content. The context encoder is closely related to autoencoders [3, 20], as it shares a similar encoder-decoder architecture. Autoencoders take an input image and try to reconstruct it after it passes through a low-dimensional âĂIJbottleneckâĂİ layer, with the aim of obtaining a compact feature representation of the scene. Unfortunately, this feature representation is likely to just compresses the image content without learning a semantically meaningful representation. In contrast,context encoder needs to solve a much harder task: to fill in large missing areas of the image, where it canâĂŹt get âĂIJhintsâĂİ from nearby pixels. This requires a much deeper semantic understanding of the scene, and the ability to synthesize high-level features over large spatial extents.oders are trained in a completely unsupervised manner. Our results demonstrate that in order to succeed at this task, a model needs to both understand the content of an image, as well as produce a plausible hypothesis for the missing parts. This task, however, is inherently multi-modal as there are multiple ways to fill the missing region while also maintaining coherence with the given context.We decouple this burden in our loss function by jointly training our context encoders to minimize both a reconstruction

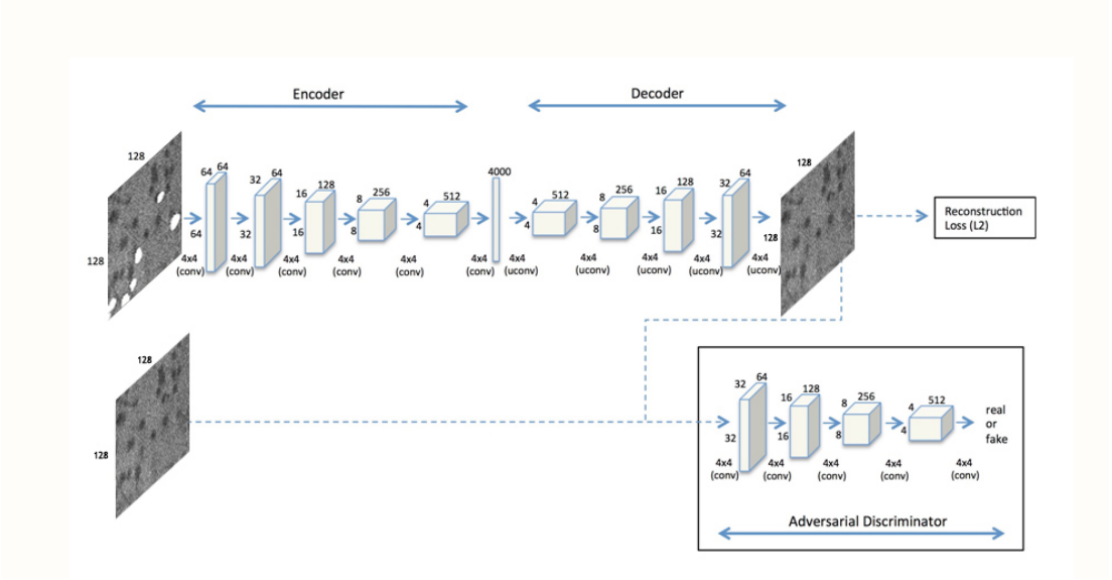

Figure 2.3: The above is the architecture for context based encoder. The first part of the architecture takes in the image patch as input and performs a set of convolution operation to finally reach a 4000×1 latent space vector. Then a set deconvolution operations follow to map the latent space vector to the inpainted image. We use both $l-2$ loss with respect to the ground truth, as well as adversarial loss to make the output more natural image like.

loss and an adversarial loss. The reconstruction (L2) loss captures the overall structure of the missing region in relation to the context, while the the adversarial loss [16] has the effect of picking a particular mode from the distribution. Figure 1 shows that using only the reconstruction loss produces blurry results, whereas adding the adversarial loss results in much sharper predictions. We define the overall loss function as

$$\mathcal{L} = \lambda_{rec}\mathcal{L}_{rec} + \lambda_{adv}\mathcal{L}_{adv} \tag{2.3}$$

We use a normalized masked L2 distance as our reconstruction loss function,$\mathcal{L}_{rec}$,where $\odot$ is the element-wise product operation. While this simple loss encourages the decoder to produce a rough outline of the predicted object, it often fails to capture any high frequency detail. This stems from the fact that the L2 loss often prefer a blurry solution, over highly accurate textures. We believe this happens because it is much âĂIJsaferâĂİ for the L2 loss to predict the mean of the distribution, because this minimizes the mean pixel-wise error, but results in a blurry averaged image. We alleviated this problem by adding an adversarial loss.

$$\mathcal{L}_{rec} = \left\| \hat{M} \odot (x - F((1-\hat{M}) \odot x)) \right\|_2^2 \tag{2.4}$$

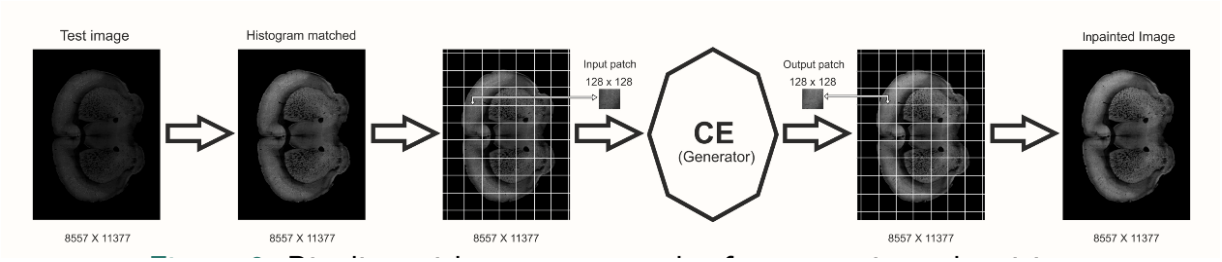

Figure 2.4: The above diagram shows the pipeline of our method. Given a test image, we first perform histogram matching on it to bring it to the intensity distribution for which our network was trained for. We then extract patches from the histogram-matched image and feed it to our network - context based encoder. The in-painted patches that are obtained from the network are then put back in their corresponding positions so as to get back our high dimensional in-painted image.

Our adversarial loss is based on Generative Adversarial Networks (GAN).To learn a generative model G of a data distribution, GAN proposes to jointly learn an adversarial discriminative model D to provide loss gradients to the generative model.G and D are parametric functions (e.g., deep networks) where $G : ZtoX$ maps samples from noise distribution Z to data distribution X.The learning procedure is a two-player game where an adversarial discriminator D takes in both the prediction of G and ground truth samples, and tries to distinguish them, while G tries to confuse D by producing samples that appear as 'real' as possible. The objective for discriminator is logistic likelihood indicating whether the input is real sample or predicted one.To customize GANs for this task, one could condition on the given context information;i.e., the mask $\hat{M}$ However, conditional GANs don't train easily for context prediction task as the adversarial discriminator D easily exploits the perceptual discontinuity in generated regions and the original context to easily classify predicted versus real samples. We thus use an alternate formulation, by conditioning only the generator (not the discriminator) on context. We also found results improved when the generator was not conditioned on a noise vector. Hence the adversarial loss for context encoders,$\mathcal{L}_{adv}$ is

$$\min_{G} \max_{D} E_{x \epsilon \chi} \quad [log(D(x)) + log(1 - D(F((1 - \hat{M}) \odot x))] \tag{2.5}$$

CHAPTER 3

RESULTS

3.1 Results for the proposed FP phase retrieval method (cDIP)

3.1.1 Dataset

Images in INRIA Holidays dataset Jégou *et al.* (2008) are used for simulating objects with uncorrelated amplitude and phase. Images are first resized to 256×256, and randomly paired together with one as an object's intensity, and the other as its phase by linearly mapping its values between 0 to 2π. These objects are further divided into training and test splits. Each object is passed through FP forward model, with parameters depending on the amount of overlap, to obtain its corresponding 64×64 sized low resolution images. Data preparation for training includes channel wise stacking of each object's low resolution images, resizing them spatially to 256×256, and performing channel wise rescaling to have values between 0 and 1. Similar scaling is also done for the ground truth high resolution intensity and phase as well. Due to the lack of publicly available FP data, we could not test our algorithm on real data.

3.1.2 Performance

We compare our results with the commonly used traditional phase retrieval algorithms such as Alternating Projections Gerchberg (1974)Fienup (1978), and Wirtinger FlowBian *et al.* (2015). Among these methods, due to lack of space, we show our comparisons only for Alternate Projections (AP) as AP's performance is better or at par with Wirtinger Flow in most cases. We also show comparisons with PtychNet Kappeler *et al.* (2017), a deep learning based phase retrieval algorithm that was originally proposed for intensity reconstruction only. Hence for comparison, we train two separate networks with PtychNet architecture, one for phase and one for intensity. We use Peak Signal to

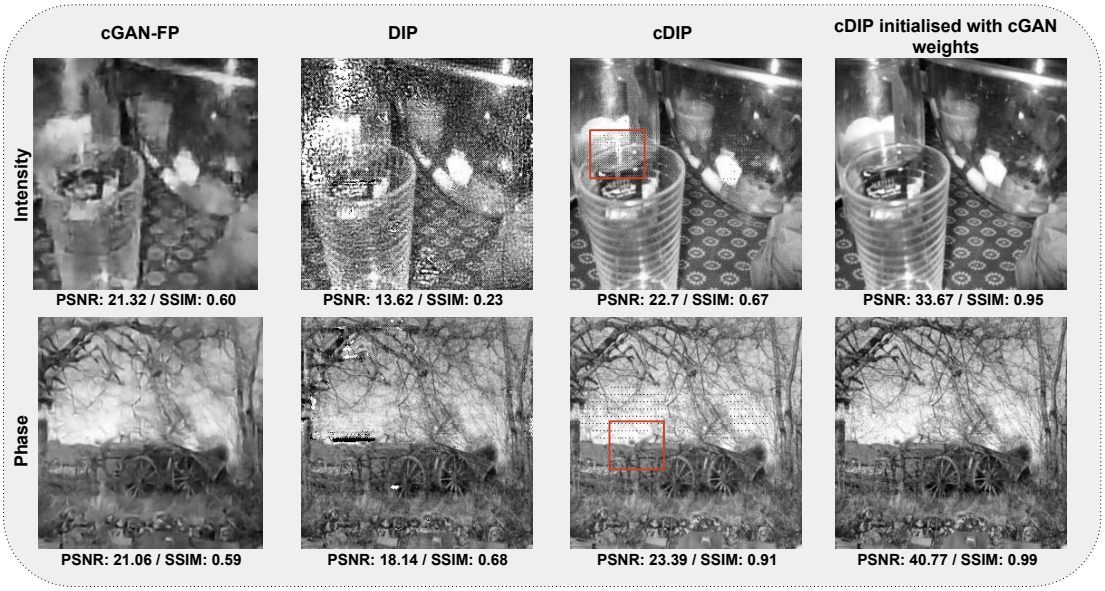

Figure 3.1: *Different choices of initialization and inputs for our generator network in the 40% overlap case:* There's significant increase in performance in terms of PSNR and SSIM when cGAN-FP is used for initializing cDIP training. However, even without any initialization cDIP gives perceptually good results with minor line artifacts, see the red bounding box. DIP Ulyanov *et al.* (2017) while being able to reconstruct sharper features as compared to cGAN-FP, suffers from considerable amount of phase-amplitude leakage. cGAN-FP on the other hand doesn't suffer much artifacts, but lack sharp features.

Noise Raio (PSNR) and Sructural Simmilarity (SSIM) as our evaluation metrics. PSNR and SSIM calculation for PtychNet results were done only on the central 240×240 region, avoiding unwanted border effects mentioned in PtychNet.

In Figure 3.1, we compare different choices of initialization and inputs for our generator network in the 40% overlap case. The first column shows cGAN-FP result trained for this 40% overlap. The second column shows result for Deep Image Prior Ulyanov *et al.* (2017) (DIP), which is a training-free approach based on minimizing the observation error. In DIP, both the input and the generator parameters are randomly initialized. Third column shows the result from conditional DIP (cDIP), where instead of giving noise as input, we use the set of observed low resolution images as input. The last column shows result for cDIP with cGAN-FP initialization where instead of a randomly initializing the generator paramters, we initialize it with the cGAN-FP parameters trained for 40% overlap. We observe significant increase in performance in terms of PSNR and SSIM when such an initialization is used. However, even without any initialization cDIP gives perceptually good results but with minor line artifacts (as

shown in red bounding box). DIP while being able to reconstruct sharper features as compared to cGAN-FP, suffers from considerable amount of phase-amplitude leakage. cGAN-FP on the other hand doesn't suffer much artifacts, but lacks sharp features.

As shown in Figure 3.2, we observe that Alternate Projection (AP) algorithm is able to achieve perceptually better intensity reconstruction than PtychNet, but suffers from phase-amplitude leakage as shown in the region bounded by red box. Also, AP's phase

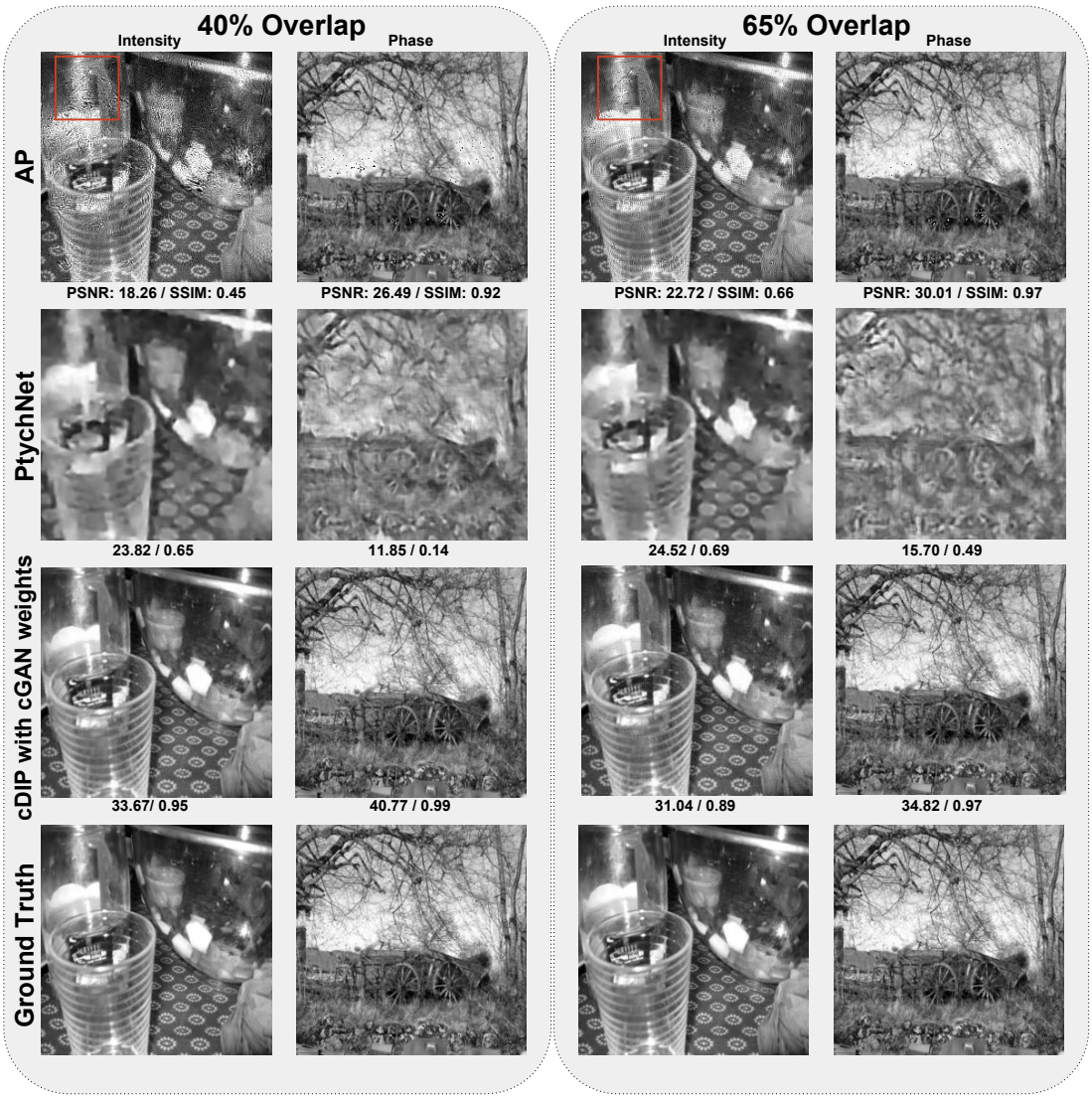

Figure 3.2: *Phase retrieval for high overlap case (40% and 65%):* Alternate Projection (AP) algorithm is able to achieve perceptually better intensity reconstruction than PtychNet, but suffers from phase-amplitude leakage as shown in the region bounded by a red box. PtychNet does not do as well as AP because it does not take into account the forward model loss during reconstruction. Our method gives the best results, as it exploits both the measurements and the image prior induced by generator network's structure.

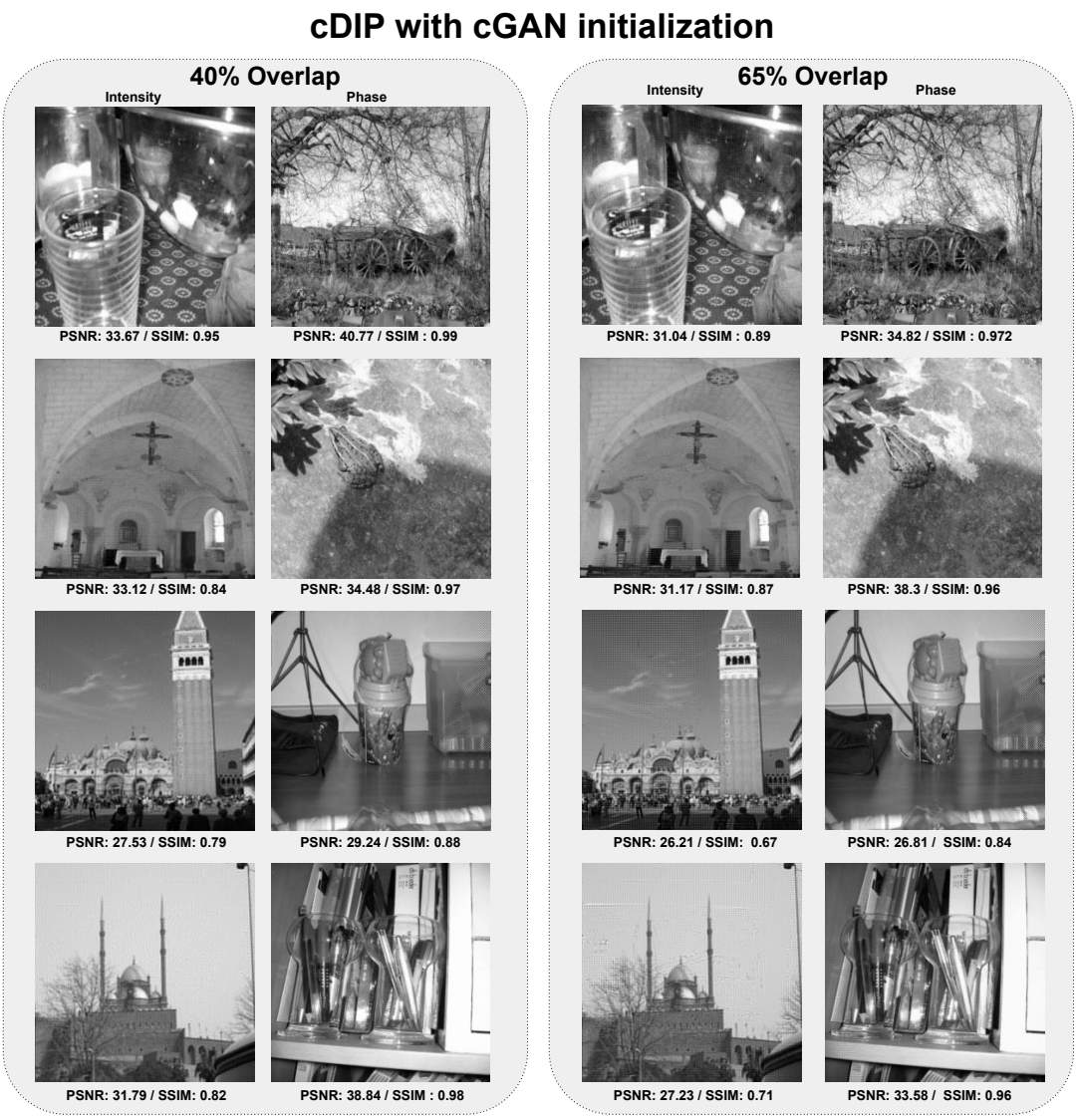

Figure 3.3: In the above figure, we show results of our algorithm for various test samples for 65% overlap case.

reconstruction is much better than PtychNet's. This is expected as PtychNet does not use the forward model loss during reconstruction. Our method gives the best results, as it exploits both the measurements and the image prior induced by the generator network's structure.

3.2 Proposed denoising method for high density gene expression in whole mouse brain images

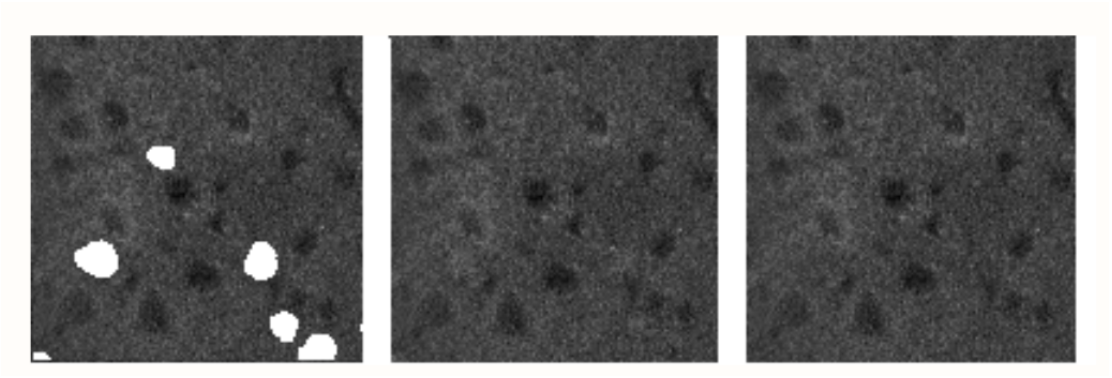

Figure 3.4: *Patch in-painting*: In the left is the patch to be in-painted, at the center is the in-painted patch we obtained using our approach, and at the right is the ground truth image.

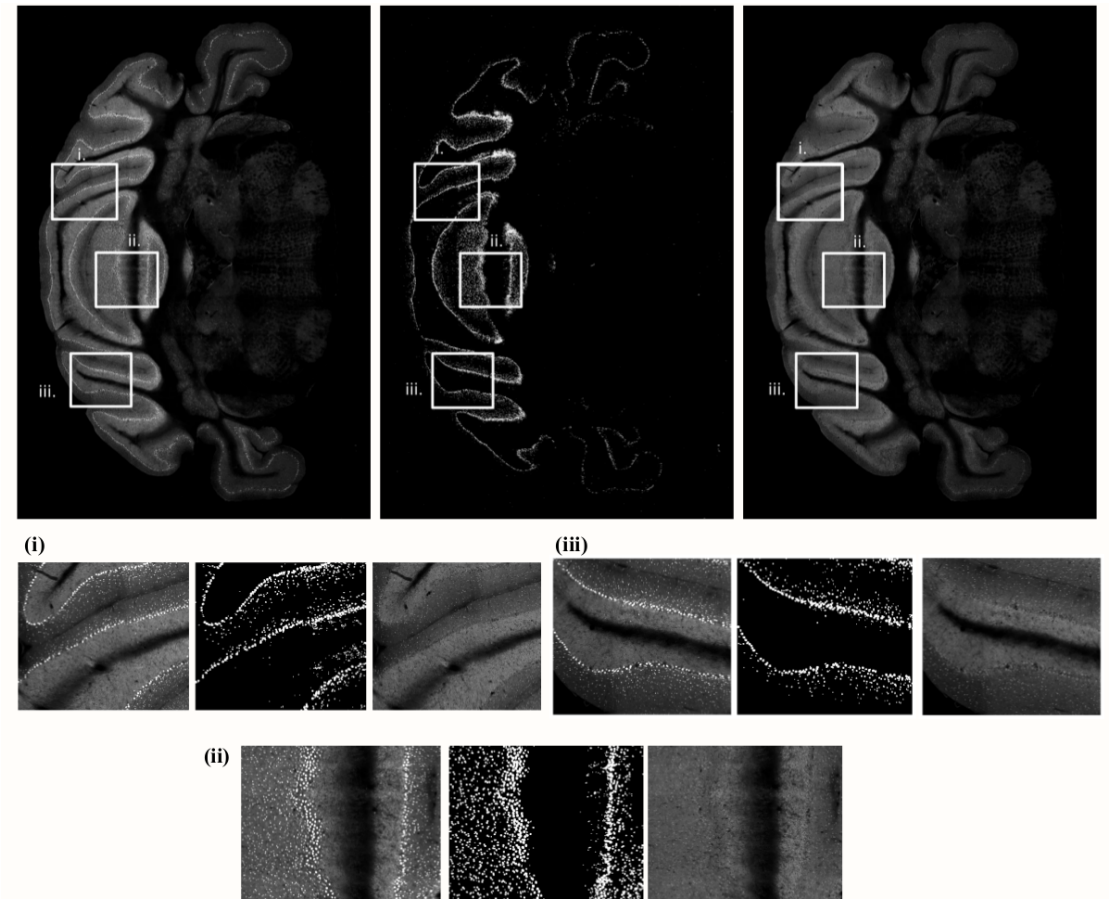

Figure 3.5: *Brain slice in-painting - example 1:* In the first row, at the left is the captured image of a brain slice. At the center is the mask that indicates the network which pixels (white dots) are to be in-painted.The image at right is the result we obtained. The bottom two rows show zoomed patches of regions shown in the first row.

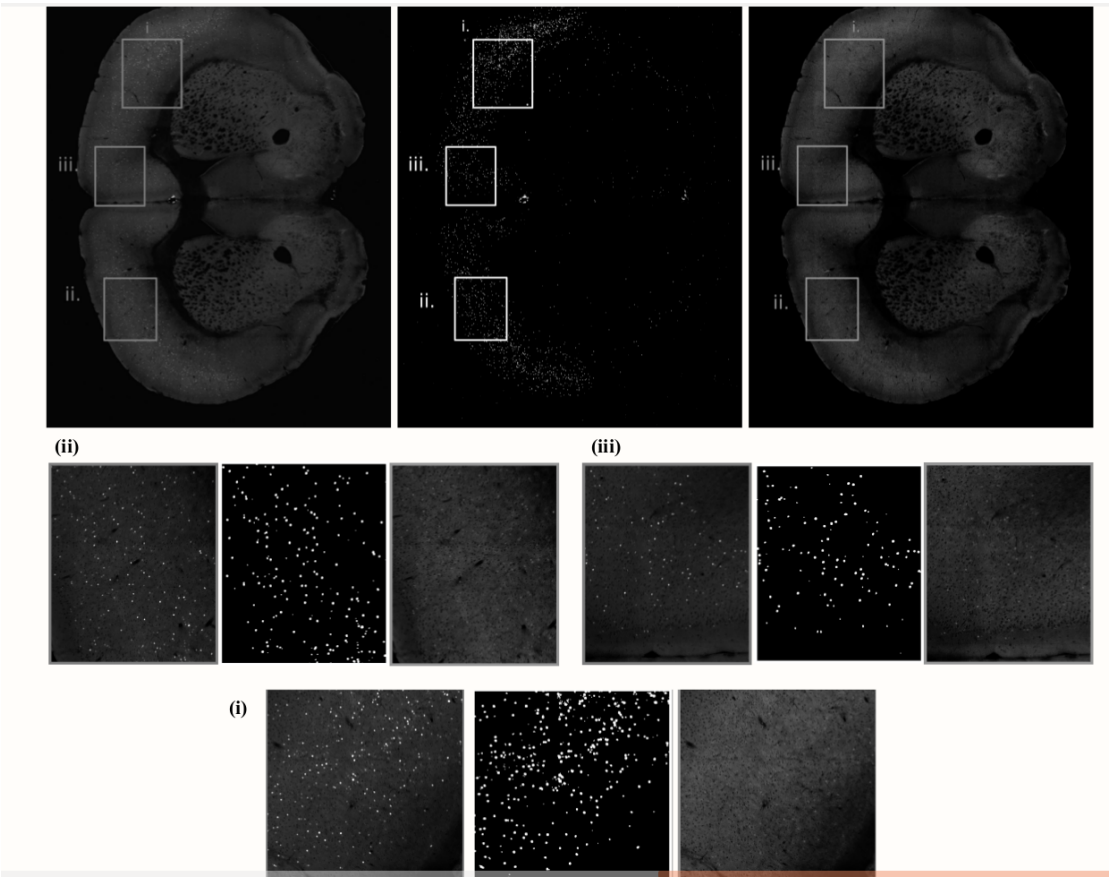

Figure 3.6: *Brain slice in-painting - example 2:* We show results for another test image as in the previous figure. In the first row, at the left is the captured image of a brain slice. At the center is the mask that indicates the network which pixels (white dots) are to be in-painted.The image at right is the result we obtained. The bottom two rows show zoomed patches of regions shown in the first row.

Figure 3.7

CHAPTER 4

Conclusion

We explored the application of deep learning for two specific problems in the area of microscopic imaging. The first one was for phase retrieval in Fourier ptychography microscopy, and the second for denoising in high density gene expression in whole mouse brain images.

For solving the problem of phase retrieval in Fourier ptychography, we propose a deep learning based method uses the same generator network that was used for cGAN-FP, but under a different setting where weights are updated based on loss with respect to measurements. Using simulations on uncorrelated phase and amplitude under both low and high overlap cases, we show that our algorithm outperforms the commonly used techniques for FP phase retrieval.

For the problem of denoising in high density gene expression in whole mouse brain images, we proposed a context-encoder based approach. Such an architecture predicts values for pixels to be in-painted, by exploiting the context information around them. The trained network was able to produce naturally looking images, and our method is currently being used for improving registration accuracy in the openSTP pipeline.

REFERENCES

1. **Bian, L.**, **J. Suo**, **G. Zheng**, **K. Guo**, **F. Chen**, and **Q. Dai** (2015). Fourier ptychographic reconstruction using wirtinger flow optimization. *Optics express*, **23 4**, 4856–66.

2. **Burer, S.** and **R. D. Monteiro** (2003). A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Mathematical Programming*, **95**(2), 329–357.

3. **Candes, E. J.**, **Y. C. Eldar**, **T. Strohmer**, and **V. Voroninski** (2015). Phase retrieval via matrix completion. *SIAM review*, **57**(2), 225–251.

4. **Candes, E. J.**, **T. Strohmer**, and **V. Voroninski** (2013). Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, **66**(8), 1241–1274.

5. **Chang, J. R.**, **C. Li**, **B. Póczos**, **B. V. K. V. Kumar**, and **A. C. Sankaranarayanan** (2017). One network to solve them all - solving linear inverse problems using deep projection models. *CoRR*, **abs/1703.09912**. URL `http://arxiv.org/abs/1703.09912`.

6. **Dave, A.**, **A. Kumar**, **Vadathya**, and **K. Mitra**, Compressive image recovery using recurrent generative model. *In 2017 IEEE International Conference on Image Processing (ICIP)*. 2017.

7. **Dave, A.**, **A. K. Vadathya**, **R. Subramanyam**, **R. Baburajan**, and **K. Mitra** (2018). Solving inverse computational imaging problems using deep pixel-level prior. *CoRR*, **abs/1802.09850**. URL `http://arxiv.org/abs/1802.09850`.

8. **Fienup, J. R.** (1978). Reconstruction of an object from the modulus of its fourier transform. *Optics letters*, **3**(1), 27–29.

9. **Gerchberg, R.** (1974). Super-resolution through error energy reduction. *Optica Acta: International Journal of Optics*, **21**(9), 709–720.

10. **Holloway, J.**, **M. S. Asif**, **M. K. Sharma**, **N. Matsuda**, **R. Horstmeyer**, **O. Cossairt**, and **A. Veeraraghavan** (2016). Toward long-distance subdiffraction imaging using coherent camera arrays. *IEEE Transactions on Computational Imaging*, **2**(3), 251–265.

11. **Holloway, J.**, **Y. Wu**, **M. K. Sharma**, **O. Cossairt**, and **A. Veeraraghavan** (2017). Savi: Synthetic apertures for long-range, subdiffraction-limited visible imaging using fourier ptychography. *Science Advances*, **3**(4), e1602564.

12. **Horstmeyer, R.**, **R. Y. Chen**, **X. Ou**, **B. Ames**, **J. A. Tropp**, and **C. Yang** (2015). Solving ptychography with a convex relaxation. *New journal of physics*, **17**(5), 053044.

13. **Jégou, H.**, **M. Douze**, and **C. Schmid** (2008). Hamming embedding and weak geometry consistency for large scale image search-extended version.

14. **Kappeler, A.**, **S. Ghosh**, **J. Holloway**, **O. Cossairt**, and **A. Katsaggelos**, Ptychnet: Cnn based fourier ptychography. *In Image Processing (ICIP), 2017 IEEE International Conference on*. IEEE, 2017.

15. **Metzler, C. A.**, **P. Schniter**, **A. Veeraraghavan**, and **R. G. Baraniuk** (2018). prdeep: Robust phase retrieval with flexible deep neural networks. *arXiv preprint arXiv:1803.00212*.

16. **Ou, X.**, **G. Zheng**, and **C. Yang** (2014). Embedded pupil function recovery for fourier ptychographic microscopy. *Optics express*, **22**(5), 4960–4972.

17. **Recht, B.**, **M. Fazel**, and **P. A. Parrilo** (2010). Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM review*, **52**(3), 471–501.

18. **Rivenson, Y.**, **Y. Zhang**, **H. Gunaydin**, **D. Teng**, and **A. Ozcan** (2017). Phase recovery and holographic image reconstruction using deep learning in neural networks. *arXiv preprint arXiv:1705.04286*.

19. **Tian, L.**, **X. Li**, **K. Ramchandran**, and **L. Waller** (2014). Multiplexed coded illumination for fourier ptychography with an led array microscope. *Biomedical optics express*, **5**(7), 2376–2389.

20. **Ulyanov, D.**, **A. Vedaldi**, and **V. S. Lempitsky** (2017). Deep image prior. *CoRR*, **abs/1711.10925**. URL `http://arxiv.org/abs/1711.10925`.

21. **Yeh, L.-H.**, **J. Dong**, **J. Zhong**, **L. Tian**, **M. Chen**, **G. Tang**, **M. Soltanolkotabi**, and **L. Waller** (2015). Experimental robustness of fourier ptychography phase retrieval algorithms. *Optics express*, **23**(26), 33214–33240.

22. **Zheng, G.**, **R. Horstmeyer**, and **C. Yang** (2013). Wide-field, high-resolution fourier ptychographic microscopy. *Nature photonics*, **7**(9), 739–745.