UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

NOVEL TECHNIQUES IN FOURIER PTYCHOGRAPHIC MICROSCOPY FOR MEDICAL IMAGING

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirement for the

Degree of

DOCTOR OF PHILOSOPHY

By

KE ZHANG
Norman, Oklahoma
2024

NOVEL TECHNIQUES IN FOURIER PTYCHOGRAPHIC MICROSCOPY FOR MEDICAL IMAGING

A DISSERTATION APPROVED FOR THE

STEPHENSON SCHOOL OF BIOMEDICAL ENGINEERING

BY THE COMMITTEE CONSISTING OF

Dr. Yuchen Qiu, Chair

Dr. Qinggong Tang

Dr. Javier Jo

Dr. Zhibo Yang

To my family and friends, for their unwavering trust and support.

Acknowledgments

I would first like to express my sincere gratitude to my advisor, Dr. Yuchen Qiu, for his guidance and support from the moment I first stepped into his office. I am grateful for the opportunity to join his lab, where I gained specialized knowledge in biomedical optics and medical image analysis, as well as a solid understanding of general research methodologies, under his mentorship. Working alongside him allowed me to observe firsthand how a researcher successfully conducts projects with problem-solving skills. These experiences have shaped me into a well-rounded engineer with enhanced research capabilities and a solution-oriented mindset, ready to tackle complex challenges in both research and daily life.

I would also like to convey my deep gratitude to my committee members, Dr. Qinggong Tang, Dr. Javier Jo, and Dr. Zhibo Yang, for their dedicated time, unwavering support, and invaluable guidance throughout my research journey. Dr. Qinggong Tang encouraged me to submit my application to the biomedical engineering PhD program, a decision that ultimately changed the course of my life. He has been a constant source of guidance, always reliable and happy to answer my questions, whether academic or personal. I deeply appreciate the solid foundation of help he has provided throughout. Dr. Javier Jo maintains high standards in research, and his insightful comments consistently inspired me to reflect deeply on experimental results and seek ways to refine them further. I am especially grateful to Dr. Zhibo Yang for his willingness to serve as my Graduate College representative close to my graduation, and he has been incredibly responsive to my requests throughout the process. Additionally, I wish to express my heartfelt gratitude to Dr.

Bin Zheng and Dr. Si Wu, who served on my committee previously. Their initial guidance and support played a crucial role in directing my research and have had a lasting impact on my academic journey. I am also deeply grateful to Dr. Rongzhu Gan, Dr. Michael Detamore, and Dr. Jay McDaniel for their invaluable support during times of difficulty.

Many thanks go to all the lab members: Dr. Xuxin Chen, Patrik Gilley, and Neman Abdoli, for their collaboration, support, and the caring environment they created. Special thanks to Xuxin for his invaluable assistance with deep learning techniques and his constant support in daily life. I am especially grateful for his patience in explaining intricate concepts to me and for always helping whenever I needed it.

Thanks to the incredible companionship and support of my friends: Dr. Feng Yan, Dr. Chen Wang, Dr. Shangyuan Jiang, Dr. Fan Zhang, Qinghao Zhang, Trisha Valerio, Jeesoo Park, Musarrat Amin, Shulong Xu, Taozhe Li, Mingwei Huang, and Yang Yan. Without them, this journey would have been far more difficult and lonelier. Their constant encouragement and genuine care kept me grounded, turning challenging days into moments of warmth.

Lastly, I want to express my deepest love and gratitude to my open-minded family for their unwavering support. My parents even began exploring books on topics like mindfulness to guide me through my dark moments. The tradition of exchanging our learning and insights every Friday night has become a cherished and unforgettable part of my life.

Table of Contents

List of Tables

List of Figures

Abstract

The emerging Fourier ptychographic microscopy (FPM) offers a highly efficient approach to enhance the throughput of conventional microscopy. By illuminating the samples with a series of incident angles, FPM computationally reconstructs the high-resolution sample images from the acquired low-resolution measurements. Thus, it overcomes the limitations of traditional optical imaging systems to simultaneously achieve large field of view and adequate spatial resolution. However, to maximize its potential, it is necessary to continuously optimize the performance and verify the clinical utility of FPM system. For this purpose, this dissertation is majorly composed of four different studies. In the first study, we initially explore the feasibility of implementing FPM to reconstruct metaphase chromosomes, aiming to enhance the imaging efficiency by mitigating the trade-off between field of view (FOV) and resolution in conventional microscopic systems. The second study focuses on improving the efficiency of data acquisition of FPM. By employing symmetric illumination and a color detector to accelerate the process, we can potentially increase the data acquisition speed up to 12 times. In the third study, we develop and evaluate a doublet based FPM prototype to further advance the design of more affordable FPM configurations. This prototype replaces the standard 4× objective lens with a commercial achromatic doublet lens of 60mm focal length. Using a 150mm tube lens and a 15×15 LED array, it achieves a theoretical equivalent NA of 0.4. The fourth study is designed to measure the DOF of FPM systems, providing essential insights for the future development of FPM based digital pathology scanners. The measurements follow the principle that DOF is the range along optical axis where the contrast value remains at or above 80% of the maximum as the focus is altered. The corresponding contrast value for

each focus position is estimated based on the specific bar pattern where the contrast value of the in-focus MTF curve drops to 0.5. The four studies on FPM investigate its applications, technological advancements, and evaluations, providing a comprehensive exploration of this new technology's potential and supporting its translation toward practical implementation.

In addition to the FPM studies, the appendix presents two projects dedicated to computer-aided diagnosis (CAD) scheme developments. These studies utilize radiomics and deep learning techniques to enhance the prediction of chemotherapy outcomes in ovarian cancer treatment with CT scans. The investigations underscore the value of integrating analytical approaches with medical imaging modalities, illustrating how CAD contributes to personalized healthcare through improved diagnostic and prognostic capabilities.

1 Introduction

1.1 Conventional Microscopy's Inherent Limits

Optical microscope aims to magnify small objects that are invisible to the naked eye, which is a highly valuable tool in a variety of fields such as biology, medicine, and material science. In contemporary optical microscopy, the infinity-corrected optical design is typically adopted [1]: an objective lens generates a parallel beam to project the specimen into infinity, which enters into a tube lens and converges as the intermediate image for the following magnification or digitization. Accordingly, the overall magnification of the whole system is determined by the ratio of the focal length of the tube lens to the focal length of the objective lens: $M = f_{tube}/f_{obj}$.

However, when operating the microscopes, the observer must frequently switch between the low and high magnification objective lenses: a low magnification objective lens offers a broad view but fewer details, whereas a high magnification lens provides adequate details but limited field of view (FOV). This tradeoff between resolving power and FOV actually results in a relatively fixed space-bandwidth-product (SBP) which counts the number of optically resolvable or effective pixels / points within the FOV. Despite the high SBP objective lens being commercially available, directly increasing the optical SBP suffers from the costly manufacture and complicated optical architecture [2].

Meanwhile, in digital pathology, the current specimen digitizer integrates a high-precision automated mechanical translation stage and dynamic focusing adjustment with a conventional microscope. During sample digitization, the translational stage moves the sample under the objective lens, capturing and storing each corresponding image. A large

number of image patches are acquired to cover the entire slide. Despite the success, the current system has a number of disadvantages including large size, high cost, and low efficiency, which limits its applications in many resource-constrained scenarios.

1.2 Fourier Ptychography Principle and Image Reconstruction

Instead of using expensive and intricate lenses or mechanical scanning, the recently developed Fourier ptychography microscopy (FPM) employs an innovative approach [3] that integrates angled illumination and computational reconstruction algorithm to achieve high resolution and large FOV imaging simultaneously.

In the frequency domain, the objective lens acts as a low pass spatial filter $P(\boldsymbol{k})$, (Figure 1-1a, red), with a cutting off frequency or radius set by the NA of the lens and the wavelength of the incident light: $2\pi NA/\lambda$ [4]. As a result, the complex sample $s(x,y)$ containing both amplitude and phase appears as a blurred image $|\mathcal{F}^{-1}[S(\boldsymbol{k})P(\boldsymbol{k})]|^2$ on the detector plane, as presented in Figure 1-1b. With (partially) coherent light illuminating the sample at an incidence angle θ_x along the *X* direction (Figure 1-1a, blue) instead of normal incidence, an additional phase factor is projected onto the sample: $s(x,y)e^{ik_x x}$, k_x=$2\pi Sin\theta_x/\lambda$, shifting corresponding spectrum to be $\boldsymbol{S}(k-k_x)$. The objective aperture then filters the spectrum and retains information within the same circular-shaped area, now centered at k_x. Therefore, by sequentially illuminating the sample at different angles of incidence, the FPM imaging system can collect frequency information from various segments of the sample's spectrum, with each segment recorded in a separate raw image (Figure 1-2). The radius of sampled spectrum can be extended to $2\pi(Sin\theta_{max}+NA)/\lambda$, with θ_{max} representing the maximum angle of incidence. Thus the cutting off frequency of the FPM system can be considered as $2\pi(Sin\theta_{max}+NA)/\lambda$.

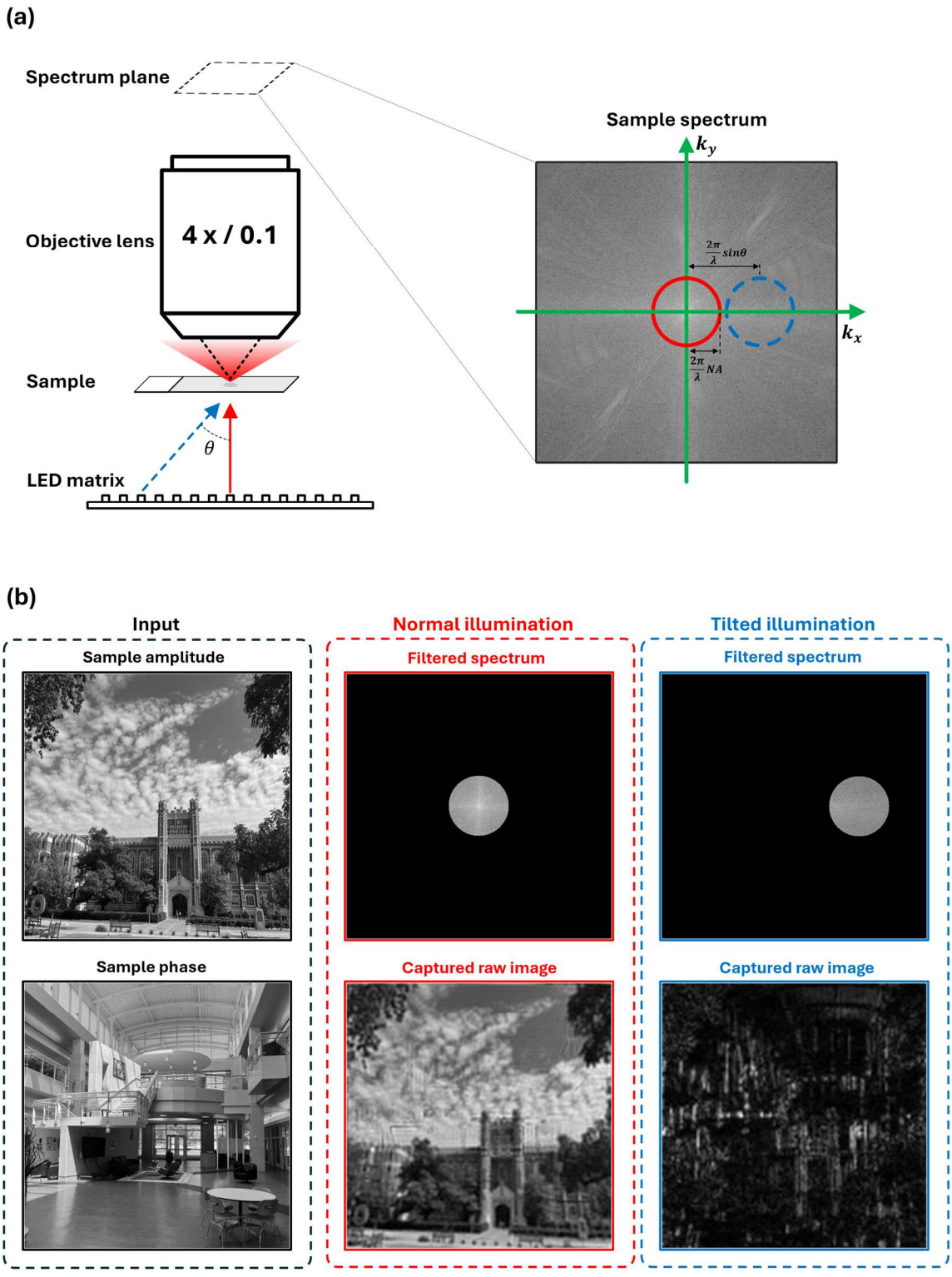


Figure 1-1: The principle of FPM technology. (a) Under normal (red) and angled (blue) illumination, the objective lens can capture different regions of the sample spectrum. (b) A

numerical simulation illustrates the captured raw images with (red) and without (blue) tilted illumination.

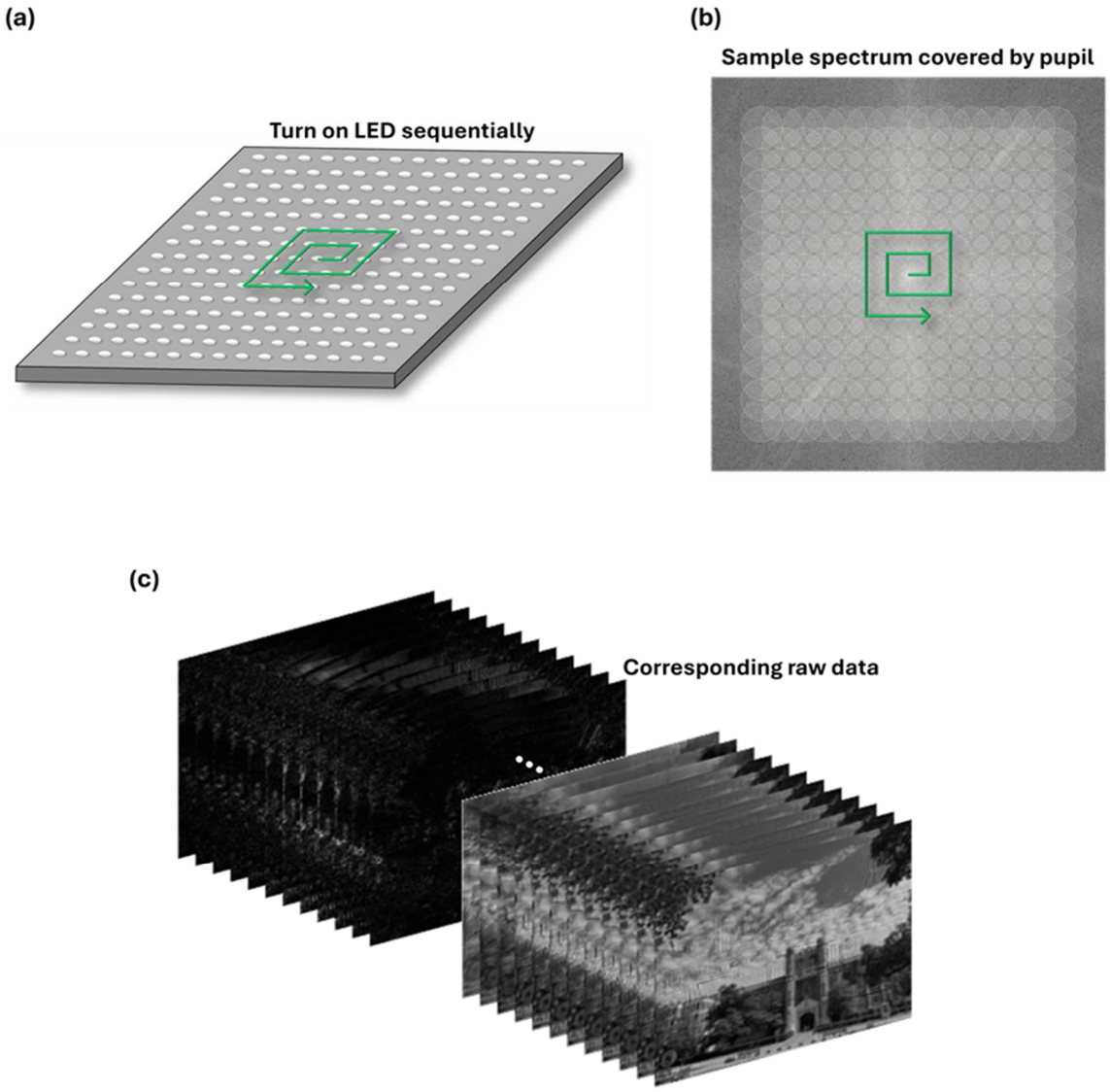


Figure 1-2: FPM captures raw images for reconstruction through tilted illumination. (a) LEDs are activated sequentially to generate tilted illumination from various angles. (b) The pupil aperture covers a larger area of the sample spectrum through shifting. (c) The captured raw images are used for reconstruction.

Accordingly, reconstructing high resolution images from the acquired raw patterns is equivalent to reconstructing the corresponding spectrum. Given that the complex spectrum requires both intensity and phase information but the detector only records the sample intensity, the missing phase information needs to be restored (phase retrieval [5]) for the spectrum reconstruction. Moreover, pupil function also needs to be estimated by the algorithm to minimize the influence of aberrations and optimize the quality of the finally reconstructed image. Therefore, in the following studies, the image reconstruction is based on a modified phase restoration algorithm embedded with pupil function recovery (EPRY) [6, 7], which iteratively updates the reconstructed spectrum until the preset accuracy tolerance is satisfied. In this method, the entire spectrum of the reconstructed image is first divided into *N* distinct but partially overlapping subregions, with each subregion corresponding to one individual captured raw pattern. During each iteration, the spectrum and the pupil functions are updated in sequence, which is performed subregion by subregion. The updating process in each iteration is composed of the following three steps. First, pupil function $P_t(\boldsymbol{k})$ is applied on the subregion spectrum $S_t^n(\boldsymbol{k}-\boldsymbol{k_n})$: $\Psi_t^n(\boldsymbol{k}) = S_t^n(\boldsymbol{k}-\boldsymbol{k_n})P_t(\boldsymbol{k})$, where t denotes the t^{th} iteration and n denotes the n^{th} subregion. In addition, $\boldsymbol{k_n}$ is defined as $(2\pi Sin\theta_{xn}/\lambda, 2\pi Sin\theta_{yn}/\lambda)$, and $(\theta_{xn}, \theta_{yn})$ are the illumination angles in *X* and *Y* directions, respectively. The spatial representation $\psi_t^n(\boldsymbol{r}) = \mathcal{F}^{-1}[\Psi_t^n(\boldsymbol{k})]$ is computed by inverse Fourier transform correspondingly. Next, the captured raw pattern $I_{raw}^n(\boldsymbol{r})$ is applied on the spatial representation as the constraint to update the amplitude of the spatial representation: $\Phi_t^n(\boldsymbol{r}) = \sqrt{I_{raw}^n(\boldsymbol{r})}\frac{\psi_t^n(\boldsymbol{r})}{|\psi_t^n(\boldsymbol{r})|}$. And the phase part remains the same values. Finally, using Levenberg-Marquardt variant of the Gaussian Newton's method, the subregion spectrum $S_t^n(\boldsymbol{k}-\boldsymbol{k_n})$ and pupil function $P_t(\boldsymbol{k})$

are updated respectively, to improve the algorithm robustness [7, 8]: $S_{t+1}(\boldsymbol{k})$=$S_t(\boldsymbol{k})$ + $\frac{|P_t(\boldsymbol{k})|[P_t(\boldsymbol{k})]^*[\phi_t^n(\boldsymbol{k})-S_t(\boldsymbol{k})P_t(\boldsymbol{k})]}{|P_t(\boldsymbol{k})|_{max}(|P_t(\boldsymbol{k})|^2+\delta_1)}$ and $P_{t+1}(\boldsymbol{k}) = P_t(\boldsymbol{k}) + \frac{|S_t(\boldsymbol{k})|[S_t(\boldsymbol{k})]^*[\phi_t^n(\boldsymbol{k})-S_t(\boldsymbol{k})P_t(\boldsymbol{k})]}{|S_t(\boldsymbol{k})|_{max}(|S_t(\boldsymbol{k})|^2+\delta_2)}$. δ_1 and δ_2 are added to ensure the method stability. To achieve a satisfactory convergence, this process is repeated for *T* times. Figure 1-3 displays the reconstruction flowchart and the reconstructed amplitude and phase images in a numerical simulation. The parameters used in the simulation are listed in Table 1-1.

Table 1-1: Parameters used in the numerical simulation.

Parameters	**Values**
Numerical aperture	0.1
Illumination wavelength	0.53 μm
LED pitch	4 mm
LED array size	15 × 15
Distance between LED array and sample	90 mm
Pixel size	2.4 μm

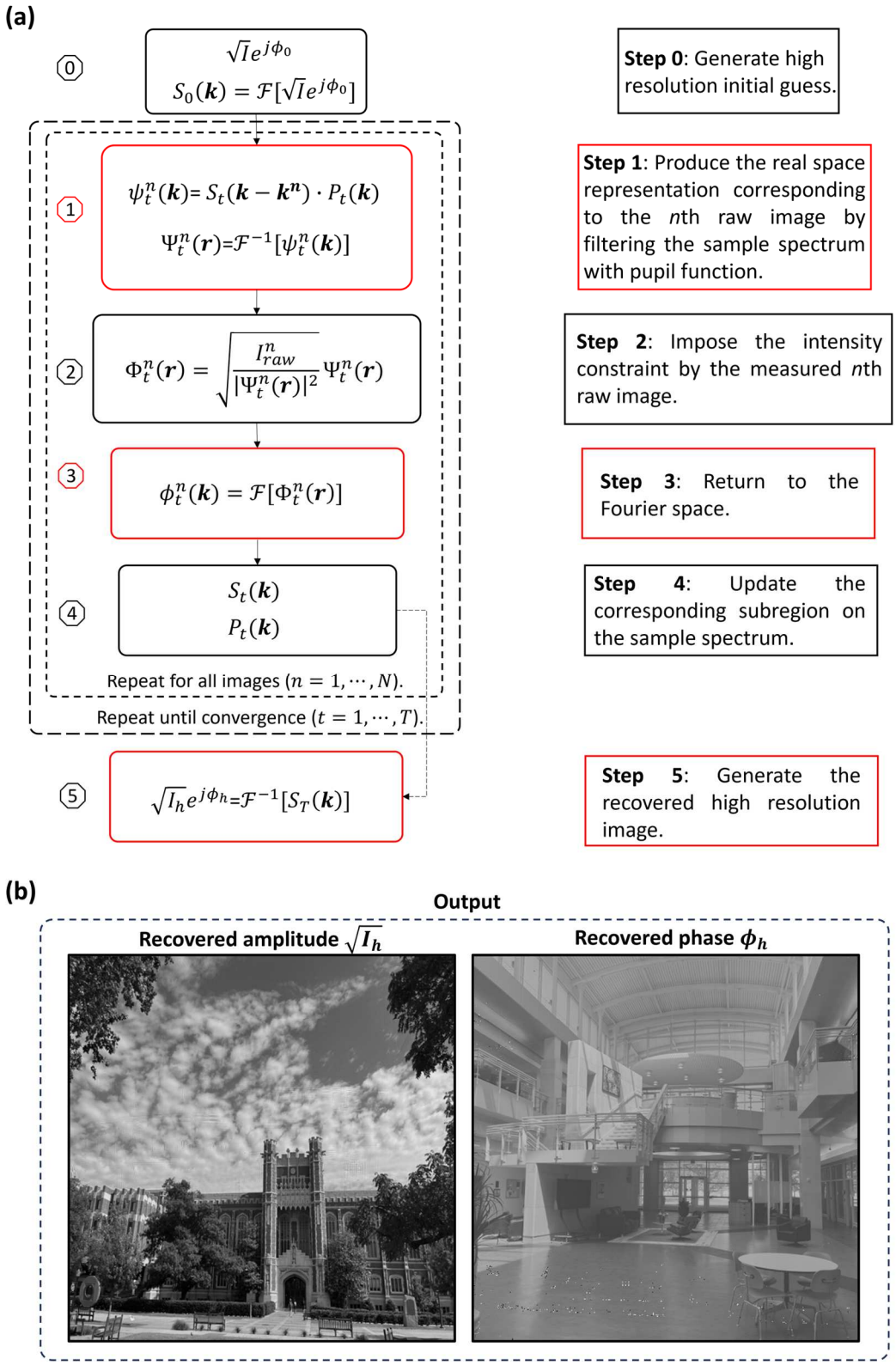


Figure 1-3: Flowchart and simulation results of the FPM reconstruction scheme. (a) The reconstruction process uses all N captured raw images to update corresponding subregions

of the spectrum in each iteration. The iteration is repeated T times to achieve convergence. (b) The reconstructed amplitude and phase results of the numerical simulation, using raw images provided in Figure 1-2c.

In summary, Fourier ptychography microscopy is one kind of aperture synthesis technology [9, 10]. By employing a low cost LED array and a low magnification lens to capture raw patterns from various illumination angles, FPM acquires richer spectral information from the sample. It computationally recovers high resolution images and overcomes the tradeoff between resolution and FOV in an optical microscope, offering a new approach to access high performance imaging without the need for expensive optics.

1.3 Outline of the Dissertation

This dissertation features 4 studies on FPM, exploring different aspects and approaches. Chapter 2 focuses on the use of FPM to achieve high speed metaphase chromosome imaging [11, 12], eliminating the need to switch between conventional objective lenses of different magnifications in chromosome searching. In Chapter 3 we propose a symmetric illumination based color FPM scheme that reduces the time required for capturing raw images and produces high quality color images [13]. Chapter 4 aims to further reduce the cost of FPM by replacing the conventional objective lens with a doublet lens. The prototype achieves comparable resolution to the conventional FPM setup and resolves the tissue details in clinical samples successfully. Chapter 5 presents a systematic evaluation of depth of field (DOF) of FPM systems [14], serving as a reference for future development of FPM based digital pathology scanners. Finally, Chapter 6 summarizes the contributions of the four FPM studies and presents the conclusions.

2 Using Fourier Ptychography Microscopy to Achieve High Resolution Chromosome Imaging: An Initial Evaluation

2.1 Introduction

Leukemia is one category of blood cancers that commonly occur in both female and male patients [15]. For the diagnosis and treatment of leukemia, chromosome karyotyping [16] is critically important, which organizes and pairs all the chromosomes in the order of decreasing length. In this process, cytogeneticists have to search the whole specimen slide and identify the analyzable metaphase chromosome cells, which is a time-intensive and tedious operation. To reduce the clinician's workload and enhance diagnostic efficiency, many research studies have been focused on the development of automated slide scanners [17] and the associated computer-aided detection schemes [18, 19], aiming to accomplish quick cell digitization and identification. Despite some encouraging progress, the performance of the current scanners is majorly hurdled by the long scanning time, which can be attributed to the limited space-bandwidth product of the objective lens [20, 21]. With few high cost exceptions [22], this limitation indicates that it is very difficult for the objective lens to simultaneously accomplish a large field of view (FOV) and high spatial resolution, which implies that high resolution objective lenses will have a much smaller FOV than the low resolution lens. Considering that the chromosome images must be obtained with high spatial resolution to ensure enough details for clinical karyotyping, the corresponding high magnification scanning will be significantly slower as each acquisition only covers a much smaller area of the specimen.

Meanwhile, the recent technology of Fourier ptychography microscopy [3, 23] is an emerging strategy to address the challenge of limited space-bandwidth products. This

method first collects a number of low magnification images under different sample illumination conditions, and then synthesizes them together to reconstruct a high resolution image. Since the data acquisition is under low magnification, it inherits the advantage of low resolution scanning systems, namely, the large FOV and large DOF (i.e., depth of field [24]). For example, in current clinical practice, the 100×/1.25 NA objective lens is used to ensure the band pattern sharpness. Given the typical objective lens with an optical field number (OFN) of 22, the corresponding FOV is approximately 0.22mm. Meanwhile, if FPM is utilized to achieve the same band pattern quality under 10×/0.25 NA lens, the FOV of the objective lens (with same OFN number) is enhanced to 2.2mm; thus the scanning efficiency can be largely enhanced. Moreover, with its large system DOF, FPM can also vastly reduce the precision requirement of the moving stage, and the total cost of the scanner can be significantly lowered. Although this new method has been applied in many different clinical scenarios, no prior research has been focused on investigating the feasibility of utilizing FPM for high resolution metaphase chromosome imaging.

In this study, we comprehensively assessed the performance of an FPM system and its clinical utility on chromosome imaging. The modulation transfer function (MTF) curves were first estimated, aiming to assess the system resolving power. Next, the analyzable metaphase chromosomes acquired from the peripheral blood and bone marrow samples of leukemia patients were imaged by FPM system, and the chromosome feature qualities were compared with the results captured by the corresponding conventional microscopes. All the experimental details are presented in the following sections.

2.2 Materials and Methods

2.2.1 Experimental Set-up

In this experiment, a Fourier ptychography based microscope was developed on the basis of Thorlabs Cerna Microscope platform (Thorlabs, New Jersey, USA), which is equipped with either 4×/0.1 NA， 10×/0.25 NA or 20×/0.40 NA objective lenses (Olympus, Tokyo, Japan) and an FL20BW CMOS camera (Tuscen, Fuzhou, China). The focal length of the platform tube lens is 150mm, thus the actual magnification is approximately 3.3×, 8.3×, and 16.7×, for the 4×/0.1 NA, 10×/0.25 NA, and 20×/0.40 NA objective lenses, respectively, as 180mm is the standard focal length for the Olympus infinity corrected optical system.

To accomplish the tilted illumination required in FPM, a 32×32 LED panel (Adafruit, New York, USA) was placed under the sample stage, with the vertical stage-sample distance being ~70mm, for the 4× and 10× objective lens configurations. The central 225 (15×15) green (4 mm spacing, central wavelength 0.53 μm) LEDs were used, which yields an illumination NA of ~0.38. Therefore, the equivalent system NA_{eq} is ~0.48 and ~0.63, respectively. For the 20× objective lens configuration, the same LED array was positioned ~47mm below the sample stage to generate green incident light. During the image collection, the 49 LEDs inside the central circular area with a radius of 48mm on the LED array were lit up in sequence to provide an equivalent system NA_{eq} of 1.11.

Our FPM system can also function as a regular microscope for 20× imaging (0.40 NA). For further comparison, high magnification imaging was conducted using a Nikon Ni transmissive microscope (Nikon, Tokyo, Japan) equipped with 60×/0.95 NA, 100×/1.25 NA, or 100×/1.45 NA objective lenses.

2.2.2 Use Standard Resolution Target and Clinical Specimen to Evaluate the Resolving Power of FPM System

To assess the resolving power, the modulation transfer function (MTF) curves of the FPM system were measured under three configurations (i.e., 4×, 10×, and 20×). As a comparison, we also measured the MTF curves of the conventional microscope equipped with 20×/0.4 NA, 60×/0.95 NA, 100×/1.25 NA, and 100×/1.45 NA objective lenses. Currently, which approach should be the standard metric to assess the performance of a coherent system has not been well determined as the object phase variations may enhance the system's power of resolving two closely spaced points [20]. Given that the standard resolution pattern does not contain any phase variance, our MTF-based performance assessment may actually underestimate the performance of the investigated FPM system. Meanwhile, it is widely accepted that the MTF curve can objectively evaluate (neither underestimate nor overestimate) the performance of incoherent imaging systems such as conventional microscopes. Therefore, when using the MTF curve, we would rather underestimate our FPM system than overestimate it, and the corresponding results of performance comparison remain solid.

Accordingly, a high resolution (up to 3649 lp/mm) USAF 1951 target (Newport, California, USA) was utilized, and the contrast values of the imaged bar patterns were measured at a sequence of discrete frequencies ranging from 0 to the maximum spatial resolution achieved by the system [25]. During the measurement, the test target was placed on the sample stage and adjusted to the in-focus position. For each pattern on the captured target, the contrast was calculated by [26]: $C = \frac{I_{max}-I_{min}}{I_{max}+I_{min}}$, where $I_{\max}$ and $I_{\min}$ are the average maximal and minimal values of the bar patterns, respectively. Based on these

estimated contrast values, a smooth normalized MTF curve was generated by a curve fitting algorithm [8].

To evaluate its clinical utility, a certain number of analyzable metaphase chromosomes were also imaged by our FPM system. All the cells were collected from the blood or bone marrow samples of leukemia patients and processed under standard clinical procedure in our medical center. The chromosomes were imaged by both the FPM system and the corresponding equivalent conventional microscopes (e.g., under 60×/0.95 NA, 100×/1.25 NA objective lenses). The chromosome band patterns depicted on each generated image were assessed and compared between these two groups of results (i.e., FPM vs conventional system).

2.3 Results

Figure 2-1 demonstrates the reconstructed images of a 1951 USAF resolution target. Figure 2-1a is directly imaged under the 4×/0.1 NA objective lens, which is able to resolve the patterns up to group 7-3 (161 lp/mm). This result generally agrees with the theoretical computation (154 lp/mm). As indicated in Figure 2-1b, when using FPM with the 4×/0.1 NA objective, the spatial resolution of the reconstructed image increases to 813 lp/mm (group 9-5), which is equivalent to the result directly captured under the 20×/0.4 NA objective lens (Figure 2-1c). Meanwhile, Figures 2-2a/b/c demonstrate the estimated MTF curves when utilizing the 4×/0.1 NA objective lens, FPM technology, and 20×/0.4 NA objective lens, respectively. As predicted by the Fourier optics theory [20], the normalized contrast value monotonically decreases as the spatial frequency increases. The MTF curve of the 4×/0.1 NA objective lens gradually declines between 0 and 150 lp/mm, followed by a sharp drop between 150 lp/mm and the cutoff frequency (approximately 237 lp/mm).

Meanwhile, the contrast value of the FPM measurement remains above 0.7 when the frequency reaches 600 lp/mm, with a cutoff frequency of 1066 lp/mm. The MTF curve of 20×/0.4 NA objective lens shows a resemblance to the FPM results, as demonstrated in Figure 2-2c.

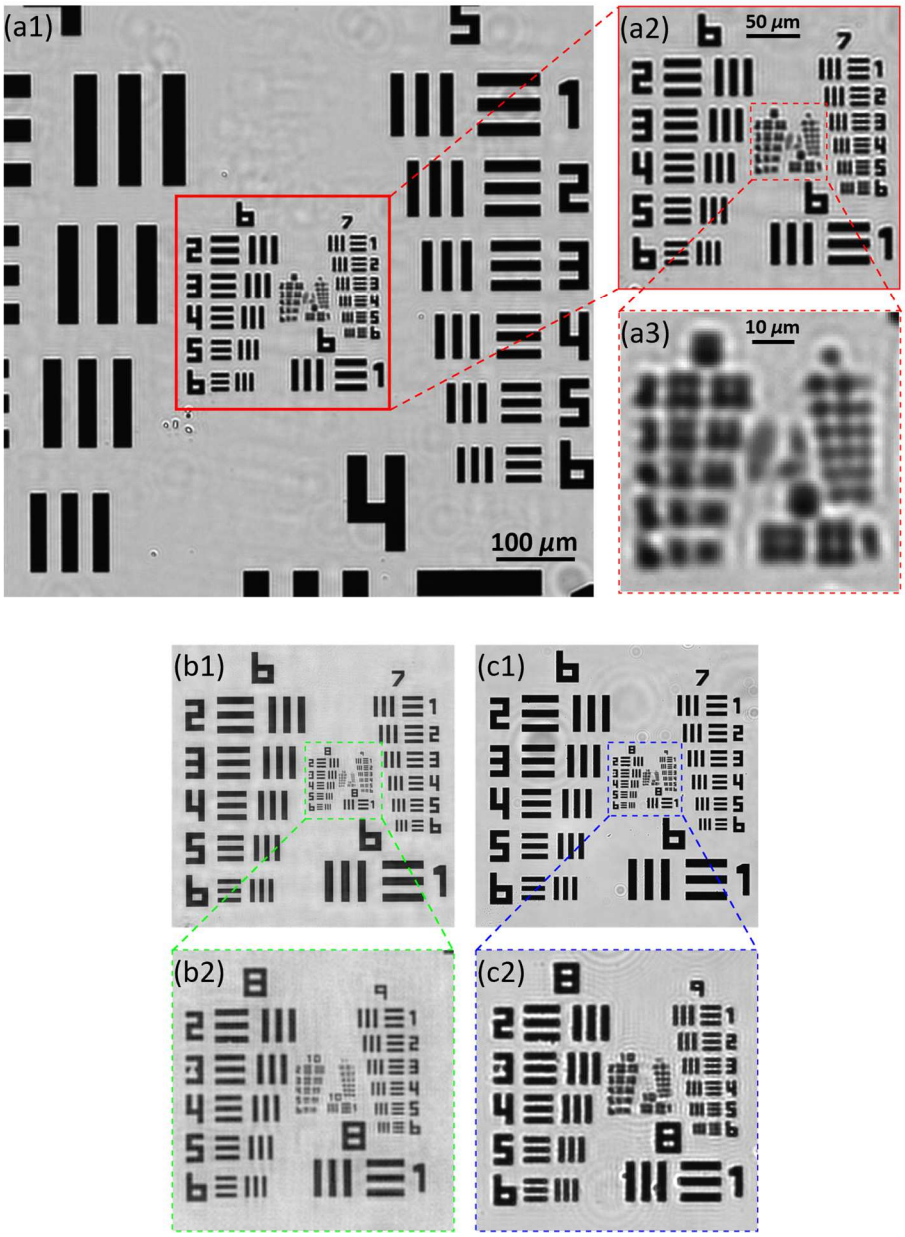


Figure 2-1: The images of a standard USAF 1951 resolution target, generated by (a) 4×/0.1 NA objective lens; (b) FPM technology with 4×/0.1 NA objective lens; and (c) 20×/0.4 NA objective lens.

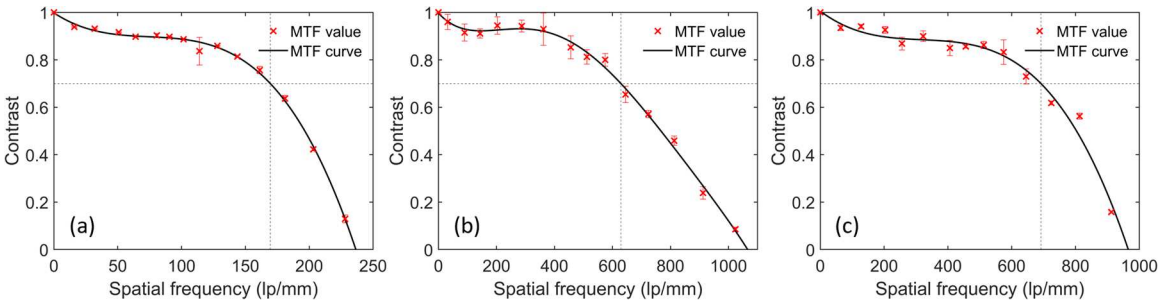


Figure 2-2: The measured MTF curves of different microscopes when using (a) 4×/0.1 NA objective lens; (b) FPM technology with 4×/0.1 NA objective lens; and (c) 20×/0.4 NA objective lens. Under these three imaging conditions, the cutoff frequencies are (a) 236.5 lp/mm, (b) 1065.83 lp/mm, and (c) 965.30 lp/mm, and the contrast values decrease to 0.7 at (a) 169.51 lp/mm, (b) 629.08 lp/mm, and (c) 691.58 lp/mm, respectively.

Figure 2-3 shows microscopic images of an analyzable metaphase chromosome acquired from the blood samples of a leukemia patient. In clinical practice, chromosome samples are stained with Giemsa dye to generate the "black and white" bands (known as G-banding) [27]. Each chromosome has a unique band pattern as its signature that is used for karyotyping or abnormality identification [27, 28]. The quality of the reconstructed images largely depends on whether these band patterns are clearly resolved. Accordingly, when the cell is imaged under the 4×/0.1 NA objective lens (Figure 2-3a), we can only distinguish this metaphase cell from the interphase cells which have a circular appearance, but we cannot recognize any details regarding the chromosome band patterns. Figures 2-3c/3d indicate the reconstructed amplitude and phase images by FPM equipped with the 4×/0.1 NA objective. Note that we rescaled the phase values to the range [0, 255] and then inverted the pixel values for the figure presentation because the raw phase image has a dark background (indicating zero phase shift). For the reconstructed amplitude/phase images (Figure 2-3c/3d), some band patterns can be distinguished, and the band sharpness is

comparable to the images captured directly by the conventional 20×/0.4 NA objective lens (Figure 2-3e).

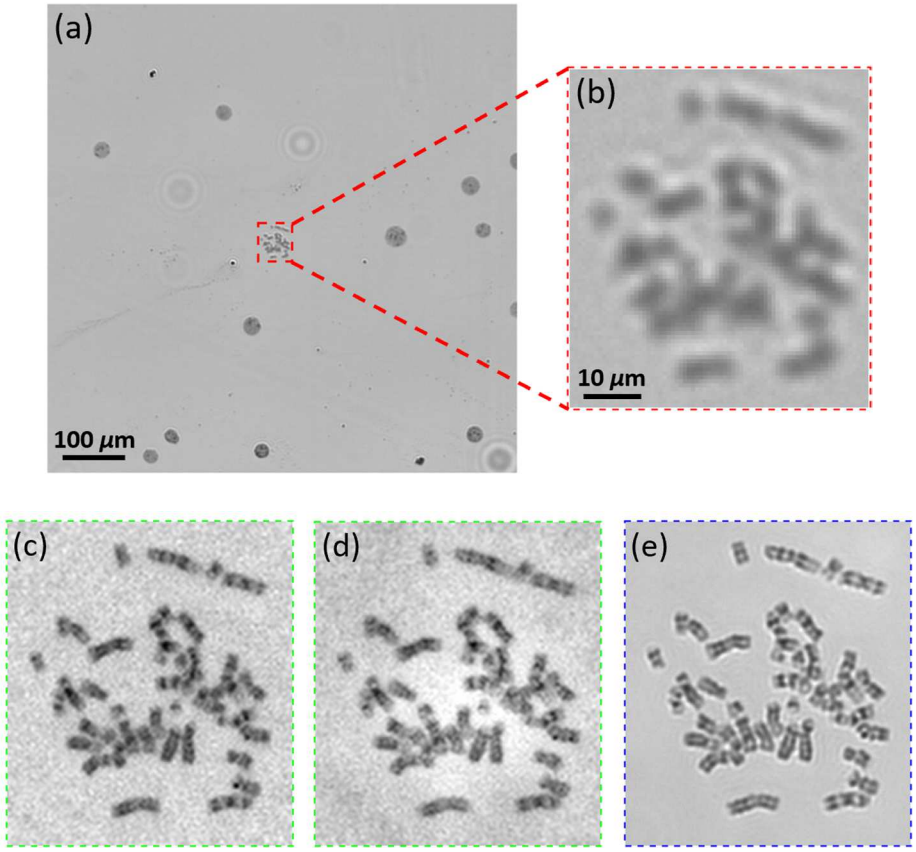
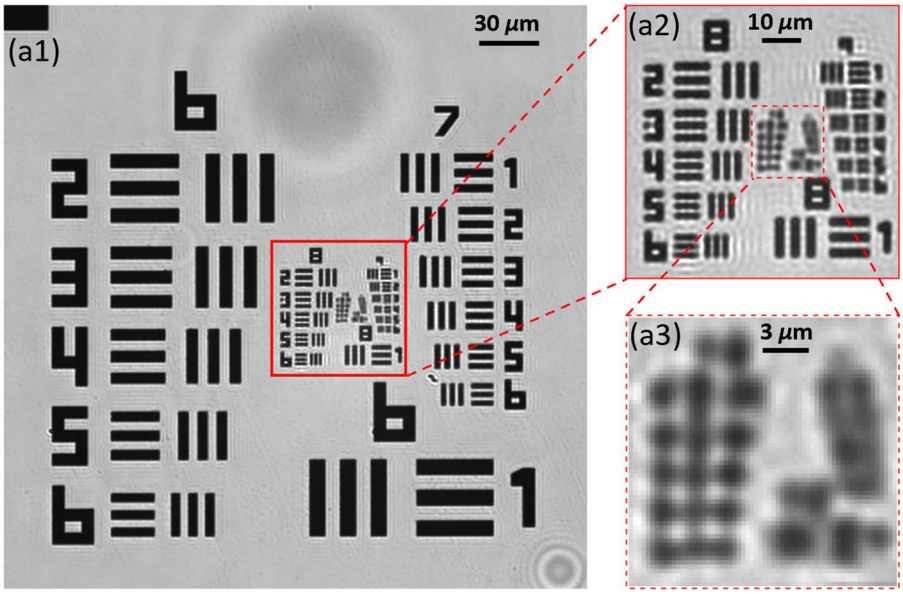


Figure 2-3: Sample images of one analyzable chromosome from the blood sample of leukemia patients, which are imaged using (a), (b) 4×/0.1 NA objective lens; (c), (d) FPM technology with 4×/0.1 NA objective lens; and (e) 20×/0.4 NA objective lens. (c) and (d) are the reconstructed amplitude and phase images, respectively.

When the raw patterns were captured using the 10×/0.25 NA objective lens, the reconstructed USAF 1951 resolution target is illustrated in Figure 2-4. Figure 2-4a is the target image directly captured by the 10×/0.25 NA objective lens; the minimum resolvable bar pattern is group 9-1 (512 lp/mm). When using the FPM technology (Figure 2-4b), the spatial resolution vastly increases to 1290 lp/mm (group 10-3), which is comparable to the

result of the 60×/0.95 NA objective lens (Figure 2-4c). Under the 100×/1.25 NA objective lens (Figure 2-4d), the target image resolves bar patterns up to group 10-6 (1825 lp/mm), which significantly surpasses the results of FPM. Meanwhile, the MTF curves of the 10×/0.25 NA objective lens (Figure 2-5a) and its corresponding FPM system (Figure 2-5b) gradually decline to 0.7 at frequency of 381 and 1055 lp/mm, respectively, before quickly falling to 0 at their respective cutoff frequencies of 681 and 1487 lp/mm. In contrast, the curves of the 60×/0.95 NA (Figure 2-5c) and 100×/1.25 NA objective lenses (Figure 2-5d) initially decrease rapidly and then slowly approach 0 at frequencies of 1642 and 2639 lp/mm, respectively.



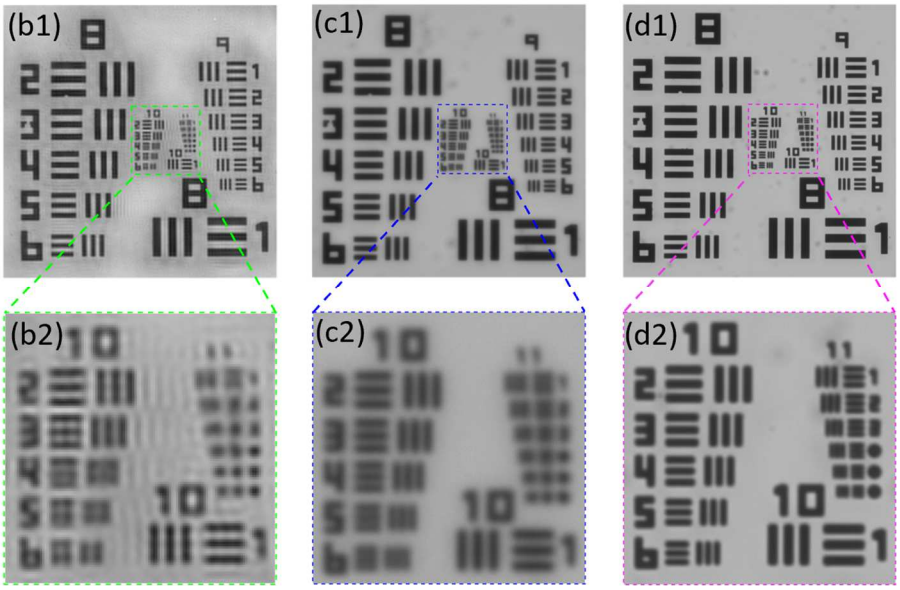
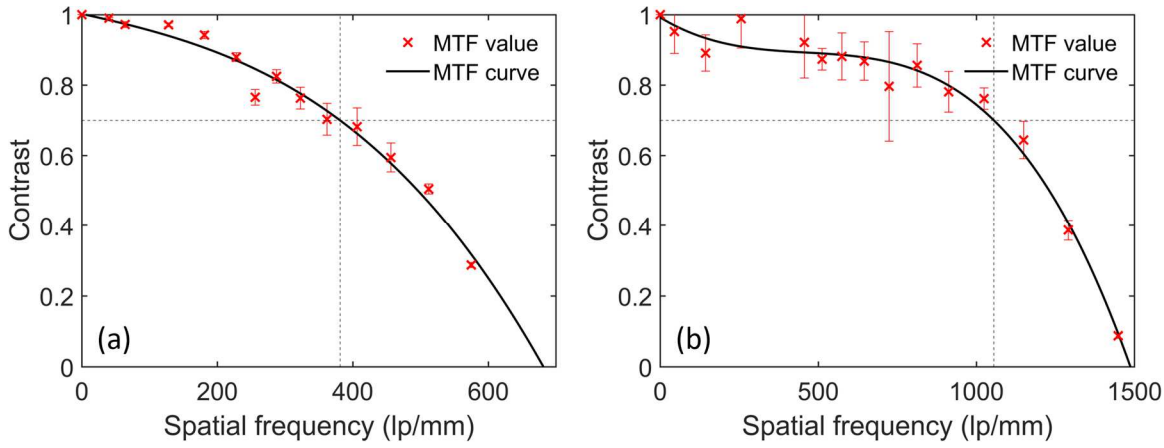


Figure 2-4: Images of the USAF1951 resolution target generated using (a) 10×/0.25 NA objective lens; (b) FPM technology with 10×/0.25 NA objective lens; (c) 60×/0.95 NA objective lens; and (d) 100×/1.25 NA objective lens.



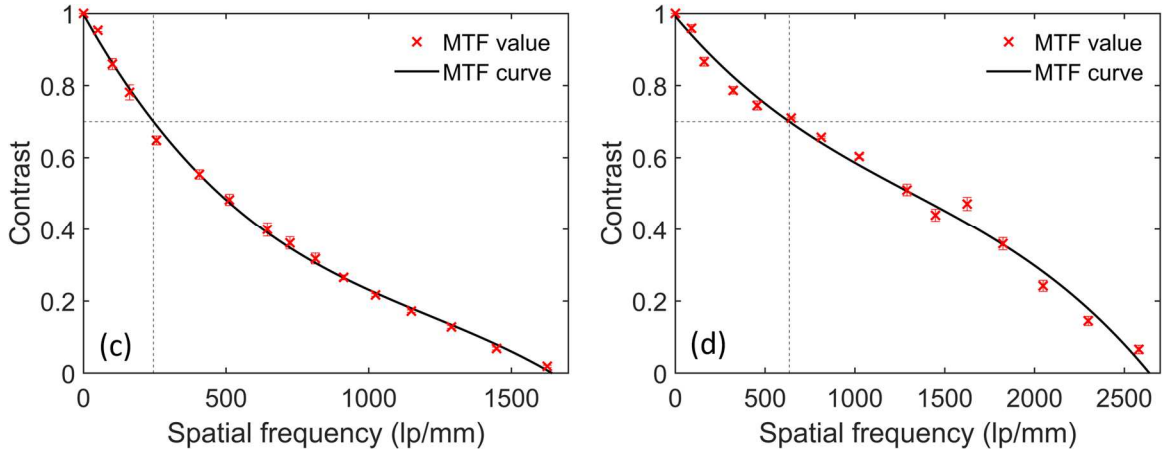


Figure 2-5: The measured MTF curves of different microscopes when using (a) 10×/0.25 NA objective lens; (b) FPM technology with 10×0.25 NA objective lens; (c) 60×/0.95 NA objective lens; and (d) 100×/1.25 NA oil immersion objective lens. Under these four imaging conditions, the cutoff frequencies are (a) 681 lp/mm, (b) 1487 lp/mm, (c) 1642 lp/mm, and (d) 2639 lp/mm, and the contrast values decrease to 0.7 at (a) 381 lp/mm, (b) 1055 lp/mm, (c) 245 lp/mm, and (d) 635 lp/mm, respectively.

Next, Figure 2-6 shows the same analyzable chromosome imaged using FPM technology and different objective lenses. For the result of 10×/0.25 NA objective lens (Figure 2-6a), the chromosome band patterns are generally recognizable, but the details remain unclear. When using FPM technology (Figure 2-6b), the quality of the band patterns is vastly enhanced, which are analyzable for the subsequent karyotyping. The reconstructed band patterns are somewhat better than the results of the 60×/0.95 NA objective lens (Figure 2-6d) but slightly inferior to the results of the 100×/1.25 NA objective lens (Figure 2-6e).

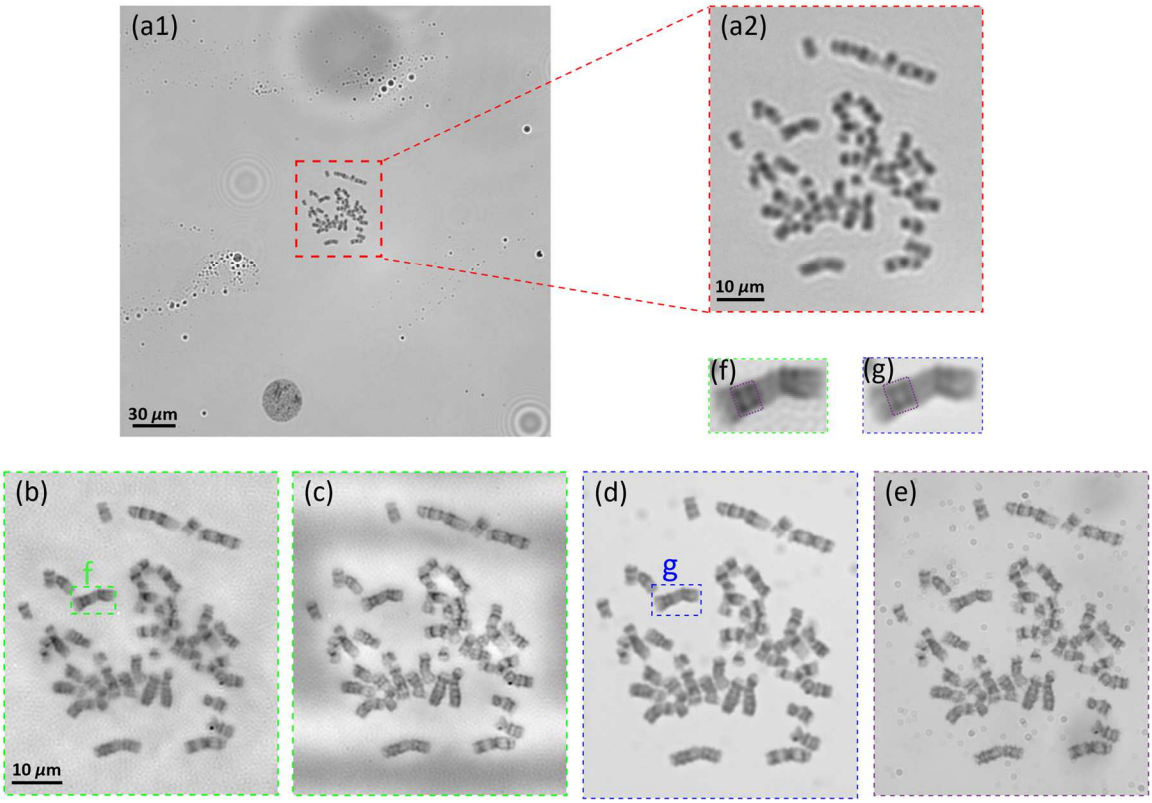


Figure 2-6: The images of a sample chromosome acquired from the patient's blood sample, which were captured or reconstructed using (a) 10×/0.25 NA objective lens; (b), (c) FPM technology with 10×/0.25 NA objective lens; (d) 60×/0.95 NA objective lens; and (e) 100×/1.25 NA oil immersion objective lens. (b) and (c) are the reconstructed amplitude and phase images, respectively.

The high-resolution images of the USAF resolution target recovered by the high NA FPM are shown in Figure 2-7. The raw image captured under the 20×/0.4 NA objective lens with normal illumination is illustrated in Figure 2-7a. The resolvable patterns reach up to group 9-5 (812.7 lp/mm), larger than the theoretical result (618.6 lp/mm). Figure 2-7b demonstrates the corresponding FPM recovered image. The spatial resolution of the system is improved to 2048 lp/mm (Group 11-1), which is inferior but comparable to the image captured with the conventional microscope's 100×/1.45 NA oil immersion objective lens

(Figure 2-7c). Accordingly, Figure 2-8 shows the measured MTF curves of the FPM-based and conventional microscopes. The MTF curve of 100×/1.45 NA oil immersion objective lens in Figure 2-8b exhibits a higher contrast than the FPM system in Figure 2-8a, leading to a higher spatial resolution of 1502.5 lp/mm at the contrast of 0.7, along with a higher cutoff frequency of 2640.2 lp/mm than those of the FPM system (1298.3 lp/mm and 2373.1 lp/mm, respectively).

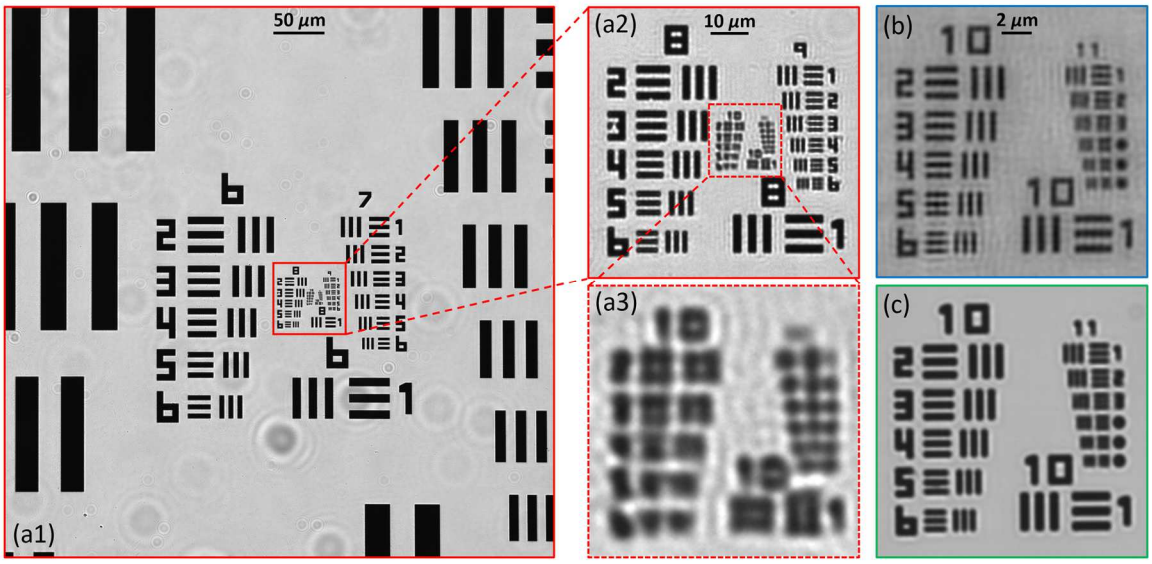


Figure 2-7: The microscopic images of a standard resolution target (USAF 1951), acquired via (a) 20×/0.4 NA objective lens with conventional illumination; (b) FPM equipped with 20×/0.4 NA objective lens; (c) conventional microscope equipped with 100×/1.45 NA oil immersion objective lens.

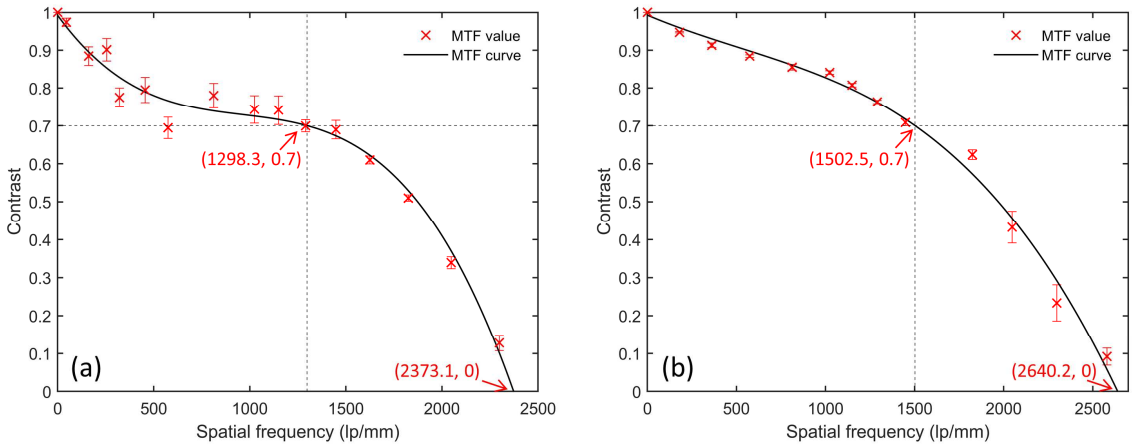


Figure 2-8: The MTF curves of (a) FPM technology equipped with 20×/0.4 NA objective lens and (b) conventional microscope equipped with 100×/1.45 NA oil immersion objective lens.

Next, one metaphase chromosome in a pathological sample of blood was imaged by the high NA FPM system and the conventional microscope, as shown in Figure 2-9. The chromosome image captured directly under the 20×/0.4 NA objective lens (Figure 2-9a) displays recognizable band patterns, but the details appear blurred and unresolvable when compared to the high resolution image captured with the 100×/1.45 NA oil immersion objective lens (Figure 2-9b). As expected, the FPM technique (under 20×/0.4 NA objective lens) significantly improves the image resolution of the chromosome in the blood sample (Figure 2-9c1 and c2). In the recovered amplitude image (Figure 2-9c1), the quality of bands is very close to the result of 100×/1.45 NA oil immersion objective lens (Figure 2-9b), despite the loss of some fine details. Similarly, the FPM system clearly resolves chromosome bands in a pathological bone marrow sample, as shown in Figure 2-10. The image quality of recovered amplitude (Figure 2-10c1) is significantly better than the raw image captured directly under 20×/0.4 NA objective lens (Figure 2-10a) and close to the

performance of the high magnification 100×/1.45 NA oil immersion objective lens (Figure 2-10b).

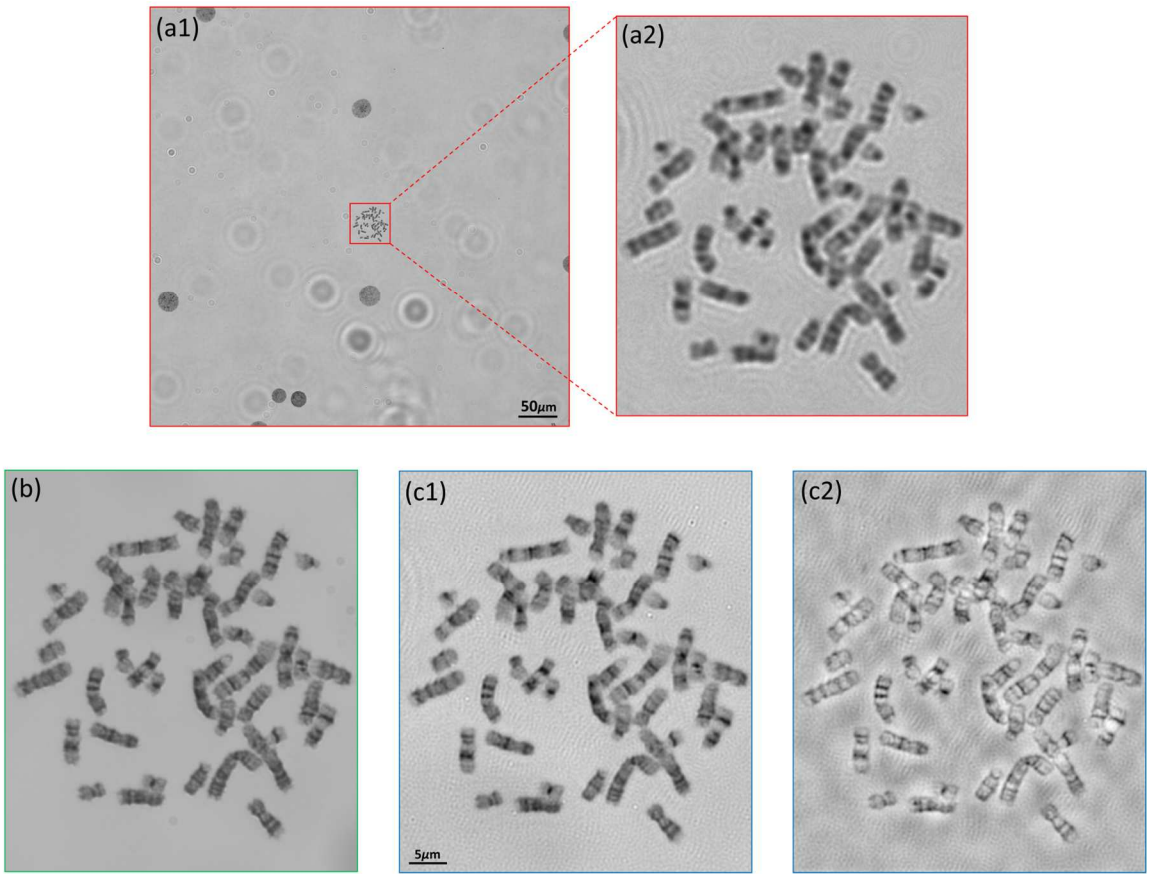


Figure 2-9: The sample images of a chromosome in the blood sample of one leukemia patient, obtained using (a) 20×/0.4 NA objective lens with conventional illumination; (b) conventional microscope equipped with 100×/1.45 NA oil immersion objective lens; (c1), (c2) FPM with 20×/0.4 NA objective lens. (c1) and (c2) are recovered amplitude and phase results, respectively.

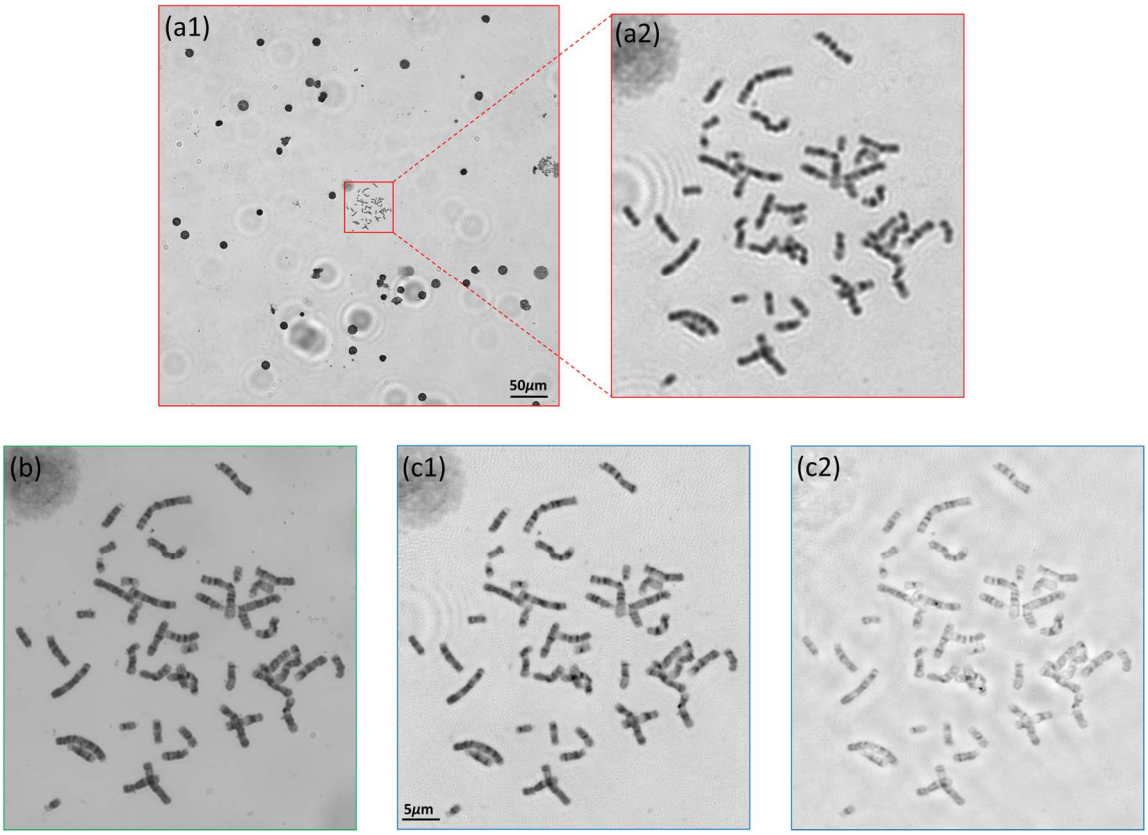


Figure 2-10: The sample images of a chromosome in the bone marrow sample of leukemia patients, obtained using (a) 20×/0.4 NA objective lens with conventional illumination; (b) conventional microscope equipped with 100×/1.45 NA oil immersion objective lens; (c1), (c2) FPM with 20×/0.4 NA objective lens. (c1) and (c2) are recovered amplitude and phase results, respectively.

2.4 Discussion

In this study, we applied the FPM technology to chromosome imaging and thoroughly investigated the band pattern quality depicted on the reconstructed images. To the best of our knowledge, this is the first study that utilizes this novel FPM technology on chromosome imaging, and it has two unique characteristics.

First, although the equivalent NA of our FPM system (raw patterns under 10×/0.25 NA， or 20×/0.4 NA objective lens) is far less than 1.25, the band quality of the

reconstructed images is close to or even comparable to the results captured by the 100×/1.25 NA or 100×/1.45 NA objective lens. In the previous research [29, 30], it is cautioned that it may not be reasonable for the direct comparison between the equivalent NA of FPM and the native NA of conventional microscopy. In this investigation, we believe this phenomenon can be majorly attributed by the two following factors. First, the effective NA of objective lens is actually higher than its rated value. For example, Figure 2-4a2 illustrates the bar patterns imaged by 10×/0.25 NA objective lens, which are resolvable up to group 9-1 with a half-pitch resolution of 0.98 μm. Accordingly, the computed NA is approximately 0.33, which is approximately 0.10 higher than 0.25. In this research, the NA of illumination is configured as 0.38, thus the final equivalent NA should be 0.71. This value generally agrees with the measured results (NA_{eq} 0.84, Figure 2-4b) [3, 31]. Second, the MTF curves are significantly different between the FPM system and an equivalent conventional microscope. The FPM system's curve initially decreases slowly and then sharply drops to zero; whereas the curve of conventional microscope decreases sharply at first and then gradually approaches zero. As a result, at a very wide range of frequency (0 to 1000 lp/mm), the FPM microscope can provide better resolving power (i.e., contrast) than the 60×/0.95 NA and 100×/1.25 NA conventional objective lenses, although both lenses have higher cutoff frequencies. Meanwhile, most chromosome band patterns are alternating bright-dark strips with spatial frequencies less than 1000 lp/mm, and the FPM system may provide better contrast on these bands, which accounts for the comparative quality of the band patterns on the reconstructed images. When the raw patterns were captured under 4× objective lens, the MTF curve of the FPM closely resembles that of the

equivalent 20×/0.4 NA objective lens; thus the band pattern qualities are comparable on the images generated by both systems.

Second, this investigation points out a new strategy to accomplish efficient slide scanning for chromosome karyotyping. For the microscopic chromosome imaging, we demonstrated that the 10×/0.25 NA FPM system can achieve a comparable result to that of the conventional 100×/1.25 NA objective lens. At the same time, the major advantage of the 10× lens remains: large FOV. This discovery provides a new possible strategy for achieving high speed chromosome digitization: acquire the raw images under 10× objective lens and then reconstruct the high resolution analyzable metaphase chromosomes for the following diagnosis work. In addition, we demonstrated that the reconstructed images from 4×/0.1 NA raw patterns achieve similar quality as compared with the 20×/0.4 NA objective lens, which is sufficient for identifying the analyzable cells. Efficiency could be further enhanced by initially capturing and reconstructing the sample under 4× lens for screening and then only the selected regions of interest (ROI) are reimaged under 10× condition to synthesize high resolution images. By adjusting the focal length of the tube lens or directly using the available lower magnification objective lens (i.e., 1.25× or 2×), it is also possible to further enlarge the FOV to achieve an even higher scanning speed without compromising the quality of the final images.

3 Using Symmetric Illumination and Color Camera to Achieve High Throughput Fourier Ptychographic Microscopy

3.1 Introduction

Fourier ptychographic microscopy (FPM) [3, 23, 32] method is an effective approach to improve both SBP (Space-bandwidth production) and DOF (Depth of field) of a conventional microscope. By using a low-cost LED panel and a phase retrieval algorithm, FPM can computationally integrate a series of low-resolution images captured under diverse illumination angles into a single high-resolution image [3, 6, 7, 12, 33, 34] while successfully inheriting the larger field of view (FOV) and DOF [11] from the adopted low-NA objective lens. Correspondingly, the SBP can be expanded by over 10 times while reducing the possibility of re-scanning due to inaccurate focusing as a result of the larger DOF.

However, despite its advantages, the FPM data acquisition process is quite time-intensive, as a large number of raw low-resolution images are required by the algorithm to successfully reconstruct a high-resolution image. Consequently, this also inhibits the application scope of FPM for tasks such as dynamic live cell imaging. This drawback can be exacerbated when applying FPM to color imaging, as images have to be acquired separately for red (R), green (G), and blue (B) image channels when relying on monochromatic cameras. Several alterations [7, 35-41] to the FPM method have been proposed to mitigate this problem by compressing the number of raw images needed. One strategy is to employ only the LEDs corresponding to the sub-regions on the spectrum

containing more information (or spectral energy) while discarding the others, which effectively reduces the data acquisition by at least half [35, 36]. Another strategy is to conduct the information multiplexing within the Fourier domain by lighting up multiple LEDs of varying incident angles (illumination multiplexing FPM) [7] or wavelengths (wavelength multiplexing FPM) [37-40] when capturing each raw image.

In this study, we built a symmetric illumination FPM based microscope and corresponding reconstruction algorithm with necessary corrections. The adopted strategy leverages the conjugate symmetry of the sample spectrum to reduce the data acquisitions by half [41], further augmented by integrating wavelength multiplexing using an RGB camera and white light illumination. To validate its performance, we first imaged a USAF 1951 resolution target and measured the system's modulation transfer function (MTF), by which the resolving power of R, G, and B channels was assessed. Then the high-resolution color images of an H&E-stained ovarian cancer tissue sample were reconstructed.

3.2 Materials and Methods

3.2.1 Experimental Set-up

To experimentally validate the performance of the symmetrical illumination FPM method, we built our FPM imaging system with the components of the Cerna Microscope platform from Thorlabs (Thorlabs, New Jersey, USA) and an FL20 color cooled CMOS camera (Tuscen, Fuzhou, China). A magnification of 3× is achieved by an equipped 4 × /0.13 NA Plan Fluor objective lens (Nikon, Tokyo, Japan) and a 150 mm achromatic doublet (AC508-150-A) used as the tube lens. At ~72mm below the sample stage, a 32×32 color LED array (Adafruit, New York, USA) was installed to provide tilted illumination. Each LED has three elements that can generate blue (0.47 μm), green (0.53 μm), and red

(0.63 μm) light respectively, and we simultaneously turned on these three elements to illuminate the sample with white light. The central 15×15 color LED matrix was utilized in our experiment, and the corresponding illumination NA was approximately ~0.36. Thus, the estimated synthetic NA of the experimental system would have been ~0.49. During image acquisition, the symmetric illumination FPM method alters the conventional approach, in which only a single LED is activated for each illumination, by simultaneously lighting up 2 symmetric LEDs (excluding the central LED). For example, when the symmetric illumination FPM method turns on the primary LED at (2, 2) on the grid it also activates the symmetric LED at (-2, -2), as illustrated in Figure 3-1a. Therefore, only 113 raw images are acquired for the following computational reconstruction, with each image (excluding the one captured by the central LED) corresponding to 2 symmetric subregions of the spectrum (Figure. 3-1b). As a comparison, the investigated samples were also imaged under 1) the conventional illumination FPM method utilizing a single LED (emitting white light) in illumination; 2) a Nikon Ni transmissive microscope (Nikon, Tokyo, Japan) with 20 × /0.4 NA and 60 × /0.95 NA objective lenses.

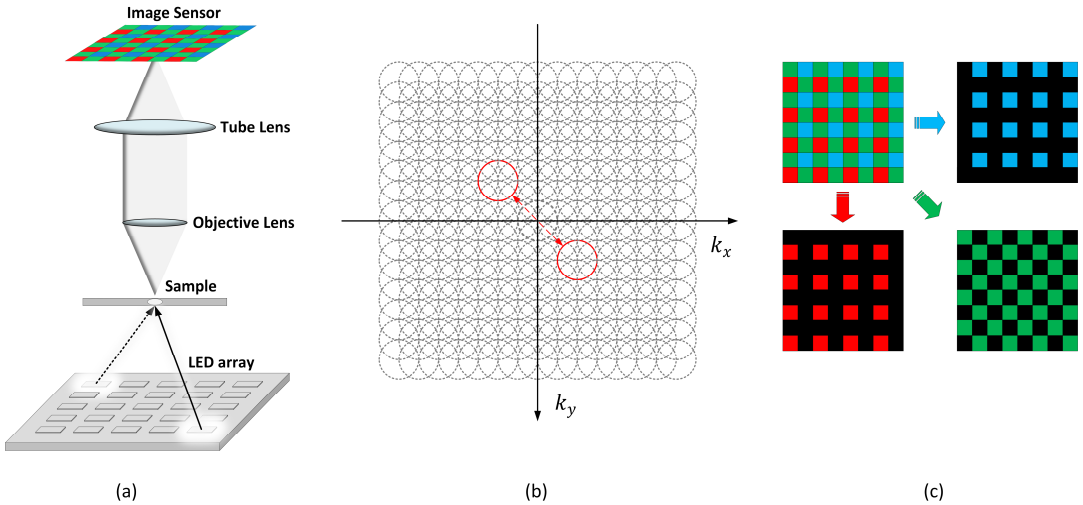


Figure 3-1: Principle of symmetrical illumination FPM. (a) Schematic of symmetrical illumination FPM illuminating the sample by a primary LED (solid arrow line) and its symmetric LED (dashed arrow line). (b) The spectrum of specimen covered by the subregions corresponding to the LED array. The red circles depict one pair of subregions of the sample spectrum corresponding to same image. (c) Color masks generated after the separation of red, green, and blue channels.

3.2.2 Symmetric Illumination FPM Image Reconstruction

The reconstruction algorithm is similar to the method detailed in section 1.2 with the following modifications. Each raw color image is split into R, G, and B channels to generate 3 single-color raw image datasets, each of which is reconstructed into a high-resolution image. Accordingly, for each channel, a high-resolution initial guess is generated by up-sampling the raw image captured under normal illumination. Second, when performing the symmetric illumination based data acquisition, each captured raw image contains information of two symmetric patterns in the spectrum. Given the assumption that the sample is sufficient thin, the phase information contained in the sample can be ignored, and the two corresponding spatial representations of the symmetric patterns in the spectrum

have same modulus [41]: $\left|\Psi_t^{n(1)}(\boldsymbol{r})\right| = \left|\Psi_t^{n(2)}(\boldsymbol{r})\right|$, where $\Psi_t^{n(i)}(\boldsymbol{r}) = \mathcal{F}^{-1}\left[S_t\left(\boldsymbol{k} - \boldsymbol{k}^{n(i)}\right) \cdot P_t(\boldsymbol{k})\right], i = 1, 2$. Thus, the equation that imposes intensity constraints can be revised as: $\Phi_t^{n(i)}(\boldsymbol{r}) = \sqrt{I_{raw}^n / (2\left|\Psi_t^{n(i)}(\boldsymbol{r})\right|^2)}\,\Psi_t^{n(i)}(\boldsymbol{r})$. In addition, given that a color camera is used, the R, G, and B channels are sub-sampled with 75%, 50%, and 75% of pixels left empty, respectively. As a result, the intensity constraint will only be applied to the masked non-empty pixels [42] (Figure 3-1c). Meanwhile, starting from the 2nd iteration, a scaling factor will be applied to the update of the real space representation. Considering that illumination intensity uncertainty and fluctuation across LEDs degrades the quality of recorded images, we used the scaling factor to alter the intensity of each low-resolution measurement during the reconstruction:[43] $c^n = \sum_{x,y} I_{est}^n / \sum_{x,y} I_{raw}^n$, where I_{est}^n (i.e., $\sum_{i=1,2}\left|\Psi_t^{n(i)}(\boldsymbol{r})\right|^2$) is the intensity of the estimated low-resolution image corresponding to the n^{th} illumination, and $\Sigma_{x,y}$ denotes the sum of all pixel values. The reconstruction pipeline is also displayed in Figure 3-2. After each channel is reconstructed, they are combined to generate a high resolution color image. The white balance was implemented by normalizing the 3 channels to achieve a uniform mean value among all channels, which is considered an effective method with relatively low computing complexity compared to other algorithms [44].

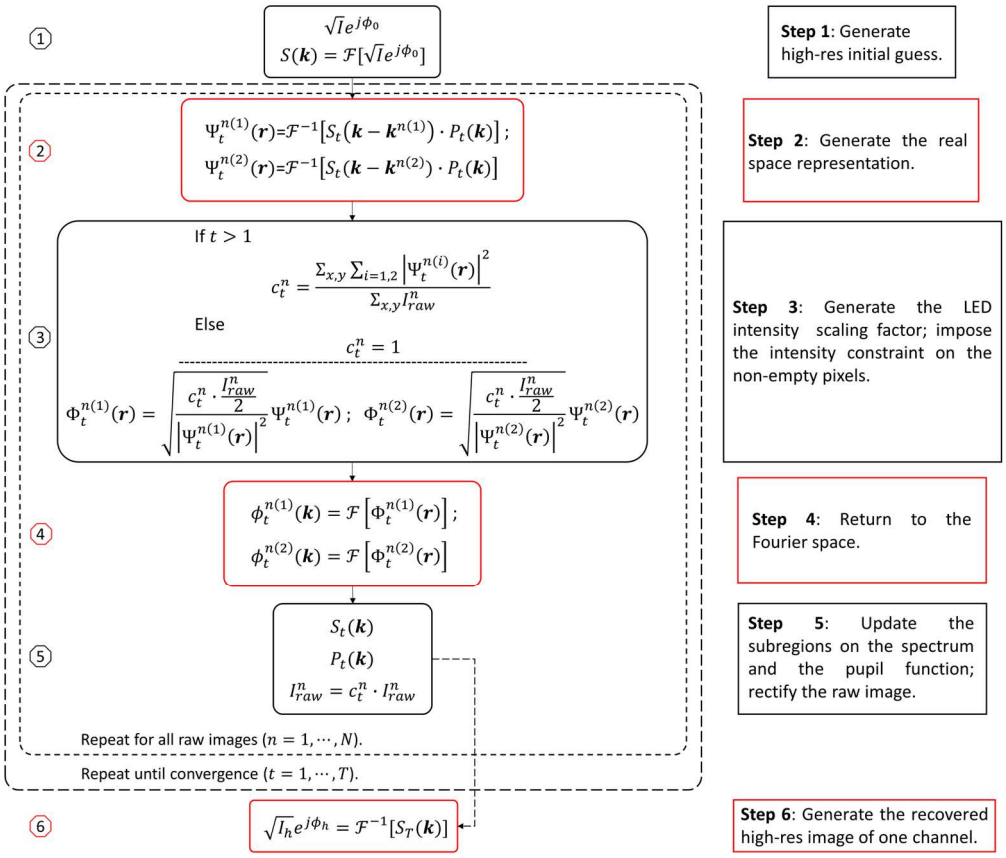


Figure 3-2: The primary reconstruction procedures of a single channel using the symmetric illumination FPM method.

3.2.3 Performance Validation of the Symmetric Illumination FPM

In order to evaluate the resolving ability of the proposed symmetric illumination FPM method, we first imaged a USAF 1951 resolution target with the highest resolution equal to 3649 lp/mm (Newport, California, USA). The recovered high-resolution target images of R, G, and B channels were subjectively compared with those of the conventional illumination FPM method as to the overall image quality and resolving limit. In addition, the MTF curves of R, G, and B channels were estimated by measuring the contrast values of a series of bar patterns on the reconstructed target images [25].

To assess its potential in clinical application, an H&E-stained ovarian cancer tissue sample was imaged by both symmetric and conventional illumination FPM methods. The sample was prepared in our medical center using a standard clinical approach. The conventional microscope equipped with 20 × /0.4 NA and 60 × /0.95 NA objective lenses was also used in the pathology slide imaging to provide more reference. The quality of the color images acquired by the above-mentioned methods was compared.

3.3 Results

First, a USAF 1951 resolution target was reconstructed using the raw data captured under a 4 × /0.13 NA objective lens with white light illumination, as shown in Figure 3-3. Figure 3-3a1 depicts the raw low-resolution image captured with the central LED activated, and Figures 3-3a2/3/4 depict enlarged images of the B, G, and R channels from a central square region of Figure 3-3a1, respectively. The results indicate that the 4 × /0.13 NA objective lens can resolve the bar patterns with frequencies up to 228.1 lp/mm (group 7-6). The B, G, and R channels of the high-resolution image recovered by the symmetric illumination FPM are shown in Figures 3-3b1/2/3. For all three channels, the spatial resolution of the 4 × /0.13 NA objective lens is improved to 724.1 lp/mm (group 9-4) when using the FPM technology, which is consistent with the calculated synthetic NA of the system. Under blue (Figure 3-3b1) or green (Figure 3-3b2) light illumination, the resolving power is slightly improved compared to the red light illumination, with the minimum resolvable bar pattern increasing to group 9-5 (812.7 lp/mm) or 9-6 (912.3 lp/mm) in the recovered images. Despite its high resolution, the recovered image of the blue channel exhibits more artifacts than the other two channels. As a comparison, Figures 3-3c1/2/3 are the B, G, and R channels of the recovered high-resolution image generated by the

conventional illumination FPM. The image quality of the symmetric illumination FPM is comparable to the results of the conventional illumination FPM in terms of background and contrast.

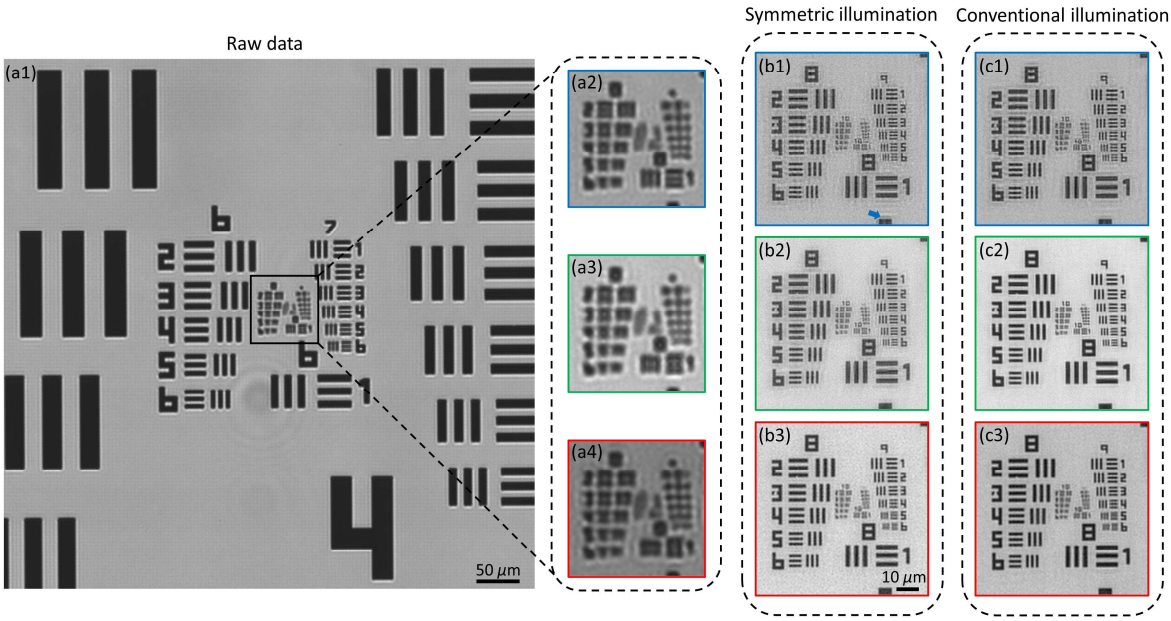


Figure 3-3: The images of a USAF 1951 resolution target. (a1) The raw image captured by a 4 × /0.13 NA objective lens with only central LED turned on. (a2)-(a4) Enlarged blue, green, and red channel images of the denoted square region in (a1), respectively. The empty pixels are filled by interpolation. (b1)-(b3) The FPM reconstructed images at blue, green, and red channels when using 2 symmetric LEDs for illumination, respectively. One artifact is pointed by a blue arrow in (b1). (c1)-(c3) The FPM reconstructed images at blue, green, and red channels when using conventional single LED illumination, respectively.

Next, Figure 3-4 illustrates the corresponding MTF curves of the symmetric and conventional illumination FPM methods for B, G, and R channels, respectively. Comparing the two FPM methods' performance across the three channels, the cutoff frequencies of both methods are quite close. However, the red channel's cutoff frequency is slightly lower than the other two channels, which aligns with our observations of the standard resolution

patterns in Figure 3-3. At the same time, the conventional illumination FPM produces generally better contrast than the symmetric illumination FPM. Within the blue channel (Figure 3-4a), the two methods show comparable performance. The conventional illumination FPM achieves larger contrast values of up to roughly 600 lp/mm while the symmetric illumination FPM catches up from behind. In contrast, the curve of the symmetric illumination FPM in the green channel drops vastly compared with that of the conventional illumination FPM (Figure 3-4b), corresponding to the observable image quality degradation regarding Figure 3-3b2 and Figure 3-3c2. A similar performance deterioration of symmetric illumination FPM is observed in the red channel (Figure 3-4c). However, the contrast value of the symmetric illumination FPM keeps a high level in the low frequency domain even though it is noticeably lower than that of the conventional illumination FPM.

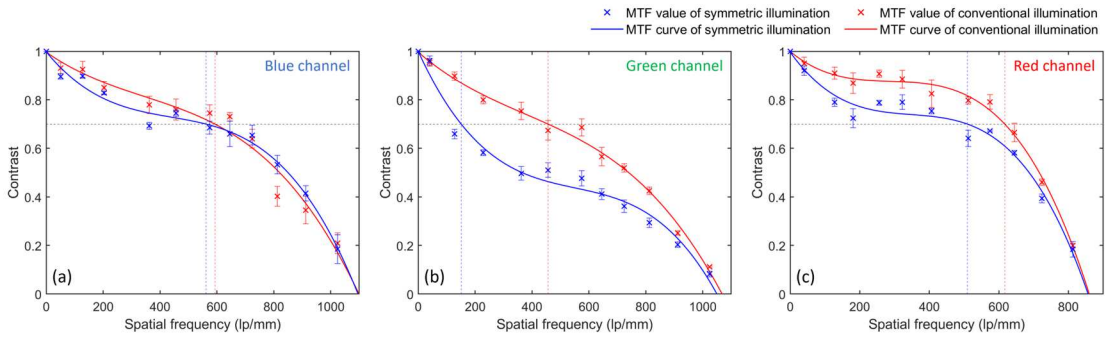


Figure 3-4: The measured MTF curves of FPM reconstructions when using symmetric and conventional illuminations, respectively. (a) Blue channel. (b) Green channel. (c) Red channel.

The results of the symmetric and conventional illumination FPM methods in reconstructing the ovarian cancer pathology slide are shown in Figure 3-5, and we used white light for illumination in this experiment. Figure 3-5a contains the raw low-resolution image of the pathology slide captured using the 4 × /0.13 NA objective lens, but we can

only differentiate the blurred and dark nuclei from the brighter surroundings without other recognizable details. Figures 3-5b1/2/3 present the B, G, and R channels of the recovered images obtained using symmetric illumination FPM method, respectively, with well-defined edges of nuclei in contrast to the raw image presented in Figure 3-5a. Among the 3 channels, the blue channel (Figure 3-5b1) has lower contrast and more artifacts compared to the green (Figure 3-5b2) and red (Figure 3-5b3) channels. Meanwhile, the red channels reconstructed by the symmetric (Figure 3-5b3) and conventional (Figure 3-5c) illumination FPM methods have similar performance in the imaging of nuclei. The final high-resolution color images generated by the symmetric illumination FPM and conventional illumination FPM are shown in Figures 3-5d and 3-5e, respectively, from which we can see that using symmetrical illumination not only preserves the image quality but also reveals more tissue outlines. By further comparison, the symmetric illumination FPM (Figure 3-5d) has a better performance than the conventional 20 × /0.4 NA objective lens (Figure 3-5f) and is close to the 60 × /0.95 NA objective lens (Figure 3-5g) as to the resolution and restored details of the tissue.

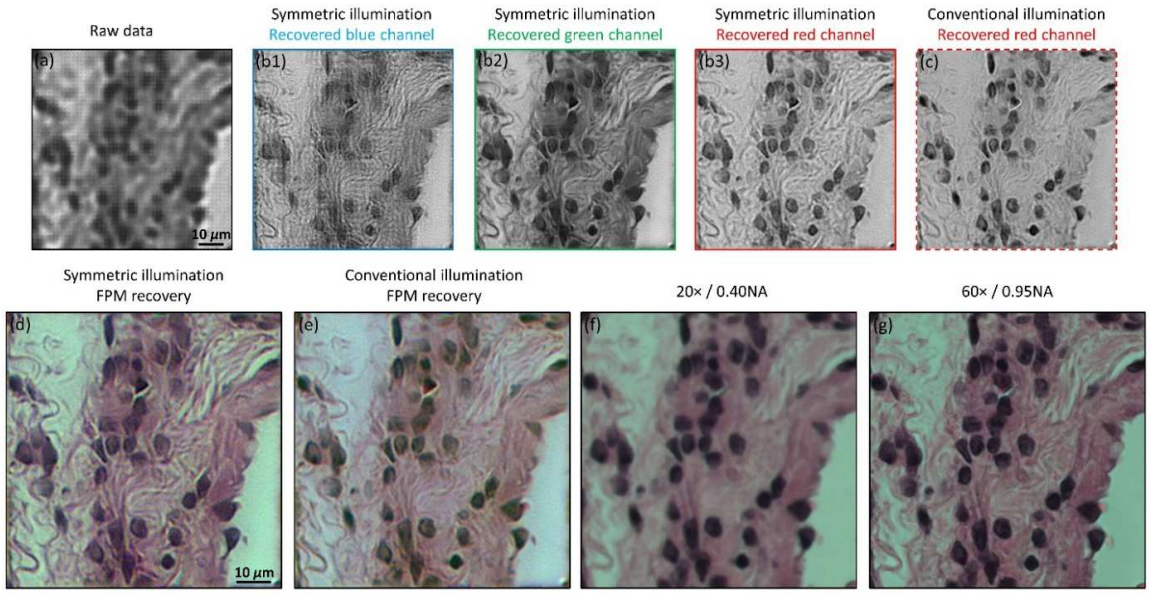


Figure 3-5: The sample images of one pathology slide from an ovarian cancer patient. (a) The raw image captured by a 4 × /0.13 NA objective lens with only central LED turned on. (b1)-(b3) The FPM reconstructed images at blue, green, and red channels when using symmetric illumination, respectively. (c) The FPM reconstructed image at red channel when using conventional illumination. (d) The FPM reconstructed color image when using symmetric illumination (by merging (b1)-(b3)). (e) The FPM reconstructed color image when using conventional illumination. (f, g) Color images captured under a conventional microscope with a 20 × /0.4 NA objective lens and a 60 × /0.95 NA objective lens, respectively.

3.4 Discussion

In this study, we developed a symmetric illumination FPM imaging prototype with a color camera and evaluated its performance using both a USAF 1951 resolution target and representative pathology slides of ovarian cancer. The reconstructed images have comparable image quality to the conventional illumination FPM and the conventional

microscope at a significant reduction in image acquisition requirements. The proposed technique accomplished this due to two unique characteristics.

First, this novel method can vastly enhance the data acquisition efficiency while maintaining the reconstructed image quality at the same level as conventional FPM. Under the thin sample assumption, the sample spectrum is conjugately symmetric; thus the symmetric illumination FPM only requires half the number of raw images that conventional illumination FPM needs. Considering that simultaneously turning on two LEDs can also double the illumination light intensity, the exposure time can be reduced by up to 50%. When using a color camera, each captured raw image pattern stores the easily separable information of the red, green, and blue channels, which further compresses the total image number by three times. Overall, the proposed symmetric illumination FPM can improve the data acquisition efficiency by a factor of up to 12 times. Meanwhile, when comparing our reconstructed images with the previously reported color FPM results [38], the quality difference for each of the three channels is not distinguishable, which indicates that the thin sample assumption is generally valid for the most representative pathology samples.

Second, the symmetric illumination color FPM does not increase the algorithmic complexity of the FPM reconstruction process. Symmetric illumination can provide a significantly larger DOF than conventional illumination FPM, which is particularly useful for color camera based data acquisition. The mechanism of the enlarged DOF for symmetric illumination FPM is thoroughly analyzed by previous investigations [40, 45]. Although the FPM optical system adopts chromatic aberration corrected objective and tube lenses, the system is still not able to focus the images of the three color channels at precisely

the same position. Thus at least one of the three channels will be out of focus to a certain extent during the image capturing. However, the symmetric illumination expands the DOF of each channel and is able to cover this out of focus range. Thus, the symmetric illumination color FPM doesn't use chromatic aberration correction, which simplifies the reconstruction algorithm [41].

4 Fourier Ptychographic Microscopy Using a Doublet as Objective Lens: An Initial Evaluation

4.1 Introduction

Fourier ptychographic microscopy (FPM) is an emerging microscopic technology which has received intensive research focus since 2013 [23]. Different from the conventional microscope, FPM acquires a series of raw images from a low resolution objective lens (e.g. 4×/0.1 NA), and computationally synthesizes one high spatial resolution image (e.g. equivalent to a 20×/0.4 NA objective lens or higher) [3, 32]. Meanwhile, the synthesized image naturally has a large field of view (FOV), which inherits from the advantages of the employed a low resolution objective lens. For example, in clinical pathology, a 20×/0.4 NA objective lens is widely adopted in conventional microscopic scanners for histopathological sample digitization, which has a typical FOV of ~0.6 mm^2 under the assumption of a typical optical field number (OFN) of 22. FPM uses a 4×/0.1 NA objective lens to achieve the same imaging performance, while the corresponding FOV will be enlarged 25 times to ~15 mm^2. Additionally, the low resolution objective lens has a significantly larger depth of field (DOF), which can vastly simplify the focusing operation. As a result, FPM is a very suitable approach for the development of a high throughput sample digitizer with affordable cost [46].

Currently, such a low cost digitizer is implemented based on either standard objective lens [47, 48] or low cost commercial camera lens [49, 50]. The standard objective lens ensures high optical quality, but they are also relatively expensive for such a cost sensitive system. The commercial cell phone camera lens is a low-cost alternative, but these lenses usually have very small diameters, which naturally limits the FOV. For instance, one

study [51] utilizes the Raspberry Pi camera module to implement an FPM system. The spatial resolving power is equivalent to that of 20×/0.5 NA conventional system, but the FOV is only 2.42 × 1.64-mm^2, which is less than 30% of a similar standard 4× objective lens based system [11, 13]. The advantage of the FPM approach is therefore largely compromised.

This investigation differs from both of these two approaches by directly utilizing a single commercially available doublet as the objective lens in the FPM system. Accordingly, a microscope prototype was developed, based on which the standard resolution target and two different clinical samples were imaged to assess its performance. All the details are presented as follows.

4.2 Materials and Methods

4.2.1 Symmetric Illumination FPM Image Reconstruction

To validate the performance of the proposed system, a doublet based FPM prototype was developed. As demonstrated in Figure 4-1, a doublet was employed as the objective lens, which has a diameter of 12.5 mm and a focal length of 60 mm (Edmund Optics, Barrington, New Jersey, USA). Acting as an infinity optical system, the objective lens projects the sample to infinity, which is subsequently converged by a 150 mm achromatic tube lens (AC254-150-A, Thorlabs, New Jersey, USA). Thus, the system achieves an overall magnification of 2.5×. Given the diameter and focal length of the doublet, the lens NA (numerical aperture) can be estimated as 0.1 [26]. To generate angled illumination, a commercial programmable LED panel (Adafruit, New York, USA) was positioned ~71 mm below the microscope stage. The central 15×15 LEDs (4 mm spacing) were sequentially activated to emit red light (wavelength λ = 0.63 μm). The illumination

NA was therefore estimated as approximately 0.37, and the equivalent NA of the entire FPM system was estimated to be ~0.47 [31]. Meanwhile, a CMOS camera with a pixel size of 2.4 μm (FL20BW, Tuscen, Fuzhou, China) was installed for image acquisition. Based on the Rayleigh criterion [26], the resolving power of the objective lens with 0.1 NA is 3.3 μm, which is magnified to 8.3 μm and then sampled by more than 3 pixel elements. This design satisfies the requirement of Nyquist sampling theorem to ensure the integrity of the original data [52]. The reconstruction algorithm is essentially the same as the method outlined in section 1.2.

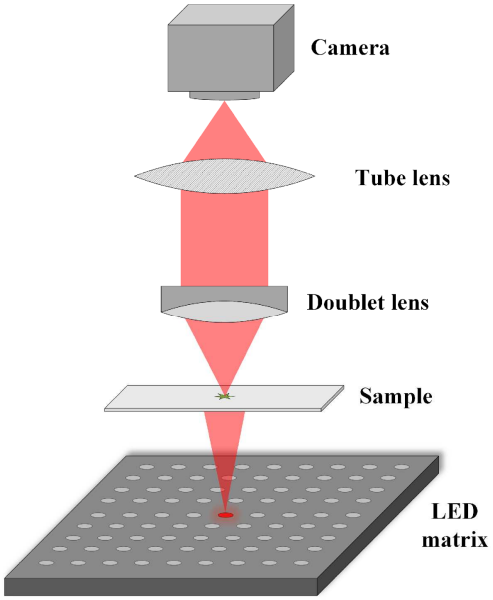


Figure 4-1: Experimental setup of the FPM imaging system using a doublet as its objective lens.

4.2.2 Performance Assessment of FPM Imaging Systems

The imaging capability of the proposed system was first evaluated by the modulation transfer function (MTF) curve, which is widely used to assess the system's capability to transfer the sample contrast to the generated image [20, 26, 53]. The MTF

was estimated using a USAF 1951 resolution target (Newport, California, USA) as the sample, which contains a number of resolution patterns covering spatial resolutions from 1 to 3649 line pairs per millimeter (lp/mm). Accordingly, during the experiment, the target was first imaged and reconstructed by our FPM system, and the contrast was then estimated on these different bar patterns, based on which the modulation transfer function (MTF) curve was generated by a curve fitting algorithm [8]. Meanwhile, to assess the clinical utility, H&E-stained ovarian cancer and gallbladder tissue samples were also imaged by our FPM system. All the investigated samples were prepared in our University Medical Center, following standard clinical protocols. The quality of the cell nuclei and cytoplasm was evaluated based on reconstructed images. For comparison purposes, the MTF curve and tissue samples were also imaged by the FPM system equipped with the standard 4× /0.1 NA lens, as well as conventional microscopic imaging with the 20× /0.4 NA lens.

4.3 Results

Figure 4-2 illustrates the images of the USAF 1951 resolution target acquired through different imaging methods. Figure 4-2a is the raw image captured by the doublet objective lens FPM imaging system under the central LED illumination. The raw image can resolve bar patterns up to group 7-2, corresponding to a spatial resolution of 143.7 lp/mm. This agrees with the theoretical estimation that the lens NA is approximately 0.1. After reconstruction, the resolving power of the FPM imaging system is significantly improved, with element 3 in group 9 (645.1 lp/mm) on the target clearly distinguished (Figure 4-2b). As a comparison, the standard 4× /0.1 NA objective lens demonstrates an equivalent level of resolving power after FPM reconstruction (Group 9-4), but with a slightly higher contrast than the doublet FPM system (Figure 4-2c). Additionally, Figure 4-

2d shows the target image captured by a conventional microscope equipped with a 20× /0.4 NA objective lens, which demonstrates a marginally higher resolving power.

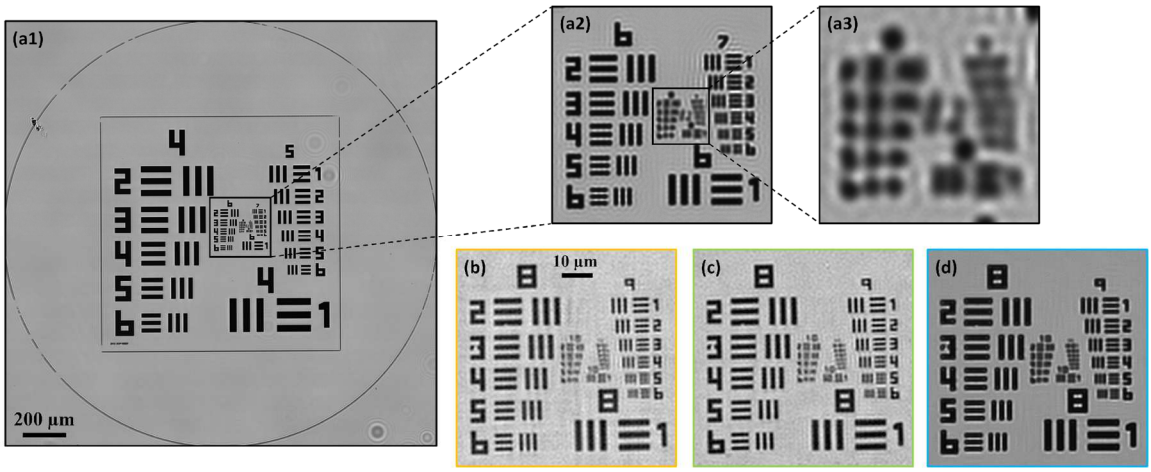


Figure 4-2: The acquired images of the USAF 1951 resolution target. (a) The captured low-resolution image under the doublet objective lens FPM imaging system. (b) The reconstruction result of the doublet objective lens FPM imaging system. c) The FPM reconstruction result of the standard 4× /0.1 NA objective lens. (d) The conventional microscopic image captured with the 20× /0.4 NA objective lens.

The measured MTF curves for each FPM imaging system are presented in Figure 4-3. The contrast value of doublet objective lens FPM is 1 at zero frequency, which consistently decreases to 0 at its cutoff frequency at 833.9 lp/mm (Figure 4-3a). Similarly, for FPM employing the standard 4×/0.1 NA objective lens, its cutting off frequency reaches 844.3 lp/mm (Figure 4-3b). But at most of the spatial frequencies, the standard lens FPM imaging system yields a relatively higher contrast than the doublet based system. Figure 4-3c depicts the MTF curve of a conventional microscope with a 20× /0.4 NA objective lens. The overall curve resembles that of the FPM with the 4× /0.1 NA objective lens, although the cutoff frequency is slightly higher (886.5 lp/mm).

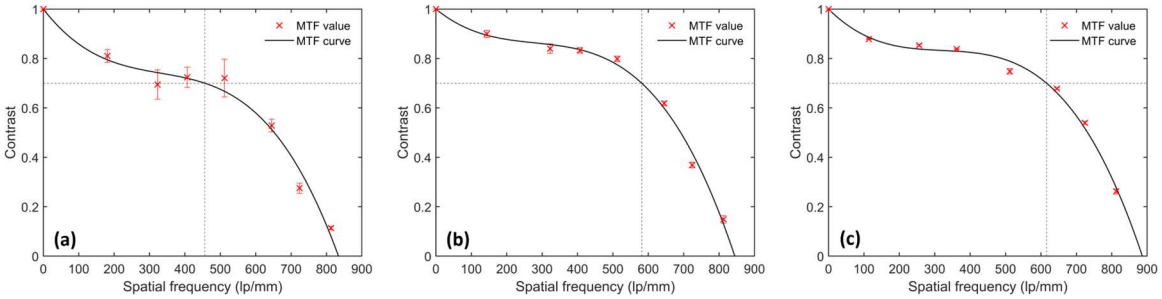


Figure 4-3: The obtained MTF curves of different FPM imaging systems. (a) FPM equipped with the doublet objective lens. (b) FPM equipped with the standard 4× /0.1 NA objective lens. (c) The 20× /0.4 NA objective lens.

Figures 4-4 and 4-5 illustrate the reconstructed images of the ovarian and gallbladder tissue samples. For the original raw image captured by the central LED (Figure 4-4a and 4-5a), the objects appear fuzzy and undistinguishable due to the limited resolving power of the doublet based objective lens. In contrast, after the reconstruction (Figure 4-4b and 4-5b), the nuclei can be distinguished with sharply defined borders, while the cytoplasm is also clearly identifiable. Accordingly, Figures 4-4c and 4-5c are reconstructed results using the standard 4× /0.1 NA objective lens, which exhibits comparable image quality and appearance. Figure 4-4d and 4-5d were obtained using the conventional 20× /0.4 NA objective lens based microscope. The conventional microscopic images captured using the 20× /0.4 NA objective lens (Figures 4-4d and 4-5d) show a resolution comparable to that of the FPM based methods, while exhibiting a higher contrast than the FPM recovered results.

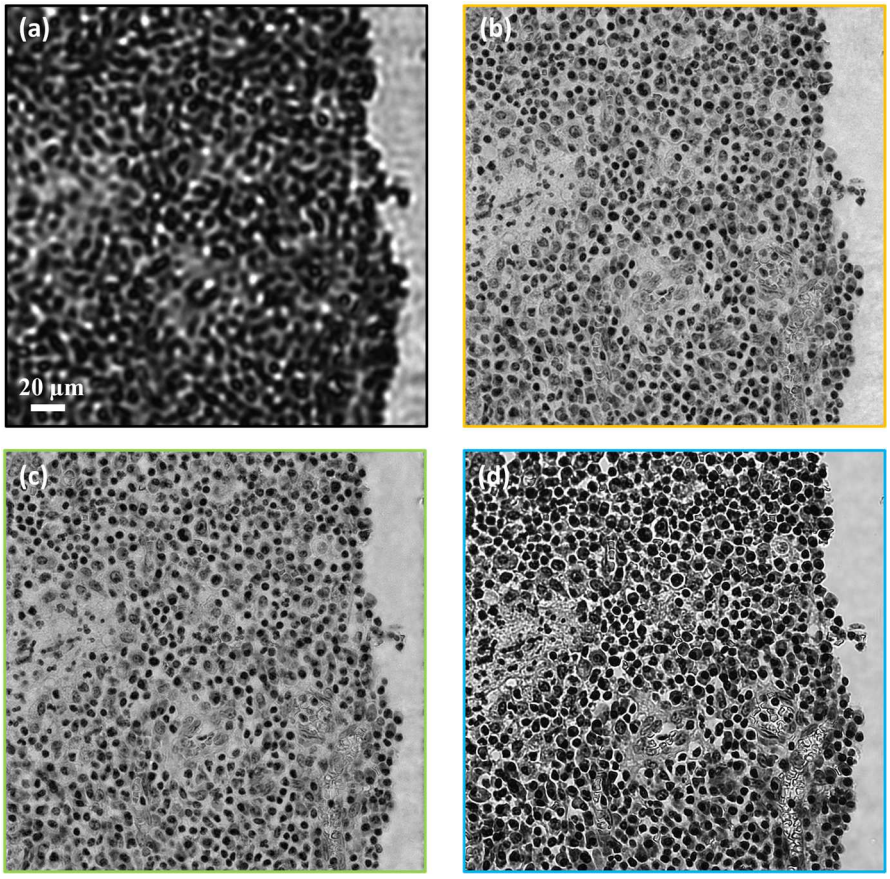


Figure 4-4: The acquired images of the ovarian cancer sample. (a) The raw image captured by the doublet objective lens FPM imaging system. (b) The reconstruction result of the doublet objective lens FPM imaging system. (c) The FPM reconstruction result of the standard 4× /0.1 NA objective lens. (d) The conventional microscopic image captured with the 20× /0.4 NA objective lens.

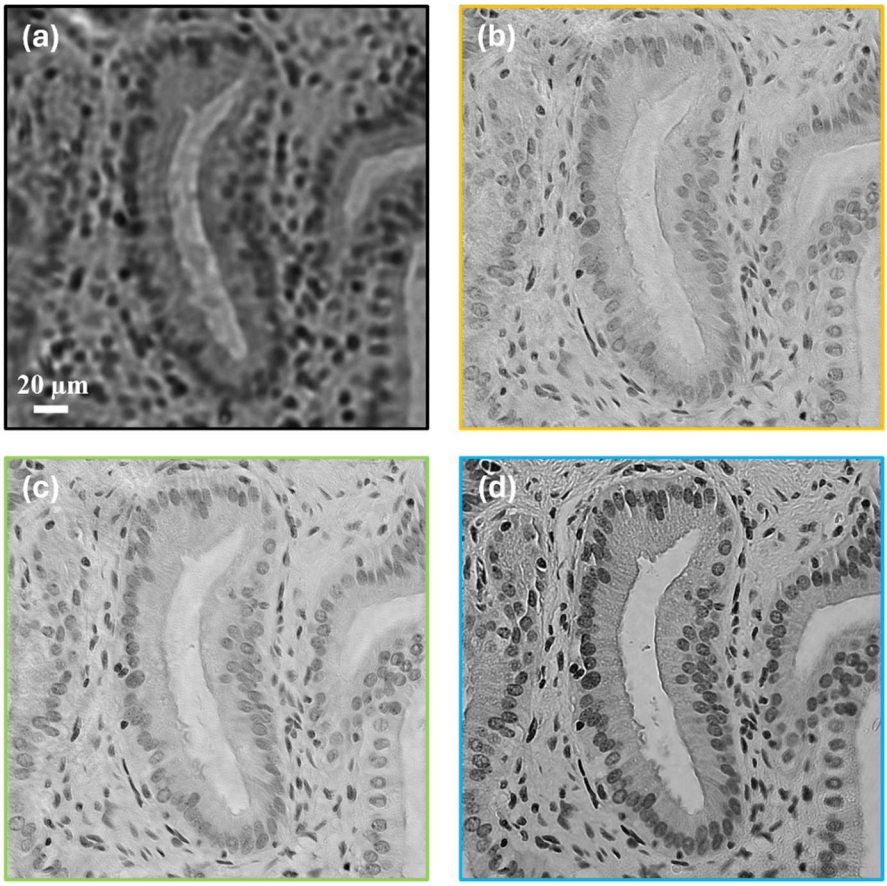


Figure 4-5: The acquired images of the gallbladder tissue sample. (a) The raw image captured by the doublet objective lens FPM imaging system. (b) The reconstruction result of the doublet objective lens FPM imaging system. (c) The FPM reconstruction result of the standard 4× /0.1 NA objective lens. (d) The conventional microscopic image captured with the 20× /0.4 NA objective lens.

4.4 Discussion

In this investigation, we developed a doublet based FPM imaging prototype, and its performance was evaluated using standard USAF1951 resolution target and clinical samples. In contrast to the conventional objective lens, employing doublet provides two major advantages: the cost effectiveness and potentially larger field of view. The cost of the doublet adopted in this study is less than $100, which is cheaper than most of the

objective lenses. The commercial objective lens utilizes multiple optical elements and a more complicated lens surface profile to minimize aberrations and achieve better performance [2]. However, our result shows that a simple doublet can also achieve similar performance in standard resolution target and clinical samples, which can be attributed to the following two reasons. First, the reconstructing scheme involves digital wave correction and system calibration algorithms, which is able to further mitigate the aberration of the doublet to a certain extent. Second, the conventional objective lens may be overqualified for this FPM application, indicating the potential for simplification to further reduce lens complexity and cost. Meanwhile, the optical aperture of the doublet in this study is significantly larger than the off-the-shelf camera lens reported in the previous studies [51], which will result in a larger field of view (FOV). In addition, the extended working distance [54] (e.g. 60 mm in this study) facilitates easier sample manipulation and the incorporation of additional equipment, such as environmental chambers.

Meanwhile, this initial study has the following limitations. First, the developed prototype only adopts the classic illumination and data acquisition strategies. Each raw image was obtained by a grayscale camera, which is under single LED illumination with a single wavelength. Thus, many new FPM techniques can be applied to improve the current system. For example, a color camera can be used to simultaneously capture the RGB channels of the raw data [37, 55], and the multiplexed illumination technique can be used to reduce the total number of raw images [7, 56]. Second, more different kinds of clinical histology/cytology [57] samples should be used to comprehensively assess the clinical utility. Third, instead of using only one doublet, it is also possible to utilize the available off-the-shelf lens elements to design a simple objective lens with a large optical aperture

and acceptable aberration for FPM. In summary, this study suggests a potential direction to develop a more cost-effective FPM system.

5 Evaluating the Depth of Field for Fourier Ptychography Microscopy

5.1 Introduction

Microscopes are routinely used in today's clinical practice, and microscopic scanners are clinically needed in modern pathologic labs to digitize the prepared samples. Currently, the performance of the available digitizers is partially limited by their space-bandwidth product [21], which cannot achieve a large field of view and high resolving power simultaneously. Fourier Ptychography Microscopy (FPM) [3, 23, 31] is an emerging technology that can overcome this limitation. This technology utilizes angled illumination to capture a series of raw images containing information of certain subregions of the sample Fourier spectrum, and these raw images were computationally synthesized [6, 7] to generate high-resolution images. Additionally, because the raw patterns are acquired under a low magnification objective lens (e.g. 4×), and the depth of the field (DOF) of the raw patterns determines the DOF of the final image, FPM offers a larger DOF compared to conventional microscopes with comparable NAs [11]. As a result, the FPM scanner eliminates the need for high precision moving stages, vastly lowering the overall cost of the system.

However, there are only few studies that focused on the experimental assessment of the DOF of FPM systems. Accordingly, in this study, we estimated the DOF of our FPM prototype, under both conventional single LED illumination and symmetric illumination conditions [40, 45, 58]. The MTF curves of the FPM system were first estimated using the standard USAF1951 resolution target, and the spatial frequency at which the MTF value drops to 0.5 was determined. Next, the bar pattern at this frequency was imaged at different

focusing positions and the contrast values were then assessed accordingly. The system's DOF was estimated as the range in which the contrast remains above 80% of the maximum value [59]. Further details are presented in the following sections.

5.2 Materials and Methods

The illumination wavelength used in this study is 0.53 μm (green), and accordingly a monochromatic camera (FL20BW) was utilized. The rest of the system setups remain consistent with what is outlined in section 3.2.1, including both conventional single LED and symmetric illumination methods. To evaluate the DOF of the FPM imaging system, we first reconstructed the images of a USAF 1951 resolution target placed on the in-focus position. The MTF curve was generated using the normalized contrast values and smoothed by a curve fitting algorithm [8], on which the spatial frequency corresponding to half of the maximum (0.5) was determined. The bar pattern corresponding to this specified frequency was utilized to evaluate the system's DOF. We varied the position of the sample stage up and down to create different defocus distances. At each position, the raw images were collected to reconstruct the chosen bar pattern. After reconstruction, the contrast values of the selected line pair were assessed, and the system DOF was estimated as the range where the measured contrast value remains above 80% of the maximum [59].

While the primary DOF measurements were conducted using the 4×/0.13 NA objective lens, complementary experiments were performed with the 4×/0.1 NA and 10×/0.25 NA objective lenses (set up exactly as in Chapter 2) on blood chromosome samples for a broader evaluation. All results are summarized and tabulated below.

5.3 Results

The reconstructed high-resolution images of the USAF1951 resolution target are

shown in Figure 5-1. Figure 5-1a depicts the raw image captured using the 4×/0.13 NA objective lens under central LED illumination, where group 7-6 (228 lp/mm) is clearly resolved, which approximately aligns with the theoretically calculated results (~201 lp/mm). Figures 5-1b and c illustrate the high-resolution images of the target recovered by the conventional illumination FPM and symmetric illumination FPM, respectively. Both methods can resolve patterns up to group 9-6 (912 lp/mm), indicating comparable resolving power.

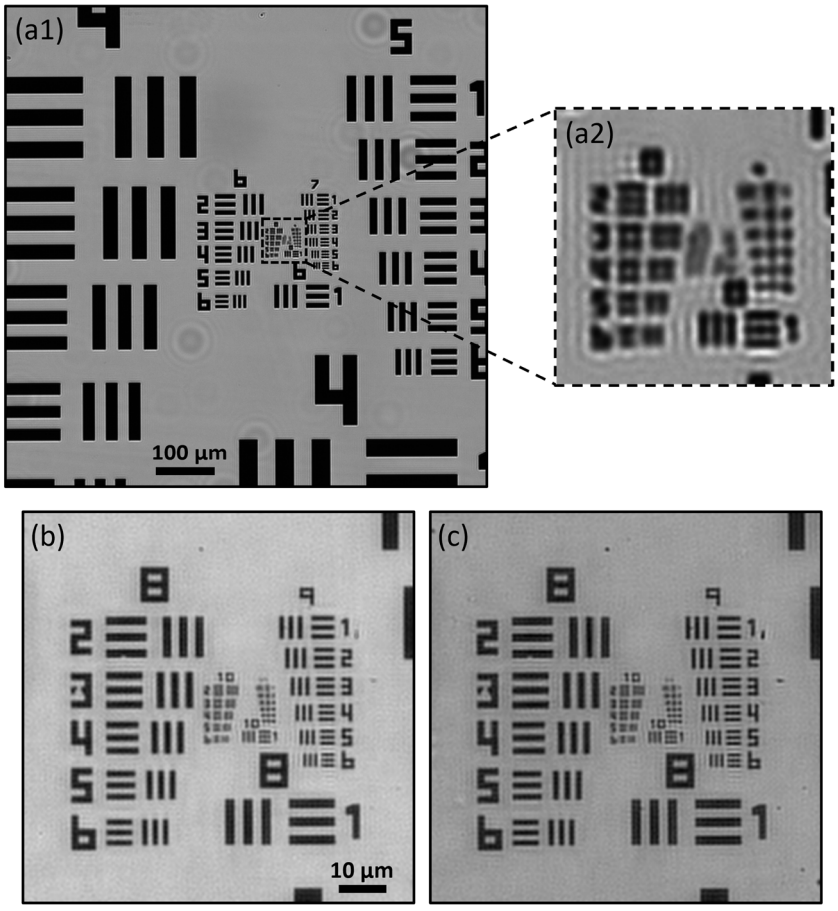


Figure 5-1: Images of the USAF resolution target. (a) One raw image acquired with the 4×/0.13 NA objective lens; (b) Reconstructed image of conventional illumination FPM; (c) Reconstructed image of symmetric illumination FPM.

Figure 5-2 illustrates the corresponding MTF curves of the conventional illumination FPM and the conventional microscope, both equipped with the 4×/0.13 NA objective lens. The two MTF curves decrease as the spatial frequency increases and drop to 0.5 at 248.9 lp/mm and 810.8 lp/mm, respectively. Therefore, the following DOFs were measured using bar patterns of group 8-1 (256 lp/mm) and group 9-5 (812 lp/mm) on the USAF1951 resolution target. Figure 5-3 demonstrates the reconstructed results of conventional and symmetric illumination FPM at different defocused positions. When the raw images were captured 10 μm away from the focal plane, the bar patterns (Figure 5-3b1) at the half-maximum frequency (group 9-5) in the symmetric illumination FPM results remain identical to the well-focused results, while the bar patterns (Figure 5-3a1) from the conventional illumination are slightly blurred. When the out-of-focus distance increases to 15 μm, the bar patterns from the symmetric illumination show slight blurring (Figure 5-3b2), whereas the conventional illumination results are significantly deteriorated (Figure 5-3a2).

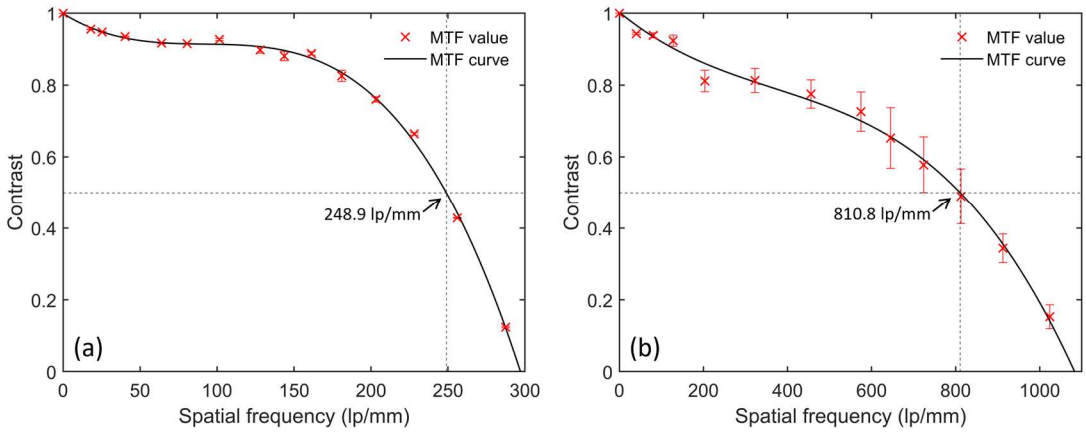


Figure 5-2: The MTF curves estimated by the image of the USAF resolution target. (a) 4×/0.13 NA objective lens; (b) Conventional illumination FPM.

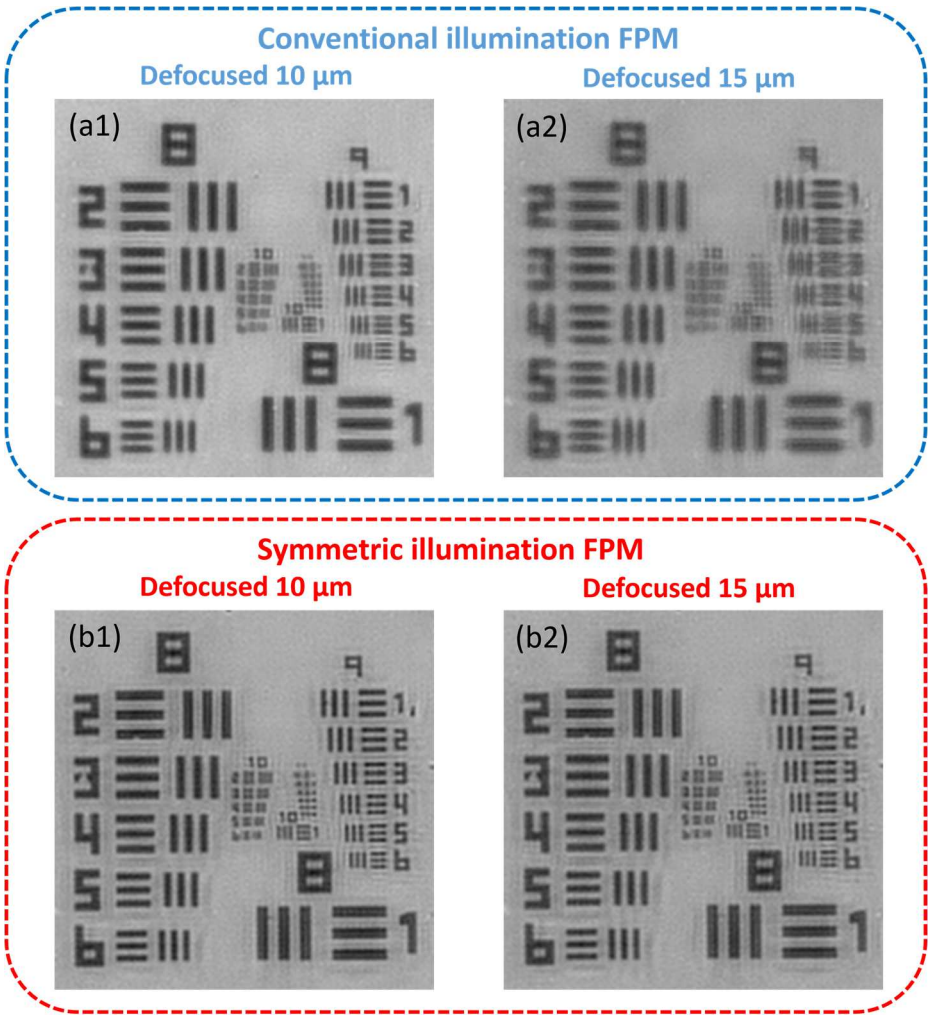
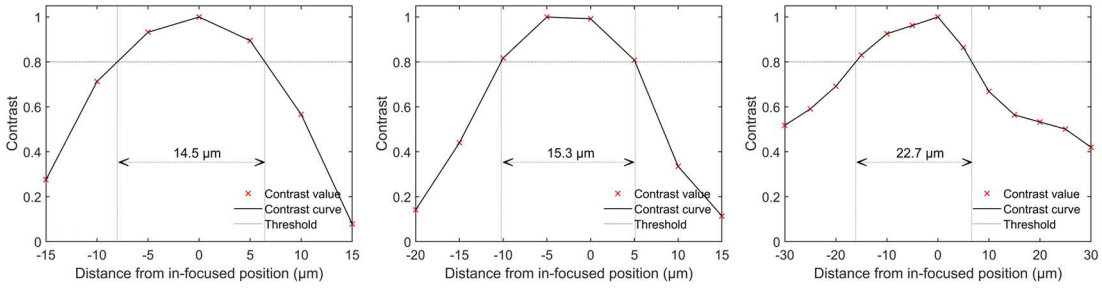


Figure 5-3: FPM reconstruction results with 10 and 15 μm defocused distances, respectively. The raw images were captured by the 4×/0.13 NA objective lens. (a) Conventional illumination FPM; (b) Symmetric illumination FPM.

Figure 5-4 indicates the estimated contrast values for the 4×/0.13 NA objective lens, conventional illumination FPM, and symmetrical illumination FPM, respectively. The estimated DOFs, in which the contrast value is higher than 0.8, are 14.5 and 15.3 μm for the 4×/0.13 NA objective lens (Figure 5-4a) and conventional illumination FPM (Figure 5-4b), respectively. Meanwhile, the estimated DOF increases to 22.7 μm for the symmetric illumination FPM (Figure 5-4c), which is substantially larger than that of the conventional

single LED illumination FPM and the 4×/0.13 NA objective lens used in this study.



(a) (b) (c)

Figure 5-4: The normalized contrast value at half-maximum spatial frequency with respect to the distance from in-focus position. (a) 4×/0.13 NA objective lens; (b) Conventional illumination FPM; (c) Symmetric illumination FPM.

Finally, Figures 5-5 and 5-6 indicate the reconstructed chromosome images which were captured at different focusing positions. For Figure 5-5a1, the raw patterns used for its reconstruction were acquired under the 4×/0.1 NA objective lens at the in-focused position. The reconstructed chromosome image, with a 10 μm defocus, still maintains a satisfactory quality as shown in Figure 5-5a2. The raw patterns of Figure 5-5a3 were captured 20 μm away from the in-focused position. Compared with Figure 5-5a1, the band patterns in Figure 5-5a3 are somewhat deteriorated but still acceptable. Meanwhile, when the chromosomes were directly captured under the 20×/0.4 NA objective lens (Figures 5-5b1/2/3), the band patterns show slight degradation when it was placed only 3 μm away from the in-focused position. Similarly, the raw patterns used for Figure 5-6a and 5-6b were captured using the 10×/0.25 NA objective lens, at in-focus position and 3.75 μm away, respectively. The band patterns in Figure 5-6b are generally as sharp as the band patterns at Figure 5-6a. However, when imaging with the 100×/1.25 NA objective lens (Figures 5-

6c/d), the chromosome bands become fuzzy when the sample was placed 2.5 μm away from the focal plane.

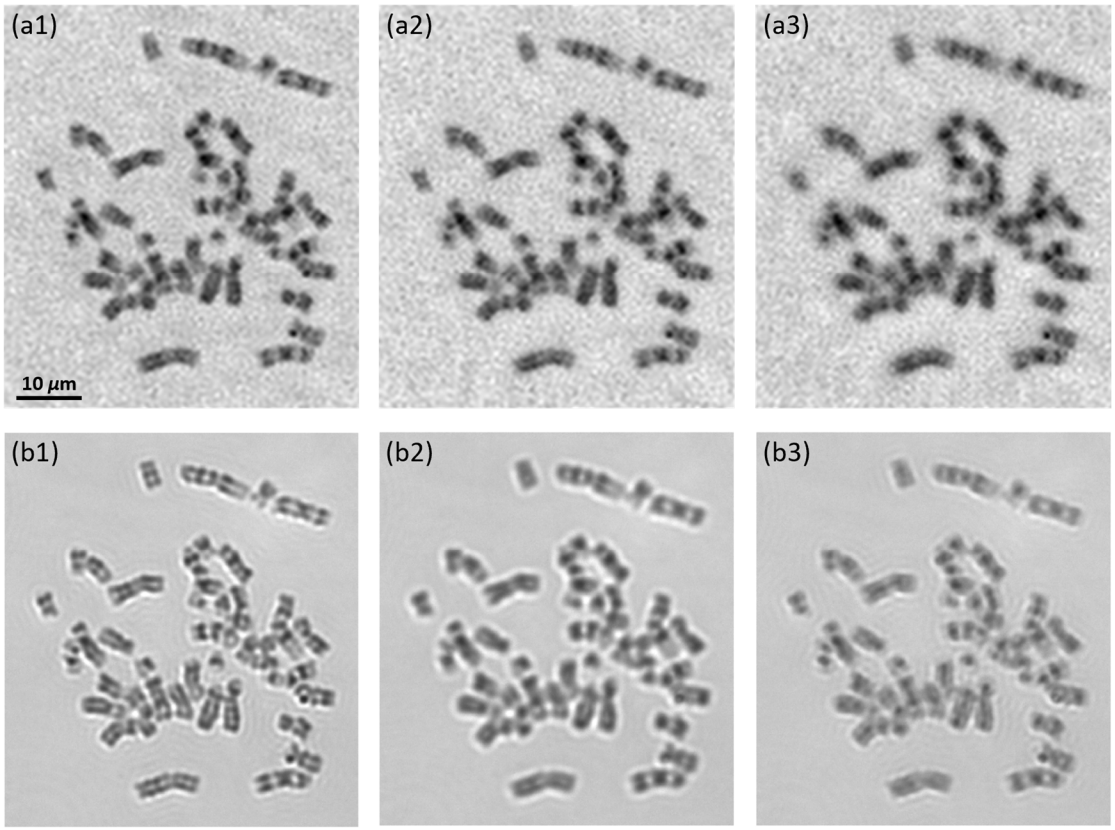


Figure 5-5: The images of a sample chromosome obtained from the patient's blood sample. The raw patterns of the chromosome were captured under the 4×/0.1 NA objective lens and reconstructed by FPM technology. The chromosome was placed at the (a1) in-focused position, (a2) 10 μm away from the in-focused position, and (a3) 20 μm away from the in-focused position. The chromosome was captured directly by the 20×/0.4 NA objective lens, which was placed at the (b1) in-focused position, (b2) 1.5 μm away from the in-focused position, and (b3) 3 μm away from the in-focused position.

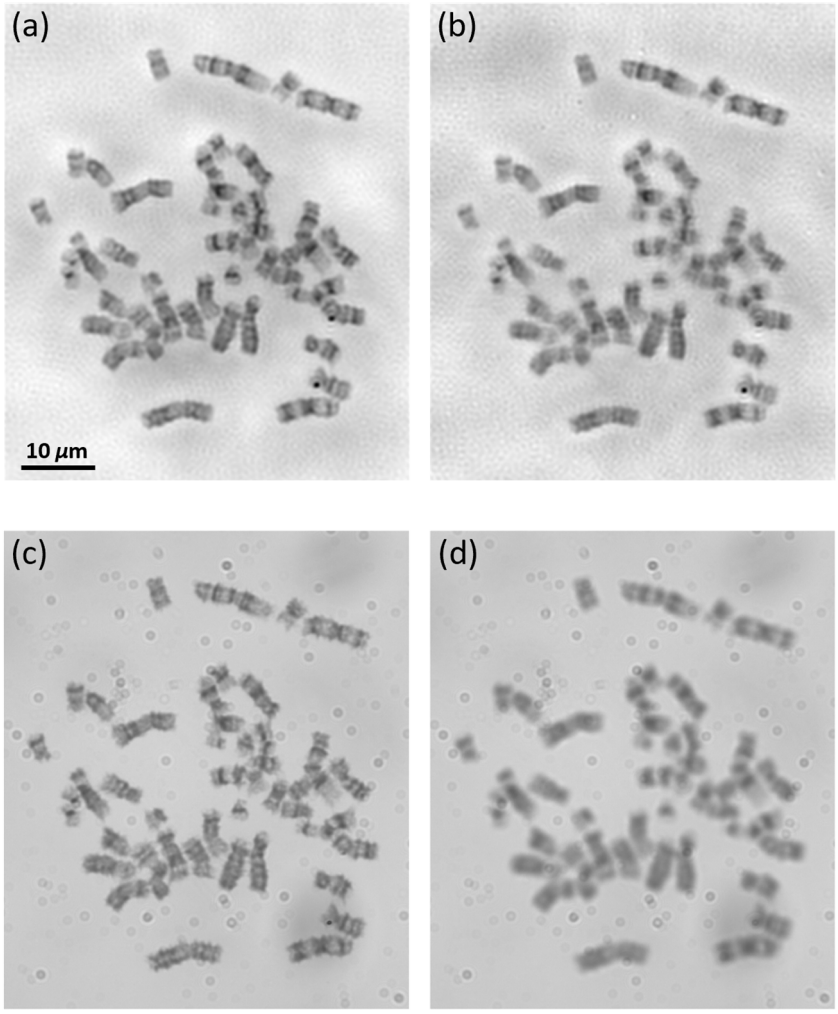


Figure 5-6: The images of a sample chromosome obtained from the patient's blood sample. The raw patterns of the chromosome were captured under the 10×/0.25 NA objective lens and reconstructed by FPM technology. The chromosome was placed at the (a) in-focused position and (b) 3.75 μm away from the in-focused position. The chromosome was captured directly by the 100×/1.25 NA objective lens, which was placed at the (c) in-focused position and (d) 2.5 μm away from the in-focused position.

5.4 Discussion and Conclusion

In this study, we experimentally investigated the system DOF for conventional and symmetric illumination FPM systems. The results indicate that the measured DOF is

approximately 15 μm for conventional illumination FPM, and 22 μm for symmetric illumination FPM. The DOF of conventional illumination FPM is approximately the same as that of the equipped objective lens (15.3μm vs 14.5μm), which agrees with the theoretical prediction. The sharpness of the reconstructed image is determined by the sharpness of the raw patterns. If the raw images are captured within the DOF, then the reconstructed image will also be within this range. The DOF is extended by symmetric illumination, as analyzed in previous research [45]. This extended DOF should be useful for the development of a low-cost sample digitizer, as it can further loosen the requirements for the mechanical focusing precision of the scanning stage. Meanwhile, it can also facilitate the development of color camera based FPM, as the focusing positions of RGB channels cannot be aligned even when using the achromatic tube lens.

Despite the significance, this study also has several limitations. The estimated value of the DOF does not agree with the previous report [45], which is based on only subject assessment of standard target or biology samples. It remains an open question on theoretical computation of the extended DOF under symmetric illumination. Second, we noticed that the measured contrast curve is not symmetric, which can be due to the fact that the strict central position is not accurately determined during the experiment. This is especially difficult for low magnification objective lenses because of the relatively large DOF. Some methods should be developed to accurately determine the well-focused location and image denoising algorithms [60, 61] should be applied to help improve image quality.

6 Conclusions and Future Work

6.1 Summary of the Dissertation

Fourier ptychographic microscopy is an emerging technology, which has the potential to be applied in future clinical practices. To achieve this, several technical challenges need to be addressed. First, we need to evaluate the quality of the clinical sample images generated by FPM and compare them with the conventional microscope having equivalent resolving power. Second, we need to explore new raw image acquisition strategies to improve data obtaining efficiency. Third, we should explore new methods to further reduce the cost of the whole FPM system. Fourth, the depth of the field should be thoroughly evaluated to ensure satisfactory image quality.

Accordingly, in this dissertation, a total of four investigations were performed to address the above technical challenges. First, under an experimental FPM prototype equipped with 4×/0.1 NA or 10×/0.25 NA objective lenses, we imaged the analyzable metaphase chromosomes, which were acquired from the peripheral blood and bone marrow samples of leukemia patients. The chromosome feature qualities of the 4×/0.1 NA FPM system are comparable to the results under the conventional 20×/0.4 NA lens, whereas the feature qualities of the 10×/0.25 NA FPM system are better than the conventional 60×/0.95 NA lens and comparable to the conventional 100×/1.25 NA lens. Second, we combined the symmetric illumination and a color detector to enhance throughput of the Fourier ptychographic microscopy (FPM) system. When using the 4×/0.13 NA objective lens, symmetric illumination FPM preserves more tissue details, which is superior to the results captured by the conventional 20 × /0.4 NA objective lens. In addition, the new approach is able to accelerate image acquisition by up to 12 times. Third, we attempted to develop and

evaluate a doublet based FPM prototype. Accordingly, the commercial 4× objective lens is replaced by one commercial achromatic doublet lens with a focal length of 60mm. The measured modulation transfer function (MTF) curves demonstrate that the cutting off frequencies are 833.9 lp/mm, which is comparable to 4×/0.1 NA objective lens based FPM system (844.3 lp/mm) or the 20×/0.4 NA objective lens on the conventional microscope (886.5 lp/mm). Fourth, we experimentally measured the DOF for our FPM prototype under different illumination conditions. During the experiment, the FPM prototype was equipped with a 4×/0.13 NA objective lens, and the results demonstrate that the DOF of the single LED illumination FPM is 15.3 μm, which is close to the DOF of the objective lens (14.5 μm). The DOF increases to 22.7 μm when symmetric illumination is adopted, which agrees with the theoretical conclusion.

6.2 Future Works

Despite the encouraging results achieved in this study, we acknowledge certain limitations that should be further addressed in future work to enhance the reliability and applicability of the findings.

First, only a limited set of clinical samples was used for imaging performance evaluation of the FPM system: The developed FPM prototypes were majorly evaluated by metaphase chromosome samples and ovarian cancer H&E tissue samples. These two kinds of clinical samples are representative but not comprehensive. A broader variety of clinical samples should be included in the future study.

Second, the proposed color detector based symmetric illumination FPM has the potential to improve the data acquisition efficiency by up to 12 times, but this has not been experimentally tested in our research. In addition, the acquisition efficiency can be further

enhanced by simultaneously activating more than two LEDs, which should be investigated in future studies.

Third, we focused on reconstructing a small region of interest as a proof of concept (e.g., 200×200). Given that the current detector can generate a very large frame (e.g., 4000×4000). The efficiency of large frame reconstruction has not been investigated in this work. Currently, reconstructing a 200×200 ROI (region of interest) needs about tens of seconds [62, 63]. When the frame size increases to 4000×4000, one potential solution is to divide the large frame into 200×200 regions and reconstruct all these regions in parallel. However, this approach may introduce stitching artifacts. Another possible solution is to develop deep learning based image reconstructing algorithms [64, 65]. As compared to model based approaches, deep learning based reconstruction is fully parallel, and can achieve up to 50× acceleration [64].

Fourth, the doublet based objective lens can be further replaced by a customized objective lens, which can provide improved SBP as compared to the commercial product. Additionally, it is also possible to use the in-stock lenses for the objective lens design [66, 67] to achieve an optimal balance between optical performance and cost.

Fifth, despite its advantages, the proposed FPM prototype is not compared with the commercially available pathology slide digitizers. For this purpose, the current prototype should be extended to an FPM scanning microscope which is able to generate whole slide images for clinical samples. Then its performance, including scanning speed and image quality, will be comprehensively evaluated and compared with commercial scanners under clinical settings.

6.3 Peer-Reviewed Publications

Journal Publications (3 as first author, 4 as co-first author, and 2 as co-author):

1) M. A. Jones, **K. Zhang***, R. Faiz, W. Islam, J. Jo, B. Zheng, Y.C. Qiu, "Pseudo color image generation for improving the performance of deep transfer learning-based computer aided diagnosis schemes in breast mass classification", Journal of Imaging Informatics in Medicine, 2024, accepted.

2) **K. Zhang**, N. Abdoli, P. Gilley, Y. Sadri, X. Chen, T. Thai, L. Dockery, K. Moore, R. Mannel, Y.C. Qiu, "Developing a Novel Image Marker to Predict the Responses of Neoadjuvant Chemotherapy (NACT) for Ovarian Cancer Patients", Computers in Biology and Medicine, 172, 108240-108247, 2024.

3) P. Gilley, **K. Zhang***, N. Abdoli, Y. Sadri, L. Adhikari, K.M. Fung, Y.C. Qiu, "Utilizing a Pathomics Biomarker to Predict the Effectiveness of Bevacizumab in Ovarian Cancer Treatment", Bioengineering, 11(7), 678-687, 2024.

4) P. Gilley, **K. Zhang***, N. Abdoli, Y. Sadri, L. Adhikari, K.M. Fung, Y.C. Qiu, "Development and Assessment of Multiple Illumination Color Fourier Ptychographic Microscopy for High Throughput Sample Digitization", Sensors, 24(14), 4505-4515, 2024.

5) **K. Zhang**, P. Gilley, X. Chen, K.M. Fung, H. Liu, B. Zheng, Y.C. Qiu, "Using symmetric illumination and color camera to achieve high throughput Fourier ptychographic microscopy", Journal of Biophotonics, 16(5), e202200303, 2023.

6) N. Abdoli, **K. Zhang***, P. Gilley, X. Chen, Y. Sadri, T. Thai, L. Dockery, K. Moore, R. Mannel, Y.C. Qiu, "Evaluating the Effectiveness of 2D and 3D Features for Predicting Tumor Response to Chemotherapy", Bioengineering, 10(11), 1334-1345, 2023.

7) X. Chen, X. Wang, **K. Zhang**, R. Zhang, K.M. Fung, T. Thai, K. Moore, R. Mannel, H. Liu, B. Zheng, Y.C. Qiu, "Recent advances and clinical applications of deep learning in medical image analysis", Medical Image Analysis, 79,102444, 2022.

8) X. Chen, **K. Zhang**, N. Abdoli, P.W. Gilley, X. Wang, H. Liu, B. Zheng, Y.C. Qiu, "Transformers improve breast cancer diagnosis from unregistered multi-view mammograms", Diagnostics, 12(7), 1549-1562, 2022.

9) **K. Zhang**, X. Lu, X. Chen, R. Zhang, K.M. Fung, H. Liu, B. Zheng, S. Li, and Y.C. Qiu, "Using Fourier ptychography microscopy to achieve high-resolution chromosome imaging: an initial evaluation". Journal of Biomedical Optics, 27(1), 016504-016504, 2022.

Conference Proceedings:

10) P. Gilley, **K. Zhang**, N. Abdoli, Y. Sadri, X. Chen, K.M. Fung, H. Liu, Y.C. Qiu, "Evaluating the imaging quality of multiplexed illumination Fourier ptychography microscopy", Proc. of SPIE 12843, Biophotonics and Immune Responses XIX, 12843061-6, 2024.

11) Y. Sadri, N. Abdoli, **K. Zhang**, P. Gilley, X. Zhao, K. Prather, H. Shakir, Y.C. Qiu, “Developing a CT-image-based clinical marker to predict vasospasm for aneurysmal subarachnoid hemorrhage patients”, Proc. of SPIE 12843, Biophotonics and Immune Responses XIX, 12843041-6, 2024.

12) X. Chen, M. Hu, **K. Zhang**, N. Abdoli, Y. Sadri, P. Gilley, O.S.V. Chekuri, F.H. Omoumi, Y.C. Qiu, B. Zheng, X. Yang, "David vs. Goliath: large foundation models are not outperforming small models in multi-view mammogram breast cancer prediction", Proc. of SPIE 12927, Medical Imaging, 2024.

13) **K. Zhang**, P. Gilley, X. Chen, K.M. Fung, H. Liu, B. Zheng, Y.C. Qiu, "Evaluating the depth of field for Fourier ptychography microscopy", Proc. of SPIE 12380, Biophotonics and Immune Responses XVIII, 123800C1-6, 2023.

14) N. Abdoli, **K. Zhang**, P. Gilley, X. Chen, Y. Sadri, T. Thai, L. Dockery, K. Moore, R. Mannel, Y.C. Qiu, "Comparing the Effectiveness of 2D and 3D Features on Predicting the Response to Chemotherapy for Ovarian Cancer Patients", Proc. of SPIE 12380, Biophotonics and Immune Responses XVIII, 123800D1-6, 2023.

15) **K. Zhang**, P. Gilley, X. Lu, X. Cheng, N. Abdoli, R. Zhang, K.M. Fung, H. Liu, B. Zheng, Y.C. Qiu, "Improving the resolution of chromosome imaging by high numerical aperture Fourier ptychography microscopy", Proc. of SPIE 11961, Biophotonics and Immune Responses XVIII, 2022.

16) X. Chen, X. Wang, **K. Zhang**, K.M. Fung, T. Thai, K. Moore, R. Mannel, H. Liu, B. Zheng, Y.C. Qiu, "Virtual adversarial training for semi-supervised breast mass classification", Proc. of SPIE 11961, Biophotonics and Immune Responses XVIII, 2022.

Appendix A Radiomics and Deep Learning Applications in Ovarian Cancer Chemotherapy Outcome Prediction

In this appendix, we present two CAD studies that utilize conventional and/or modern algorithms to address various challenges in decision-making during cancer treatment. The first study focuses on the radiomics approach, in which we developed a novel image marker to predict the clinical outcomes of ovarian cancer patients undergoing neoadjuvant chemotherapy (NACT), based on support vector machines (SVM) [68]. The second study fine-tuned the CLIP (Contrastive Language-Image Pre-training) [69] model using LoRA (Low Rank Adaptation) [70] to predict chemotherapy response in patients diagnosed with advanced-stage ovarian cancer. We hope these efforts will facilitate the clinical-grade CAD development for personalized and effective cancer treatment.

A.1 Study 1: Radiomic Image Marker for Predicting Clinical Outcomes in Ovarian Cancer Neoadjuvant Chemotherapy (NACT)

A.1.1 Background

Being the most lethal gynecological malignancy, ovarian carcinomas are anticipated to result in approximately 19,680 new cases and 12,740 deaths within the US during 2024 [71, 72]. Among all the diagnosed patients, approximately 90% are categorized as epithelial carcinoma, which can be further divided into four subtypes, namely, serous, endometrioid, clear cell, and mucinous carcinomas [73]. Meanwhile, the remaining 10% of patients are diagnosed with non-epithelial ovarian cancers, which include germ-cell or sex cord stromal tumors [73]. Given that there are not effective early stage screening approaches, most ovarian cancer patients are diagnosed at an advanced stage (III or IV) [74, 75]. Currently, the standard of care is to first perform primary debulking surgery (PDS)

to remove the visible tumors. Then adjuvant chemotherapy (ACT) is followed up to control the residual invisible tumor cells. Recently, an alternative approach has emerged, suggesting the initiation of neoadjuvant chemotherapy (NACT) [76] before cytoreductive surgery (i.e. interval debulking surgery [IDS]) and additional follow-up chemotherapy. Some investigations have indicated its advantages for certain groups of patients who are not well-suited for PDS [77, 78].

However, due to tumor heterogeneity, responses to NACT therapy vary significantly across different patient subcategories [79]. As a result, the well responsive patients can receive optimal debulking in IDS with either no visible residual tumors or residual tumors measuring less than 1 cm in diameter, leading to the best overall survival after the following treatment [80-82]. Nevertheless, the remaining patients experiencing partial response may have to undergo suboptimal debulking treatment, leaving residual tumors larger than 1 cm in diameter. Therefore, to achieve an optimal treatment outcome, it is of crucial importance to develop a prognostic model which is capable of identifying patients who are likely to receive suboptimal debulking surgery after NACT therapy. To date, different approaches have been explored to predict the outcome of ovarian cancer treatments, which are based on FIGO (International Federation of Gynecology and Obstetrics) staging system [83], demographic information [84], radiology findings [85, 86], CA-125 (cancer antigen 125) [87, 88], HE4 (Human Epididymis Protein 4) [89], BRCA 1/2 gene mutations [90], and so forth. However, due to the limited reliability and robustness of these models, they are not yet widely recognized [85, 91] in clinical practice.

Among all those technologies used in the prognostic assessment, the computed tomography (CT) scan [92] has its unique superiority over the others due to its advantages

of wide availability and low operating cost [93, 94]. Meanwhile, radiomics [95-97], a recent emerging approach, can extract and quantify the tumor characteristics from CT images, aiming to uncover the underlying biological mechanism associated with patient's diagnosis or prognosis. However, few studies have been conducted on using these novel techniques for clinical outcome prediction in NACT treatment for ovarian cancer patients. In this investigation, we aim to develop a radiomics based CT image marker to accurately identify the patients who may not be able to receive optimal cytoreduction after NACT treatment.

A.1.2 Materials and Methods

A.1.2.1 Database

This study was approved by the Institutional Review Board (IRB 13649) at the University of Oklahoma. Given that all the patients' data are de-identified and collected retrospectively, the patients' consents are not needed. This dataset contains 42 advanced stage ovarian cancer patients who received NACT regimens. Among them, 28 patients underwent optimal debulking in the following IDS, while the remaining patients did not. All the selected patients were diagnosed with recurrent epithelial ovarian cancer (EOC). Their treatments were performed at the University of Oklahoma Health Sciences Center. The patients' histology types include endometrial, serous, clear cell, and mucinous. For each patient, we collected the pre-therapy CT images following the standard acquisition protocols. Specifically, the image obtainment was conducted on a GE LightSpeed VCT 64-detector or GE Discovery 600 16-detector CT machines. X-ray tube current was set to be from 100 to 600 mA to fit the various body sizes, with a tube power of 120 kVp. The examination was conducted by the following procedures: The contrast agent (Isovue 370,

100 cc) was first injected into the patient intravenously at a rate of 2–3 cc per second. The initial scan started 60 seconds after injection began, while the second scan started 5 minutes after the completion of injection. CT images were reconstructed to 1.25 mm in axial axis, and 2.5 mm in sagittal and coronal directions.

A.1.2.2 Image Feature Computation

Prior to the image feature computation, we performed tumor segmentation on each case to generate the 3D volume of interest (VOI). An experienced radiologist identified each tumor on the acquired CT images and attached annotation to the image showing the largest tumor area. With the annotated reference images, all the slices containing metastatic tumors were accurately identified and processed by a previously developed segmentation algorithm [27], which combines a multilayer topographic region growth algorithm with adaptive thresholds and a dynamic edge tracking method. Due to the heterogeneity of metastatic tumors, the automated segmentation results may not be sufficiently accurate for the following process. Therefore, the automatically segmented results were visually reviewed by experienced researchers and the tumor contours were manually adjusted as needed.

Next, a large amount of radiomic features were computed from the segmented tumors in the original CT images, using the open-source platform Pyradiomics v3.0.1 [28]. Three categories of features were computed: shape, first order, and texture features. The shape features characterize the 3D size and geometry of the lesion from diverse perspectives, such as voxel volume and elongation. First order features depict the intensity distribution of the voxels inside the segmented VOI. The texture features include gray level co-occurrence matrix (GLCM) features [29], gray level dependence matrix (GLDM)

features [30], gray level run length matrix (GLRLM) features [31], gray level size zone matrix (GLSZM) features [32], and neighboring gray tone difference matrix (NGTDM) features [33]. These features enable a quantitative depiction of the tumor textures by assessing the spatial distribution and variation of voxel values. For instance, GLDM features are generated from a matrix that describes the appearance of qualified neighboring voxels surrounding the central voxel to measure the dependency relationships across intensity levels. Furthermore, we also obtained extra first order and texture features from several filtered image types: exponential, gradient, local binary pattern 3D (LBP3D), logarithm, square, square root, and wavelet filters. In summary, a total of 1373 features extracted from the smallest tumor in each case were utilized for training the model, as illustrated in Table A-1.

Table A-1: The number of features of different image types and feature types used in this study.

Image Type	Feature Type							Total
	Shape	First order	GLCM	GLDM	GLRLM	GLSZM	NGTDM	
Original	14	18	24	14	16	16	5	107
Exponential	0	16	0	7	8	3	0	34
Gradient	0	18	24	14	16	16	5	93
LBP3D	0	45	21	14	16	16	5	117
Logarithm	0	18	24	14	16	16	5	93
Square	0	17	24	14	16	16	5	92
Square root	0	18	24	14	16	16	5	93
Wavelet	0	144	192	112	128	128	40	**744**
Total	14	294	**333**	203	232	227	70	1373

A.1.2.3 Develop a Machine Learning based Model to Predict the Clinical Outcome

Using the established feature pool, we developed a classification model to predict whether the disease can be optimally debulked in the IDS (Figure A-1). At the onset of the model, all the features were first scaled using the formula: $\mathbf{x}' = (\mathbf{x} - \mu)/\sigma$, where $\mathbf{x}$ is a

vector containing original values of a feature, and μ and σ are the corresponding mean and standard deviation calculated from the training set, respectively. Given the imbalanced nature of the dataset, SMOTE (synthetic minority over-sampling technique) was employed to augment data in the minority class [98], which is based on the imbalanced-learn library v0.9.0 [99]. In this method, it first selects a sample $\mathbf{x}'_r$ from the randomly shuffled minority class. Using $\mathbf{x}'_r$ as a reference point, the algorithm locates its k nearest neighbors among the minority group. A new sample $\mathbf{x}'_s$ is then synthesized using the formula: $\mathbf{x}'_s = \mathbf{x}'_r + \beta(\mathbf{x}'_i - \mathbf{x}'_r)$, where $\mathbf{x}'_i$ is one of the k nearest neighbors, and β is a value randomly selected from the interval (0, 1). This synthetic process continues until the minority dataset is augmented to a balanced level.

Next, feature dimension reduction was conducted using principal component analysis (PCA) [100] on the matrix $\mathbf{X}'$ containing features of all training instances. As a popular method, PCA can effectively reduce the dimension while retaining variance as much as possible [101], especially when the dataset is relatively small [102]. Accordingly, PCA generates the principal components (PCs) based on the co-variance matrix $\mathbf{M} = \mathbf{X}'^{\mathrm{T}}\mathbf{X}'$. Each PC is an eigenvector of the matrix M representing a direction in the redefined feature space, while the variance explained by this PC is represented by its corresponding eigenvalue. Based on the spectral theorem [103], these variances quantify the importance of each PC to the matrix. Since the variance decreases significantly, it becomes feasible to use only a small portion of the most important PCs to approximate the original feature vectors with satisfactory accuracy [100]. We used PCA to reduce the original feature set to less than 10 features prior to classification.

After that, a support vector machine (SVM) classifier [68] was implemented using the scikit-learn v1.0.2 library [104], which aims to differentiate between sub-optimally debulked and optimally debulked cases within the dataset. As a robust classification algorithm, SVM serves well in other related studies of prognostic modeling of NACT [105-107] and has demonstrated its effectiveness in small size dataset based classification tasks [108]. In binary classification tasks, rather than just finding one hyperplane that separates the two classes, SVM seeks an optimal hyperplane that has the largest margin between the two classes, relying on the samples closest to the hyperplane (i.e., support vectors). The SVM classifier with either a linear kernel (Linear-SVM) or a Gaussian radial basis function kernel (RBF-SVM) was adopted in the classification task. The regularization parameter C for both kernels was searched among a logarithmic array: 10^{-5}, 10^{-4}, …, 10^{2}. For the RBF kernel specifically, its kernel coefficient gamma was chosen from another logarithmic array: 10^{-3}, 10^{-2}, …, 10. The pipeline of this machine learning model is depicted in Figure A-1a.

Finally, we employed a nested leave-one-out cross-validation (LOOCV) [109, 110] to assess the classification performance of the SVM-based clinical outcome prediction models (Figure A-1b). LOOCV is a unique type of cross-validation technique, in which the testing set only contains one case and all the remaining cases are used to train the model. This approach can maximally mitigate the influence of random partitioning of dataset, ensuring an almost unbiased evaluation result [111]. However, since hyperparameter tuning was also implemented within the LOOCV process, it incorporates testing data into model selection, which may lead to an overestimation of the model's generalization performance. Thus, we adopted the nested LOOCV with 2 levels: The outer LOOCV is sorely responsible

for the model evaluation; the inner LOOCV is embedded within the outer LOOCV to fine-tune the hyperparameters of each module in the model, including SMOTE, PCA, and SVM. This nested structure effectively decouples the testing data from training data, leading to a significant decrease in evaluation bias, and bringing error estimation close to the true level. The model performance was assessed by the receiver operating characteristic (ROC) curve [112], and various metrics such as the area under the curve (AUC) [113] value, overall accuracy (ACC), positive predictive value (PPV), and negative predictive value (NPV). Overall, the model employed in this preliminary study has limited complexity given the small dataset size, which aims to minimize the risk of learning from noise rather than the actual patterns. We utilized PCA for dimension reduction and SVM for prognosis classification, recognizing their simplicity and robustness. Similar methodologies have been adopted successfully in prior radiomics based classification studies [114, 115].

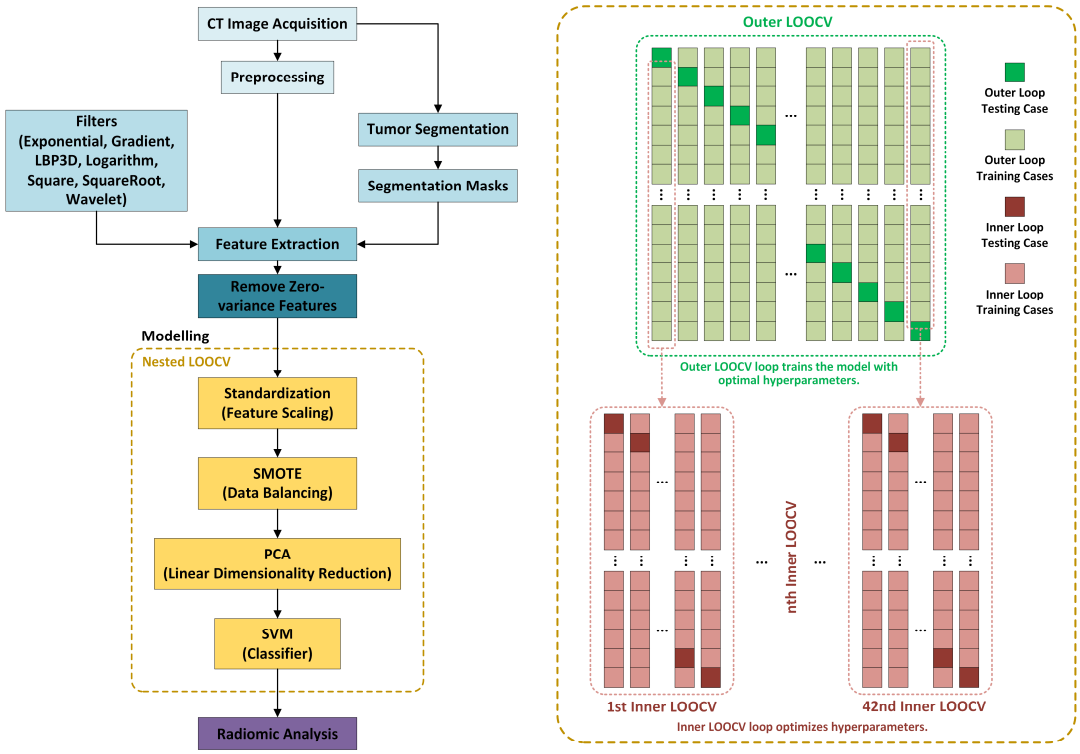


Figure A-1: Pipeline of the radiomic analysis in this study.

A.1.3 Results

Figure A-2 demonstrates one sample case of our tumor segmentation results as well as 6 of the extracted 1373 features from this tumor. All the CT images containing this tumor are displayed from left to right in the first row, and the outline of the ROI tightly enclosing the tumor tissue in each image is identified and marked by red color. The tumor in each CT image is isolated and illustrated in the second row which contains all the information that will be extracted for this case. The 3D shape of this tumor is displayed in the third row, along with 6 of its logarithm and gradient features.

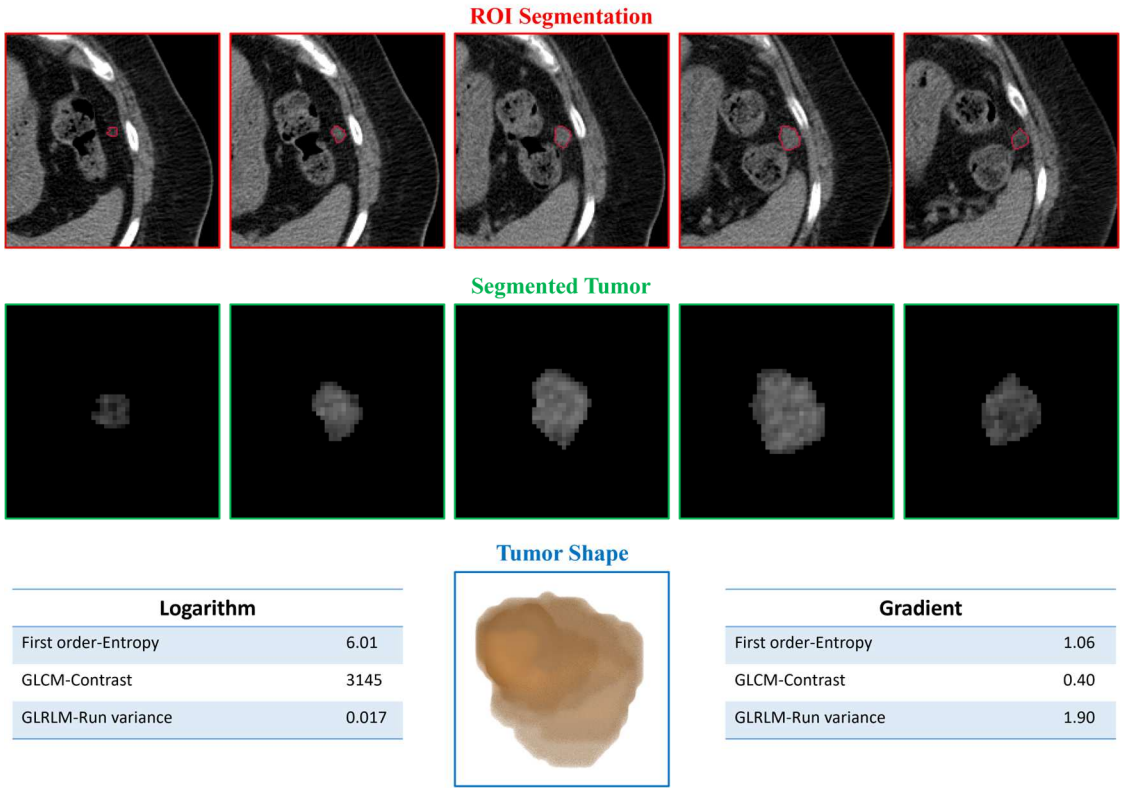


Figure A-2: An example of tumor segmentation and extracted features.

The hierarchically clustered heatmap [116] of Pearson's correlation coefficients for the extracted 1373 radiomic features is demonstrated in Figure A-3, with the values ranging from -1 to 1. The negative value is represented using a brown color, while the positive value is represented by a green color. Darker colors indicate a higher absolute value of the

correlation coefficient. The entire heatmap shows that the feature correlations are relatively low, and the dendrogram indicates that the associations occur on the features from different categories. Figure A-4 provides the distributions of correlation values among features of the same feature type or image type. From Figure A-4a, we can observe that the features within most subgroups have low association, while the only exception is the shape features. When sorting the features based on image types (Figure A-4b), it becomes evident that certain features extracted from the same image type exhibit relatively higher correlations with each other. For example, the median correlation coefficient for features derived from gradient filter processed images is around 0.38, while this value is only 0.23 for the wavelet filter subgroup. Figure A-5 plots the histogram of the absolute values of all the calculated correlation coefficients. It can be noted that 90.2 % of the correlation coefficient values are less than or equal to 0.5. The results indicate that the extracted 1373 raw features contain comprehensive information describing tumor attributes with negligible redundancy.

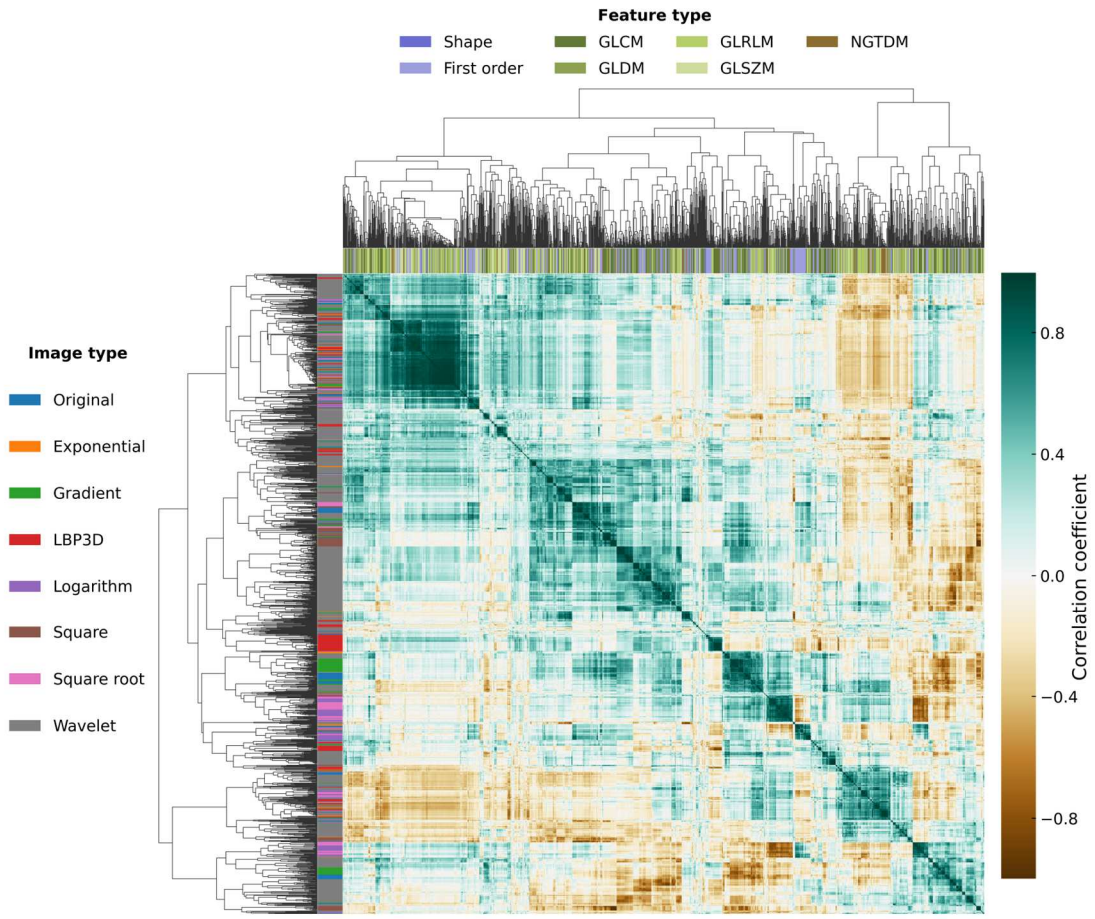


Figure A-3: The hierarchically clustered heatmap of Pearson's correlation coefficients between the extracted 1373 features.

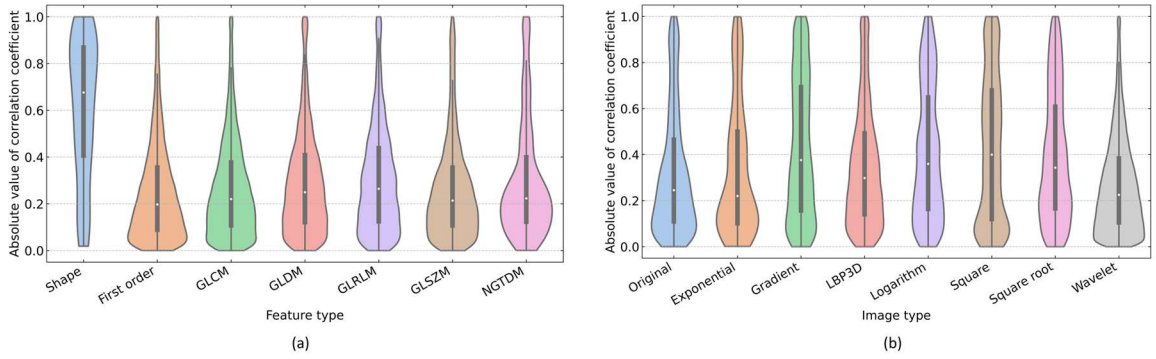


Figure A-4: The distribution of the absolute value of Pearson's correlation coefficients between features of the same feature type or image type.

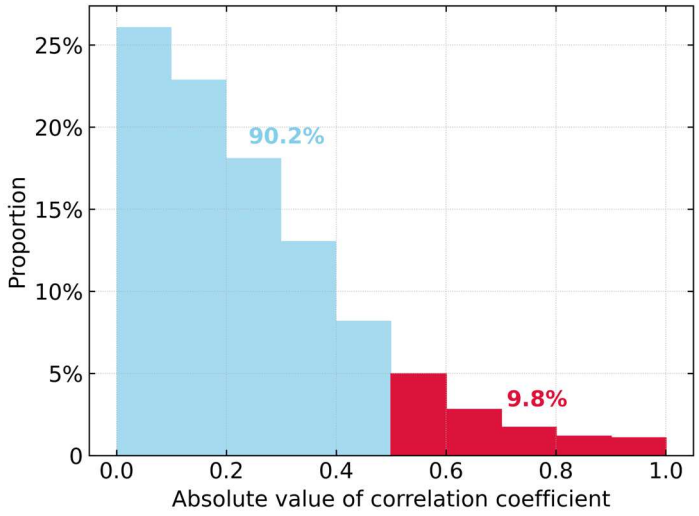


Figure A-5: The histogram of the absolute value of correlation coefficients between the extracted 1373 features.

Table A-2 summarizes the Linear-SVM based models' performance when only using one type of feature in modeling. The GLSZM features perform best among all different categories, yielding an AUC value of 0.638 ± 0.094 and an ACC value of 71.4 %. Meanwhile, the GLCM features achieve the second-highest performance, with AUC and ACC values of 0.633 ± 0.094 and 64.3 %, respectively. Similarly, the shape and first order features can still distinguish the positive cases from negative cases to a certain extent, resulting in AUC values of 0.617 ± 0.095 and 0.559 ± 0.096, respectively. In contrast, the remaining feature classes did not show significant discriminative powers. Figure A-6 demonstrates the averaged weight of different feature types applying on the PCA-derived new features when all the 1373 features are used as the input of the Linear-SVM based model. It reflects that the features from different subcategories have approximately equal contributions, with the GLDM and GLRLM contributing slightly more than the others.

Table A-2: The prediction performances of the Linear-SVM based models trained with different types of features.

Feature Type	AUC	ACC
Shape	0.617 ± 0.095	64.3 %
First order	0.559 ± 0.096	66.7 %
GLCM	0.633 ± 0.094	64.3 %
GLDM	0.408 ± 0.092	52.4 %
GLRLM	0.454 ± 0.094	59.5 %
GLSZM	**0.638 ± 0.094**	**71.4 %**
NGTDM	0.523 ± 0.096	38.1 %

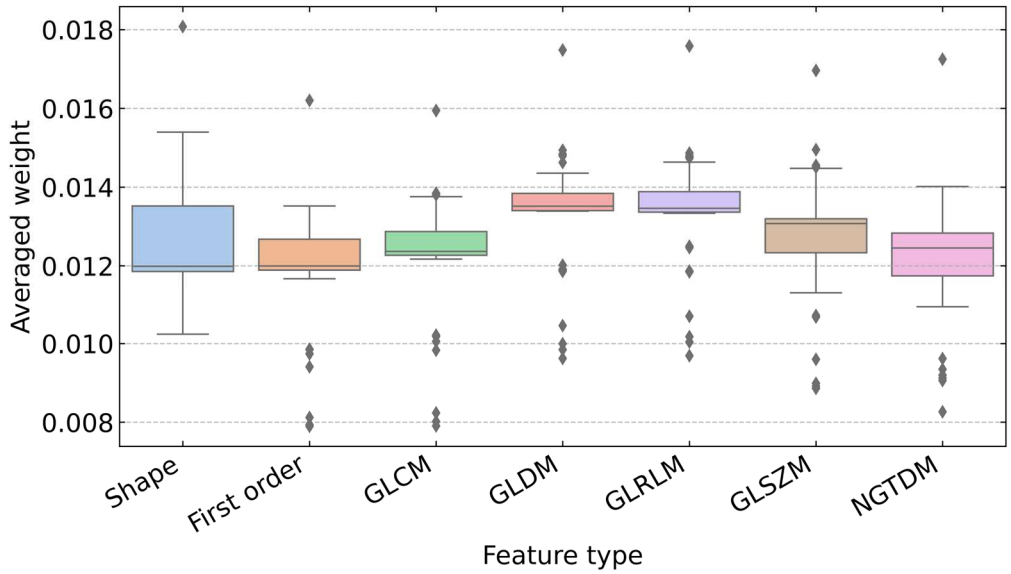


Figure A-6: Averaged weight of each feature type on the PCA-derived new features when training the Linear-SVM based model with the 1373 features.

The ROC curves used to evaluate the performance of the models trained with 1373 features are plotted in Figure A-7. The corresponding AUC value of the ROC curve for Linear-SVM model is 0.745 ± 0.086, which was produced by an ROC curve fitting program based on the maximum likelihood estimation method (ROCKIT, http://metz-roc.uchicago.edu/, University of Chicago). In contrast, the RBF-SVM based model produced an ROC curve that is generally higher than that of the Linear-SVM, enhancing

its AUC value to 0.806±0.078. Figure A-8 shows the confusion matrices for both models. As can be seen from Figure A-8(a), out of the 10 cases predicted as "Suboptimal" by the Linear-SVM based model, 7 of them were truly sub-optimally debulked in IDS. This yields a positive predictive value (PPV) of 70% (7/10). On the other hand, among the 32 cases predicted to be "Optimal" by the Linear-SVM, 7 cases were confirmed receiving suboptimal debulking, corresponding to a negative prediction value (NPV) of 78.1% (25/32). Meanwhile, the RBF-SVM model (Figure A-8b) offered PPV and NPV values of 81.8% (9/11) and 83.9% (26/31), respectively. Based on the above calculations, the total accuracy (ACC) values of the linear and RBF kernel based SVM models reached 76.2% (32/42) and 83.3% (35/42), respectively. This performance enhancement can be attributed to the presence of nonlinear prognostic information in the dataset, a characteristic that nonlinear RBF-SVM captures more efficiently than Linear-SVM.

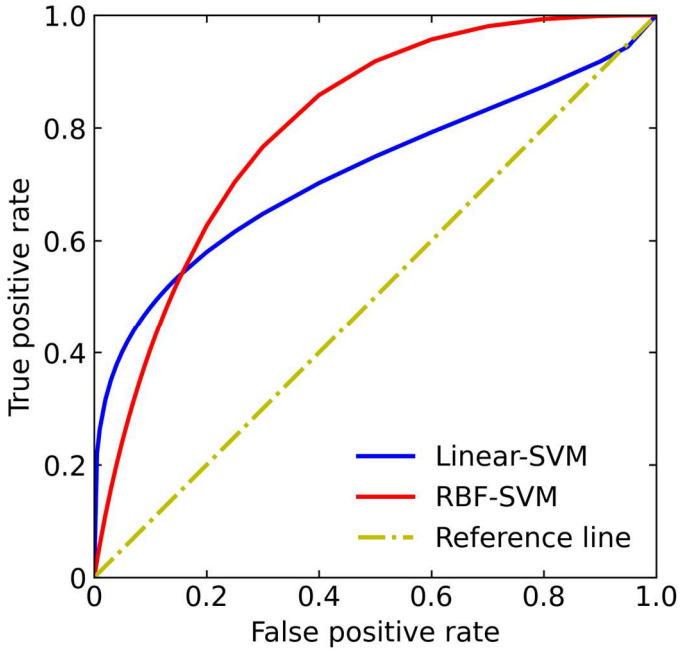


Figure A-7: The ROC curves for Linear-SVM and RBF-SVM based models trained with 1373 features, respectively.

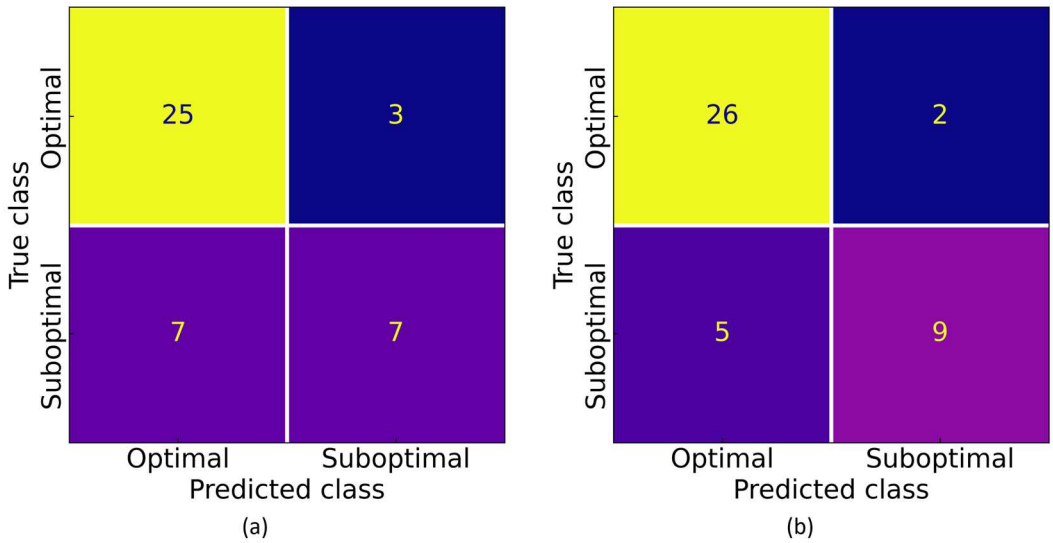


Figure A-8: The confusion matrices for (a) Linear-SVM and (b) RBF-SVM based models trained with 1373 features, respectively.

A.1.4 Discussion

In this investigation, a novel image marker was initially developed and verified to categorize patients into optimal and suboptimal debulking groups prior to the NACT treatment. The major significance is this marker's potential to accurately stratify patients at an early stage. If subsequent research successfully validates this strategy, it could assist gynecologic oncologists in identifying patients who may not benefit from the planned NACT treatment. Thus, alternative treatments can be chosen before chemotherapy is administered, allowing patients to avoid ineffective treatments and minimize the risk of severe side effects. Besides its significance, this study has the following unique characteristics.

Different from the conventional image based NACT prognostic models, our proposed model utilized radiomics features to comprehensively quantify the tumor characteristics. The conventional models rely on either visual findings on the presence of malignancy sites (e.g. pleural effusion or mesentery deposits [85]) or some simple

measurement depicted on the acquired images (e.g. tumor diameters [117, 118]). Despite the association between these features and cytoreduction, this method suffers from inter- and intra-radiologist variability and time-varied diagnostic criteria [86, 119], which causes low robustness and unsatisfactory performance. In our study, the copious tumor information was thoroughly conveyed by the highly varied radiomics features. Followed by the PCA and SVM, this information was further identified and synthesized to generate a single clinical marker for outcome prediction. In a dataset containing 42 ovarian cancer patients, the marker achieved an overall accuracy of 76.2% and 83.3% when utilizing linear and RBF kernel based SVM algorithms, respectively. The results indicate the effectiveness of our proposed method. In addition, as compared to the other investigated molecular markers, our proposed method will not create additional financial burdens on the patients as CT imaging is a routine examination for ovarian cancer patients.

Furthermore, to the best of our knowledge, the computer feature pool in this study is by far one of the largest feature pools for this specific or similar medical imaging task. We computed a total of 1373 radiomics features for NACT prognosis prediction, which can be divided into three main categories (e.g., shape, intensity, and texture). The prognostic capacities of various feature types were evaluated by training a Linear-SVM based model with each type individually. The result shows the five subcategories of texture features have varying levels of performance on this task: only GLSZM and GLCM indicate significant discriminative power in classifying optimal and suboptimal debulking in IDS (Table A-2). Given that these two texture feature subgroups describe the connected or co-occurrence zones with the same grayscale level, it implies the NACT prognosis is more sensitive to such an intensity change. However, in the final feature cluster, all these features

demonstrate an approximately equal contribution (Figure A-6), which may be attributed to the fact that the features from different groups provide complementary information to the final marker. This observation also implies the necessity of the radiomics method: each of these computed 1373 features characterizes the tumor heterogeneity uniquely, which is highly desirable in generating the final marker.

Despite the encouraging results, we acknowledge that the study has the following limitations. First, the dataset only contains 42 patients from a single institution. The performance and robustness of our proposed algorithm should be further verified on a comprehensive multiple-institution database with diversified patients. Second, the model is developed based on the conventional SVM algorithm, and our study only used radiomics features. Some emerging technologies such as deep learning [120-122] were not adopted. One possible improvement of this study would be to combine radiomics and deep learning techniques together to further enhance the prediction accuracy [123]. Third, the feature repeatability and reproducibility were not investigated in the current study. Utilizing radiomic features that remain consistent across various scanning setups, such as scanning equipment and acquisition software, can effectively decrease the risk of type I error and ensure a broader validity of the model [124]. Fourth, given the limited size of the dataset, we did not conduct a model parameter uncertainty study. Despite these limitations, this investigation initially demonstrates the efficacy of using radiomic features for predicting NACT outcomes in ovarian cancer treatment and lays a solid foundation for advancing precision treatment in future research.

A.2 Study 2: Fine-tunning CLIP (Contrastive Language-Image Pretraining) for Predicting Tumor Response in Ovarian Cancer Chemotherapy

A.2.1 Background

Ovarian cancer is commonly detected at an advanced stage and is highly aggressive, posing significant challenges in delivering effective treatment. As discussed in Appendix A.1, the typical treatment of ovarian cancer involves primary debulking surgery (PDS) followed by adjuvant chemotherapy. However, the heterogeneity of ovarian cancer leads to varying responses to chemotherapy among patients [125]. To improve patient outcomes and minimize the side effects of ineffective treatments by enabling personalized treatment plans for each individual, it is crucial to accurately predict the response of residual tumors to chemotherapy following PDS.

Computed tomography (CT) based Radiomics methods [126-134] have been extensively explored for tumor response predictions in ovarian cancer chemotherapy treatments. More recently, deep learning (DL) methods utilizing CT images have begun to emerge in this area [135-137]. Nevertheless, in contrast to the extensive research conducted in radiomics, DL studies for predicting chemotherapy responses in ovarian cancer remain limited, which might be attributed to the restricted availability of high quality datasets. In this regard, we propose to investigate the fine-tuning of CLIP (Contrastive Language-Image Pretraining) [69], a foundation model pretrained on a vast and diverse range of datasets, to predict ovarian cancer chemotherapy response. Unlike the conventional DL models, such as ResNet [138], which are typically trained on well labeled datasets (e.g., ImageNet [139]), CLIP was trained using contrastive learning to align text with their paired image within a unified embedding space [69], without being tied to any specific task. Thus,

CLIP is able to perform zero-shot learning [140] in which it identifies objects or concepts it has never encountered as examples during training. In this study, by leveraging CLIP's outstanding generalization capabilities with Low-Rank Adaptation (LoRA), we aim to mitigate the challenges posed by limited data availability.

A.2.2 Materials and Methods

A.2.2.1 Database

This study received approval from the Institutional Review Board (IRB13649) at the University of Oklahoma. The dataset used in this study was retrospectively collected from the CT series of 182 patients diagnosed with ovarian cancer. All patients underwent systematic chemotherapy following PDS to manage recurrent high-grade ovarian, peritoneal, or tubal carcinoma. Out of the 182 patients, 124 were classified as responders, achieving a 6-month progression-free (PFS) [141] of “Yes”, while 58 patients were identified as non-responders showing minimal response to treatment. The CT images were collected under standardized clinical protocols after PDS, using either a GE Light Speed VCT 64 or a GE Discovery 600 16-row detector scanner, with a tube voltage of 120 kVp. The current was set according to the patient's body size, ranging from 100 to 600 mA.

Considering that different cases have varying numbers of tumors, for each case we selected only the CT slice that contains the largest tumor area from the largest tumor. The tumors were initially segmented from slices by a hybrid algorithm [142] first, followed by a visual review and manual correction by senior researchers. Further details about the segmentation are provided in the study [132]. The segmented tumor was then shifted to the center of the image and cropped using a 200×200 square window, ensuring that the entire tumor fits within the window for all cases. The process is illustrated in Figure A-9.

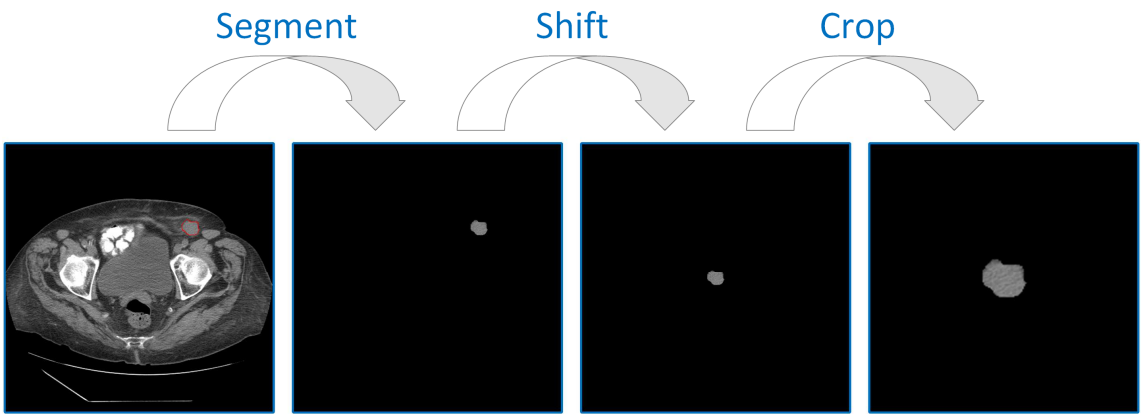


Figure A-9: Preprocessing pipeline.

A.2.2.2 Model Architecture of CLIP

There are two main separate building blocks within the CLIP model [69] to conduct contrastive learning: a vision encoder and a text encoder (Figure A-10). Those two encoders were trained on 400 million mutually matched image-text pairs acquired from the Internet to extract image and text embedding representations, respectively. The extracted embeddings are then projected into a shared embedding space. In this space, the image and text feature vectors are of the same dimensionality and are aligned with their corresponding pairs by maximizing their similarity, while simultaneously reducing the similarity between unmatched pairs to enhance the model's discrimination power.

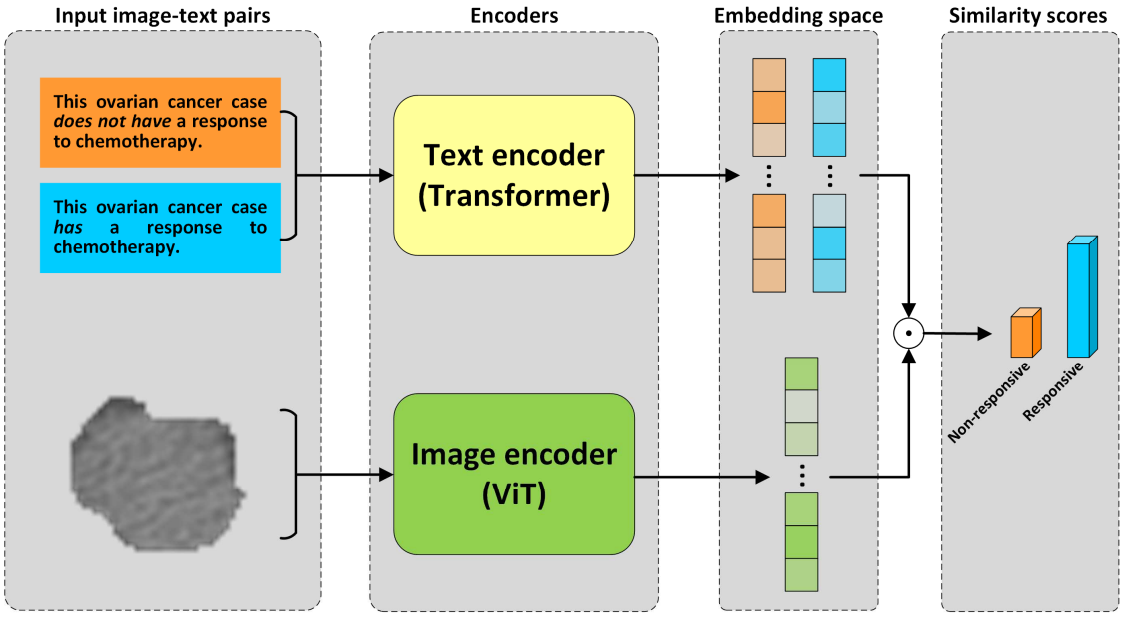


Figure A-10: Basic structure of CLIP.

In CLIP, the image encoder is responsible for compressing the image into a compact representation, preserving rich high-level semantic features that are indispensable for subsequent tasks. In this study, we selected two CLIP models utilizing ViT-B/16 and ViT-B/32 [143] as their vision encoders, respectively. These models first divide the input image into a sequence of small patches, which are then processed through 12 Transformer [144] blocks. Each Transformer block is composed of two major components: multi-head self-attention (MHSA), which enables the network to put more focus on the key parts of input, and feed-forward network (FFN), as shown in Figure A-11. The self-attention operation within MHSA first projects the input matrix $\boldsymbol{X} \in \mathbb{R}^{n \times d}$, where n represents the input sequence length and d denotes the dimensionality, into three attention components: the query $\boldsymbol{Q}$, key $\boldsymbol{K}$, and value $\boldsymbol{V}$. Those are calculated using trainable weight matrices $\boldsymbol{W}_q \in \mathbb{R}^{d \times d_k}$, $\boldsymbol{W}_k \in \mathbb{R}^{d \times d_k}$, $\boldsymbol{W}_v \in \mathbb{R}^{d \times d_v}$, respectively, as follows: $\boldsymbol{Q} = \boldsymbol{X}\boldsymbol{W}_q$, $\boldsymbol{K} = \boldsymbol{X}\boldsymbol{W}_k$, and $\boldsymbol{V} = \boldsymbol{X}\boldsymbol{W}_v$. Then the queries are compared to keys to calculate attention scores that are used

as weights to combine corresponding values: $\text{Attention}_{self}(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = \text{softmax}(\boldsymbol{Q}\boldsymbol{K}^{\text{T}}/\sqrt{\boldsymbol{d_k}})\boldsymbol{V}$. Rather than relying on a single self-attention operation, MHSA adopts multiple attention heads to process the data $\boldsymbol{X}$ parallelly. The outputs generated from each head are concatenated, and then projected back to the input dimension d: $\text{MHSA}(\boldsymbol{X}) = Concat(\text{Attention}_{self}^{1}, \cdots, \text{Attention}_{self}^{h})\boldsymbol{W_b}$, where $\boldsymbol{W_b} \in \mathbb{R}^{(hd_v)\times d}$, and h represents the number of heads. This enables the model to focus on various types of relationships within the data simultaneously, as each attention head works independently, thereby enhancing the network's capacity to handle intricate data. Alongside the image encoder, the text encoder in the selected CLIP configurations also contains 12 concatenated Transformer blocks, comprising a parallel architecture.

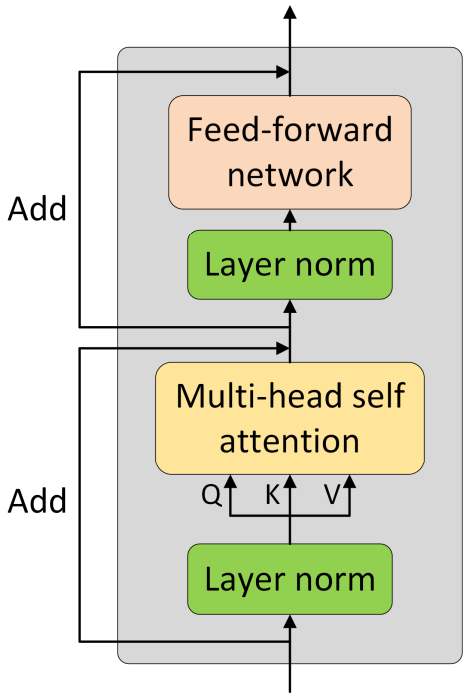


Figure A-11: A Transformer block.

A.2.2.3 Fine-tuning CLIP with LoRA (Low-Rank Adaptation)

With the emergence of large pre-trained models, such as BERT [145] and GPT-3 [146], which continuously grow in size, fine-tuning these models for downstream tasks becomes increasingly challenging. To mitigate overfitting and leverage the embedded knowledge in the large models when training on restricted downstream task datasets, a series of parameter-efficient fine-tuning (PEFT) methods [147] have been proposed. These methods are characterized by fine-tuning only a small portion of current parameters or integrating a minimal number of new parameters, aimed at achieving optimal generalization performance for a specific task. Low-Rank Adaptation (LoRA) [70] is one of the PEFT methods that modify the data flow passing through a pretrained weight matrix $\boldsymbol{W_T} \in \mathbb{R}^{m\times n}$ by introducing a trainable weight matrix $\Delta\boldsymbol{W}$ derived from the product of two low-rank matrices, $\boldsymbol{W}_A \in \mathbb{R}^{r\times n}$ and $\boldsymbol{W}_B \in \mathbb{R}^{m\times r}$: $\boldsymbol{h} = \boldsymbol{x}(\boldsymbol{W_T} + \boldsymbol{W_B W}_A) = \boldsymbol{x}(\boldsymbol{W_T} + \Delta\boldsymbol{W})$. During fine-tuning, only the introduced $\boldsymbol{W}_A$ and $\boldsymbol{W}_B$ are trained while all original parameters are frozen, as depicted in Figure A-12. Although LoRA can be implemented across any portion of the network, in this study the application of LoRA is focused on the attention mechanisms: $\boldsymbol{W}_q$, $\boldsymbol{W}_k$, and $\boldsymbol{W}_v$ in specified Transformer blocks, for reduced complexity and high performance. Different small rank values r were adopted to evaluate their impact on performance, including 1, 2, and 3.

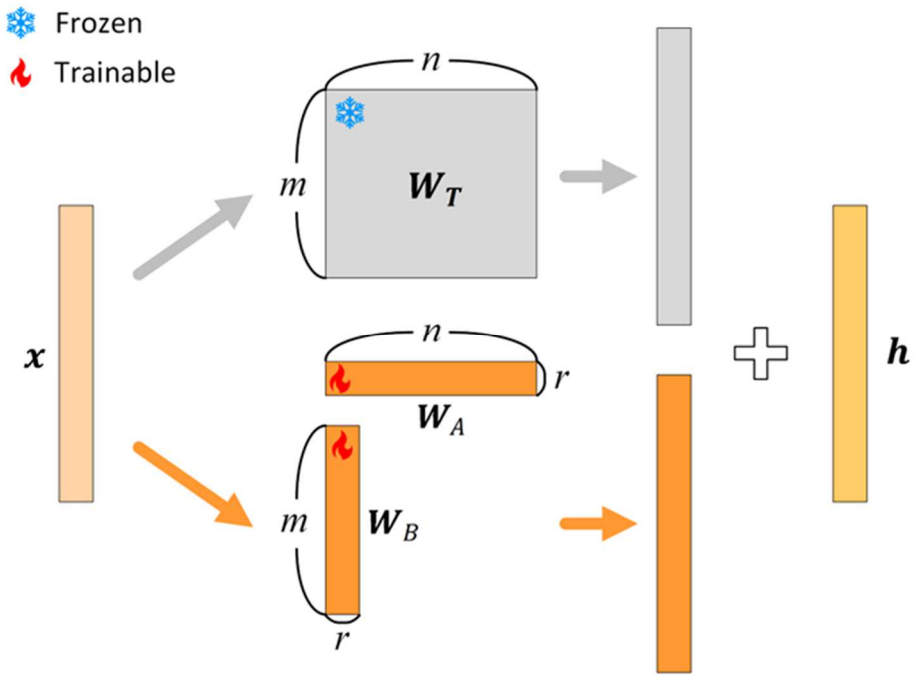


Figure A-12: Schematic of LoRA.

Given that there are 12 Transformer blocks for both image and text encoders in the ViT-B/16 and ViT-B/32 based CLIP models, we investigated training multiple models, each with a unique set of Transformer blocks fine-tuned with LoRA. To be specific, those models were fine-tuned or trained independently. For instance, one model had only its lowest block ($N = 1$) fine-tuned, while another model had both the first and second blocks fine-tuned; this pattern continues up to a model that had all 12 blocks fine-tuned (Figure A-13, marked as "Low"). These CLIP models are designated as ViT-B/16-Low or ViT-B/32-Low. Additionally, we also adopted an inverted approach by starting with a model in which only the highest block ($N = 12$) was fine-tuned, and then involving lower blocks in separate models (Figure A-13, marked as "High"). These CLIP models are designated as ViT-B/16-High or ViT-B/32-High.

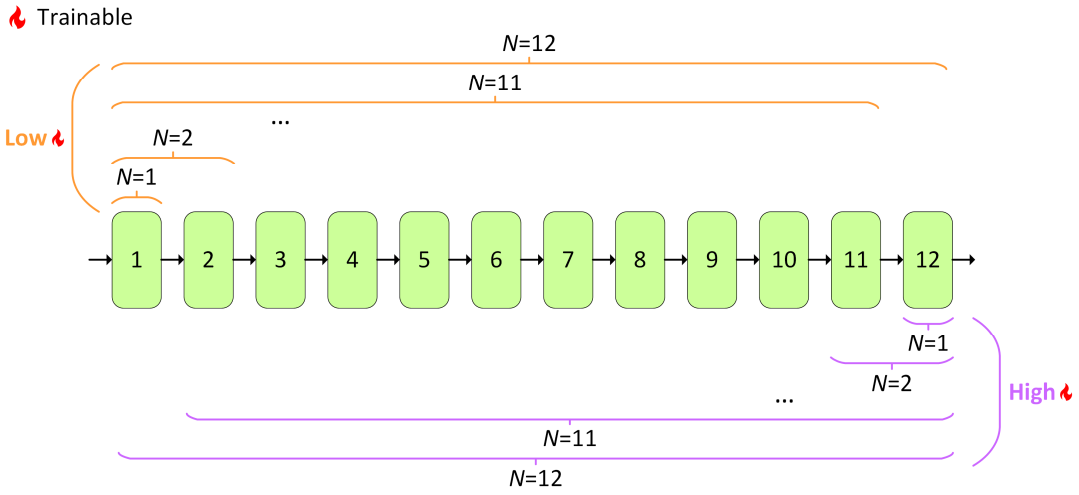


Figure A-13: Apply LoRA to a different set of Transformer blocks.

A.2.2.4 Training Implementations and Evaluation Metrics

During the training process, all pretrained parameters of the models were kept frozen, with only the low-rank matrices in LoRA being fine-tuned. The training utilized the Adam optimizer [148] with $\beta_1 = 0.9$, $\beta_2 = 0.99$, and a learning rate of 7×10^{-4}. The models were trained for a total of 1000 epochs, with the first 70 steps used for warm up followed by a cosine decay schedule for the remaining steps. We also applied weighted sampling [149] in training by assigning probabilities inversely proportional to class counts. In this way, the instance in the underrepresented class, specifically the non-responsive group, has improved likelihood of being selected for training in each iteration, thereby mitigating the impact of the significant imbalance in the dataset. Data augmentation strategies, including random 90-degree rotation, and horizontal and vertical flips, were applied.

To make comparison, we fine-tuned all original parameters of the CLIP model without utilizing LoRA. Additionally, we employed conventional fine-tuning strategies by removing the projection layer of CLIP image encoder (without using the text encode of

CLIP), as well as classification heads of Vision Transformer [143] (ViT-B/16 and ViT-B/32) and ResNet [138] (ResNet-50) models pretrained on ImageNet-1K [139]. These were substituted with randomly initialized one-layer classification heads that have two outputs, with earlier layers remaining frozen during training. The experiments were implemented with PyTorch [150] on the NVIDIA RTX 6000Ada GPU, providing 48GB of VRAM (video random access memory).

We conducted stratified 5-fold cross-validation to assess the models' performance, maintaining a consistent proportion of the two classes in each fold. This stratification helps reduce evaluation bias by training and testing on datasets that are close to true class distributions. The average classification Accuracy (ACC) and AUC (area under the receiver operating characteristic [ROC] curve) values are reported, offering a comprehensive examination of models' capabilities.

A.2.3 Results

Table A-3 presents a detailed comparison of the performance and training configurations of various CLIP models and conventionally fine-tuned deep learning models, evaluated with 5-fold cross-validation. The CLIP model ViT-B/16-Low, fine-tuned using LoRA with $N = 5, r = 1$, demonstrated the highest performance with an AUC of 0.785 ± 0.039 and an ACC of 0.780 ± 0.032. The second-best performance was achieved by another LoRA fine-tuned CLIP model, ViT-B/16-High with $N = 8, r = 2$, underscoring the effectiveness and adaptability of LoRA. In comparison, the full fine-tuning on the CLIP model (ViT-B/16 based) did not exhibit any advantage, especially in its low ACC performance. For zero-shot evaluation, the original CLIP model consistently classified all cases as non-responders, resulting in an ACC of 0.319. This suggests that the pretrained

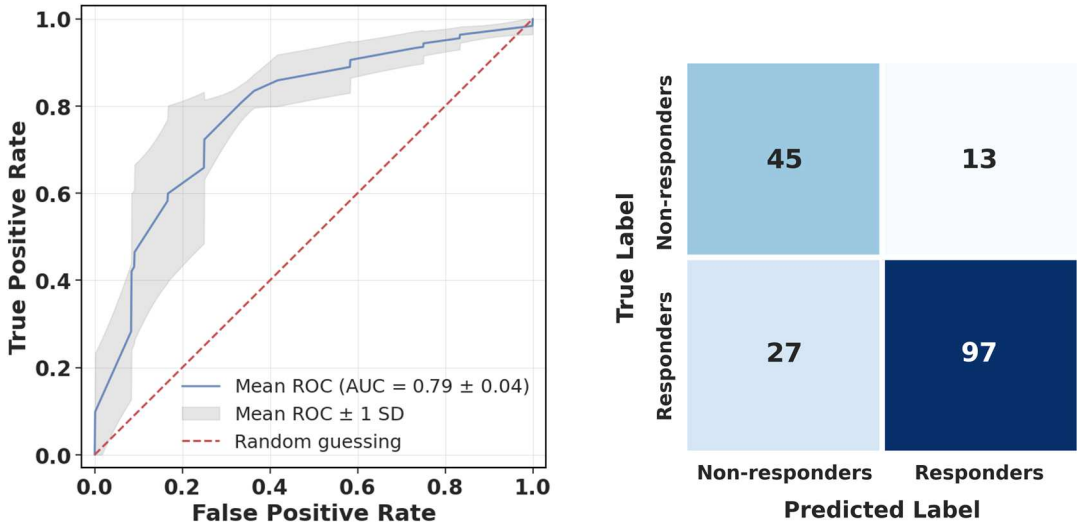
CLIP model lacks sufficient domain knowledge to differentiate between the two classes in this study. When employing the conventional fine-tuning method for the CLIP model, which involves using only its pretrained image encoder and substituting the projection layer with a trainable classification head (head-only), the model exhibited the lowest performance in both AUC and ACC among all fine-tuning approaches. This performance was notably worse than that of conventional ViT and ResNet models which were pretrained on ImageNet-1K and fine-tuned using the same head-only method. The ResNet-50 model, as a representative convolution neural network (CNN), outperformed both two ImageNet-1K pretrained ViT models, with AUC and ACC of 0.754 ± 0.089 and 0.725 ± 0.110, respectively.

Table A-3: Performance and training setup comparison between CLIP models and conventionally fine-tuned deep learning models.

Base Model	Model Variants	Fine-tuned Layers	Fine-tuning Methods	**AUC**	**ACC**
CLIP (400 million image-text pairs) **[69]**	ViT-B/16	Lower layers of both image and text encoders	LoRA(N = 5, r = 1)	**0.785 ± 0.039**	**0.780 ± 0.032**
	ViT-B/32		LoRA(N = 7, r = 1)	0.771 ± 0.041	0.747 ± 0.057
	ViT-B/16	Higher layers of both image and text encoders	LoRA(N = 8, r = 2)	0.777 ± 0.085	0.758 ± 0.070
	ViT-B/32		LoRA(N = 8, r = 3)	0.741 ± 0.100	0.747 ± 0.081
	ViT-B/16	All parameters	Full fine-tuning	0.768 ± 0.028	0.725 ± 0.052
	ViT-B/16	Zero shot			0.319
	ViT-B/16	Classification head	Head-only fine-tuning	0.659 ± 0.067	0.620 ± 0.110
ResNet (ImageNet-1K) **[138]**	ResNet-50			0.754 ± 0.089	0.725 ± 0.110
ViT (ImageNet-1K) **[143]**	ViT-B/16			0.707 ± 0.079	0.720 ± 0.091
	ViT-B/32			0.664 ± 0.063	0.709 ± 0.084

The ROC curve and confusion matrix of LoRA fine-tuned CLIP model ViT-B/16-Low (N = 5, r = 1) are shown in Figure A-14. Overall, the ROC curve (Figure A-14a) is close to the top-left corner, indicating the model's effectiveness in distinguishing between

classes. Figure A-14b displays the corresponding confusion matrix, showing that 97 out of 124 true responders were classified correctly, yielding a sensitivity of 0.782 (97/124). Meanwhile, among the 58 non-responders, 45 cases were successfully predicted to have no response to chemotherapy, resulting in a specificity of 0.776 (45/58).



(a) (b)

Figure A-14: ROC curve (a) and confusion matrix (b) for the CLIP model ViT-B/16-Low, fine-tuned using LoRA with $N = 5, r = 1$.

To investigate the impact of the number of Transformer blocks *N* involved in LoRA on the performance of CLIP models, Figure A-15 displays the AUC and ACC values of ViT-B/16-Low and ViT-B/16-High with *N* varying from 1 to 12. There is a general trend indicating that the performance of CLIP models improves when more blocks are fine-tuned, regardless of whether they are the lower or higher level blocks. However, it does not imply that the best performing models have to incorporate all blocks in LoRA, as the models tuned with 5-10 blocks typically achieve the highest performance. Another notable observation when *N* is within 5-10 range is that the fine-tuned ViT-B/16-Low CLIP models with a low rank value *r* of 1 dominate in terms of AUC performance, indicating a superior

classification capability. Conversely, when conducting fine-tuning by involving more later Transformer blocks closer to the output, a higher rank value *r* of 2 or 3 tends to yield better performance.

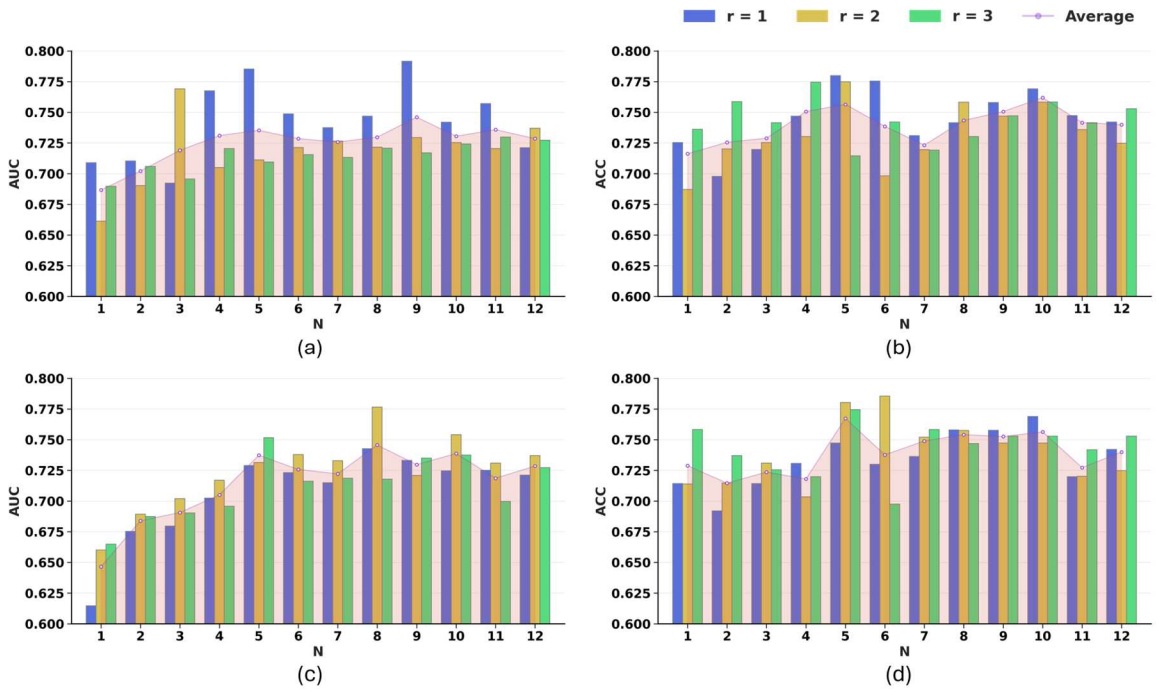
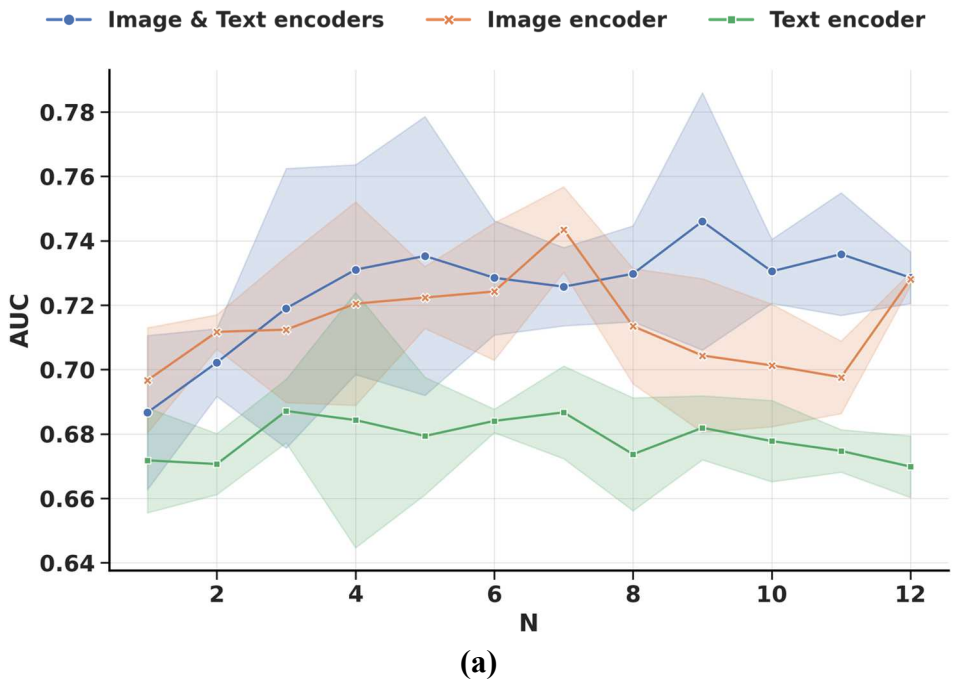


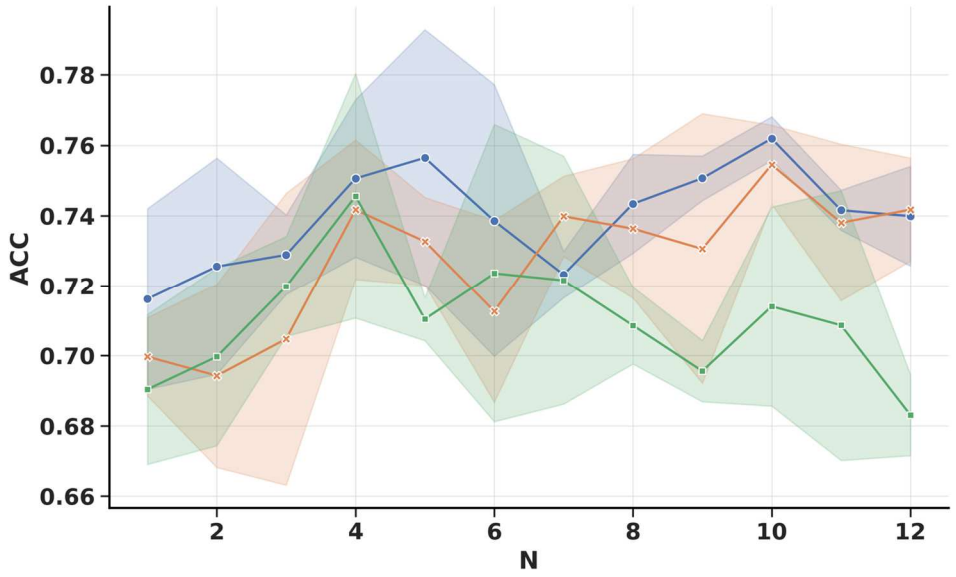
Figure A-15: Performance comparison of CLIP models: (a,b) ViT-B/16-Low and (c,d) ViT-B/16-High, which are fine-tuned by LoRA on various numbers of Transformer blocks N with different rank value r.

Figure A-16 displays the AUC and ACC values of the CLIP model ViT-B/16-Low, when LoRA is applied to the image encoder, the text encoder, or both encoders. From the AUC curves shown in Figure A-16a, the CLIP model with only the image encoder fine-tuned has a performance more closely aligned with the model fine-tuned on both encoders, compared to the fine-tuning only involving the text encoder. This implies that the image encoder has a more significant impact on CLIP models' overall discriminative capabilities. However, the interaction between the image and the text encoders appears to be more

complex, as indicated by the entangled curves shown in Figure A-16b. This complexity is particularly evident when the number of fine-tuned blocks *N* falls within the range of 5 to 10.



(a)



(b)

Figure A-16: Performance comparison of CLIP model ViT-B/16-Low, when fine-tuning by LoRA on the image encoder, the text encoder, or both encoders.

A.2.4 Discussion

In this study, we preliminarily investigate the feasibility of using CLIP, a highly generalized foundation model pretrained on a vast dataset, to predict ovarian cancer chemotherapy response after the PDS treatment using CT images. By adopting LoRA, which is an efficient method for large model fine-tuning, we compared its performance against the conventional transfer learning method focusing on training the classification head alone. The results suggest that the LoRA method can assist the CLIP model in learning the domain-specific knowledge of chemotherapy response effectively, enabling predictions with remarkable performance. This study offers a novel pathway to predict chemotherapy responses by employing one of the most current deep learning algorithms, CLIP, as opposed to the prevailing radiomics and conventional deep learning methods. In addition to the primary findings discussed above, there are several noticeable patterns that need further discussion, as outlined below.

First, in Table A-3 we compared various models and training strategies for classifying chemotherapy response. The ViT-B/16 models consistently outperformed the ViT-B/32 models in distinguishing responders from non-responders. This difference in performance remains regardless of whether CLIP models or conventional ViT models pretrained on ImageNet-1K were utilized, and whether LoRA fine-tuning or classification head-only fine-tuning was adopted. The smaller patch size in ViT-B/16 allows it to extract finer spatial features and perform more intricate attention operations, meanwhile, our CT images contain many subtle texture patterns that are likely out of the scope of ViT-B/32 models. This is probably the reason why ViT-B/16 outperforms ViT-B/32 in our experiments.

Second, another finding regarding the CLIP model fine-tuning is that LoRA was demonstrated to be notably more effective than the head-only fine-tuning approach which only achieved an AUC of 0.659. This performance is even lower than that of the head-only fine-tuned conventional ViT models, which indicates that the multimodal feature integration capability is of crucial importance to the CLIP model. Relying on a classification head is far inadequate for enabling the CLIP model to adapt to a new domain. Therefore, exploring PEFT approaches [147] that are specifically designed for CLIP fine-tuning is advisable, particularly when limited data is available.

Third, fine-tuning low level Transformer blocks of CLIP models with LoRA appears to yield better outcomes compared to fine-tuning the high level blocks, as indicated in Table A-3. Additionally, the LoRA requires a lower rank r in low level blocks fine-tuning. It is well-known that within a deep neural network, the low-level layers are responsible for learning basic features, such as edges, and textures, while the high-level layers try to extract more abstract features by combining those provided by low-level layers to make decisions [151]. This suggests that when fine-tuning CLIP models with size-limited and task-specific datasets, low level Transformer blocks are more efficient at adapting to new domains with minimal adjustments. In contrast, deep level blocks are more complex and accordingly require more resources for fine-tuning [152]. Therefore, although LoRA can efficiently adapt the CLIP model to new domains, fine-tuning all Transformer blocks in CLIP models with restricted data may not yield optimal results.

Last, in most cases, fine-tuning the image encoder alone yields higher performance than fine-tuning the text encoder alone, as evidenced in Figure A-16. This can be attributed to the fact that in our study, the image encoder can access more diverse image data, while

the text encoder is constrained to a single pair of inputs with minimal variation during the embedding alignment. Thus, the characteristics of this task emphasize the advantages of the image encoder. However, fine-tuning the text encoder alongside the image encoder can lead to higher performance by providing more flexibility in aligning image and text embeddings.

Though this study offers valuable insights, there are several limitations that should be resolved in future study. First, in current study, only a single image of the largest tumor was utilized for each case, while one case could have multiple tumors, each depicted by several CT slides. Incorporating this additional discriminative information into CLIP model fine-tuning is likely to enhance its capabilities in chemotherapy response predictions [153-155]. Second, we utilized the segmented tumors as the input for the CLIP model, without exploring the impact of using unsegmented tumors as the input, although it is beyond the scope of this study. Third, there are many other well developed PEFT methods, such as Adapters [156], and BitFit [157]. Comparing those techniques and even employing different methods for image and text encoders, with consideration of their distinct input types, may further improve the performance and efficiency. Last but not least, the lack of a second dataset to test our model prevents a more comprehensive performance evaluation. In spite of these limitations, this study serves as an initial performance assessment of CLIP models fine-tuned with LoRA, providing useful perspectives on the application of currently fast-evolving foundation models in predicting chemotherapy responses.

References

1. Greb, C., *Infinity Optical Systems.* Optik & Photonik, 2016. **11**(1): p. 34-37.
2. Park, J., et al., *Review of bio-optical imaging systems with a high space-bandwidth product.* Advanced Photonics, 2021. **3**(4): p. 044001.
3. Zheng, G., R. Horstmeyer, and C. Yang, *Wide-field, high-resolution Fourier ptychographic microscopy.* Nature Photonics, 2013. **7**(9): p. 739-745.
4. Greenbaum, A., et al., *Wide-field computational color imaging using pixel super-resolved on-chip microscopy.* Optics Express, 2013. **21**(10): p. 12469-12483.
5. Fienup, J.R., *Phase retrieval algorithms: a comparison.* Applied Optics, 1982. **21**(15): p. 2758-2769.
6. Ou, X., G. Zheng, and C. Yang, *Embedded pupil function recovery for Fourier ptychographic microscopy.* Optics Express, 2014. **22**(5): p. 4960-4972.
7. Tian, L., et al., *Multiplexed coded illumination for Fourier Ptychography with an LED array microscope.* Biomedical Optics Express, 2014. **5**(7): p. 2376-2389.
8. Bradie, B., *A Friendly Introduction to Numerical Analysis*. 2006, New Jersey: Pearson Education.
9. Brouw, W.N. *Aperture Synthesis*. in *Image Processing Techniques in Astronomy*. 1975. Dordrecht: Springer Netherlands.
10. Alexandrov, S.A., et al., *Synthetic Aperture Fourier Holographic Optical Microscopy.* Physical Review Letters, 2006. **97**(16): p. 168102.
11. Zhang, K., et al., *Using Fourier ptychography microscopy to achieve high-resolution chromosome imaging: an initial evaluation.* Journal of Biomedical Optics, 2022. **27**(1): p. 016504.

12. Zhang, K., et al., *Improving the resolution of chromosome imaging by high numerical aperture Fourier ptychography microscopy*. SPIE BiOS. Vol. 11961. 2022: SPIE.

13. Zhang, K., et al., *Using symmetric illumination and color camera to achieve high throughput Fourier ptychographic microscopy.* Journal of Biophotonics, 2023. **16**(5): p. e202200303.

14. Zhang, K., et al., *Evaluating the depth of field for Fourier ptychography microscopy*. SPIE BiOS. Vol. 12380. 2023: SPIE.

15. Siegel, R.L., et al., *Cancer Statistics, 2021.* CA: A Cancer Journal for Clinicians, 2021. **71**(1): p. 7-33.

16. Piper, J., *Automated Cytogenetics in the Study of Mutagenesis and Cancer*, in *Advances in Mutagenesis Research*, G. Obe, Editor. 1990, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 127-153.

17. Wang, X., et al., *Development and evaluation of automated systems for detection and classification of banded chromosomes: current status and future perspectives.* Journal of Physics D: Applied Physics, 2005. **38**(15): p. 2536.

18. Abid, F. and L. Hamami, *A survey of neural network based automated systems for human chromosome classification.* Artificial Intelligence Review, 2018. **49**(1): p. 41-56.

19. Qin, Y., et al., *Varifocal-Net: A Chromosome Classification Approach Using Deep Convolutional Networks.* Ieee Transactions on Medical Imaging, 2019. **38**(11): p. 2569-2581.

20. Goodman, J.W., *Introduction to Fourier Optics*. 2005, Austin, Texas: Roberts and Company.

21. Lohmann, A.W., et al., *Space–bandwidth product of optical signals and systems.* Journal of the Optical Society of America A, 1996. **13**(3): p. 470-473.

22. Fan, J., et al., *Video-rate imaging of biological dynamics at centimetre scale and micrometre resolution.* Nature Photonics, 2019. **13**(11): p. 809-816.

23. Zheng, G., et al., *Concept, implementations and applications of Fourier ptychography.* Nature Reviews Physics, 2021. **3**(3): p. 207-223.

24. Bass, M., *Handbook of Optics: Volume I-Geometrical and Physical Optics, Polarized Light, Components and Instruments*. 2010: McGraw-Hill Education.

25. Qiu, Y., et al., *Impact of the optical depth of field on cytogenetic image quality.* J Biomed Opt, 2012. **17**(9): p. 96017-1.

26. Smith, W.J., *Modern Optical Engineering: The Design of Optical Systems*. 4th ed. 2008, New York: McGraw-Hill Education.

27. Yunis, J.J. and O. Sanchez, *G-banding and chromosome structure.* Chromosoma, 1973. **44**(1): p. 15-23.

28. Wang, X., et al., *Development and assessment of an integrated computer-aided detection scheme for digital microscopic images of metaphase chromosomes.* Journal of Electronic Imaging, 2008. **17**(4): p. 043008.

29. Cole, R.W., T. Jinadasa, and C.M. Brown, *Measuring and interpreting point spread functions to determine confocal microscope resolution and ensure quality control.* Nature Protocols, 2011. **6**(12): p. 1929-1941.

30. Wicker, K. and R. Heintzmann, *Resolving a misconception about structured illumination.* Nature Photonics, 2014. **8**(5): p. 342-344.

31. Ou, X., et al., *High numerical aperture Fourier ptychography: principle, implementation and characterization.* Optics Express, 2015. **23**(3): p. 3472-3491.

32. Pan, A., C. Zuo, and B. Yao, *High-resolution and large field-of-view Fourier ptychographic microscopy and its applications in biomedicine.* Reports on Progress in Physics, 2020. **83**(9): p. 096101.

33. Ou, X., et al., *Quantitative phase imaging via Fourier ptychographic microscopy.* Optics Letters, 2013. **38**(22): p. 4845-4848.

34. Sun, J., et al., *Resolution-enhanced Fourier ptychographic microscopy based on high-numerical-aperture illuminations.* Scientific Reports, 2017. **7**(1): p. 1187.

35. Zhang, Y., et al., *Self-learning based Fourier ptychographic microscopy.* Optics Express, 2015. **23**(14): p. 18471-18486.

36. Guo, K., et al., *Optimization of sampling pattern and the design of Fourier ptychographic illuminator.* Optics Express, 2015. **23**(5): p. 6171-6180.

37. You, Z., et al., *Fourier ptychographic microscopy using wavelength multiplexing.* Journal of Biomedical Optics, 2017. **22**(6): p. 1-10.

38. Dong, S., et al., *Spectral multiplexing and coherent-state decomposition in Fourier ptychographic imaging.* Biomedical Optics Express, 2014. **5**(6): p. 1757-1767.

39. Sun, J., et al., *Single-shot quantitative phase microscopy based on color-multiplexed Fourier ptychography.* Optics Letters, 2018. **43**(14): p. 3365-3368.

40. Zhang, M., D. Yang, and Y. Liang, *Color Fourier ptychographic microscopy based on symmetrical illumination and wavelength multiplexing.* Journal of Optics, 2020. **22**(6): p. 065604.

41. Zhou, A., et al., *Analysis of Fourier ptychographic microscopy with half of the captured images.* Journal of Optics, 2018. **20**(9): p. 095701.

42. Dong, S., et al., *Sparsely sampled Fourier ptychography.* Optics Express, 2014. **22**(5): p. 5455-5464.

43. Bian, Z., S. Dong, and G. Zheng, *Adaptive system correction for robust Fourier ptychographic imaging.* Optics Express, 2013. **21**(26): p. 32400-32410.

44. Lukac, R., *Single-Sensor Imaging: Methods and Applications for Digital Cameras*. 2018: CRC Press.

45. Zhang, M., et al., *Symmetrical illumination based extending depth of field in Fourier ptychographic microscopy.* Optics Express, 2019. **27**(3): p. 3583-3597.

46. Jiang, S., et al., *Spatial- and Fourier-domain ptychography for high-throughput bio-imaging.* Nature Protocols, 2023. **18**(7): p. 2051-2083.

47. Breslauer, D.N., et al., *Mobile Phone Based Clinical Microscopy for Global Health Applications.* PLOS ONE, 2009. **4**(7): p. e6320.

48. Chung, J., et al., *Counting White Blood Cells from a Blood Smear Using Fourier Ptychographic Microscopy.* PLOS ONE, 2015. **10**(7): p. e0133489.

49. Switz, N.A., M.V. D'Ambrosio, and D.A. Fletcher, *Low-Cost Mobile Phone Microscopy with a Reversed Mobile Phone Camera Lens.* PLOS ONE, 2014. **9**(5): p. e95330.

50. Lee, K.C., et al., *A Smartphone-Based Fourier Ptychographic Microscope Using the Display Screen for Illumination.* ACS Photonics, 2021. **8**(5): p. 1307-1315.

51. Aidukas, T., et al., *Low-cost, sub-micron resolution, wide-field computational microscopy using opensource hardware.* Scientific Reports, 2019. **9**(1): p. 7457.

52. Oppenheim, A.V., A.S. Willsky, and S. Hamid, *Signals and systems, Processing series*. 1997, Prentice Hall Upper Saddle River.

53. Fischer, R., B. Tadic-Galeb, and P. Yoder, *Optical system design*. 2008: McGraw-Hill Education.

54. Zhang, Y. and H. Gross, *Systematic design of microscope objectives. Part I: System review and analysis.* Advanced Optical Technologies, 2019. **8**(5): p. 313-347.

55. Gilley, P., et al., *Evaluating the imaging quality of multiplexed illumination Fourier ptychography microscopy*. SPIE BiOS. Vol. 12843. 2024: SPIE.

56. Lin, B., et al., *Efficient multiplexed illumination and imaging approach for Fourier ptychographic microscopy.* Journal of the Optical Society of America A, 2022. **39**(5): p. 883-896.

57. Liang, M., et al., *All-in-focus fine needle aspiration biopsy imaging based on Fourier ptychographic microscopy.* Journal of Pathology Informatics, 2022. **13**: p. 100119.

58. Zhang, K., et al., *Using symmetric illumination and color camera to achieve high throughput Fourier ptychographic microscopy.* Journal of Biophotonics. **n/a**(n/a): p. e202200303.

59. Reginald, H.H.H.a.B.C., *The frequency response of a defocused optical system.* Proc. R. Soc. Lond. A, 1955. **231**(1184): p. 91-103.

60. Fu, B., et al., *A blind medical image denoising method with noise generation network.* Journal of X-Ray Science and Technology, 2022. **30**: p. 531-547.

61. Cui, X., et al., *Artifact-Assisted multi-level and multi-scale feature fusion attention network for low-dose CT denoising.* Journal of X-Ray Science and Technology, 2022. **30**: p. 875-889.

62. Cao, R., C. Shen, and C. Yang, *High-resolution, large field-of-view label-free imaging via aberration-corrected, closed-form complex field reconstruction.* Nature Communications, 2024. **15**(1): p. 4713.

63. Tian, L., et al., *Computational illumination for high-speed in vitro Fourier ptychographic microscopy.* Optica, 2015. **2**(10): p. 904-911.

64. Nguyen, T., et al., *Deep learning approach for Fourier ptychography microscopy.* Optics Express, 2018. **26**(20): p. 26470-26484.

65. Jiang, S., et al., *Solving Fourier ptychographic imaging problems via neural network modeling and TensorFlow.* Biomedical Optics Express, 2018. **9**(7): p. 3306-3319.

66. Sun, L., et al., *Lens Factory: Automatic Lens Generation Using Off-the-shelf Components.* ArXiv, 2015. **abs/1506.08956**.

67. König, G., et al., *Automated lens design for optical systems consisting of stock lenses.* Optics Express, 2021. **29**(24): p. 39027-39041.

68. Awad, M. and R. Khanna, *Support Vector Machines for Classification*, in *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*, M. Awad and R. Khanna, Editors. 2015, Apress: Berkeley, CA. p. 39-66.

69. Radford, A., et al. *Learning Transferable Visual Models From Natural Language Supervision*. in *International Conference on Machine Learning*. 2021.

70. Hu, E.J., et al., *LoRA: Low-Rank Adaptation of Large Language Models.* CoRR, 2021. **abs/2106.09685**.

71. Siegel, R.L., A.N. Giaquinto, and A. Jemal, *Cancer statistics, 2024.* CA: A Cancer Journal for Clinicians, 2024. **74**(1): p. 12-49.

72. Siegel, R.L., et al., *Cancer statistics, 2023.* CA: A Cancer Journal for Clinicians, 2023. **73**(1): p. 17-48.

73. Matulonis, U.A., et al., *Ovarian cancer.* Nature Reviews Disease Primers, 2016. **2**(1): p. 16061.

74. Slatnik, C.L.P. and E. Duff, *Ovarian cancer: Ensuring early diagnosis.* The Nurse Practitioner, 2015. **40**(9): p. 47-54.

75. Hennessy, B.T., R.L. Coleman, and M. Markman, *Ovarian cancer.* The Lancet, 2009. **374**(9698): p. 1371-1382.

76. Sato, S. and H. Itamochi, *Neoadjuvant chemotherapy in advanced ovarian cancer: latest results and place in therapy.* Therapeutic advances in medical oncology, 2014. **6**(6): p. 293-304.

77. Mahner, S., et al., *Neoadjuvant chemotherapy in ovarian cancer revisited.* Annals of Oncology, 2016. **27**: p. i30-i32.

78. Vergote, I., et al., *Neoadjuvant Chemotherapy or Primary Surgery in Stage IIIC or IV Ovarian Cancer.* New England Journal of Medicine, 2010. **363**(10): p. 943-953.

79. Cowan, R., et al., *Primary Surgery or Neoadjuvant Chemotherapy in Advanced Ovarian Cancer: The Debate Continues….* American Society of Clinical Oncology Educational Book, 2016(36): p. 153-162.

80. Rauh-Hain, J.A., et al., *Primary Debulking Surgery Versus Neoadjuvant Chemotherapy in Stage IV Ovarian Cancer.* Annals of Surgical Oncology, 2012. **19**(3): p. 959-965.

81. Vergote, I., et al., *Neoadjuvant Chemotherapy or Primary Surgery in Stage IIIC or IV Ovarian Cancer.* New England Journal of Medicine, 2010. **363**(10): p. 943-953.

82. Chi, D.S., et al., *Improved optimal cytoreduction rates for stages IIIC and IV epithelial ovarian, fallopian tube, and primary peritoneal cancer: a change in surgical approach.* Gynecologic Oncology, 2004. **94**(3): p. 650-654.

83. Tajik, P., et al., *The FIGO Stage IVA Versus IVB of Ovarian Cancer: Prognostic Value and Predictive Value for Neoadjuvant Chemotherapy.* International Journal of Gynecologic Cancer, 2018. **28**(3): p. 453-458.

84. Cioffi, R., et al., *Neoadjuvant chemotherapy in high-risk ovarian cancer patients: Role of age.* Tumori Journal, 2019. **105**(2): p. 168-173.

85. Patel, A., et al., *Emerging Trends in Neoadjuvant Chemotherapy for Ovarian Cancer.* Cancers, 2021. **13**(4): p. 626.

86. Suidan, R.S., et al., *A multicenter prospective trial evaluating the ability of preoperative computed tomography scan and serum CA-125 to predict suboptimal cytoreduction at primary debulking surgery for advanced ovarian, fallopian tube, and peritoneal cancer.* Gynecologic Oncology, 2014. **134**(3): p. 455-461.

87. Zhang, D., et al., *Serum CA125 levels predict outcome of interval debulking surgery after neoadjuvant chemotherapy in patients with advanced ovarian cancer.* Clinica Chimica Acta, 2018. **484**: p. 32-35.

88. Oikonomopoulou, K., et al., *Prediction of ovarian cancer prognosis and response to chemotherapy by a serum-based multiparametric biomarker panel.* British Journal of Cancer, 2008. **99**(7): p. 1103-1113.

89. Shen, Y. and L. Li, *Serum HE4 superior to CA125 in predicting poorer surgical outcome of epithelial ovarian cancer.* Tumor Biology, 2016. **37**(11): p. 14765-14772.

90. Huang, Y.-W., *Association of BRCA1/2 mutations with ovarian cancer prognosis: An updated meta-analysis.* Medicine, 2018. **97**(2): p. e9380.

91. Liu, Y.L., et al., *Characteristics and survival of ovarian cancer patients treated with neoadjuvant chemotherapy but not undergoing interval debulking surgery.* jgo, 2019. **31**(1): p. 0-0.

92. Fass, L., *Imaging and cancer: A review.* Molecular Oncology, 2008. **2**(2): p. 115-152.

93. Kyriazi, S., S.B. Kaye, and N.M. deSouza, *Imaging ovarian cancer and peritoneal metastases—current and emerging techniques.* Nature Reviews Clinical Oncology, 2010. **7**(7): p. 381-393.

94. Qiu, Y., et al., *Early prediction of clinical benefit of treating ovarian cancer using quantitative CT image feature analysis.* Acta Radiologica, 2016. **57**(9): p. 1149-1155.

95. Mayerhoefer, M.E., et al., *Introduction to Radiomics.* Journal of Nuclear Medicine, 2020. **61**(4): p. 488-495.

96. Lambin, P., et al., *Radiomics: the bridge between medical imaging and personalized medicine.* Nature Reviews Clinical Oncology, 2017. **14**(12): p. 749-762.

97. van Timmeren, J.E., et al., *Radiomics in medical imaging—"how-to" guide and critical reflection.* Insights into Imaging, 2020. **11**(1): p. 91.

98. Fernandez, A., et al., *SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary.* Journal of Artificial Intelligence Research, 2018. **61**: p. 863-905.

99. Lemaître, G., F. Nogueira, and C.K. Aridas, *Imbalanced-learn: a python toolbox to tackle the curse of imbalanced datasets in machine learning.* J. Mach. Learn. Res., 2017. **18**(1): p. 559–563.

100. Abdi, H. and L.J. Williams, *Principal component analysis.* WIREs Computational Statistics, 2010. **2**(4): p. 433-459.

101. Salih Hasan, B.M. and A.M. Abdulazeez, *A Review of Principal Component Analysis Algorithm for Dimensionality Reduction.* Journal of Soft Computing and Data Mining, 2021. **2**(1): p. 20-30.

102. Van Der Maaten, L., E.O. Postma, and H.J. van den Herik, *Dimensionality reduction: A comparative review.* Journal of Machine Learning Research, 2009. **10**(66-71): p. 13.

103. Olivier, C., S. Bernhard, and Z. Alexander, *Spectral Methods for Dimensionality Reduction*, in *Semi-Supervised Learning*. 2006, MIT Press. p. 293-308.

104. Pedregosa, F., et al., *Scikit-learn: Machine Learning in Python.* J. Mach. Learn. Res., 2011. **12**(null): p. 2825–2830.

105. Roy, S., et al., *Co-clinical FDG-PET radiomic signature in predicting response to neoadjuvant chemotherapy in triple-negative breast cancer.* European Journal of Nuclear Medicine and Molecular Imaging, 2022. **49**(2): p. 550-562.

106. Jayaprakasam, V.S., et al., *MRI radiomics features of mesorectal fat can predict response to neoadjuvant chemoradiation therapy and tumor recurrence in patients with locally advanced rectal cancer.* European Radiology, 2022. **32**(2): p. 971-980.

107. Hu, Y., et al., *Computed tomography-based deep-learning prediction of neoadjuvant chemoradiotherapy treatment response in esophageal squamous cell carcinoma.* Radiotherapy and Oncology, 2021. **154**: p. 6-13.

108. Cervantes, J., et al., *A comprehensive survey on support vector machine classification: Applications, challenges and trends.* Neurocomputing, 2020. **408**: p. 189-215.

109. Varma, S. and R. Simon, *Bias in error estimation when using cross-validation for model selection.* BMC Bioinformatics, 2006. **7**(1): p. 91.

110. Rao, R.B., G. Fung, and R. Rosales, *On the Dangers of Cross-Validation. An Experimental Evaluation*, in *Proceedings of the 2008 SIAM International Conference on Data Mining (SDM)*. p. 588-596.

111. Molinaro, A.M., R. Simon, and R.M. Pfeiffer, *Prediction error estimation: a comparison of resampling methods.* Bioinformatics, 2005. **21**(15): p. 3301-3307.

112. Park, S.H., J.M. Goo, and C.-H. Jo, *Receiver Operating Characteristic (ROC) Curve: Practical Review for Radiologists.* Korean J Radiol, 2004. **5**(1): p. 11-18.

113. Bradley, A.P., *The use of the area under the ROC curve in the evaluation of machine learning algorithms.* Pattern Recognition, 1997. **30**(7): p. 1145-1159.

114. Danala, G., et al., *Applying Quantitative Radiographic Image Markers to Predict Clinical Complications After Aneurysmal Subarachnoid Hemorrhage: A Pilot Study.* Annals of Biomedical Engineering, 2022. **50**(4): p. 413-425.

115. Peng, A., et al., *A machine learning model to precisely immunohistochemically classify pituitary adenoma subtypes with radiomics based on preoperative magnetic resonance imaging.* European Journal of Radiology, 2020. **125**: p. 108892.

116. Wilkinson, L. and M. Friendly, *The History of the Cluster Heat Map.* The American Statistician, 2009. **63**(2): p. 179-184.

117. Vallius, T., et al., *18F-FDG-PET/CT based total metabolic tumor volume change during neoadjuvant chemotherapy predicts outcome in advanced epithelial ovarian cancer.* European Journal of Nuclear Medicine and Molecular Imaging, 2018. **45**(7): p. 1224-1232.

118. Eisenhauer, E.A., et al., *New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1).* European Journal of Cancer, 2009. **45**(2): p. 228-247.

119. Krasovitsky, M., et al., *Interobserver and intraobserver variability of RECIST assessment in ovarian cancer.* International Journal of Gynecologic Cancer, 2022. **32**(5): p. 656-661.

120. Chen, X., et al., *Recent advances and clinical applications of deep learning in medical image analysis.* Medical Image Analysis, 2022. **79**: p. 102444.

121. Li, J., et al., *Transforming medical imaging with Transformers? A comparative review of key properties, current progresses, and future perspectives.* Medical Image Analysis, 2023. **85**: p. 102762.

122. Altan, A. and S. Karasu, *Recognition of COVID-19 disease from X-ray images by hybrid model consisting of 2D curvelet transform, chaotic salp swarm algorithm and deep learning technique.* Chaos, Solitons & Fractals, 2020. **140**: p. 110071.

123. Özçelik, Y.B. and A. Altan, *Overcoming Nonlinear Dynamics in Diabetic Retinopathy Classification: A Robust AI-Based Model with Chaotic Swarm Intelligence Optimization and Recurrent Long Short-Term Memory.* Fractal and Fractional, 2023. **7**(8): p. 598.

124. Traverso, A., et al., *Repeatability and Reproducibility of Radiomic Features: A Systematic Review.* International Journal of Radiation Oncology*Biology*Physics, 2018. **102**(4): p. 1143-1158.

125. Veneziani, A.C., et al., *Heterogeneity and treatment landscape of ovarian carcinoma.* Nature Reviews Clinical Oncology, 2023. **20**(12): p. 820-842.

126. Zargari, A., et al., *Prediction of chemotherapy response in ovarian cancer patients using a new clustered quantitative image marker.* Phys Med Biol, 2018. **63**(15): p. 1361-6560.

127. Qiu, Y.C., et al., *Early prediction of clinical benefit of treating ovarian cancer using quantitative CT image feature analysis.* Acta Radiologica, 2016. **57**(9): p. 1149-1155.

128. Danala, G., et al., *Applying Quantitative CT Image Feature Analysis to Predict Response of Ovarian Cancer Patients to Chemotherapy.* Acad Radiol, 2017. **24**(10): p. 1233-1239.

129. Lu, H., et al., *A mathematical-descriptor of tumor-mesoscopic-structure from computed-tomography images annotates prognostic- and molecular-phenotypes of epithelial ovarian cancer.* Nature Communications, 2019. **10**(1): p. 764.

130. Rizzo, S., et al., *Radiomics of high-grade serous ovarian cancer: association between quantitative CT features, residual tumour and disease progression within 12 months.* European Radiology, 2018. **28**(11): p. 4849-4859.

131. Wei, W., et al., *A Computed Tomography-Based Radiomic Prognostic Marker of Advanced High-Grade Serous Ovarian Cancer Recurrence: A Multicenter Study.* Frontiers in Oncology, 2019. **9**.

132. Abdoli, N., et al., *Evaluating the Effectiveness of 2D and 3D CT Image Features for Predicting Tumor Response to Chemotherapy.* Bioengineering, 2023. **10**(11): p. 1334.

133. Abdoli, N., et al., *Comparing the effectiveness of 2D and 3D features on predicting the response to chemotherapy for ovarian cancer patients*. SPIE BiOS. Vol. 12380. 2023: SPIE.

134. Zhang, K., et al., *Developing a novel image marker to predict the clinical outcome of neoadjuvant chemotherapy (NACT) for ovarian cancer patients.* Computers in Biology and Medicine, 2024. **172**: p. 108240.

135. Avesani, G., et al., *CT-Based Radiomics and Deep Learning for BRCA Mutation and Progression-Free Survival Prediction in Ovarian Cancer Using a Multicentric Dataset.* Cancers, 2022. **14**(11): p. 2739.

136. Huang, X., et al., *Predicting prognosis for epithelial ovarian cancer patients receiving bevacizumab treatment with CT-based deep learning.* npj Precision Oncology, 2024. **8**(1): p. 202.

137. Wang, S., et al., *Deep learning provides a new computed tomography-based prognostic biomarker for recurrence prediction in high-grade serous ovarian cancer.* Radiotherapy and Oncology, 2019. **132**: p. 171-177.

138. He, K., et al. *Deep Residual Learning for Image Recognition*. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016.

139. Jia, D., et al. *ImageNet: A large-scale hierarchical image database*. in *Computer Vision and Pattern Recognition, IEEE Conference on*. 2009.

140. Romera-Paredes, B. and P.H.S. Torr, *An embarrassingly simple approach to zero-shot learning*, in *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37*. 2015, JMLR.org: Lille, France. p. 2152–2161.

141. Fallowfield, L.J. and A. Fleissig, *The value of progression-free survival to patients with advanced-stage cancer.* Nature Reviews Clinical Oncology, 2012. **9**(1): p. 41-47.

142. Danala;, G., et al. *Improving efficacy of metastatic tumor segmentation to facilitate early prediction of ovarian cancer patients' response to chemotherapy*. in *Proc. of SPIE*. 2017.

143. Dosovitskiy, A., et al., *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale.* ArXiv, 2020. **abs/2010.11929**.

144. Vaswani, A., et al. *Attention is All you Need*. in *Neural Information Processing Systems*. 2017.

145. Devlin, J., *Bert: Pre-training of deep bidirectional transformers for language understanding.* arXiv preprint arXiv:1810.04805, 2018.

146. Brown, T.B., et al., *Language models are few-shot learners*, in *Proceedings of the 34th International Conference on Neural Information Processing Systems*. 2020, Curran Associates Inc.: Vancouver, BC, Canada. p. Article 159.

147. Lialin, V., V. Deshpande, and A. Rumshisky, *Scaling down to scale up: A guide to parameter-efficient fine-tuning.* arXiv preprint arXiv:2303.15647, 2023.

148. Kingma, D. and J. Ba, *Adam: A Method for Stochastic Optimization.* International Conference on Learning Representations, 2014.

149. Scholz, D., et al. *Imbalance-aware loss functions improve medical image classification*. in *Medical Imaging with Deep Learning*. 2024.

150. Paszke, A., et al., *PyTorch: an imperative style, high-performance deep learning library*, in *Proceedings of the 33rd International Conference on Neural Information Processing Systems*. 2019, Curran Associates Inc. p. Article 721.

151. Alzubaidi, L., et al., *Review of deep learning: concepts, CNN architectures, challenges, applications, future directions.* Journal of Big Data, 2021. **8**(1): p. 53.

152. Chen, Y., A. Yuille, and Z. Zhou. *Which layer is learning faster? a systematic exploration of layer-wise convergence rate for deep neural networks*. in *The Eleventh International Conference on Learning Representations*. 2023.

153. Chang, R., et al., *Predicting chemotherapy response in non-small-cell lung cancer via computed tomography radiomic features: Peritumoral, intratumoral, or combined?* Frontiers in Oncology, 2022. **12**.

154. Chang, R., et al., *Deep radiomic model based on the sphere–shell partition for predicting treatment response to chemotherapy in lung cancer.* Translational Oncology, 2023. **35**: p. 101719.

155. Chen, X., et al., *Transformers Improve Breast Cancer Diagnosis from Unregistered Multi-View Mammograms.* Diagnostics, 2022. **12**(7): p. 1549.

156. Houlsby, N., et al. *Parameter-efficient transfer learning for NLP*. in *International conference on machine learning*. 2019. PMLR.

157. Zaken, E.B., S. Ravfogel, and Y. Goldberg, *Bitfit: Simple parameter-efficient fine-tuning for transformer-based masked language-models.* arXiv preprint arXiv:2106.10199, 2021.