UNDERSTANDING AND MANIPULATION OF OPTICAL SPECKLES

ZHAO QI

PhD

The Hong Kong Polytechnic University

2024

The Hong Kong Polytechnic University

Department of Biomedical Engineering

Understanding and manipulation of optical speckles

Zhao Qi

A thesis submitted in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

July 2024

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____(Signed)

_____ZHAO QI_____(Name of student)

ABSTRACT

Optical technologies have experienced significant advancements, leading to various scientific and technological applications. However, the endeavor to probe deeper into inhomogeneous media is hindered by the inherent challenge of optical scattering. Notably, at depths exceeding 1 mm beneath the skin, the multiple scattering phenomena transform potential imaging clarity into optical speckles. This thesis addresses the critical need for innovative methods to circumvent optical scattering and elucidate information concealed within optical speckles. It presents a comprehensive exploration of optical speckles, offering novel insights into overcoming, understanding, and utilizing optical speckles. The body of this thesis is methodically divided into five chapters, each contributing to the overarching narrative of advancing biomedical optics.

In Chapter 2, a parameter-free algorithm has been proposed for iterative wavefront shaping to overcome speckles, aiming to focus lasers through scattering media and setting the foundation for subsequent chapters. This innovative approach avoids the time-consuming and experience-dependent parameter tuning process, which is inevitable for existing iterative algorithms. Experimental validation, employing ground glass and multi-mode fibers, substantiates the algorithm’s efficacy. These results showcase its robust capability to achieve laser focusing

across various scattering environments.

In Chapter 3, the focusing capabilities of iterative wavefront shaping are extended to image retrieval from speckles. We introduce a Generative Adversarial Network (GAN) to effectively tackle the challenge of spatiotemporal decorrelation in optical speckles, enabling the retrieval of images from speckles that have decorrelated to unknown statuses, rather than neglecting intervals between acquiring the training and test datasets as most other research. The GAN framework, notable for its broad generalizability, is trained to retrieve high-fidelity human face images from decorrelated speckles. The ability has been demonstrated even under conditions where the scattering medium has significantly decorrelated, such as after the optical system has been inactive for extended periods (up to 37 hours in experiments) before being reactivated. The experiments mark a significant stride in broadening the applications of learning-based methodologies in speckle imaging.

Chapter 4 further advances the theoretical understanding of speckles and delves into the information delocalization within optical speckles, employing learning-based models and information entropy for an in-depth analysis. The study also examines the speckle sampling condition for high-fidelity information retrieval from optical speckles, an important but previously unexplored research question. Experimental observations disclose a uniform

dispersion of information among fully developed optical speckles, ensuring the integrity of information retrieval is maintained irrespective of the spatial positions of optical speckles. A theoretical framework emerges from a synthesis of physical models and empirical data, postulating that neural networks can be trained to retrieve information with high fidelity from optical speckles, provided the entropy of the speckle autocorrelation exceeds the entropy of the target autocorrelation.

Chapter 5 represents a synthesis of the practical algorithmic advancements and theoretical insights from the previous chapters. Speckle Transformer, a cutting-edge vision transformer-based model, has been designed to harness the delocalized information within optical speckles for target classification. The proposed model directly extracts speckle features for classification, surpassing traditional methods where classification follows information retrieval.

Chapter 6 signifies a departure from the primary focus on speckle imaging to embracing the natural randomness of optical speckles and exploring the utilization of optical speckles in the realm of security. An innovative speckle-based optical cryptosystem has been proposed to achieve a straightforward yet highly effective encryption mechanism. The proposed cryptosystem is distinguished by its robust security, rapid encryption, and cost-efficiency. Within this framework, a piece of ground glass serves as the physical secret key, enabling the

encryption of face images through the scattering medium of seemingly chaotic optical speckles at the speed of light. Subsequently, these images are decrypted from the optical speckles using a pre-trained neural network, ensuring that the retrieved face images retain high fidelity and are recognizable by face recognition algorithms. To the best of our knowledge, this is the first demonstration of a speckle-based optical cryptosystem for face recognition.

In summary, these chapters illustrate a comprehensive journey from overcoming the challenges posed by scattering, understanding delocalized information in speckles, to harnessing the properties of optical speckles for diverse applications. Accordingly, this thesis significantly advances our comprehension of delocalized information within optical speckles and charts a new course for speckle-related research and applications. The research and experimental outcomes not only elucidate the underlying principles but also herald the advent of a transformative paradigm in deep tissue optics, which promises to extend penetration depths and augment resolution, thereby broadening the scope and efficacy of optical applications in biomedical research.

PUBLICATIONS ARISING FROM THE THESIS

First or co-first authored papers

([†]equal contribution; [#]correspondence author)

1. Lai, P.[†,#], **Zhao, Q.**[†], Zhou, Y., Cheng, S., Woo, C. M., Li, H., Yu, Z., Huang, X., Yao, J., Pang, W., Li, H., Huang, H., Li, W., Zheng, Y., Wang, Z., Yuan, C., & Zhong, T.[#] (2024). Deep-tissue optics: Technological development and applications. *Chinese Journal of Lasers*, *51*(1), 0107003.

2. Huang, Hao.[†], **Zhao, Q**.[†], Li, H.[†], Zheng, Y., Yu, Z., Zhong, T., Cheng, S., Woo, C. M., Gao, Y., Liu, H., Zheng, Y., Tian, J.[#], & Lai P.[#] (2024). DeepSLM: Speckle-licensed modulation via deep adversarial learning for authorized optical encryption and decryption. *Advanced Intelligent Systems, 9*(1), 202400150.

3. Li, H.[†], Yu, Z.[†], **Zhao, Q**.[†], Luo, Y., Cheng, S., Zhong, T., Woo, C. M., Liu, H., Wang, L. V.[#], Zheng, Y.[#], & Lai, P.[#] (2023). Learning-based super-resolution interpolation for sub-Nyquist sampled laser speckles. *Photonics Research*, *11*(4), 631-642.

4. **Zhao, Q**.[†], Li, H.[†], Yu, Z.[†], Woo, C. M., Zhong, T., Cheng, S., Zheng, Y., Liu, H., Tian, J.[#], & Lai, P.[#] (2022). Speckle-based optical cryptosystem and its application for human face recognition via deep learning. *Advanced Science*, *9*(25), 2202407.

5. Woo, C. M.[†], **Zhao, Q.**[†], Zhong, T., Li, H., Yu, Z., & Lai, P.[#] (2022). Optimal efficiency of focusing diffused light through scattering media with iterative wavefront shaping. *APL Photonics*, *7*(4).

6. **Zhao, Q.**[†]**,** Woo, C. M.[†], Li, H., Zhong, T., Yu, Z.[#], & Lai, P.[#] (2021). Parameter-free optimization algorithm for iterative wavefront shaping. *Optics Letters*, *46*(12), 2880-2883.

7. **Zhao, Q.**[†], Li, H.[†], Zhong, T.[†], Cheng, S., Huang, H., Yao, J., Li, W., Li, H., Woo, C. M., Gong, L., Zheng, Y.[#], Yu, Z.[#], & Lai, P.[#] (2023). Extended learning generalizability for high-fidelity human face imaging from spatiotemporally decorrelated speckles. *Under review*.

8. **Zhao, Q.**[†], Li, H.[†,#], Yu, Z.[†], Li, H.[†], Cheng, S., Huang, H., Zhong, T., Woo, C. M., Wang, Z., Zheng, Y., Liu, H.[#], & Lai, P.[#] (2024). Delocalized information in optical speckles: a learning-based study. *In preparation*.

9. **Zhao, Q.**[†], Li, H.[†], Yu, Z.[†], Li, H., Cheng, S., Huang, H., Zhong, T., Woo, C. M., Wang, Z., & Lai, P.[#] (2024). Speckle transformer: classification through scattering media with limited information. *In preparation*.

Non-first authored papers

1. Yu, Z.[†], Li, H.[†], Zhao, W.[†], Huang, P., Lin, Y., Yao, J., Li, W., Zhao, Q., Wu, P., Li, B., Genevet, P.[#], Song, Q.[#] & Lai, P.[#] (2024). High-security learning-based optical encryption assisted by disordered metasurface. *Nature Communications*, *15*(1), 2607.

2. Yu, Z.[†], Li, H.[†], Zhong, T.[†], Park, J.[†], Cheng, S., Woo, C. M., Zhao, Q., Yao, J., Zhou, Y., Huang, X., Pang, W., Yoon, H., Shen, Y., Liu, H., Zheng, Y., Park, Y.[#], Wang, L. V.[#], & Lai P.[#] (2022). Wavefront shaping: A versatile tool to conquer multiple scattering in multidisciplinary fields. *The Innovation*, *3*(5), 15.

3. Li, H., Yu, Z., Zhao, Q., Zhong, T., & Lai, P.[#] (2022). Accelerating deep learning with high energy efficiency: from microchip to physical systems. *The Innovation*, *3*(4), 2.

4. Cheng, S.[†], Zhong, T.[†], Woo, C. M., Zhao, Q., Hui, H., & Lai, P.[#] (2022). Long-distance pattern projection through an unfixed multimode fiber with natural evolution strategy-based wavefront shaping. *Optics Express*, *30*(18), 32565-32576.

5. Woo, C. M.[†], Li, H.[†], Zhao, Q., & Lai, P.[#] (2021). Dynamic mutation enhanced particle swarm optimization for optical wavefront. *Optics Express*, *29*(12), 18420-18426.

6. Wang, Z.[†], Li, H.[†], Zhao, Q., RV, V., Yuan, X., Chen, Z., Lai, P.[#], & Pu, J.[#] (2023). Single-shot information retrieval from speckles with spatially multiplexed points detection

(SMPD). *Under review*.

Patents

1. Lai, P., Zhao, Q., Li, H., & Yu, Z. (2021). Speckle-based optical cryptosystem for encrypted face recognition (*Chinese Patent* No. ZL202111273688.7).

Conference Presentations

1. Zhao, Q., Li, H., Yu, Z., Li, H., & Lai, P.[#] (2024, August 4 - 7). *Learning-based study on spatially dispersed information in optical speckles*. The 11th WACBE World Congress on Bioengineering, Hong Kong, China (Poster).
2. Zhao, Q., Li, H., Yu, Z., Li, H., & Lai, P.[#] (2024, June 1). *Delocalized information in optical speckles and its learning-based demonstration*. The 5th Advanced Photonics Forum, Shen Zhen, Guangdong, China (Poster).
3. Zhao, Q., Li, H., Yu, Z., & Lai, P.[#] (2023, August 19 - 20). *Learning-based study on spatially dispersed imaging from optical speckles*. Visible Light Communication and Optical Computing 2023, Shenzhen, Guangdong, China (Poster).

4. Zhao, Q., Li, H., Yu, Z., & Lai, P.[#] (2023, May 26 - 28). *Spatially dispersed information retrieved from optical speckles: a learning-based study*. PhotoniX Forum 2023, Hangzhou, Zhejiang, China (Poster).

5. Zhao, Q., Li, H., Yu, Z., & Lai, P.[#] (2023, January 28 - February 03). *Speckle-based optical cryptosystem for face recognition*. SPIE Photonics West 2023, San Francisco, California, United States (Oral).

6. Zhao, Q., Li, H., Yu, Z., & Lai, P.[#] (2022, November 1 - 4). *Encrypting face recognition using optical speckles*. IEEE TENCON 2022, Wan Chai, Hong Kong, China (Oral).

7. Zhao, Q., Woo, C. M., Li, H., Zhong, T., Yu, Z., & Lai, P.[#] (2021, October 10 - 12). *Integration of genetic and bat algorithms towards a parameter-free optimization scheme in iterative wavefront shaping*. SPIE Photonics Asia 2021, Online, China (Oral).

8. Zhao, Q., Woo, C. M., Li, H., Zhong, T., Yu, Z., & Lai, P.[#] (2021, July 27- 30). *Innovative heuristic algorithm in wavefront shaping - parameter-free algorithm*. International Conference on Biomedical Health Informatics 2021, Online, Korea (Oral).

ACKNOWLEDGEMENT

As I stand on the precipice of completing my doctoral journey, I am filled with a profound sense of gratitude. The year 2020 holds special significance in my heart, marking a pivotal moment in my academic path. Despite the challenges that halted my master's research, I was fortunate to receive an extraordinary opportunity from Prof. Puxiang Lai to pursue my studies at The Hong Kong Polytechnic University. This marked the beginning of a transformative four-year odyssey at PolyU, indelibly shaped by the global pandemic, COVID-19.

The pandemic presented formidable challenges, significantly altering lives and academic endeavors alike. It disrupted student mobility, research collaborations, and academic exchanges. Yet, it also provided a unique opportunity for deeper reflection, exploration, and scholarly engagement. Amidst these trying times, I delved into the realms of biomedical photonics, computational optics, wavefront shaping, and deep learning—knowledge that I am confident will propel my future career in biomedical imaging.

This period underscored the critical importance of resilience and adaptability. I am immensely thankful for the support of the Biophotonics Lab members, who were instrumental in both my research and personal growth. My deepest appreciation goes to my mentor, Prof. Puxiang Lai, whose steadfast guidance, unwavering support, and infinite patience have been the bedrock of

my academic journey. His profound knowledge, rich experience, and passion for research have not only augmented my scholarly pursuits but also served as a beacon of inspiration in my life. I also extend special thanks to our collaborators, including Prof. Jie Tian, Prof. Yuanjin Zheng, and Prof. Honglin Liu, whose expertise significantly enhanced my studies and publications.

Additionally, the contributions of my lab mates, including Dr. Huanhao Li, Dr. Zhipeng Yu, Dr. Yingying Zhou, Dr. Xiazi Huang, Dr. Tianting Zhong, Dr. Fei Cao, Ms. Chi Man Woo, Mr. Shengfu Cheng, Mr. Jing Yao, Ms. Weiran Pang, Mr. Haoran Li, Mr. Haofan Huang, Mr. Wenzhao Li, Mr. Yuandong Zheng, Mr. Zhiyuan Wang, Ms. Chuqi Yuan, Ms. Siyang Zheng, Mr. Xian Hu, Ms. Xiaozhou Xiao, Mr. Kan Chen, Ms. Yuzhen Li, and Mr. Weibing Cai have been invaluable to my academic and personal lives, offering generous assistance and support.

Last but not least, my heartfelt gratitude also extends to my mother, Ms. Qiong Shi, my father, Mr. Tao Zhao, and my girlfriend, Ms. Yilin Yang, whose understanding and encouragement have always been the pillars of my existence. Their unwavering love and support have been the cornerstone of my achievements.

In closing, I am confident that the experiences and knowledge gained at The Hong Kong Polytechnic University will be instrumental in my future pursuits. I look forward to the opportunities that lie ahead and am excited to embark on the next chapter of my journey.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Abbreviations

Term	Full-term
ACO	Ant colony optimization
AES	Advanced encryption standards
AO	Adaptive optics
BA	Bat algorithm
BCE	Binary cross entropy
CelebA	Large-scale CelebFaces attributes dataset
CFES	Compression friendly encryption scheme
CUDA	Compute unified device architecture
DL	Deep learning
DF	Decay factor of the mutation rate in genetic algorithm
DM	Disordered metasurface
DMA	Dynamic mutation algorithm
DMD	Digital micromirror device
DMR	Dynamic mutation rate

DNN	Deep neural networks
DOPC	Digital optical phase conjugation
DRPE	Double random phase encryption
EBE	Electron beam evaporation
EBL	Electron beam lithography
FFHQ	Flickr faces high quality dataset
FOV	Field of view
GA	Genetic algorithm
GAN	Generative adversarial network
GD	Ground glass
imPCC	Retrieved image Pearson correlation coefficient
MD5	Message digest algorithm
MHA	Multi-head attention
MLP	Multi-layer perceptron
MMF	Multimode fiber
MNIST	Modified national institute of standards and technology

MR	Mutation rate in genetic algorithms
MSE	Mean square error
OCT	Optical coherence tomography
OPC	Optical phase conjugation
PBR	Peak-to-background ratio
PCC	Pearson correlation coefficient
PER	Pulse emission rate
PFA	Parameter-free algorithm
PMMA	Polymethyl methacrylate
PSF	Point spread function
PSNR	Peak signal to noise ratio
PSO	Particle swarm optimization
RIE	Reactive ion etching
RM	Reflection matrix
ROI	Region of interest
RSA	Rivest-Shamir-Adleman encryption

SA	Simulated annealing algorithm
SBP	Speckle background Pearson correlation coefficient
SD	Standard deviation
SGD	Stochastic gradient descent
SLM	Spatial light modulator
SMFP	Scattering mean free path
SSIM	Structural similarity index measure
TM	Transmission matrix
WFS	Wavefront shaping

LIST OF NOTATIONS

Symbol	Description
$Autocorr$	Autocorrelation
c	Ratio of correct units in the parameter-free algorithm
D	Wavefront matrix on the DMD
E	Incident light field
I	Intensity
L	Loss function in neural network training
P	Dynamic mutation rate in the parameter-free algorithm
R	Reflection matrix
T	Transmission matrix
U	Output light field
y	Ground truths in the loss function
$\hat{y}$	Neural network outputs in the loss function
η	Focusing efficiency in the parameter-free algorithm

1 INTRODUCTION

Part of contents in Sections 1.1-1.4 are extracted from a published peer-reviewed paper:

Lai, P.[†,#], **Zhao, Q.**[†], Zhou, Y., Cheng, S., Woo, C. M., Li, H., Yu, Z., Huang, X., Yao, J., Pang, W., Li, H., Huang, H., Li, W., Zheng, Y., Wang, Z., Yuan, C., & Zhong, T.[#] (2024). Deep-tissue optics: Technological development and applications. *Chinese Journal of Lasers*, *51*(1), 0107003.

The exploration of optical speckles stands at the forefront of this thesis. Our primary objective is to demystify and harness the potential of optical speckles. To this end, we employ wavefront shaping, a technique that strategically modifies the phase of light waves, to effectively navigate through optical scattering. Furthermore, we integrate deep learning-based models, which serve as sophisticated tools to decipher and retrieve delocalized information embedded within speckles. As a prelude to our detailed research findings, the following sections will present a comprehensive overview, including progress and innovative techniques that have shaped the field of speckle-related research.

1.1 Scattering in Optics

Optics, a pivotal sub-discipline of physics, delves into the phenomena, properties, and applications of light. Over time, optics has evolved into an independent discipline, and optical imaging plays a crucial role in scientific research. By utilizing the phenomena and properties of light to record images of objects, optical imaging has extensive applications [1]. For instance, optical imaging offers high resolution that is free from ionizing radiation, making it safer than X-rays or gamma rays that pose potential risks of oncogenicity [2]. Additionally, optical imaging shows flexibility in configuration to provide rich target information based on the amplitude, phase, wavelength, polarization, and other characteristics of light [3-4]. Furthermore, the application of contrast agents further enhances imaging specificity and contrast, thereby improving the visualization of desired targets and opening new avenues for disease diagnosis and treatment [5].

These advantages have inspired the development of various high-resolution optical imaging technologies, such as confocal microscopy [6], multiphoton microscopy [7], photoacoustic microscopy [8], optical coherence tomography [9], *etc*. Whilst encouraging, these implementations have encountered fundamental challenges in living biological tissues. The limitation stems from the strong scattering of light due to the inherent inhomogeneous spatial

distribution of the refractive index of scattering media, which encompasses diverse constituents and functions [10-12]. The inhomogeneous refractive indices have various effects on the amplitudes and phases of wavefronts. As a result, the light beam spreads quickly and is accompanied by the accumulated scattering of light (approximately one scattering event per 0.1 mm optical path length at visible wavelengths), generating optical speckles rather than clear spots or imaging outside the medium [13], as shown in Figure 1-1a. In combination, these result in an intrinsic trade-off between spatial resolution and penetration depth for optical techniques in biological tissues [14].

The main purpose of this thesis is to understand and manipulate optical speckles. To achieve this goal, wavefront shaping is employed to address optical scattering through scattering media. Additionally, deep learning-based methods are proposed to analyze and retrieve delocalized information from speckles, and the conditions under which information can be retrieved with high fidelity are investigated, leveraging the information entropy. From another perspective, considering their inherent randomness, optical speckles can be utilized as ciphertexts in optical cryptosystems to protect private data. Subsequent sections will elaborate on the backgrounds of related research.

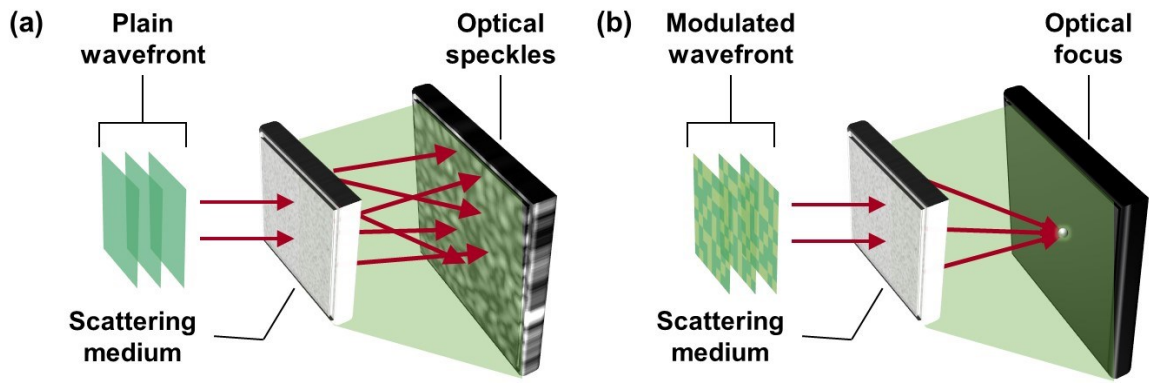

Figure 1-1 (a) A plain (unmodulated) wavefront passes through a scattering medium, resulting in optical speckles, rather than an optical focus. (b) With wavefront shaping, the modulated wavefront passes through the same scattering medium, generating a clear optical focus. This figure is reproduced from Ref. [15-16].

1.2 Computational Approaches

Despite extensive research efforts to overcome scattering challenges, traditional physics-based methods have limitations, particularly in scenarios characterized by strong scattering [12]. In recent years, the advent of computational optics has marked a significant shift in research related to optical speckles, rapidly evolving into an interdisciplinary domain that integrates optical theories with computational algorithms. This fusion of disciplines strives to harness the combined strengths of physics and computational science to realize applications beyond the

reach of conventional optics [17]. Predominant computational optics strategies employed in deep tissue optics include digital optical phase conjugation (DOPC), iterative wavefront shaping, and transmission/reflection matrix (TM/RM) [15].

Traditional optical phase conjugation (OPC) employs a phase conjugate mirror, such as a phase conjugate crystal, to correct the distorted wavefront by the scattering medium, achieving time reversal and refocusing light back to the guide star [18]. In contrast, DOPC utilizes a suite of digital tools comprising a digital camera, computer, spatial light modulator, and algorithms to supplant the phase-conjugation mirror [19]. These components synergistically ascertain and produce the phase-conjugated wavefront, effectively compensating for phase aberrations induced by optical scattering, thereby facilitating optical focus through a scattering medium [20]. Despite its promise, DOPC faces challenges with the stringent pixel alignment required between the wavefront sensor and the wavefront modulator. Coupled with intricate optical design and the necessity for guide stars, these factors hinder the widespread adoption of DOPC in other optical instruments and its application in deep tissue imaging.

Exploring cost-effective and straightforward approaches, iterative wavefront shaping has emerged as a promising computational method. Pioneered by Vellekoop *et al*. in 2007, this technique introduced an iterative optimization-based approach to wavefront compensation,

marking a milestone in achieving optical focusing through dense scattering media and revolutionizing deep tissue optical focusing and imaging [16]. Iterative wavefront shaping adjusts the phase of the incident light wavefront using feedback signals, with the optical field being refined iteratively to enhance the focus brightness or feedback signals over successive iterations [21], as depicted in Figure 1-1b. The feedback signals can take various forms, such as focal intensity, peak-to-background ratio (PBR) in the captured speckles [22], and photoacoustic signal strengths [23]. The choice of heuristic optimization algorithm is critical, significantly influencing the focusing performance [24]. Given that wavefront optimization is inherently a non-convex problem, the selected algorithm must navigate past local optima, converge swiftly, and robustly withstand environmental perturbations [25].

Various algorithms have been proposed to address these challenges. For instance, Li *et al*. introduced a dynamic mutation algorithm (DMA) in 2021 to facilitate focusing amidst non-stationary scattering media [26], while Woo *et al*. combined genetic algorithms with ant colony optimization (GA-ACO) in 2022 to optimize focusing efficiency [27]. Nonetheless, these algorithms typically necessitate optimization durations ranging from seconds to minutes. The primary bottleneck is the sluggish response of spatial light modulators (Table 1-1) and cameras used for capturing feedback signals, compounded by communication delays between hardware

components [28]. Consequently, achieving real-time focusing through scattering media remains elusive, underscoring the need for hardware capable of higher speeds.

Table 1-1 Commonly used spatial light modulators in wavefront shaping.

	Modulation	Refresh rate	Control units
Liquid crystal-based spatial light modulator [29]	Phase	0.1 kHz	~10^6
Digital micromirror device [30]	Amplitude	23 kHz	~10^6
Grating light valve [31]	Amplitude	350 kHz	~10^3

While iterative wavefront shaping has proven effective for optical focusing through scattering media, it inherently requires a time-intensive optimization process for each focal point. Nevertheless, the establishment of a mathematical model for the scattering medium offers a solution and enables the pre-calculation of wavefront compensation patterns necessary for focusing on various spatial locations [32-33]. In the context of the transmission matrix (TM), a linear mathematical matrix model describes the relations between the incident and scattered output wavefronts to characterize the scattering medium. Within this framework, speckles can be conceptualized as the cumulative effect of diverse wavefronts [34].

If the incident light field is E and the output light field is U, the transmission matrix T can

be represented as [35]:

$$U = T \cdot E = \begin{bmatrix} t_{11} & \cdots & t_{1n} & \cdots & t_{1N} \\ \vdots & \ddots & & & \vdots \\ t_{m1} & & t_{mn} & & t_{mN} \\ \vdots & & & \ddots & \vdots \\ t_{M1} & \cdots & t_{Mn} & \cdots & t_{MN} \end{bmatrix} \cdot E. \quad (1\text{-}1)$$

The intensity I_{out} of resultant optical speckles is the amplitude function of the scattered output light field:

$$I_{out} = |U^2|. \quad (1\text{-}2)$$

In the process of measuring the TM, a diverse array of modulated wavefronts, such as those based on the Hadamard set, are projected onto the scattering medium. The corresponding outputs are then recorded to ascertain T in Equation (1-1) [36-37].

With the TM determined, it becomes possible to derive wavefronts for optical focus, project patterns through scattering media, or retrieve images from speckles. Notably, Boniface *et al*. achieved non-invasive focusing through scattering media by employing a fluorescent-based TM approach in 2020 [38]. More recently, in 2023, Cheng *et al*. successfully conducted phase optimization using an alternating projection method derived from TM, culminating in the projection of specific patterns through scattering media [39].

While the advancements in iterative wavefront shaping and transmission matrix models are

commendable, their reliance on feedback signals presents a challenge. In the context of deep tissue imaging, the practicality of designating guide stars or obtaining feedback signals within tissue samples is often unfeasible [40]. This limitation constrains the broader application of TM models. The situation motivates the development of the reflection matrix (RM, or R) that shares a conceptual similarity with the TM, yet establishes the relations between the incident light field E and the reflected light field $U_{reflected}$ [41]:

$$U_{reflected} = R \cdot E = \begin{bmatrix} r_{11} & \cdots & r_{1n} & \cdots & r_{1N} \\ \vdots & \ddots & & & \vdots \\ r_{m1} & & r_{mn} & & r_{mN} \\ \vdots & & & \ddots & \vdots \\ r_{M1} & \cdots & r_{Mn} & \cdots & r_{MN} \end{bmatrix} \cdot E, \tag{1-3}$$

where the introduction of a R addresses the challenge by utilizing a reflected wavefront instead of a transmitted wavefront.

Compared with TM, the RM approach offers a strategic advantage by positioning both the incident and reflected light detectors on the same side of the scattering medium. This configuration obviates the necessity for feedback signals across the medium [42]. In a groundbreaking study, Kang *et al*. successfully measured the RM of a scattering medium, significantly amplifying the energy of the incident light [43]. Building on this, Cao *et al*. introduced an RM-based Optical Coherence Tomography (RM-OCT), achieving remarkable focusing at a depth of 9.6 times the average scattering mean free path (SMFP) [44]. Despite

these advances, existing RM-based methods still need multiple measurements, posing a challenge to keeping pace with the rapid decorrelation characteristic of scattering media.

1.3 Deep Learning-based Approaches

Efforts of the aforementioned computational optics are based on complex physical models that simulate wavefront propagation through scattering media, facilitating optical focusing and imaging of rudimentary targets like letters, numbers, and basic patterns. Yet, these methods encounter efficiency roadblocks when faced with unstable scattering media or intricate imaging targets. Under such conditions, the efficacy of computational optical techniques diminishes, posing significant challenges to the successful recovery of images or the precise control of wavefronts.

The advent of artificial intelligence, particularly deep learning (DL), has revolutionized the approach to complex speckle-related problems [45]. DL has served as a formidable tool in deciphering the intricacies of scattering media [46]. By training Deep Neural Networks (DNNs) with established data pairs, which include ground-truth images alongside their corresponding speckles, it's possible to extract multifaceted speckle features to retrieve valuable information,

as illustrated in Figure 1-2. The prowess of a pre-trained neural network lies in its ability to retrieve high-fidelity images, such as human faces, directly from speckles. This negates the necessity for optical focusing and raster scanning typically required in iterative wavefront shaping [28].

Additionally, DNNs have further promoted the application of speckle imaging through non-stationary scattering media [47]. With their ability to discern features and structures at various levels, DNNs exhibit remarkable generalization capabilities. They build the input-output wavefront relationship and enable accurate retrieval of original images from speckles [48]. Further, DNNs can learn from speckles across different statuses of a disturbed scattering medium, enhancing their adaptability to changes within the scattering medium [49]. A notable development was the semi-supervised learning model introduced by Fan *et al*. in 2021, which was designed to counteract the time-varying nature of multimode fibers (MMF) [50]. More recently, in 2024, Li *et al*. developed a multiscale memory dynamic-learning network that updated parameters through online training and transmitted video data over a one-kilometer-long MMF [51]. These underscore the superior robustness of DL-based methods over traditional approaches.

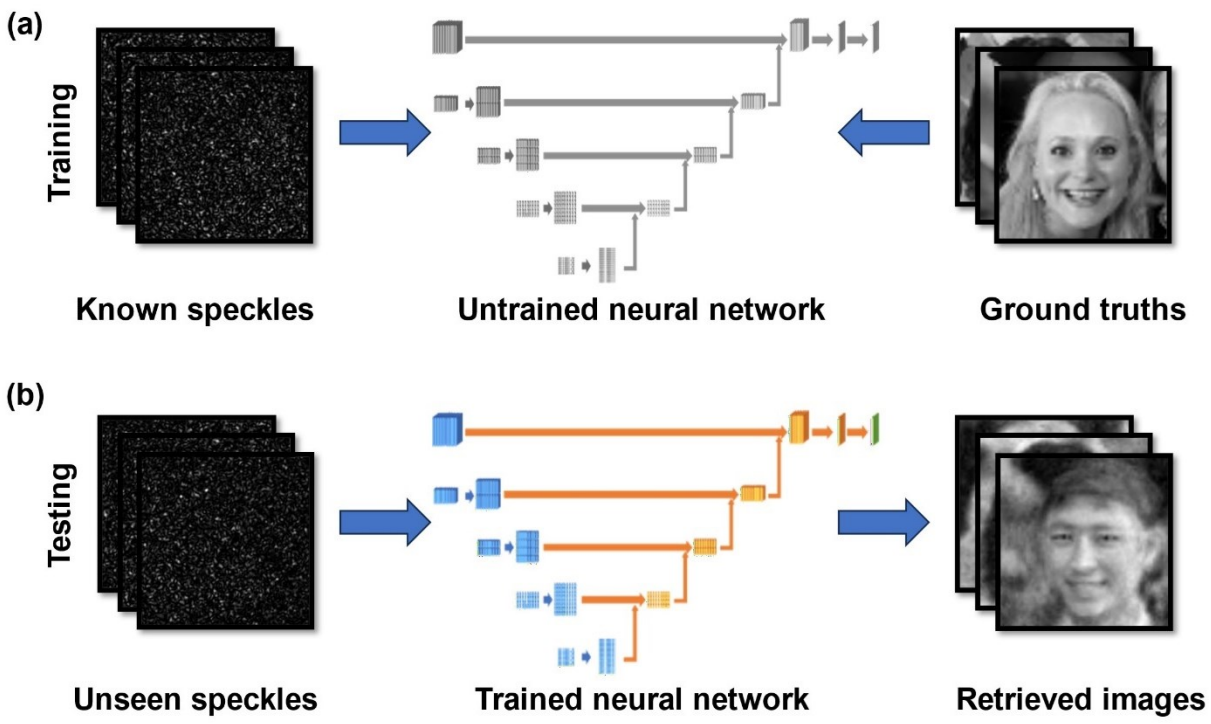

Figure 1-2 Training neural networks: known speckles and the corresponding ground truths are used to train the designed neural network, whose outputs are gradually tuned to approximate the ground truths. (b) Testing neural networks: unseen speckles are used to test the trained neural network, whose outputs are retrieved images from speckles. This figure is reproduced from Ref. [52].

1.4 Multimode Fibers in Imaging

The advancements discussed thus far have introduced noninvasive techniques aimed at mitigating or suppressing scattering in deep tissues. Yet, the scattering properties of living

biological tissues, coupled with their millisecond-scale decorrelation time, pose significant challenges to the deployment of these technologies for high-resolution focusing and imaging within living tissues. Recognizing the limitations of current noninvasive methods, some researchers are pivoting towards minimally invasive strategies that leverage ultrathin optical MMFs [53].

It is important to note that conventional endoscopes can also bypass the interference caused by tissue scattering, facilitating the use of high-resolution optical technologies in deep tissue applications. However, the endoscopes typically incorporate traditional optical elements, such as image sensors and optical lenses. Their substantial sizes not only make them cumbersome but also lead to insertion trauma [54]. In contrast, MMFs present a host of benefits: they are minimally invasive, with diameters ranging from 100 to 200 μm (comparable to an adult's hair strand), offering both flexibility and cost-effectiveness [55]. Yet, the inherent mode dispersion and coupling within MMFs result in an optical field output that resembles speckles produced by scattering media, complicating the direct interpretation of the transmitted spatial information [56].

To surmount this obstacle, fiber optics and computational optics have converged to innovate optical endoscopy. Through wavefront shaping, output wavefronts of MMFs can be precisely

controlled. That is, leveraging computational optics, iterative wavefront shaping can channel light into an MMF and facilitate optical focus within deep tissues, as illustrated in Figure 1-3 [57]. Moreover, by determining the TM/RM of the fiber, raster scanning of the focused light can be performed within the field of view of the MMF to generate or retrieve high-resolution images [58].

Alternatively, neural networks can also be trained in advance to establish a correlation between the input wavefronts and the resultant speckles through an MMF, thereby retrieving images from MMF speckles [59]. That said, it must be highlighted that the delicate nature of MMFs, which is susceptible to deformation, calls for neural networks with enhanced generalizability. A strategy to improve generalizability involves incorporating diverse statuses of the MMF into the training dataset, ensuring that the pre-trained network can accurately retrieve images even post-deformation of the MMF.

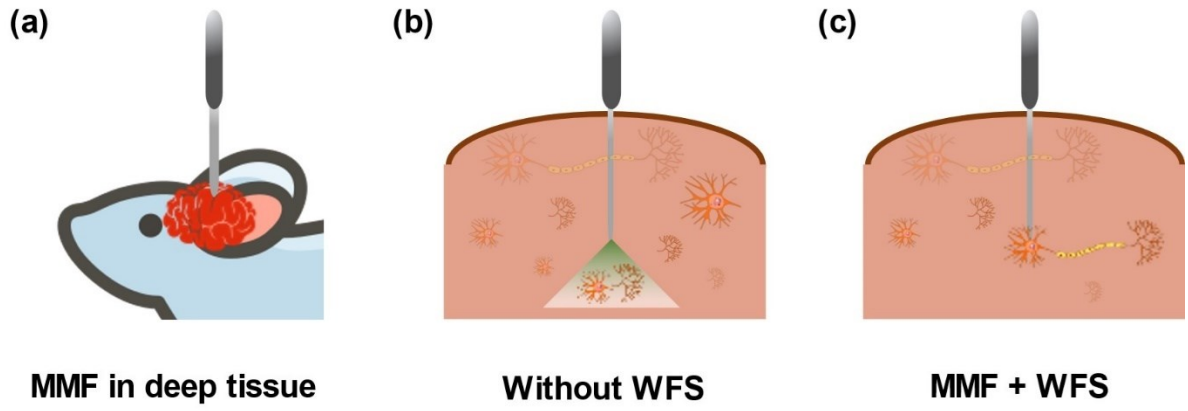

Figure 1-3 (a) Illustration of a lensless MMF being inserted into deep tissue with minimal invasion. (b) Without wavefront shaping (WFS), light output from the MMF is scrambled, forming a diffused speckle-like pattern. (c) With wavefront shaping, the light output from the MMF can be high-resolution focused and raster scanned, allowing for high-resolution optical manipulation and imaging in deep tissue. This figure is reproduced from Ref. [60].

1.5 Speckles for Encryption

Apart from overcoming optical scattering and retrieving information from speckles, the inherent randomness of speckles, characterized by bright and dark spots, can serve as candidates for encryption in hardware-based cryptosystems [61]. Subsequently, research related to the utilization of speckles for encryption will be introduced.

Among the hardware-based cryptosystems, optical cryptosystems have garnered considerable attention with the progress of optical computing and computational imaging [62-63]. The optical cryptosystems offer several benefits over traditional software-based cryptosystems. First, optical cryptosystems typically demonstrate faster speeds, higher security, and lower costs in comparison to commonly employed software-based cryptosystems [64]. Additionally, optical encryption enables much longer secret key lengths, enhancing the overall system security [61]. Last but not least, optical encryption achieves cost efficiency [64]. On the contrary, software-based cryptosystems often require expensive high-performance computers to attain comparable security levels.

Due to these advantages, two main types of optical cryptosystems have been developed, namely: double random phase encryption (DRPE) [65-66] and speckle-based optical cryptosystems [67-68]. Regarding DRPE, two phase masks are utilized at the input plane, and plaintexts are encrypted on the Fourier plane [65]. Accordingly, the optical setup of DRPE is complex due to the interferometric design, and DRPE is challenging to integrate with other systems [69]. As for speckle-based optical cryptosystems, optical speckles are directly employed as ciphertexts to encrypt plaintexts, eliminating the need for interferometric design and attracting considerable research interests [68]. In optical speckle-based cryptosystems, scattering media are utilized as

the physical secret key, resulting in optical speckles with randomly distributed bright and dark spots. The optical speckles are the ciphertexts and can be captured by regular digital cameras for further processing. Accordingly, speckle encryption is flexible for integration with existing systems due to its straightforward implementation, as shown in Figure 1-4 [52]. Additionally, the random nature of optical speckles creates nearly infinite information channels, yielding a high level of security and information protection with extremely long physical secret key lengths and enhancing the security of the speckle-based cryptosystem [70].

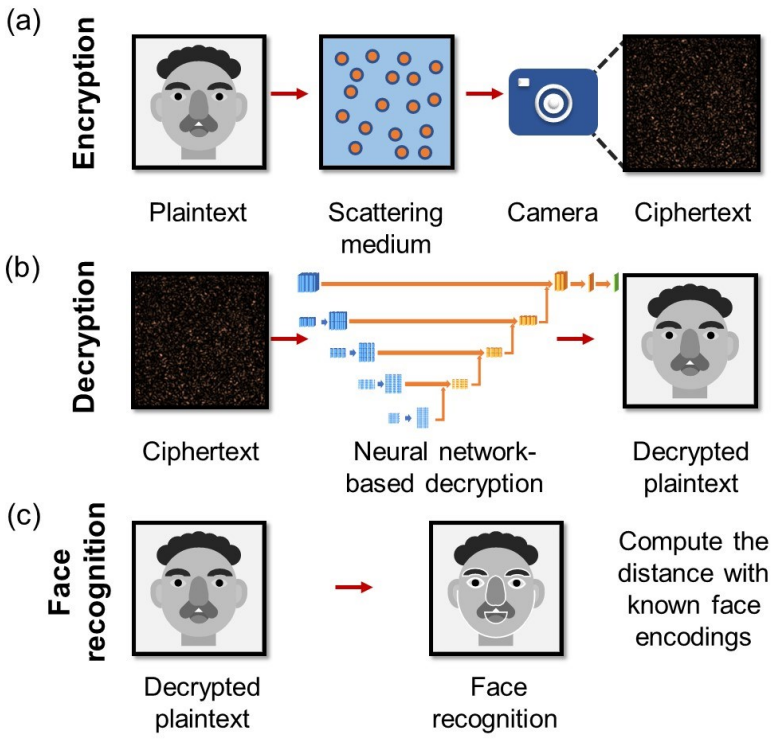

Figure 1-4 Conceptual illustration of the speckle-based optical cryptosystem: a ground glass is exploited as the physical secret key to encrypt face images via random optical

speckles at light speed; a well-trained neural network can decrypt speckles to face images for recognition. This figure is reproduced from Ref. [52].

Furthermore, although speckle encryption is simple to implement with optical setups, decrypting speckles demands elaborate algorithms for retrieving plaintexts from speckles. From another perspective, decrypting speckles is equivalent to retrieving image information from speckles, and methods for speckle imaging can be applied for decryption, including transmission matrix-based methods [70] and deep learning-based approaches [67]. For encrypting simple-structured images, the two methods perform well in retrieving digits, characters, or simple patterns from optical speckles [70-71]. Consequently, the applications of speckle-based cryptosystems have mainly focused on encrypting simple images (*e.g.*, characters, clothes, animals, *etc.*) with high security and fast-speed encryption [65-70]. However, speckle-based optical cryptosystems for complex tasks, such as encrypted face recognition, remain rarely explored. Here, challenges encompass handling rapidly changing optical speckles for decryption with high fidelity to preserve key facial features and detailed structures. Additionally, transmission matrix-based methods and deep learning-based approaches all require lots of speckle data during training processes to build the complex

relationships between speckles and face images. Accordingly, high fidelity image retrieval can help face recognition from decrypted images with high accuracy for practical applications.

1.6 Motivation

In this thesis, efforts have been devoted to overcoming, understanding, and utilizing optical speckles. The scope of research encompasses the algorithm for iterative wavefront shaping, the retrieval of information from speckles, the concept of delocalized information within speckles, the process of classification through scattering media, and the utilization of optical speckles in cryptographic systems. The motivations behind these studies are expounded as follows:

(1) Parameter-free algorithm: Iterative wavefront shaping is a critical technique for directing light through or within scattering media. A notable challenge in this domain is the fine-tuning of numerous parameters to secure robust and optimal focusing, which is inevitable for existing iterative algorithms. Additionally, the parameter-tuning process can be laborious and highly dependent on the particularities of the scattering samples and experimental setups. To address this, the thesis introduces a novel parameter-free algorithm designed to autonomously calibrate parameters utilizing real-time feedback. This

innovation is driven by the desire to bolster the algorithm's adaptability to environmental fluctuations and a diversity of scattering conditions.

(2) Spatiotemporally decorrelated speckles: Deep learning-based strategies hold considerable promise for retrieving images directly from speckles. Yet, most of the related research neglects the time intervals between acquiring the training and test datasets, and data from the training and test sets are highly correlated. Accordingly, these neural networks face hurdles in non-stationary scattering media, where the testing data may significantly decorrelate from the training data, thus constraining the practicality of deep learning applications. To surmount this challenge, this thesis aims to enhance the generalizability of neural networks. The goal is to equip deep learning models with the ability to adapt to unknown statuses of the scattering medium that deviate from the initial training data.

(3) Delocalized information in speckles: While a range of methods for information retrieval from speckles have been developed, the distribution of information within speckles warrants deeper exploration, including the extent of delocalization in speckles and the sampling condition for information retrieval from speckles. This thesis delves into the foundational theories of speckle imaging from the perspective of information entropy. It examines the delocalized distribution of information in speckles and identifies the speckle

sampling condition for high-fidelity image retrieval. The concept of delocalized information is investigated by training neural networks to extract information from speckle regions of interest (ROIs) of varying sizes and locations. Furthermore, experimental evidence supports the retrieval of delocalized information with high fidelity when the sampled speckle ROIs contain sufficient information.

(4) Classification through scattering media: Despite the burgeoning body of research in speckle imaging, the task of image retrieval from speckles when only limited information is available remains a challenge, as underscored in the discourse on delocalized information within speckles. In response to this quandary, this thesis introduces a neural network tailored for direct classification from speckles using sparse information. This model is designed to obviate the necessity for exhaustive speckle data prior to image retrieval, thereby facilitating classification predicated on the intrinsic characteristics of speckles.

(5) Speckle-based optical cryptosystem: Beyond the scope of mitigating optical speckles, this thesis also ventures into the realm of their applications within cryptographic systems. Owing to their intrinsic randomness, optical speckles serve as exemplary candidates for ciphertexts, providing the dual benefits of enormously long secret keys and rapid encryption capabilities. Although speckle-based cryptosystems have been established for encoding

simple targets, speckle-based optical cryptosystems for complex tasks remain largely unexplored, such as encrypted face recognition. The primary challenge lies in decrypting images from rapidly changing optical speckles and recognizing faces from the decrypted images. Additionally, to achieve high accuracy in face recognition, it's crucial to maintain high fidelity in key facial features and detailed structures of the decrypted face images. This thesis introduces a straightforward yet remarkably effective speckle-based optical cryptosystem, and demonstrates its application in the field of encrypted human face recognition.

1.7 Thesis Outline

This thesis is structured into five studies in accordance with the aforementioned motivations, as outlined in Figure 1-5 and the following contents:

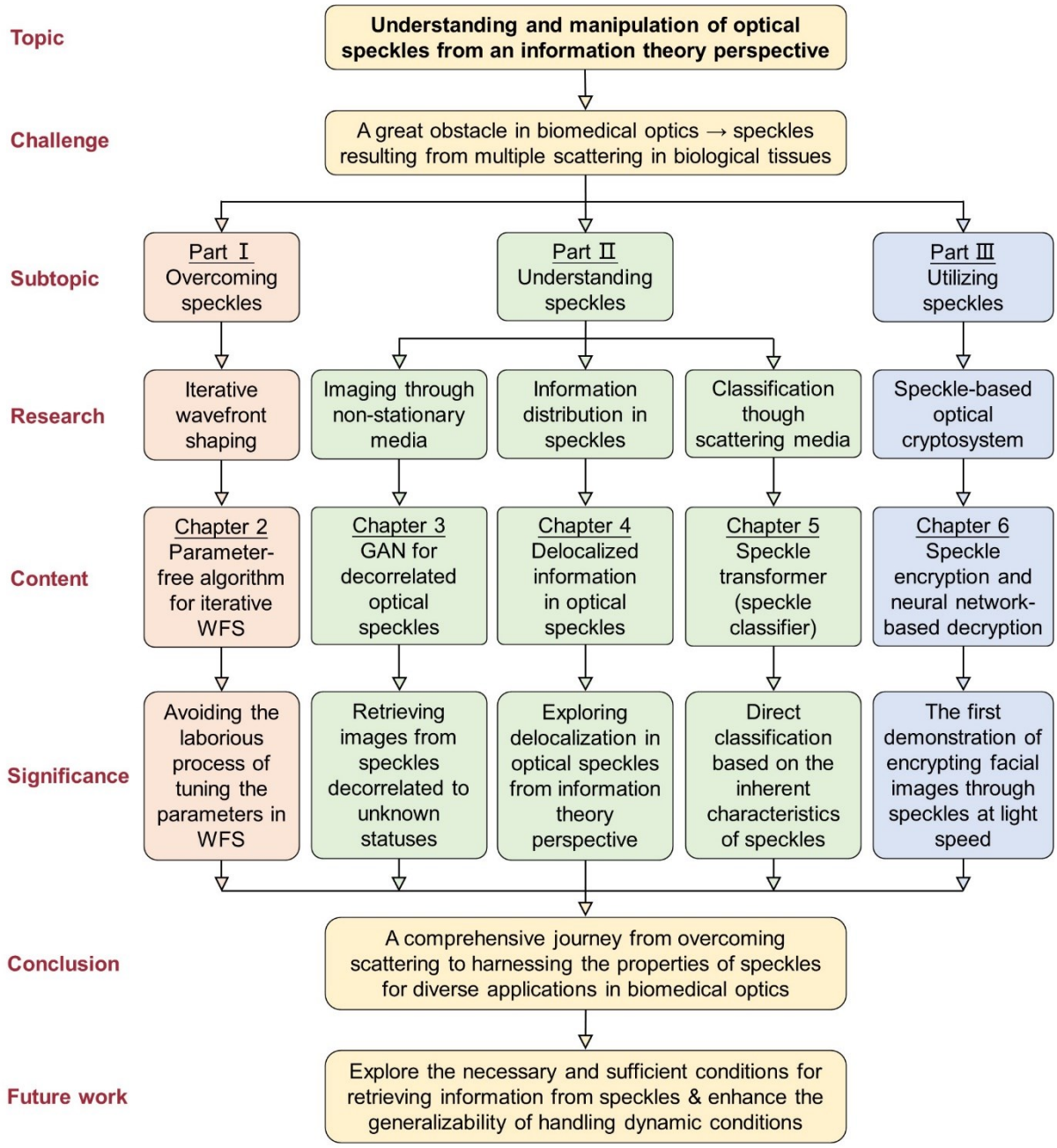

Figure 1-5 Thesis outline.

Chapter 2 introduces a parameter-free algorithm for iterative wavefront shaping, aiming to manipulate wavefronts to focus through scattering media and setting the foundation for subsequent chapters. This algorithm amalgamates Genetic, Bat, and Dynamic Mutation

Algorithms to facilitate automatic optimization of parameters throughout the iterative wavefront optimization process. This approach is designed to circumvent the labor-intensive and expertise-reliant parameter-tuning procedure, which is inevitable for existing iterative algorithms. A series of experiments employing ground glass and multimode fibers substantiate the algorithm's effectiveness and repeatability.

Chapter 3 delves into the phenomenon of spatiotemporal decorrelation in optical speckles. It introduces a Generative Adversarial Network (GAN)-based framework with enhanced generalizability, capable of addressing the non-stationary scattering media and the resultant decorrelation between training and testing datasets. The ability to retrieve images from decorrelated speckles is an extension of the focusing capabilities developed in Chapter 2, showcasing the application of advanced computational methods to enhance image clarity. Furthermore, we separate training and testing datasets with different time intervals in experiments, so that training and testing data acquisition windows do not overlap in time. The results demonstrate that the proposed GAN framework can be trained to retrieve high-fidelity face images from speckles decorrelated to statuses not encountered during training, even after the optical system has been inactive for an extended period (up to 37 hours in experiments) and subsequently reactivated. This pre-emptive training capability, which preserves network

performance despite decorrelation, is pivotal for broadening neural networks in the realm of speckle imaging.

Chapter 4 investigates the concept of delocalized information within optical speckles through the prism of information entropy, which further advances the understanding of information distribution in optical speckles. This chapter complements the practical applications explored in Chapters 2 and 3, providing a deeper theoretical understanding that informs the development of speckle-based techniques. Contrasting with ballistic imaging where light travels unimpeded, light is subject to multi-path scattering in a scattering medium. The multi-path scattering results in the spatial delocalization of a single point from the object across multiple points within the speckle, and conversely. Utilizing deep learning models, the research examines the distribution of image information within optical speckles. The experiments suggest that information is dispersed across the speckle field, enabling the retrieval of target information from speckles located in various spatial regions and of differing sizes. Subsequent to the analysis of physical models and experimental data, it has been empirically found that if the entropy of speckle autocorrelation exceeds that of the target autocorrelation, it is feasible to train neural networks to extract information from speckles with high fidelity, as indicated by a Pearson correlation coefficient greater than 0.9. The speckle sampling condition is crucial as it indicates that neural

networks can effectively retrieve information from speckles with high fidelity, as evidenced by the Pearson correlation coefficient (PCC) greater than 0.9, provided that the sampled speckle ROIs contain sufficient information. Otherwise, speckle ROIs are expected to be extended to include more information for high-fidelity information retrieval. Accordingly, this research paves the way for future developments in speckle-related research, and has the potential to inspire new applications and methodologies for biomedical imaging.

In Chapter 5, delocalized information within speckles is further explored for its application in direct classification through scattering media rather than solely for imaging, which is a synthesis of the practical algorithmic advancements and theoretical insights from the previous chapters. Speckle Transformer, a model based on the vision transformer, is designed to harness delocalized information for accurately classifying targets. This method extracts features from speckles and bypasses the need for comprehensive speckle data prior to image retrieval. Consequently, it facilitates classification predicated on the intrinsic characteristics of speckles, even when information is sparse, and entropy analyses further highlight the influence of delocalized speckle information on classification accuracies. Notably, this approach surpasses the accuracy of methods that classify after image retrieval, marking a significant advancement in speckle-based classification techniques.

In Chapter 6, the focus shifts from overcoming speckles to harnessing them for a high-security optical cryptosystem, which signifies a departure from the primary focus on imaging to embracing the natural randomness of speckles and expanding the scope of speckle-based applications in the realm of security. A straightforward yet highly effective speckle-based cryptosystem is introduced, and its application in human face recognition is demonstrated. The cryptosystem leverages scattering ground glass as a physical secret key, encrypting face images into seemingly random optical speckles at the speed of light. Subsequently, a pre-trained U-Net-based neural network decrypts images from speckles, enabling remarkable-fidelity retrieval for analysis by the subsequent face recognition algorithm. The high security, rapid processing, and cost-efficiency of this speckle-based cryptosystem make it a powerful tool for practical applications and the advancement of high-security cryptographic systems. Additionally, to the best of our knowledge, this is the first demonstration of a speckle-based optical cryptosystem for face recognition, which can be applied to other types of biometric data and optical data storage.

In the final chapter (Chapter 7), the thesis culminates by encapsulating the key contributions and envisioning the trajectory for future research endeavors. Together, these chapters illustrate a comprehensive journey from overcoming the challenges posed by scattering to harnessing the

properties of speckles for diverse applications. Each chapter contributes a piece to the challenge of optical scattering, with the collective work pushing the boundaries of optical imaging and expanding its potential applications. Ultimately, this thesis aspires to make a meaningful impact on the wider research community through a comprehensive exploration of speckle analysis and its practical applications, paving the way for novel research pathways and potential advancements in the field of biomedical optics.

References

1. Luker, G. D., & Luker, K. E. (2008). Optical imaging: current applications and future directions. *Journal of Nuclear Medicine, 49*(1), 1-4.
2. Balas, C. (2009). Review of biomedical optical imaging: a powerful, non-invasive, non-ionizing technology for improving in vivo diagnosis. *Measurement science and technology, 20*(10), 104020.
3. Rotter, S., & Gigan, S. (2017). Light fields in complex media: mesoscopic scattering meets wave control. *Reviews of Modern Physics, 89*(1), 015005.
4. de Aguiar, H. B., Gigan, S., & Brasselet, S. (2017). Polarization recovery through scattering

media. *Science Advances, 3*(9), e1600743.

5. Dzik-Jurasz, A. S. K. (2003). Molecular imaging in vivo: an introduction. *The British journal of radiology, 76*(suppl_2), S98-S109.

6. Elliott, A. D. (2020). Confocal microscopy: principles and modern practices. *Current protocols in cytometry, 92*(1), e68.

7. Benninger, R. K., & Piston, D. W. (2013). Two-photon excitation microscopy for the study of living cells and tissues. *Current protocols in cell biology, 59*(1), 4-11.

8. Cho, S. W., Park, S. M., Park, B., Lee, T. G., Kim, B. M., Kim, C., ... & Kim, C. S. (2021). High-speed photoacoustic microscopy: a review dedicated on light sources. *Photoacoustics, 24*, 100291.

9. Pircher, M., & Zawadzki, R. J. (2017). Review of adaptive optics OCT (AO-OCT): principles and applications for retinal imaging. *Biomedical optics express, 8*(5), 2536-2562.

10. Sheng, P., & van Tiggelen, B. (2007). Introduction to wave scattering, localization and mesoscopic phenomena.

11. Yoon, S., Kim, M., Jang, M., Choi, Y., Choi, W., Kang, S., & Choi, W. (2020). Deep optical imaging within complex scattering media. *Nature Reviews Physics, 2*(3), 141-158.

12. Lai, P., Zhao, Q., Zhou, Y., Cheng S., Woo, C. M., Li, H., Yu, Z., Huang, X., Yao, J., Pang,

W., Li, H., Huang, H., Li, W., Zheng, Y., Wang, Z., Yuan, C., and Zhong, T., Deep-Tissue Optics: Technological Development and Applications. *Chinese Journal of Lasers 51*(1) 0107003, 2024.

13. Cao, H., Mosk, A. P., & Rotter, S. (2022). Shaping the propagation of light in complex media. *Nature Physics, 18*(9), 994-1007.
14. Park, J. H., Yu, Z., Lee, K., Lai, P., & Park, Y. (2018). Perspective: Wavefront shaping techniques for controlling multiple light scattering in biological tissues: toward in vivo applications. *APL photonics, 3*(10).
15. Yu, Z., Li, H., Zhong, T., Park, J. H., Cheng, S., Woo, C. M., ... & Lai, P. (2022). Wavefront shaping: a versatile tool to conquer multiple scattering in multidisciplinary fields. *The Innovation, 3*(5).
16. Vellekoop, I. M., & Mosk, A. P. (2007). Focusing coherent light through opaque strongly scattering media. *Optics letters, 32*(16), 2309-2311.
17. Xiang, M., Liu, F., Liu, J., Dong, X., Liu, Q., & Shao, X. (2024). Computational optical imaging: challenges, opportunities, new trends, and emerging applications. *Frontiers in Imaging*, *3*, 1336829.
18. He, G. S. (2002). Optical phase conjugation: principles, techniques, and applications.

Progress in Quantum Electronics, 26(3), 131-191.

19. Wang, Y. M., Judkewitz, B., DiMarzio, C. A., & Yang, C. (2012). Deep-tissue focal fluorescence imaging with digitally time-reversed ultrasound-encoded light. *Nature communications, 3*(1), 928.

20. Si, K., Fiolka, R., & Cui, M. (2012). Fluorescence imaging beyond the ballistic regime by ultrasound pulse guided digital phase conjugation. *Nature photonics, 6*(10), 657-661.

21. Vellekoop, I. M. (2008). Controlling the propagation of light in disordered scattering media. *arXiv preprint arXiv*:0807.1087.

22. Horstmeyer, R., Ruan, H., & Yang, C. (2015). Guidestar-assisted wavefront-shaping methods for focusing light into biological tissue. *Nature photonics, 9*(9), 563-571.

23. Lai, P., Wang, L., Tay, J. W., & Wang, L. V. (2015). Photoacoustically guided wavefront shaping for enhanced optical focusing in scattering media. *Nature photonics, 9*(2), 126-132.

24. Vellekoop, I. M., & Mosk, A. P. (2008). Phase control algorithms for focusing light through turbid media. *Optics communications, 281*(11), 3071-3080.

25. Fayyaz, Z., Mohammadian, N., Reza Rahimi Tabar, M., Manwar, R., & Avanaki, K. (2019). A comparative study of optimization algorithms for wavefront shaping. *Journal of*

innovative optical health sciences, 12(04), 1942002.

26. Li, H., Woo, C. M., Zhong, T., Yu, Z., Luo, Y., Zheng, Y., ... & Lai, P. (2021). Adaptive optical focusing through perturbed scattering media with a dynamic mutation algorithm. *Photonics Research, 9*(2), 202-212.

27. Woo, C. M., Zhao, Q., Zhong, T., Li, H., Yu, Z., & Lai, P. (2022). Optimal efficiency of focusing diffused light through scattering media with iterative wavefront shaping. *APL Photonics, 7*(4).

28. Tzang, O., Niv, E., Singh, S., Labouesse, S., Myatt, G., & Piestun, R. (2019). Wavefront shaping in complex media with a 350 kHz modulator via a 1D-to-2D transform. *Nature Photonics, 13*(11), 788-793.

29. Yang, Y., Forbes, A., & Cao, L. (2023). A review of liquid crystal spatial light modulators: devices and applications. *Opto-Electronic Science*, *2*(8), 230026-1.

30. Ren, Y. X., Lu, R. D., & Gong, L. (2015). Tailoring light with a digital micromirror device. *Annalen der physik*, *527*(7-8), 447-470.

31. Tzang, O., Niv, E., Singh, S., Labouesse, S., Myatt, G., & Piestun, R. (2019). Wavefront shaping in complex media with a 350 kHz modulator via a 1D-to-2D transform. *Nature Photonics*, *13*(11), 788-793.

32. Mujumdar, S. (2023). Transmission matrices go nonlinear. *Nature Physics, 19*(11), 1563-1564.

33. Popoff, S. M., Lerosey, G., Carminati, R., Fink, M., Boccara, A. C., & Gigan, S. (2010). Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media. *Physical review letters, 104*(10), 100601.

34. Conkey, D. B., Caravaca-Aguirre, A. M., & Piestun, R. (2012). High-speed scattering medium characterization with application to focusing light through turbid media. *Optics express, 20*(2), 1733-1740.

35. Wang, Z., Wu, D., Huang, G., Luo, J., Ye, B., Li, Z., & Shen, Y. (2021). Feedback-assisted transmission matrix measurement of a multimode fiber in a referenceless system. *Optics Letters, 46*(22), 5542-5545.

36. Yoon, J., Lee, K., Park, J., & Park, Y. (2015). Measuring optical transmission matrices by wavefront shaping. *Optics Express, 23*(8), 10158-10167.

37. Yu, H., Lee, K., & Park, Y. (2017). Ultrahigh enhancement of light focusing through disordered media controlled by mega-pixel modes. *Optics express, 25*(7), 8036-8047.

38. Boniface, A., Dong, J., & Gigan, S. (2020). Non-invasive focusing and imaging in scattering media with a fluorescence-based transmission matrix. *Nature communications,*

11(1), 6154.

39. Cheng, S., Zhong, T., Woo, C. M., & Lai, P. (2023). Alternating projection-based phase optimization for arbitrary glare suppression through multimode fiber. *Optics and Lasers in Engineering, 161*, 107368.

40. He, Y., Wu, D., Zhang, R., Cao, Z., Huang, Y., & Shen, Y. (2021). Genetic-algorithm-assisted coherent enhancement absorption in scattering media by exploiting transmission and reflection matrices. *Optics Express, 29*(13), 20353-20369.

41. Yu, H., Park, J. H., & Park, Y. (2015). Measuring large optical reflection matrices of turbid media. *Optics Communications, 352*, 33-38.

42. Yu, Z., Li, H., Zhong, T., & Lai, P. (2022). Enhancing spatiotemporal focusing of light deep inside scattering media with Time-Gated Reflection Matrix. *Light: Science & Applications, 11*(1), 167.

43. Kang, S., Jeong, S., Choi, W., Ko, H., Yang, T. D., Joo, J. H., ... & Choi, W. (2015). Imaging deep within a scattering medium using collective accumulation of single-scattered waves. *Nature Photonics, 9*(4), 253-258.

44. Lee, S. Y. (2022). Imaging through optical multimode fiber: towards ultra-thin endoscopy (Doctoral dissertation, Massachusetts Institute of Technology).

45. Cheng, S., Li, H., Luo, Y., Zheng, Y., & Lai, P. (2019). Artificial intelligence-assisted light control and computational imaging through scattering media. *Journal of innovative optical health sciences, 12*(04), 1930006.

46. Horisaki, R., Takagi, R., & Tanida, J. (2017). Learning-based focusing through scattering media. *Applied optics, 56*(15), 4358-4362.

47. Borhani, N., Kakkava, E., Moser, C., & Psaltis, D. (2018). Learning to see through multimode fibers. *Optica, 5*(8), 960-966.

48. Caramazza, P., Moran, O., Murray-Smith, R., & Faccio, D. (2019). Transmission of natural scene images through a multimode fibre. *Nature communications, 10*(1), 2029.

49. d'Arco, A., Xia, F., Boniface, A., Dong, J., & Gigan, S. (2022). Physics-based neural network for non-invasive control of coherent light in scattering media. *Optics Express, 30*(17), 30845-30856.

50. Li, Z., Zhou, W., Zhou, Z., Zhang, S., Shi, J., Shen, C., ... & Dai, Q. (2024). Self-supervised dynamic learning for long-term high-fidelity image transmission through unstabilized diffusive media. *Nature Communications*, *15*(1), 1498.

51. Fan, P., Ruddlesden, M., Wang, Y., Zhao, L., Lu, C., & Su, L. (2021). Learning enabled continuous transmission of spatially distributed information through multimode fibers.

Laser & Photonics Reviews, *15*(4), 2000348.

52. Zhao, Q., Li, H., Yu, Z., Woo, C. M., Zhong, T., Cheng, S., ... & Lai, P. (2022). Speckle-based Optical Cryptosystem and its Application for Human Face Recognition via Deep Learning. *Advanced Science*, *9*(25), 2202407.

53. Choi, Y., Yoon, C., Kim, M., Yang, T. D., Fang-Yen, C., Dasari, R. R., ... & Choi, W. (2012). Scanner-free and wide-field endoscopic imaging by using a single multimode optical fiber. *Physical review letters, 109*(20), 203901.

54. Ohayon, S., Caravaca-Aguirre, A., Piestun, R., & DiCarlo, J. J. (2018). Minimally invasive multimode optical fiber microendoscope for deep brain fluorescence imaging. *Biomedical optics express, 9*(4), 1492-1509.

55. Caravaca-Aguirre, A. M., & Piestun, R. (2017). Single multimode fiber endoscope. *Optics express, 25*(3), 1656-1665.

56. Shin, J., Bosworth, B. T., & Foster, M. A. (2017). Compressive fluorescence imaging using a multi-core fiber and spatially dependent scattering. *Optics letters, 42*(1), 109-112.

57. Tzang, O., Caravaca-Aguirre, A. M., Wagner, K., & Piestun, R. (2018). Adaptive wavefront shaping for controlling nonlinear multimode interactions in optical fibres. *Nature Photonics, 12*(6), 368-374.

58. Zhong, T., Yu, Z., Li, H., Li, Z., Li, H., & Lai, P. (2019). Active wavefront shaping for controlling and improving multimode fiber sensor. *Journal of innovative optical health sciences, 12*(04), 1942007.

59. Caravaca-Aguirre, A. M., Niv, E., Conkey, D. B., & Piestun, R. (2013). Real-time resilient focusing through a bending multimode fiber. *Optics express, 21*(10), 12881-12887.

60. Zhong, T., Qiu, Z., Wu, Y., Guo, J., Li, H., Yu, Z., ... & Lai, P. (2022). Optically selective neuron stimulation with a wavefront shaping-empowered multimode fiber. *Advanced Photonics Research*, *3*(3), 2100231.

61. Jeon, S. H., & Gil, S. K. (2019). Secret-key-sharing Cryptosystem Using Optical Phase-shifting Digital Holography. *Current Optics and Photonics, 3*(2), 119-127.

62. Liu, S., Guo, C., & Sheridan, J. T. (2014). A review of optical image encryption techniques. *Optics & Laser Technology, 57*, 327-342.

63. Javidi, B., Carnicer, A., Yamaguchi, M., Nomura, T., Pérez-Cabré, E., Millán, M. S., ... & Markman, A. (2016). Roadmap on optical security. *Journal of Optics, 18*(8), 083001.

64. Zhao, A., Jiang, N., Liu, S., Zhang, Y., & Qiu, K. (2021). Physical layer encryption for WDM optical communication systems using private chaotic phase scrambling. *Journal of Lightwave Technology, 39*(8), 2288-2295.

65. Unnikrishnan, G., Joseph, J., & Singh, K. (2000). Optical encryption by double-random phase encoding in the fractional Fourier domain. *Optics letters, 25*(12), 887-889.

66. Chen, W., Javidi, B., & Chen, X. (2014). Advances in optical security systems. Advances in optical security systems. *Advances in Optics and Photonics, 6*(2), 120-155.

67. Zhou, L., Xiao, Y., & Chen, W. (2020). Learning complex scattering media for optical encryption. *Optics Letters, 45*(18), 5279-5282.

68. Liu, Y., Yu, P., Li, Y., & Gong, L. (2020). Exploiting light field imaging through scattering media for optical encryption. *OSA Continuum, 3*(11), 2968-2975.

69. Hou, J., & Situ, G. (2022). Image encryption using spatial nonlinear optics. elight, 2(1), 3.

70. Qu, G., Yang, W., Song, Q., Liu, Y., Qiu, C. W., Han, J., ... & Xiao, S. (2020). Reprogrammable meta-hologram for optical encryption. *Nature communications, 11*(1), 5484.

71. Clemente, P., Durán, V., Tajahuerce, E., & Lancis, J. (2010). Optical encryption based on computational ghost imaging. *Optics letters, 35*(14), 2391-2393.

2 PARAMETER-FREE ALGORITHM FOR ITERATIVE WAVEFRONT SHAPING

This chapter is modified from the following published peer-reviewed paper:

Zhao, Q.[†], Woo, C. M.[†], Li, H., Zhong, T., Yu, Z.[#], & Lai, P.[#] (2021). Parameter-free optimization algorithm for iterative wavefront shaping. *Optics Letters*, *46*(12): 2880-2883.

Optical focusing through scattering media has a significant impact on optical applications in biological tissues. Accordingly, iterative wavefront shaping has been successfully used to focus light through or inside scattering media, and various heuristic algorithms have been introduced to improve the performance. While encouraging, in most heuristic algorithms for iterative wavefront shaping, lots of efforts might be needed to tune parameters towards robust and optimum optimization. Moreover, optimal parameters might differ for different scattering samples and experimental conditions, and parameter-tuning is a time and experience consuming

work. In this chapter, a parameter-free algorithm is proposed for iterative wavefront shaping. The parameter-free algorithm combines a traditional genetic algorithm with a bat algorithm, and the mutation rate can be automatically calculated through real-time feedback. Using the proposed parameter-free algorithm in iterative wavefront shaping, robust and optimum performance can be achieved without a parameter-tuning process.

2.1 Introduction

Light, serving as a basic tool, has been widely used in imaging, diagnosis, therapy, stimulation, *etc*. [1]. One of the key research questions in these fields is to address strong scattering of light within thick biological samples [2]. Because of the inhomogeneous refractive index distribution (wavelength-scale), photons suffer from multiple scattering events, and the carried information is scrambled [3], limiting the viability of high-resolution optical techniques within 1mm beneath the tissue surface [4]. Iterative wavefront shaping (iterative WFS) was first proposed in 2007 to compensate for phase distortions and achieve controllable optical delivery through complex media [5]. In the past decade, iterative WFS has seen remarkable progresses in deep optical imaging [6]. Various algorithms, such as a genetic algorithm (GA) [7], simulated

annealing algorithm (SA) [8], particle swarm optimization [9], and bat algorithm (BA) [10], have been introduced to improve the performance. For all these algorithms, it is generally necessary to conduct a parameter tuning procedure to obtain optimum parameters for satisfactory focusing performance inside or through scattering media. Moreover, optimized parameters may vary for different samples and environmental conditions, such as noise levels [11]. Efforts are required to obtain optimized parameters for a particular application, which poses an obstacle to newcomers and hinders the extension of iterative WFS into new research fields [12-13].

In this chapter, a parameter-free algorithm (PFA) will be introduced, aiming to achieve smart iterative WFS using a digital micromirror device (DMD). The PFA is a combination of the BA and the GA, integrating the merits of rapid convergence of the BA with the high robustness of the GA. The framework of the PFA inherits the framework of the GA. The solution of the PFA to find the optimum wavefronts in each iteration is similar to that of the BA. A dynamic mutation method is used to automatically obtain the mutation rate (MR) [14]. The only parameter that needs to be preset is the number of bats, which are the candidate solutions in each iteration. That said, through the experiment, it is found that the variation in the number of bats has a trivial influence on the performance of the PFA, which is systematically better than

that of the BA and the GA. As a result, using our method in iterative WFS, one can achieve a better enhancement ratio without a parameter tuning process. The PFA can be further used in other research fields in which iterative WFS can contribute, such as controlling multimode interactions in optical fibers [15], manipulating multi-dimensional characteristics of the fiber laser [16], and focusing optical laser pulses through scattering media [17].

2.2 Methods

In iterative WFS, the resultant optical focus within or through scattering media is the result of the superposition of all output channels from corresponding control units on a spatial light modulator, such as a DMD [18]. The theoretical value of the enhancement ratio (also referred to as the peak-to-background intensity ratio) is achieved when every output channel contributes positively to the optical focus. The most important step in the PFA is to calculate the MR in each measurement, since it is in direct proportion to the number of units negatively contributing to the optical focus. In each measurement, the key is to adjust the units which negatively contribute to the enhancement ratio of the optical focus.

To address this question, a transmission matrix model is introduced, where a transmission

matrix T is used to bridge the relationship between the input field E and the output field U [19]:

$$U = [u_1 \cdots u_m \cdots u_M]^\dagger = T \cdot E$$

$$= \begin{bmatrix} t_{11} & \cdots & t_{1n} & \cdots & t_{1N} \\ \vdots & \ddots & & & \vdots \\ t_{m1} & & t_{mn} & & t_{mN} \\ \vdots & & & \ddots & \vdots \\ t_{M1} & \cdots & t_{Mn} & \cdots & t_{MN} \end{bmatrix} \cdot [d_1 \cdots d_n \cdots d_N]^\dagger e \quad (2\text{-}1)$$

$$d_i = \begin{cases} 0 \\ 1 \end{cases} for \text{ DMD}, \quad (2\text{-}2)$$

where $U = [u_1 \cdots u_m \cdots u_M]^\dagger$ contains M output channels, and u_m is the m^{th} output channel on the output plane. $D = [d_1 \cdots d_n \cdots d_N]^\dagger$ is a binary wavefront matrix on the DMD, which contains N control units. The character e represents the electric field of a uniform plane wave projected onto the DMD. $E = [d_1 \cdots d_n \cdots d_N]^\dagger e$ is the input field, which is reflected from the DMD.

For a randomly picked output channel u_m serving as the target output channel to generate an optical focus, it is the product of m^{th} row T_m in the transmission matrix T and the wavefront matrix D. Here we realign T_m into two parts. The first K^{th} units ($K \leq N$) are correctly set, and the remaining units are incorrectly set:

$$u_m = T_m De = \left(\begin{array}{c} \underbrace{t_{m,1}d_{m,1} + \cdots + t_{m,K}d_{m,K}}_{correct\ \text{units}} + \\ \underbrace{t_{m,K+1}d_{m,K+1} + \cdots + t_{m,N}d_{m,N}}_{wrong\ units} \end{array} \right) e.$$

$$u_m = T_m De = \left(\begin{array}{c} \underbrace{t_{m,1}d_{m,1} + \cdots + t_{m,K}d_{m,K}}_{correct\ units} + \\ \underbrace{t_{m,K+1}d_{m,K+1} + \cdots + t_{m,N}d_{m,N}}_{wrong\ units} \end{array} \right) e. \quad (2\text{-}3)$$

According to Ref. [20], u_m can be calculated as

$$|u_m| = (2c - 1) \cdot N \cdot \frac{\sigma}{\sqrt{2\pi}} \cdot |e|, \quad (2\text{-}4)$$

where σ is the standard deviation of the distribution of the real part of T, and $c = K/N$ is the ratio of correct units, and it should be larger than 0.5, as only less than half of the units will be changed in each measurement to increase the robustness of the algorithm.

The maximum value of $|u_m|$ can be achieved when all units on the DMD are correctly adjusted: $|u_{m\_max}| = |u_m||_{r=1} = N \cdot \frac{\sigma}{\sqrt{2\pi}} \cdot |e|$. Thus, the focusing efficiency can be expressed by

$$\eta = \frac{I_m}{I_{m\_max}} = \frac{|u_m|^2}{|u_{m\_max}|^2} = (2c - 1)^2, \quad (2\text{-}5)$$

where I_m is the intensity of the instantaneous focus, and I_{m_max} is the maximum focal intensity that can be achieved [20]. Thus, c can be obtained through η directly:

$$c = \frac{1 + \sqrt{\eta}}{2} \quad (2\text{-}6)$$

In this chapter, the dynamic mutation rate of the newly generated bats P is related to c and the bat number h, which is defined as

$$P = \frac{1-c}{h} = \frac{1-\sqrt{\eta}}{2h}, \tag{2-7}$$

where h is the predetermined number of bats in the PFA. During each iteration, the newly generated bats will be dynamically mutated according to the dynamic mutation rate P. To be specific, we should randomly select $\lceil P \times N \rceil$ units on the newly generated pattern and revise them ($\lceil\ \rceil$ represents rounding up).

The dynamic mutation rate P is then applied in the PFA, which is a combination of the GA and BA, to achieve reinforced performance [21]: the PFA inherits the MR of the GA, and its strategy to search for the optimum solution is similar to that of the BA. To start with, the flow charts of the BA and GA are shown in Figure 2-1a and Figure 2-1b, respectively. The BA is a heuristic optimization algorithm inspired by the echolocation behavior of bats. A population of bats (*i.e.*, h bats) fly randomly with different velocities v, positions x, and frequencies Q to search for prey. Besides, when approaching the prey, bats tend to decrease the loudness L and increase the pulse emission rate PER of ultrasound waves, which can be referred to as the traditional BA [22]. In each iteration (one iteration contains h measurements), new solutions (*i.e.*, DMD patterns) are generated by flying around the best-performing bat or just flying

randomly, according to PER of the bat. New solutions are evaluated according to the feedback function (*i.e.*, enhancement ratio). Then better-performing solutions will be adopted. After many iterations, the optimal solution can be achieved. As for the GA, the new offsprings are the cross-fertilization of their parents with a mutation rate [23].

For the PFA, we combine the flying around the best-performing bat in the BA and the mutation rate of the GA, as shown in Figure 2-1c. In each measurement, the current bat is cross-fertilized with the best-performing bat and a random binary template to generate a new bat (cross-fertilization rate = 0.5). After cross-fertilization, the dynamic mutation rate P of the current bat is calculated according to Equation (2-7), and some pixels of the new solution are mutated based on the dynamic mutation rate P. If the new bat performs better than the current one, the current bat will be replaced by the new one. An optimal solution will be achieved after many iterations.

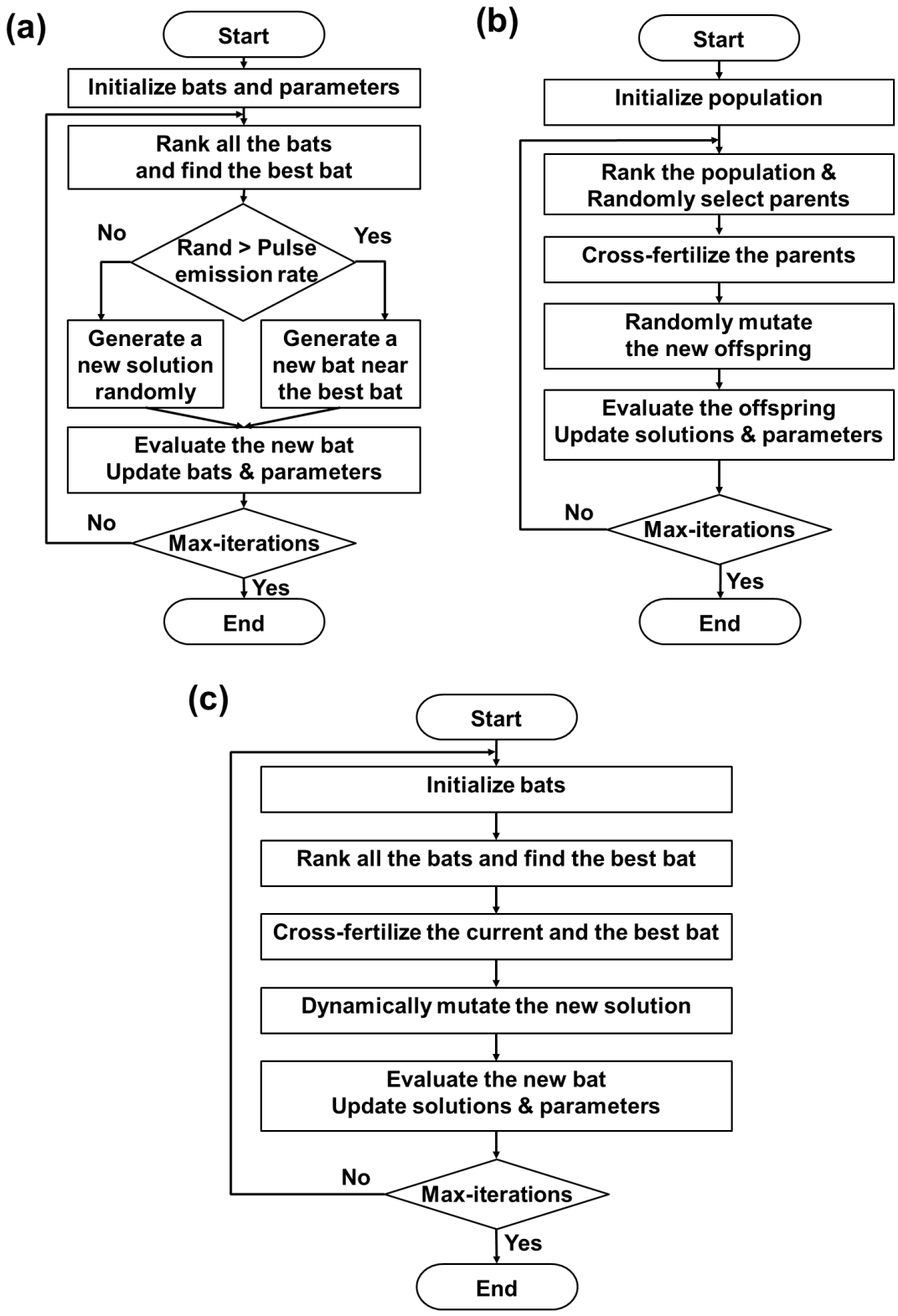

Figure 2-1 Flow charts of different optimization algorithms used in iterative WFS: (a) BA, (b) GA, and (c) PFA.

2.3 Results

2.3.1 Simulations

The enhancement ratios of optical foci achieved by three algorithms (PFA, BA, and GA) in iterative WFS are then numerically simulated with parameters listed in Table 2-1. The control unit number on the DMD is 1024, and the transmission matrix output mode number is 400 in simulation. As shown in Figure 2-2a, there is a big difference between the enhancement ratios for BAs with well-tuned and improper parameters, even though only one parameter (*i.e.*, PER) is different, highlighting the significance and necessity of having appropriate parameters in these iterative WFS methods. Moreover, the enhancement ratio using the PFA is comparatively higher than that using the BA or the GA of equal bats or offsprings. In the PFA, the only parameter that needs to be set is the number of bats. To further test how this number affects the performance, simulations using different numbers of bats for the PFA are conducted. The results, as shown in Figure 2-2b, exhibit weak correlations with the bat numbers in a large range of 20 to 100. This, again, suggests that the proposed algorithm is really smart and free from any prior knowledge about the parameters.

Table 2-1 Parameters used in simulations for different algorithms.

Algorithm	PFA	BA with well-tuned parameters	BA with improper parameters	GA with well-tuned parameters
Parameters used in simulations	N/A	$Q_{Min} = 0.08$ $Q_{Max} = 0.35$ PER = 0.035 L = 0.99	$Q_{Min} = 0.08$ $Q_{Max} = 0.35$ PER = 0. 35 L = 0.99	$MR_{Initial} = 0.022$ $MR_{Final} = 0.008$ DF = 1000

Q_{Min}: Minimum frequency; Q_{Max}: Maximum frequency; PER: Pulse emission rate; L: Loudness; $MR_{Initial}$: Initial mutation rate; MR_{Final}: Final mutation rate; DF: Decay factor of the mutation rate. N/A means that no parameter is needed to be preset. The numbers of bats in PFA and BA are 40. The numbers of population and offsprings in GA are 40.

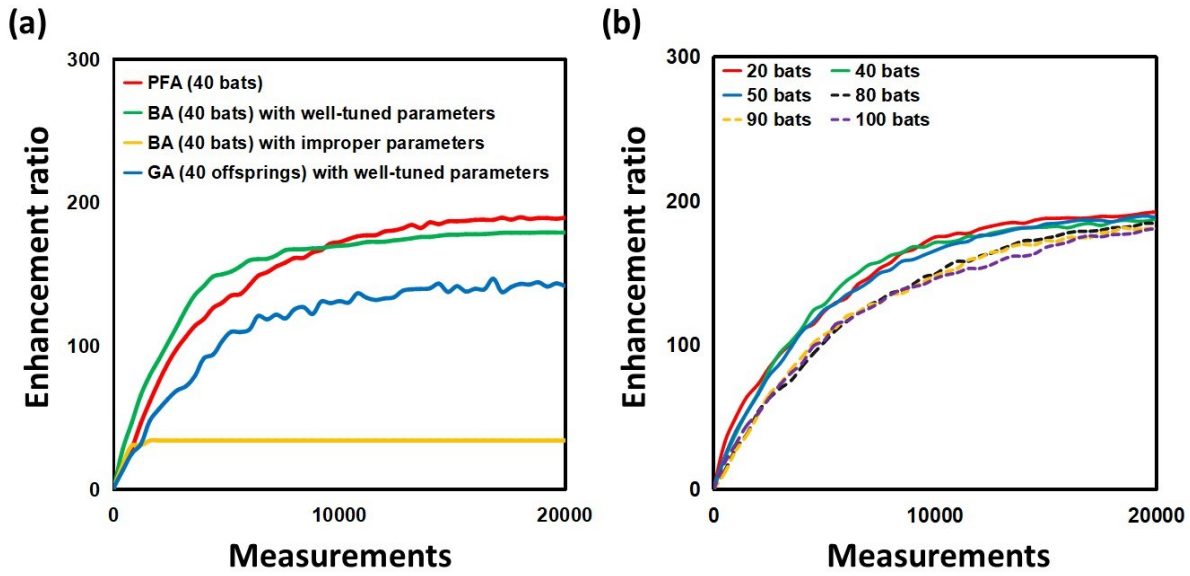

Figure 2-2 Simulation results: (a) different algorithms with 40 bats or offsprings. For a regular BA and GA, well-tuned parameters are adopted as provided in Table 2-1. (b) Results of PFAs with 20, 40, 50, 80, 90, and 100 bats.

2.3.2 Experiments

After simulation, we built an optical setup to assess the PFA, which is depicted in Figure 2-3. A continuous-wave laser source (EXLSR-532-300-CDRH, Spectra Physics, U.S.) with a wavelength of 532 nm and maximum energy of 300 mW is used as the light source. The laser output is expanded by a pair of convex lenses (L1 and L2) to illuminate the DMD (DLP4100, Texas Instruments Inc., U.S.). The DMD serves as a binary amplitude spatial light modulator, with 1024 independent units being utilized in experiments. After being modulated and reflected by the DMD, the light is shrunk by another pair of lenses (L3 and L4), and then focused onto a ground glass (DG10-220-MD, Thorlabs, U.S., diameter of 2.54 cm, 220 grits) by using an objective lens (UIS 2 Plan N 10×/0.25, Olympus, Japan). The resultant speckles from the ground glass are recorded by a CMOS camera (Blackfly S BFS-U3-04S2M-CS, FLIR, Canada). Each measurement took about 100 ms in experiments.

The increases of the focal intensity enhancement ratio with the number of iterations are shown in Figure 2-4a. The parameters used for different algorithms in experiments are listed in Table 2-2. Very different performances are observed for the BA with well-tuned and improper parameters, which agrees well with simulations. The maximum enhancement ratio obtained with the PFA is 121, which is ~25% and ~20% higher than that of the GA and the BA,

respectively. Additionally, Figure 2-4b shows the normalized speckles after optimization in experiments. As seen, the focus achieved by the PFA has the highest peak intensity due to the best optimization performance.

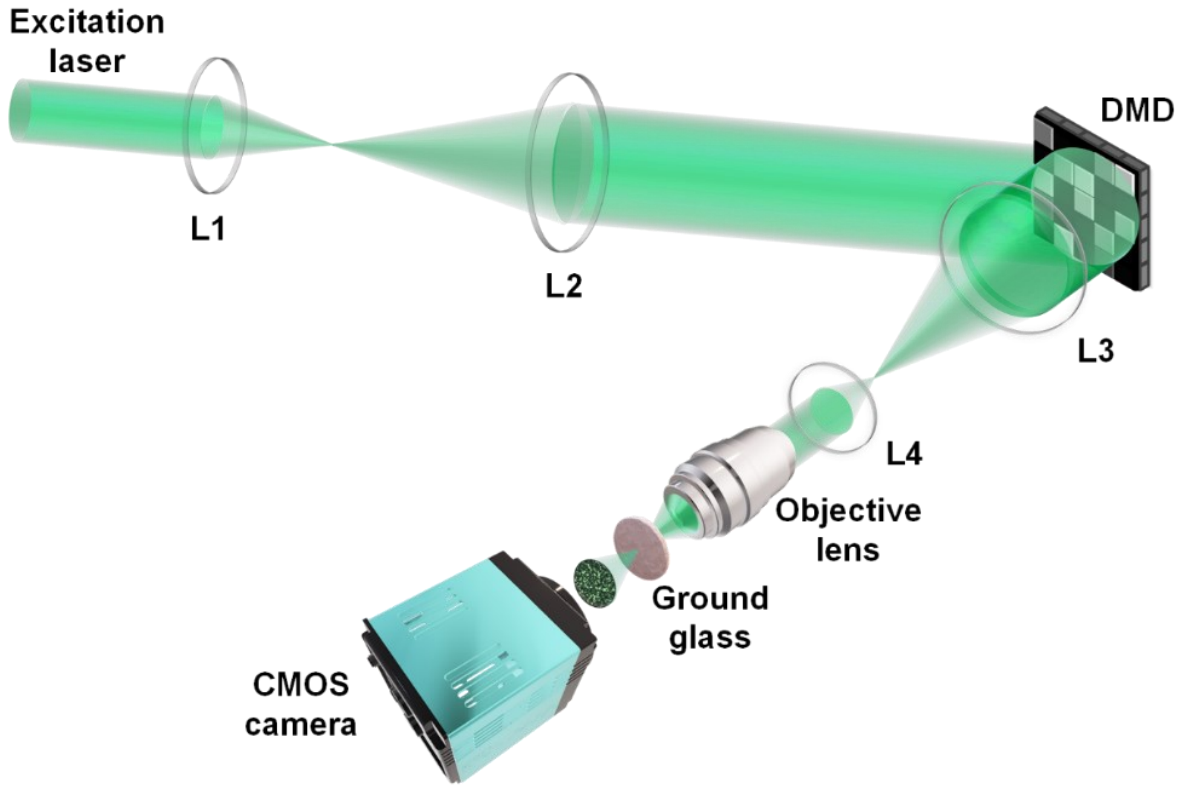

Figure 2-3 Experimental setup. Laser source: 532 nm laser; L1, f = 60 mm; L2 and L3, f = 250 mm; L4, f = 50 mm. DMD, digital micromirror device.

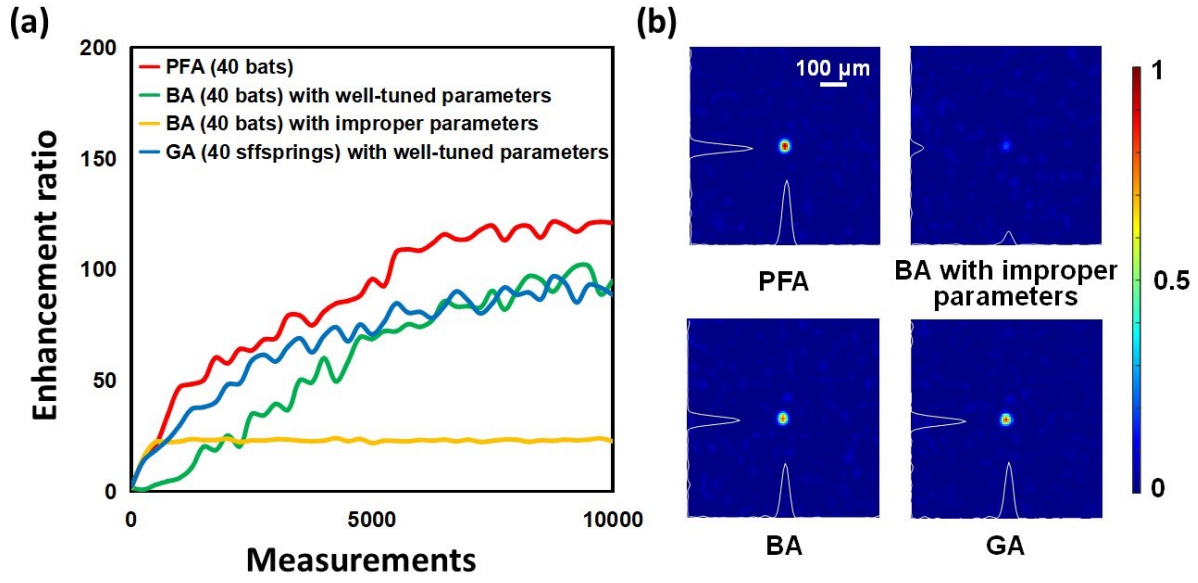

Figure 2-4 Experimental results for different algorithms. (a) Different algorithms with 40 bats or offsprings. (b) Focal speckles after optimization.

Table 2-2 Parameters used in experiments for different algorithms.

Algorithm	PFA	BA with well-tuned parameters	BA with improper parameters	GA with well-tuned parameters
Parameters used in experiments	N/A	Q_{Min}=0.08 Q_{Max}=0.35 PER=1 L=0.99	Q_{Min}=0.08 Q_{Max}=0.35 PER=0. 5 L=0.99	$MR_{Initial}$=0.008 MR_{Final}=0.002 DF=2000

Q_{Min}: Minimum frequency; Q_{Max}: Maximum frequency; PER: Pulse emission rate; L: Loudness; $MR_{Initial}$: Initial mutation rate; MR_{Final}: Final mutation rate; DF: Decay factor of the mutation rate. N/A means that no parameter is needed to be preset. The numbers of bats in PFA and BA are 40. The numbers of population and offsprings in GA are 40.

To verify the universality of the PFA under different conditions, such as different scattering media, we utilized a multimode fiber (MMF, SUH200, Xinrui, China, core diameter = 200μm, NA = 0.22, length = 0.5m) as the scattering medium for comparison. The results of the ground glass and the MMF are shown in Figure 2-5. As seen, although the absolute enhancement ratios differ between the ground glass and the MMF experiments, the independence of performance on the number of bats is consistent.

These results confirm that the proposed parameter-free algorithm can be used to obtain superior performance without priori knowledge or parameter tuning for iterative WFS. In general, to get a satisfying enhancement ratio in iterative WFS, a careful parameter tuning procedure should be applied to fit for diverse experiment conditions. With the PFA, researchers can be set free from this time-consuming procedure. Even a beginner can steer iterative WFS optimization that outperforms the GA and BA.

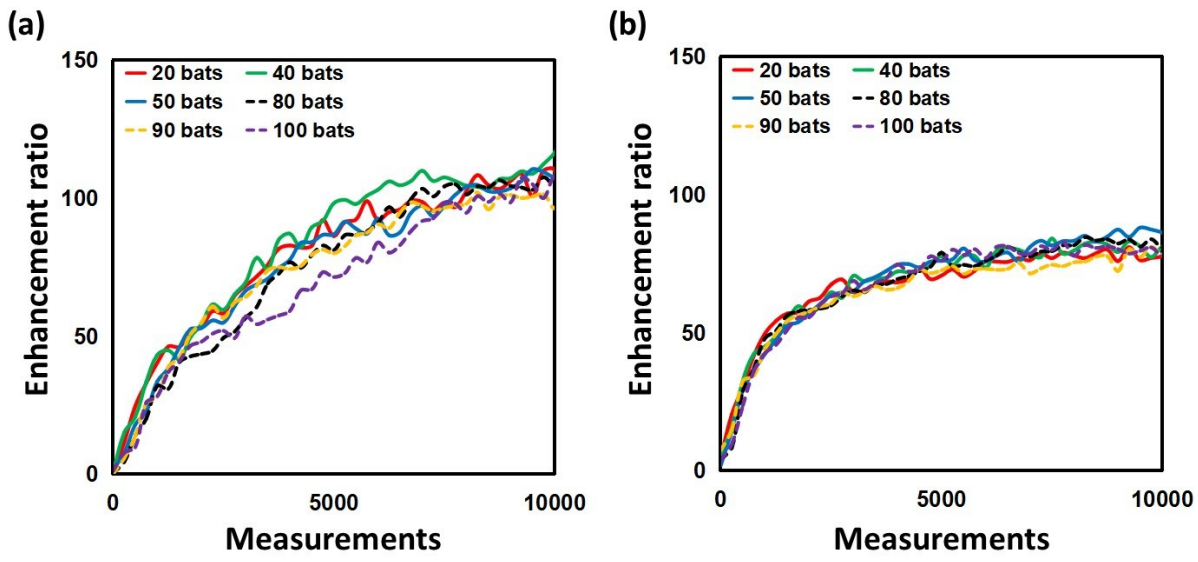

Figure 2-5 Experimental results of PFAs with 20, 40, 50, 80, 90, and 100 bats for (a) ground glass and (b) MMF.

2.4 Discussions

At last, the performance of the PFA in the existence of noises and solutions to further improve the enhancement ratio is discussed. The PFA is derived from the dynamic mutation algorithm, which has demonstrated high adaptability against perturbations [20], and the PFA inherits this feature. A field programmable gate array-based iterative WFS system can be implemented to achieve real-time optical focus and overcome environmental noises [24]. Moreover, this fast iterative WFS version allows more input units to be used to improve the enhancement ratio in a short duration [25]. Further, the super-pixel encoding method can be implemented to achieve

phase modulation and improve the focusing efficiency [26]. With that, η in Equation (2-5), Equation (2-6), and Equation (2-7) should be adjusted to the ratio between the instant enhancement ratio and the theoretical enhancement ratio (*i.e.*, $N \times \pi/4$ for phase modulation) [18].

2.5 Conclusion

In conclusion, this chapter introduces a Parameter-free Algorithm (PFA) for iterative wavefront shaping that draws inspiration from both the Bat Algorithm (BA) and the Genetic Algorithm (GA). A standout feature of PFA is its ability to autonomously calculate the mutation rate via real-time feedback, obviating the need for the laborious and expertise-reliant parameter tuning process that plagues existing iterative algorithms. This innovation simplifies the optimization process and holds significant promise for researchers, particularly those new to the field. Furthermore, PFA demonstrates superior performance over traditional algorithms, and PFA's versatility is underscored by its potential applicability, including optical focusing and imaging through or within scattering media.

References

1. Yun, S. H., & Kwok, S. J. (2017). Light in diagnosis, therapy and surgery. *Nature biomedical engineering, 1*(1), 0008.

2. Wiersma, D. S. (2013). *Disordered photonics. Nature Photonics, 7*(3), 188-196.

3. Cheng, S., Li, H., Luo, Y., Zheng, Y., & Lai, P. (2019). Artificial intelligence-assisted light control and computational imaging through scattering media. *Journal of innovative optical health sciences, 12*(04), 1930006.

4. Horstmeyer, R., Ruan, H., & Yang, C. (2015). Guidestar-assisted wavefront-shaping methods for focusing light into biological tissue. *Nature photonics, 9*(9), 563-571.

5. Vellekoop, I. M., & Mosk, A. P. (2007). Focusing coherent light through opaque strongly scattering media. *Optics letters, 32*(16), 2309-2311.

6. Yoon, S., Kim, M., Jang, M., Choi, Y., Choi, W., Kang, S., & Choi, W. (2020). Deep optical imaging within complex scattering media. *Nature Reviews Physics, 2*(3), 141-158.

7. Conkey, D. B., Brown, A. N., Caravaca-Aguirre, A. M., & Piestun, R. (2012). Genetic algorithm optimization for focusing through turbid media in noisy environments. *Optics express, 20*(5), 4840-4849.

8. Fayyaz, Z., Mohammadian, N., Salimi, F., Fatima, A., Tabar, M. R. R., & Avanaki, M. R. (2018). Simulated annealing optimization in wavefront shaping controlled transmission. *Applied optics, 57*(21), 6233-6242.

9. Li, B. Q., Zhang, B., Feng, Q., Cheng, X. M., Ding, Y. C., & Liu, Q. (2018). Shaping the wavefront of incident light with a strong robustness particle swarm optimization algorithm. *Chinese Physics Letters, 35*(12), 124201.

10. Wang, Z. Q., Zhao, Q., Yu, P. P., Yang, J. M., Li, Y. M., & Gong, L. (2019). Bat algorithm-enabled binary optimization for scattered light focusing. *Applied Physics Express, 12*(10), 102002.

11. Lai, P., Wang, L., Tay, J. W., & Wang, L. V. (2015). Photoacoustically guided wavefront shaping for enhanced optical focusing in scattering media. *Nature photonics, 9*(2), 126-132.

12. Wu, D., Luo, J., Li, Z., & Shen, Y. (2019). A thorough study on genetic algorithms in feedback-based wavefront shaping. *Journal of Innovative Optical Health Sciences, 12*(04), 1942004.

13. Fayyaz, Z., Mohammadian, N., Reza Rahimi Tabar, M., Manwar, R., & Avanaki, K. (2019). A comparative study of optimization algorithms for wavefront shaping. *Journal of innovative optical health sciences, 12*(04), 1942002.

14. Eiben, Á. E., Hinterding, R., & Michalewicz, Z. (1999). Parameter control in evolutionary algorithms. *IEEE Transactions on evolutionary computation, 3*(2), 124-141.

15. Tzang, O., Caravaca-Aguirre, A. M., Wagner, K., & Piestun, R. (2018). Adaptive wavefront shaping for controlling nonlinear multimode interactions in optical fibres. *Nature Photonics, 12*(6), 368-374.

16. Wei, X., Jing, J. C., Shen, Y., & Wang, L. V. (2020). Harnessing a multi-dimensional fibre laser using genetic wavefront shaping. *Light: Science & Applications, 9*(1), 149.

17. Katz, O., Small, E., Bromberg, Y., & Silberberg, Y. (2011). Focusing and compression of ultrashort pulses through scattering media. *Nature photonics, 5*(6), 372-377.

18. Vellekoop, I. M. (2015). Feedback-based wavefront shaping. *Optics express, 23*(9), 12189-12206.

19. Popoff, S. M., Lerosey, G., Fink, M., Boccara, A. C., & Gigan, S. (2011). Controlling light through optical disordered media: transmission matrix approach. *New Journal of Physics, 13*(12), 123021.

20. Li, H., Woo, C. M., Zhong, T., Yu, Z., Luo, Y., Zheng, Y., ... & Lai, P. (2021). Adaptive optical focusing through perturbed scattering media with dynamic mutation algorithm. *Photonics Research, 9*(2), 202-212.

21. Luo, Y., Yan, S., Li, H., Lai, P., & Zheng, Y. (2020). Focusing light through scattering media by reinforced hybrid algorithms. *APL photonics, 5*(1).

22. Yang, X. S. (2010). A new metaheuristic bat-inspired algorithm. In *Nature Inspired Cooperative Strategies for Optimization (NICSO, 2010)* (pp. 65-74). Berlin, Heidelberg: Springer Berlin Heidelberg.

23. Varnosfaderani, M. H. H., Mozaffarzadeh, M., UpputurI, P. K., & Pramanik, M. (2019, February). Genetic algorithm for feedback-based wavefront shaping in optical imaging. In *Photons Plus Ultrasound: Imaging and Sensing 2019* (Vol. 10878, pp. 582-587). SPIE.

24. Tzang, O., Niv, E., Singh, S., Labouesse, S., Myatt, G., & Piestun, R. (2019). Wavefront shaping in complex media with a 350 kHz modulator via a 1D-to-2D transform. *Nature Photonics, 13*(11), 788-793.

25. Yu, H., Lee, K., & Park, Y. (2017). Ultrahigh enhancement of light focusing through disordered media controlled by mega-pixel modes. *Optics express, 25*(7), 8036-8047.

26. Goorden, S. A., Bertolotti, J., & Mosk, A. P. (2014). Superpixel-based spatial amplitude and phase modulation using a digital micromirror device. *Optics express, 22*(15), 17999-18009.

3 SPATIOTEMPORALLY DECORRELATED SPECKLES

This chapter is modified based on the following manuscript that is currently under review:

Zhao, Q.[†], Li, H.[†], Zhong, T.[†], Cheng, S., Huang, H., Yao, J., Li, W., Li, H., Woo, C. M., Gong, L., Zheng, Y.[#], Yu, Z.[#], & Lai, P.[#] (2023). Extended learning generalizability for high-fidelity human face imaging from spatiotemporally decorrelated speckles. *Under review*.

Computational imaging through scattering media has long been a coveted yet challenging pursuit. Researchers have made strides in extracting target information from speckles, primarily through iterative wavefront shaping (as detailed in Chapter 2), calibrating the transmission matrix of the scattering medium, or employing neural networks. These methods effectively quantify the relationship between the targets and the corresponding speckles. However, the fidelity of the retrieved images is significantly compromised when the medium's status changes due to intrinsic motion or external perturbations. Such variability leads to a decorrelation between the data used for training and that used for validation, which has impeded the practical

applications of these frameworks.

In this chapter, we introduce a generative adversarial network (GAN)-based framework with enhanced generalizability, designed to address the spatiotemporal instabilities of scattering media and the resultant decorrelation between training and testing datasets. Experiments demonstrate that the proposed GAN framework can be trained to retrieve face images from speckles, even when the scattering medium has undergone decorrelation to unknown statuses after network training. Notably, the proposed GAN outperforms existing learning-based methods by non-holographically retrieving images from unstable scattering media and effectively managing speckle decorrelation, even after the optical system has been inactive for an extended period (up to 37 hours in experiments) and subsequently reactivated. This remarkable capability paves the way for broad applications where networks can be pre-trained and maintain their effectiveness for data acquired at a later time. Such resilience is pivotal for broadening the scope of learning-based methodologies in speckle imaging, encompassing applications like imaging through deep tissues and target sensing under extreme environmental conditions.

3.1 Introduction

Optical imaging stands as a cornerstone in the exploration of the microscopic and macroscopic realms, ranging from cellular structures to whole bodies, and this field has continually evolved, offering scalable resolution and expansive fields of view [1-2]. In environments like biological tissues or analogous media such as fog, dusty air, and turbid water, scattering predominantly governs light-matter interactions, curtailing imaging depth and compromising resolution. Coherent light exacerbates this challenge, producing speckles with stark contrasts that obscure clear imaging and interpretation [3]. Seminal efforts [4-6], particularly the transmission matrix approach, have addressed static scenarios by linearly correlating input and output fields [7]. While this enables computational retrieval of images from speckles via the inverse transmission matrix, the linear model's limitations often yield suboptimal imaging quality due to the insufficient accuracy of the linear TM model [8].

The advent of deep learning has catalyzed a paradigm shift, with deep neural networks (DNNs) significantly refining optical system performance. These networks account for the inherent nonlinearity in light-matter interactions and perturbations [9-15], thereby elevating imaging fidelity [16-19].

Despite their success in stationary scattering environments [16], DNNs face challenges when the scattering medium or optical system is non-stationary due to motion, perturbations, or vibrations. Such non-stationarity leads to speckle decorrelation, undermining modulation or imaging performance [20-23]. Statistic invariants of speckles are therefore explored among a set of diffusers with the same macroscopic parameters or a multimode fiber under different configurations [24-27]. While these studies show promise, thus far the DNNs primarily operate with simple objects (digits, letters, and binary patterns) trained on extensive datasets reflecting different statuses of the scattering medium. Crucially, these datasets are often collected concurrently, resulting in a temporal overlap that maintains the correlation between training and testing data. This correlation allows for low generalizability to unknown scattering medium statuses.

In practical scenarios, however, significant time lapses between training and testing are common, necessitating the use of data collected under altered medium conditions. This temporal gap introduces decorrelation, diminishing the pre-trained networks' efficacy. Despite training on diverse data representing various medium statuses, DNN performance degrades when confronted with changed conditions. Thus, in real-world applications where the scattering medium deviates from previously encountered statuses, the pre-trained networks struggle to

adapt to the decorrelated speckles.

In this chapter, we introduce a generative adversarial network (GAN) framework capable of training to recover complex face images from a dynamic scattering medium. The framework employs a U-Net-based generator to discern intrinsic speckle features, enabling the high-fidelity retrieval of images. Concurrently, a multi-layer convolution-based discriminator assesses the retrieved images, providing critical feedback to refine the generator's outputs [28-30]. Notably, the GAN's generator harnesses speckle features to retrieve images, while the discriminator guides the generator in enhancing the retrieved images.

Experimental findings affirm that the GAN retains its ability to retrieve face images from speckles, even when the intervals between training and testing datasets are extended, resulting in minimal correlation. This evidences the GAN's remarkable temporal generalizability. The framework's robustness is further corroborated through experiments utilizing a non-stationary metasurface as the scattering medium. Here, the network is initially trained with collected data, after which the optical system is deactivated for 37 hours (or longer if necessary) and subsequently reactivated to gather testing data, thereby simulating real-world conditions. To our knowledge, this represents the inaugural instance of a neural network being trained to retrieve high-fidelity face images from speckles gathered on different days from training data.

In summation, the proposed GAN showcases exceptional temporal generalizability, surmounting the challenges associated with imaging through a non-stationary scattering medium of unknown statuses. This opens a door to opportunities for non-holographically retrieving intricate images from decorrelated speckles, indicating a significant advancement in the field.

3.2 Methods

3.2.1 Optical Setups

The optical system for acquiring speckles is depicted in Figure 3-1a. The light source is a continuous-wave 532-nm laser (EXLSR-532-300-CDRH, Single mode, 300 mW, Spectra-Physics, USA). The laser beam from the light source is first expanded by a 4-f system (L1 and L2) to cover the entire aperture of the spatial light modulator (SLM, HOLOEYE PLUTO VIS056 1080p, German). Then the image information displayed on the SLM modulates the wavefront. The information to modulate the input wavefront is the face image from thumbnails in the Flickr Faces High Quality (FFHQ) database [31]. Here, the 128×128 thumbnails are up-sampled to 640×640 and displayed at the center of the SLM. After SLM modulation, the

modulated light is shrunk by another 4-f system (L3 and L4) and then focused by an objective lens (RMS20X, Olympus, Japan). The focused laser travels through a scattering medium, transforming into random optical speckles, which are captured by a CMOS camera (FL3-U3-32S2M-CS, PointGrey, Canada). The scattering medium used in experiments includes a ground glass diffuser and a disordered metasurface. During the data collection, ambient perturbations (*e.g.*, moving people, other on-going experiments on the same optical table, running scientific instruments in other rooms, *etc*.) could mechanically shift the scattering medium, and laser coherence depends on the working stability of the laser (not intentionally controlled), leading to decorrelated speckles.

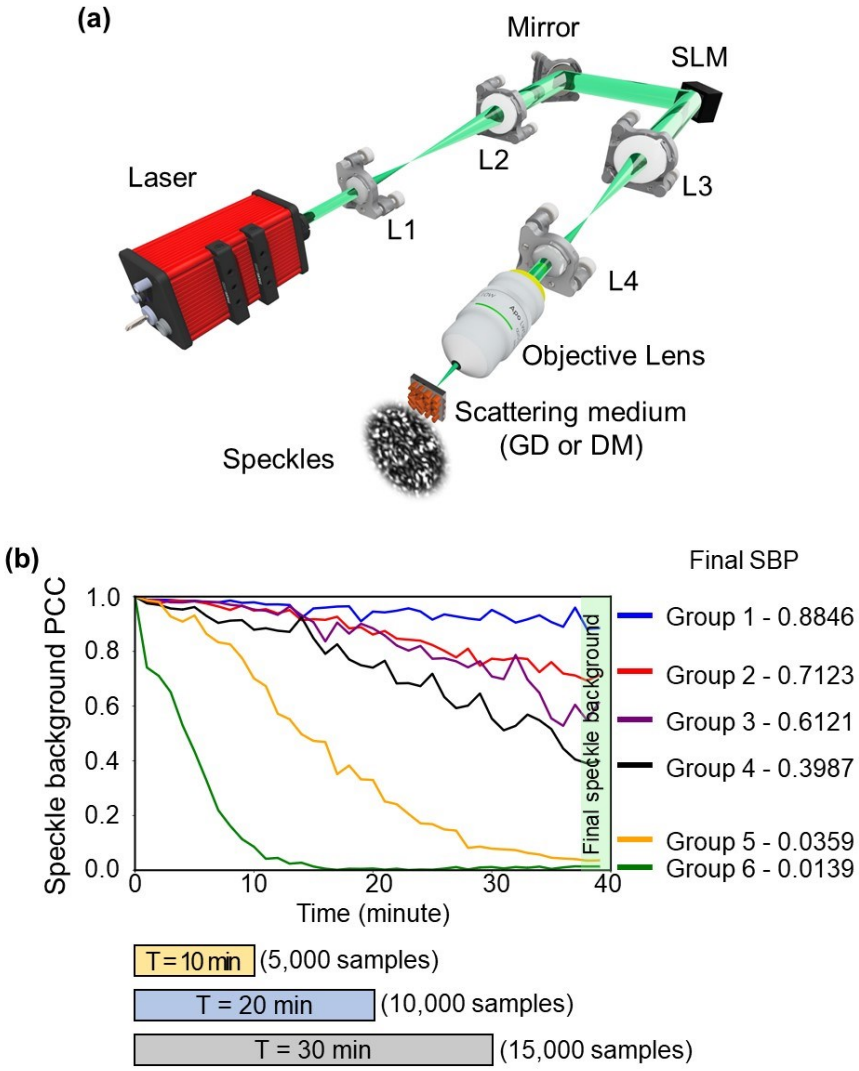

Figure 3-1 (a) Diagram of the optical setup for acquiring speckles. L1 and L2: the first 4-f system to expand the laser beam. SLM: spatial light modulator. L3 and L4: the second 4-f system to shrink the laser beam. Scattering medium: ground glass (GD) or disordered metasurface (DM). (b) Speckle background PCC (SBP) during six 40-min ground glass experiments. Lower SBP corresponds to a larger deviation from the initial status and lower stability. Final SBP is the SBP of each data group at the end (marked in green on the right Y axis).

3.2.2 Data Acquisition and Speckle Instability

Experimentally, to generate data at different levels of instability, a scattering medium (*i.e.*, ground glass) is constantly disturbed by surrounding perturbations, including air flow and platform vibrations, *etc*. The corresponding instability is characterized by variations of background speckles generated by loading a uniform phase pattern (all phase elements set to 2π on the SLM) every minute (*i.e.*, every 500 captured speckles).

In the ground glass experiments, the duration for the collection of each data group is up to 40 minutes. Every captured background speckle is compared with the initial background speckle for the calculation of Pearson correlation coefficient (PCC), which is termed the speckle background PCC (SBP). As shown in Figure 3-1b, the SBP of six groups of data continuously decays with time in experiments due to environmental perturbations (including air flow, platform vibrations, *etc*.), and a lower SBP corresponds to a higher deviation from the initial status and hence lower stability. With monotonic variations, we can mark SBP at the end of each group (*i.e.*, Final SBP) to distinguish different groups. For example, for Group 1, Final SBP is 0.8846 (the highest among six groups), indicating relatively stationary conditions; for Group 6, SBP drops down more quickly with Final SBP of $0.0139 < 1/e \approx 0.3678$, indicating speckles have become totally decorrelated and the dataset includes significantly perturbed

information [18,25]. In the following experiments, these six groups of datasets are utilized to train and test the proposed GAN framework.

3.2.3 Neural Networks

The proposed GAN includes a generator and a discriminator, as shown in Figure 3-2a, and the related Python code is available on GitHub (https://github.com/863zq/863zq.github.io/blob/main/Code/Main_complex_GAN.py).

The generator is based on U-Net, which is extensively employed in speckle imaging [5,13,19,23]. Compared with traditional U-Net [29], the primary difference is that convolution layers here are all based on complex algebra, *i.e.*, inputs, outputs, and parameters in convolutions are all complex-valued, in order to more accurately mimic the random scattering process as modeled by the transmission matrix theory. Here, the input of the U-Net-based generator is the speckle with 128×128 pixels, which is transferred into the complex domain with zero phase. Then, the input speckles can be processed by the generator's encoders (four down-sampling paths, blue filters in Figure 3-2b) and decoders (four up-sampling paths, red filters in Figure 3-2b). After that, the final layer converts the complex feature map into real

numbers and outputs the retrieved images with 64×64 pixels.

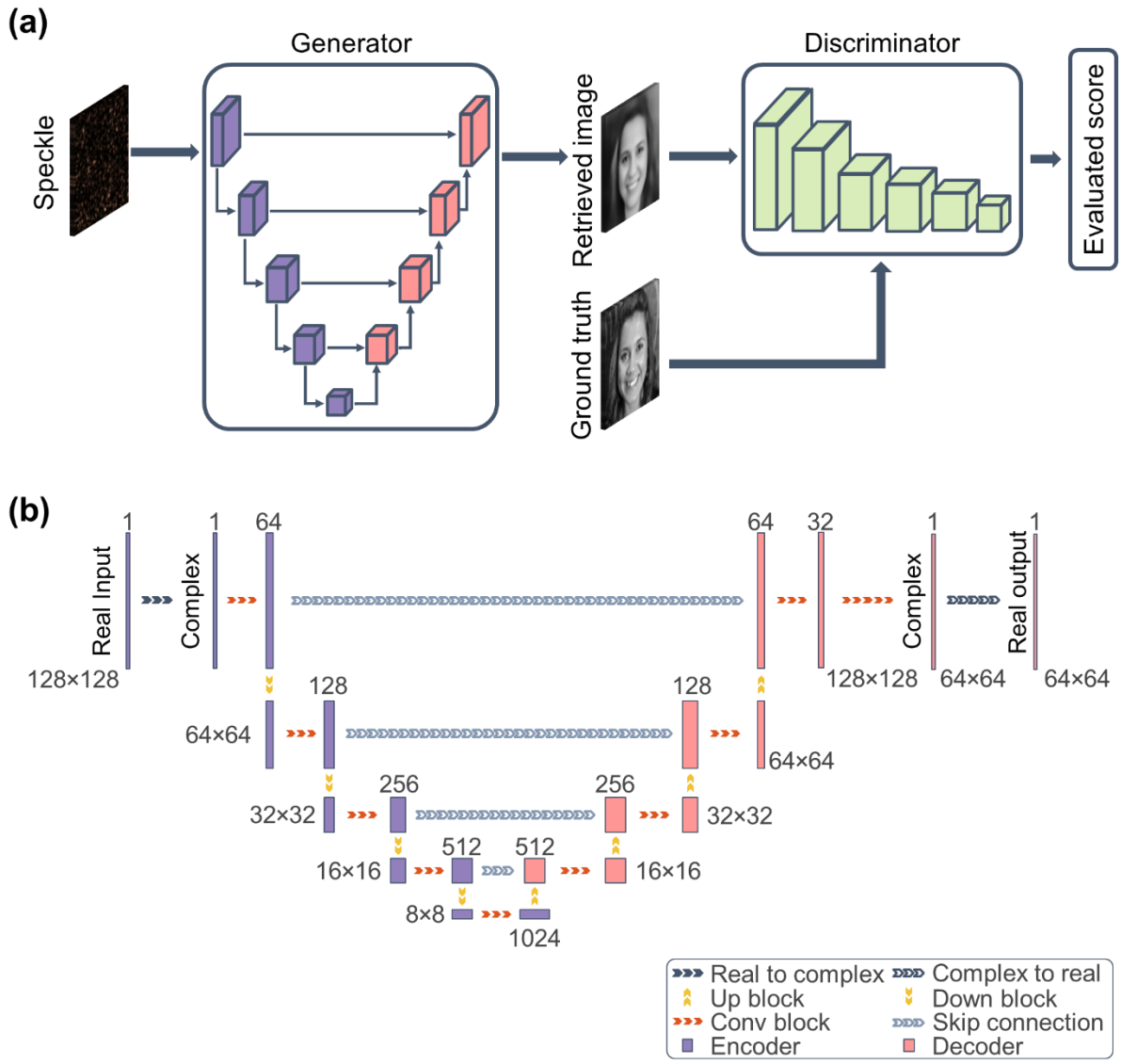

Figure 3-2 Schematic of the proposed GAN framework. (a) GAN structure: the generator is based on U-Net, with speckle as the input and retrieved images as the output; the discriminator is based on six convolutional layers and one linear layer, with retrieved images or ground truth images as input and evaluated loss as output. Ground truth image:

Copyright 2010, appnight-122, by Existence Church, Flickr (https://www.flickr.com/photos/sandiegochurch/4379311601/); the original images are converted to grayscale, under terms of the CC-BY 2.0 license. (b) Detailed structures of the generator: the encoders are highlighted in blue, and the decoders are highlighted in red. The dimensions of the feature maps are specified next to each block.

In addition to the generator, a discriminator based on six convolutional layers is designed to evaluate the retrieved images from the generator, as shown in Figure 3-2a. With the retrieved image or the ground truth as the input, the discriminator's output is the evaluated loss to discriminate between the retrieved images and the corresponding ground truth images, and the generator tries to cheat the discriminator so that the retrieved image is refined to approach the corresponding ground truth image [30]. After network training, the so-called Nash equilibrium is eventually reached, and the generator then successfully retrieves images from speckles with high fidelity [29].

3.2.4 Network Training

The loss function of the generator L_G combines the image loss $L_{G,image}$ and the adversarial loss $L_{G,adv}$. Here, the image loss combines the mean square error MSE and the Pearson correlation coefficient PCC to reveal differences between the retrieved image $G(x)$ (*i.e.*, the generator output when x is the input speckle) and the ground truth y. The adversarial loss $L_{G,adv}$ is MSE between the discriminator output with the generator output as the input and the discriminator output with the corresponding ground truth as the input. Then, the final loss function for the generator L_G is a weighted average of the adversarial loss $L_{G,adv}$ and the image loss $L_{G,image}$. The weights in L_G here are empirically tuned to improve the performance of the generator:

$$L_{G,image} = MSE[G(x), y] - PCC[G(x), y] \tag{3-1}$$

$$L_{G,adv} = MSE\big[D\big(G(x)\big), D(y)\big] \tag{3-2}$$

$$L_G = 0.2 \times L_{G,adv} + 0.8 \times L_{G,image} \tag{3-3}$$

$$MSE(x, y) = \langle (x - y)^2 \rangle \tag{3-4}$$

$$PCC(x, y) = \frac{\langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle}{\sigma_x \sigma_y} \tag{3-5}$$

In Equation (3-5), $\langle\ \rangle$ and σ denote the average operation and standard deviation, respectively.

The proposed discriminator is designed to assess the generator output, and the discriminator output is expected to be image loss. The proposed discriminator is different from discriminators in traditional GAN to generate new images, which discriminate images and noises. In the proposed discriminator, the loss function L_D is MSE between the predicted output of the discriminator, *i.e.*, $D\big(G(x)\big)$, and the real image loss $L_{G,image}(G(x), y)$:

$$L_D = MSE\left(D\big(G(x)\big), L_{G,image}(G(x), y)\right) \tag{3-6}$$

During training and testing, each group (Group 1-6 in Figure 3-1b) contains 20,000 image-speckle pairs, and six neural networks are individually trained with only one group of data. Human face images are collected from thumbnails in the FFHQ dataset [31], from which 128×128 thumbnails are down-sampled to 64×64 as the ground truth images. The dimensions of the speckles fed into the generator are 128×128. During network training, GAN is trained for 20 epochs using Adam optimizers with batch size = 32, and the initial learning rate is 0.0001 with cosine annealing. Furthermore, in each training epoch (as illustrated in Figure 3-3), the discriminator is trained five times, while the generator is trained once to enhance the discrimination between retrieved images and ground truth images [32].

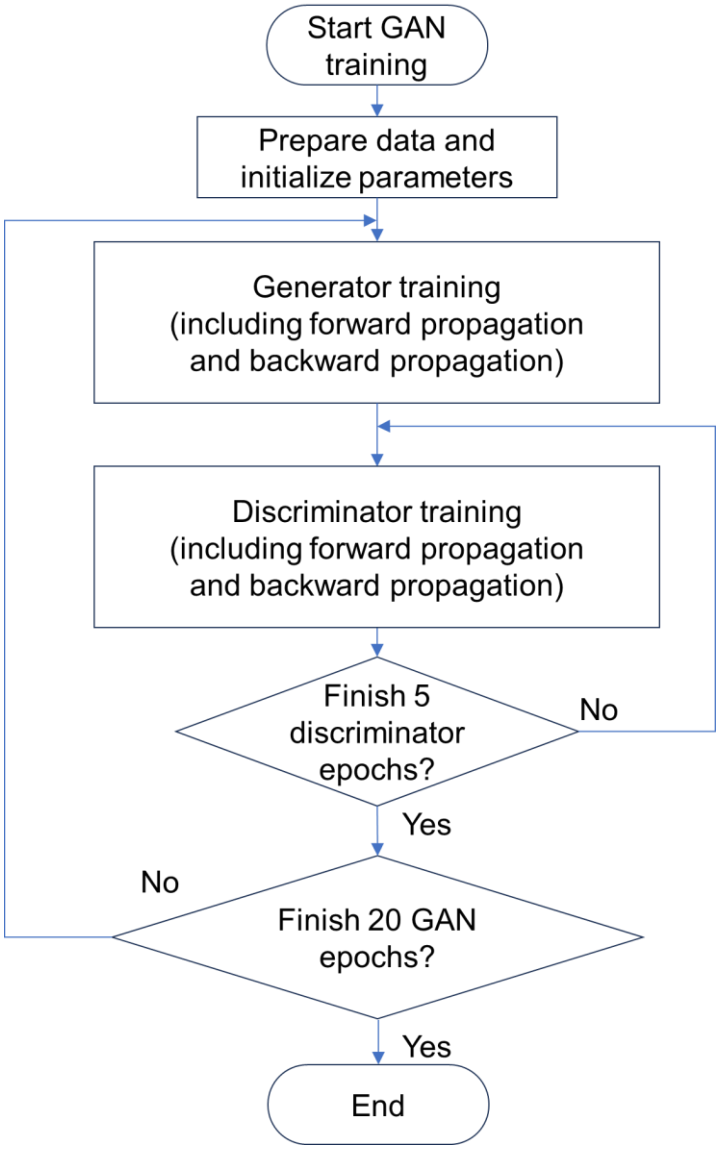

Figure 3-3 Flow chart of GAN training: during each GAN training epoch, the generator is trained only once, but the discriminator is trained five times to improve the convergence and network performance.

Notably, speckles used during network testing are sampled after the training dataset collection is finished. As for different experiments, the acquisition time intervals between training and

testing datasets vary, in order to test the network performance at different scattering medium statuses. As for the software framework, we utilize PyTorch 2.0.1 with CUDA 12.1 and Python 3.11.4, which is implemented on a Dell Precision Tower 5810 with Intel Xeon E5-1650 V3 CPU, 64 GB RAM, and a Nvidia GeForce RTX 3090 GPU. During network training, one epoch takes about 20 minutes for a training dataset with 15,000 samples, and the entire training process with 20 epochs lasts for about 6 hours.

3.3 Results

3.3.1 Imaging through a Non-Stationary Diffuser

Evaluation for the proposed GAN is first based on a ground glass diffuser (220-grit ground glass), which is continuously subjected to environmental vibrations and random perturbations during data acquisition.

3.3.1.1 Qualitative Analysis

In order to evaluate the applicability of the proposed GAN in practical applications, we separate training and testing datasets with different time intervals, so that training and testing data

acquisition windows do not overlap in time; the scattering medium statuses corresponding to the testing dataset (red points in Figure 3-4a, including 32 speckles) vary and decorrelate from those for the training dataset (corresponding to the training dataset durations T in Figure 3-4a). Three representative sets, Groups 1, 4, and 6, are analyzed here, as shown in Figure 3-4a. The entire dataset duration (T) is 40 min, and the data that cover the first 1/4 (T = 10 min), 1/2 (T = 20 min), and 3/4 (T = 30 min) of the full dataset are respectively used as the training dataset, whose training sample amounts are 5,000, 10,000, and 15,000 speckle-image pairs, respectively. Then, the testing dataset is sampled with different time intervals (Δt) after the training dataset (red points in Figure 3-4a), such as 1 min, 5 min, and 10 min, *etc*. Qualitative results from representative sets, *i.e.*, Groups 1, 4, and 6, are taken as examples in Figure 3-4.

Performance with T = 30 min (15,000 training samples) is illustrated in Figure 3-4b. As for Group 1 (Final SBP = 0.8846), PCC between the retrieved image and the corresponding ground truth in the testing dataset (called imPCC for simplicity) can reach 0.9665 when Δt = 1 min, and fine features, such as eyes, eyebrows, noses, ears, mouths, and cheeks, can be clearly seen with high fidelity. As Δt increases to 5 min and 10 min, the visualized results are still retrieved with high fidelity (imPCC is 0.9526 and 0.9499, respectively), although SBP drops from 0.9493 (Δt = 1 min) to 0.9295 (Δt = 5 min) and 0.9009 (Δt = 10 min). With stronger decorrelation, *i.e.*,

Group 4 with Final SBP = 0.3987, the retrieval performance (Group 4 in Figure 3-4b) is comparable to Group 1, while the retrieved image PCC of Group 4 is just slightly lower than that of Group 1 due to non-stationary medium statuses. Under highly non-stationary conditions, *i.e.*, Group 6 with Final SBP = 0.0139, the performance deteriorates with larger time intervals, and imPCC is significantly lower than that in Groups 1 and 4.

With shorter training durations, *e.g.*, T = 20 min (10,000 training samples), although visually discernable, the overall imPCCs for Δt = 1, 5, and 10 min in Figure 3-4c are systematically lower than their counterparts in Figure 3-4b. As for Figure 3-4c IV and V, when the time intervals are respectively extended to 15 min and 20 min, the testing datasets decorrelate more from the training datasets. As a result, the retrieved images appear blurrier than before. Especially for Group 6 in Figure 3-4c IV and V, imPCC is as low as 0.3818 and 0.2620, and important facial features (such as eyes and cheeks) become more obscure. When the training duration is even shortened to 10 min with merely 5,000 training samples, as shown in Figure 3-4d, the proposed GAN for Groups 1 and 4 can mostly retain imPCCs above 0.84 and even 0.95. While for Group 6, imPCC significantly drops below 0.31 when $\Delta t \geq 10$ min (SBP reduces to 0.0064). The testing dataset from Group 6 can be regarded to be totally decorrelated from the training dataset.

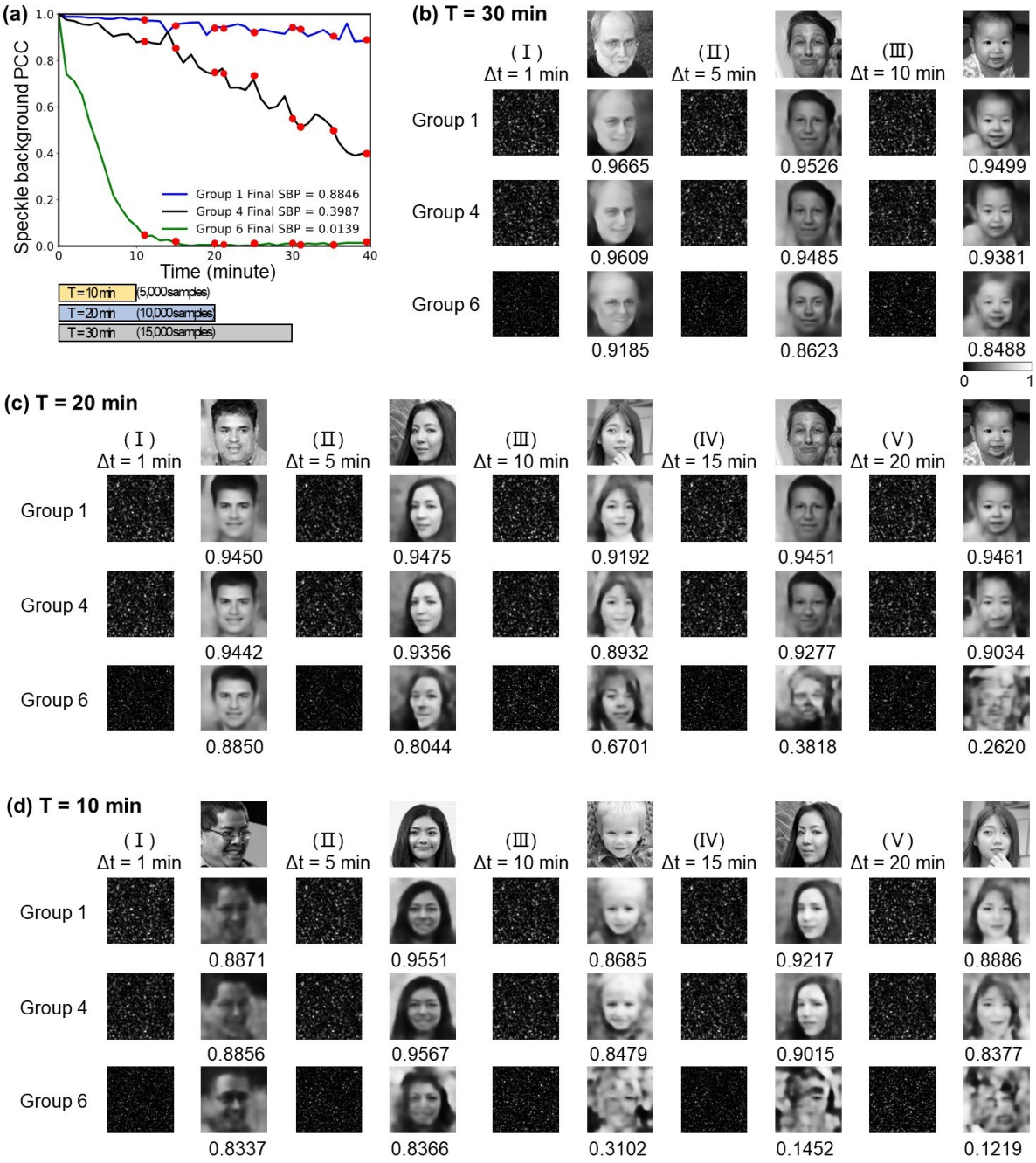

Figure 3-4 Retrieved images from speckles with different training durations (T) and different time intervals (Δt) between training and testing datasets. (a) Training datasets (Group 1, Group 4, and Group 6) are divided according to training dataset durations (T), including 10 min, 20 min, and 30 min. The speckle background PCC of testing datasets

are marked as red points. (b) T = 30 min, Δt = 1, 5, 10 min. (c) T = 20 min, Δt = 1, 5, 10, 15, 20 min. (d) T = 10 min, Δt = 1, 5, 10, 15, 20 min. The top row of each column is the ground truth image. On other rows, the right columns represent the corresponding retrieved images by inputting the speckles in the left columns into the generator in the GAN. The numbers under retrieved images are PCCs between the retrieved images and the corresponding ground truth images (*i.e.*, imPCC). The ground truth images are selected from the Flickr Faces High Quality (FFHQ) database dataset [31].

3.3.1.2 Quantitative Analysis

Quantitatively, performances from six groups of experiments (*i.e.*, imPCC) with different time-varied divisions based on the training data duration T are investigated in Figure 3-5. Generally, a larger time interval between training and testing datasets (Δt) corresponds to a lower instant SBP as the background speckles change continuously for each group. The influences of imPCCs with respect to time intervals and SBP are individually investigated below.

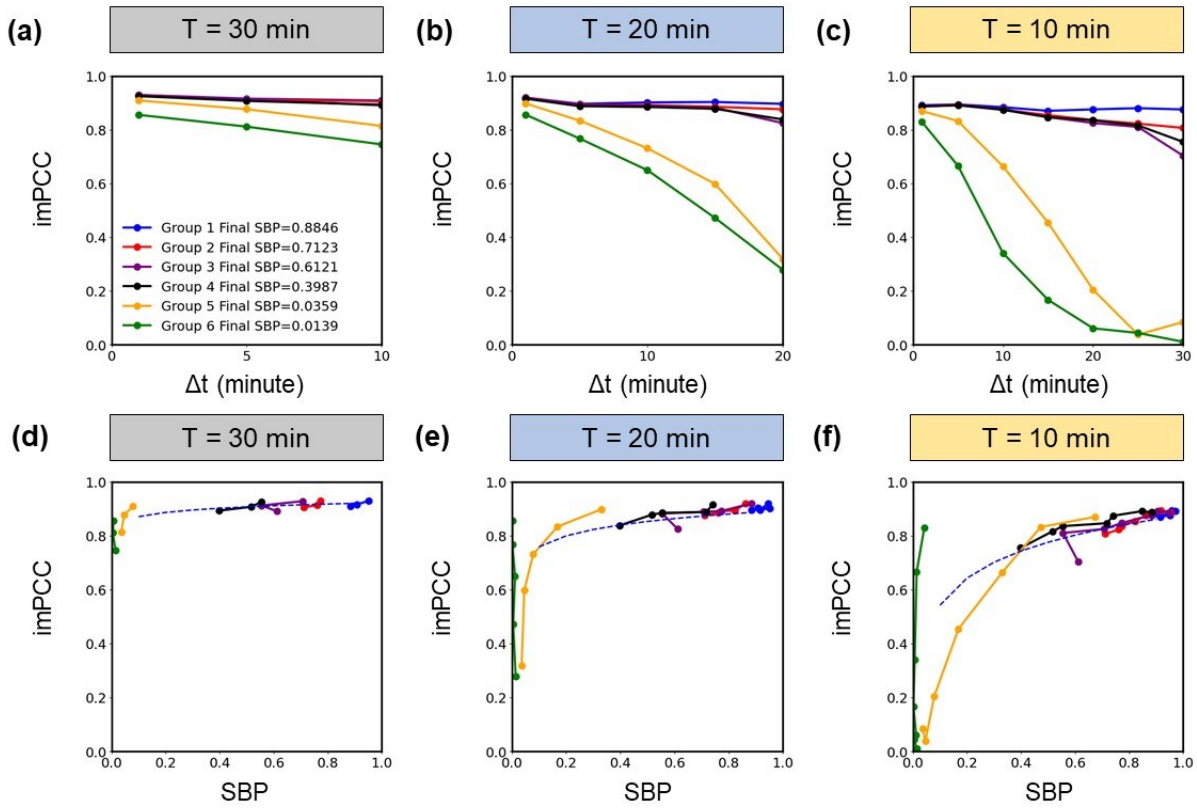

Figure 3-5 Average PCC between the retrieved images and the corresponding ground truths (imPCC) from speckles with different training durations (T) and different time intervals (Δt) between the training datasets and testing datasets: (a-c) imPCC versus different time intervals for training dataset duration T = 30, 20, and 10 min, respectively. (d-f) imPCC versus different SBP for T = 30, 20, and 10 min, respectively (the curves are fitted using a logarithmic function and shown in blue dashed curves).

Temporal generalization of the proposed GAN is revealed in Figure 3-5a, Figure 3-5b, and Figure 3-5c. For Groups 1-4 with Final SBPs above 0.3987, the proposed GAN is able to

retrieve images from speckles with high fidelity. When $\Delta t \leq 10$ min, for T = 30 min with 15,000 image-speckle pairs (Figure 3-5a), imPCCs are relatively stable (above 0.8910) and can hardly be differentiated. Similar observations can also be discovered in Figure 3-5b (T = 20 min with 10,000 training samples) and Figure 3-5c (T = 10 min with 5,000 training samples), though their overall imPCCs are lower than their counterparts in Figure 3-5a. With larger time intervals, *i.e.*, $\Delta t \geq 10$ min, mild descending trends are seen in Figure 3-5b and Figure 3-5c, and imPCCs for Groups 3 and 4 decrease more rapidly than those for Groups 1 and 2. For more severely decorrelating datasets, *i.e.*, Groups 5 and 6 (Final SBP = 0.0359 and 0.0139, respectively), imPCCs drop much more apparently than other groups.

On the other hand, varying statuses of the scattering medium can be essentially characterized by the instant SBP, which describes the extent of decorrelation of the medium. From the perspective of SBP, it is interesting to discover that the performance tendencies vary a lot compared to the time-interval based plots: imPCCs for six groups are well differentiated with a time interval basis in parallel (Figure 3-5a, Figure 3-5b, and Figure 3-5c), but they almost collapse into one curve with the SBP basis in series (Figure 3-5d, Figure 3-5e, and Figure 3-5f). A clearly increasing tendency can be observed, and imPCC increases dramatically from SBP = 0 and saturates gradually. Additionally, we fit the curves in Figure 3-5d, Figure 3-5e, and Figure

3-5f, as represented by the blue dashed curve. As for T = 30, 20, and 10 min, imPCC becomes 0.9004, 0.8353, and 0.7315, even if SBP drops to 0.3678 (corresponding to $1/e$). This suggests that the well-trained GAN is able to adapt to new medium statuses and retrieve images from unknown speckles, even when the scattering medium statuses in the testing datasets vary considerably from the training datasets. Overall, the generalizability of the trained neural network is improved by learning more scattering medium statuses, leading to better retrieved image qualities after speckle decorrelation.

3.3.2 Imaging through a Disordered Metasurface

In the above ground glass-based experiments, the proposed framework can overcome moderate decorrelation of the scattering medium, with the longest time interval (Δt) between training and testing datasets being up to 30 min. Next, we will test whether the proposed framework is able to generalize its ability for even longer periods. A disordered metasurface with random phase profiles is used for investigation. The fabrication process of the metasurface involves a combination of electron beam lithography (EBL) and reactive ion etching (RIE):

1. Use electron beam evaporation (EBE) to deposit a thin film (thickness = 600 nm) of TiO_2

on a smooth glass (SiO_2) substrate.

2. Spin-coat a layer of polymethyl methacrylate (PMMA) onto the TiO_2 layer. After EBL, the PMMA resist is developed.

3. Use EBE to deposit a Cr film on the PMMA pattern. Then, wash off the PMMA resist.

4. Use RIE to selectively remove the TiO_2 material to build the desired pattern.

5. Use chemical etching to remove the Cr mask.

Furthermore, the results of the metasuface experiments are illustrated in Figure 3-6a. As seen, although SBP on Day 1 drops down to 0.6 after 3 hours, the overall stability (SBP) fluctuates from 0.6 to 0.7 on Day 2, even if the optical system (including the laser source and all electronic devices) is shut off for 37 hours. It is worth noting that the network is trained based on the data acquired on Day 1, but the proposed framework can still effectively retrieve the face images from speckles obtained on Day 2. As shown in Figure 3-6c, the human face features can be clearly recognized with all imPCCs above 0.83. As for the results in Figure 3-5d, Figure 3-5e, and Figure 3-5f, imPCC highly depends on SBP. This also applies to a non-diffuser and decorrelating metasurface. As we can see, with stable SBP on Day 2, the corresponding imPCCs in Figure 3-6b remain stable without obvious increasing or decreasing tendency. All this

indicates the temporal generalization capability of the proposed framework.

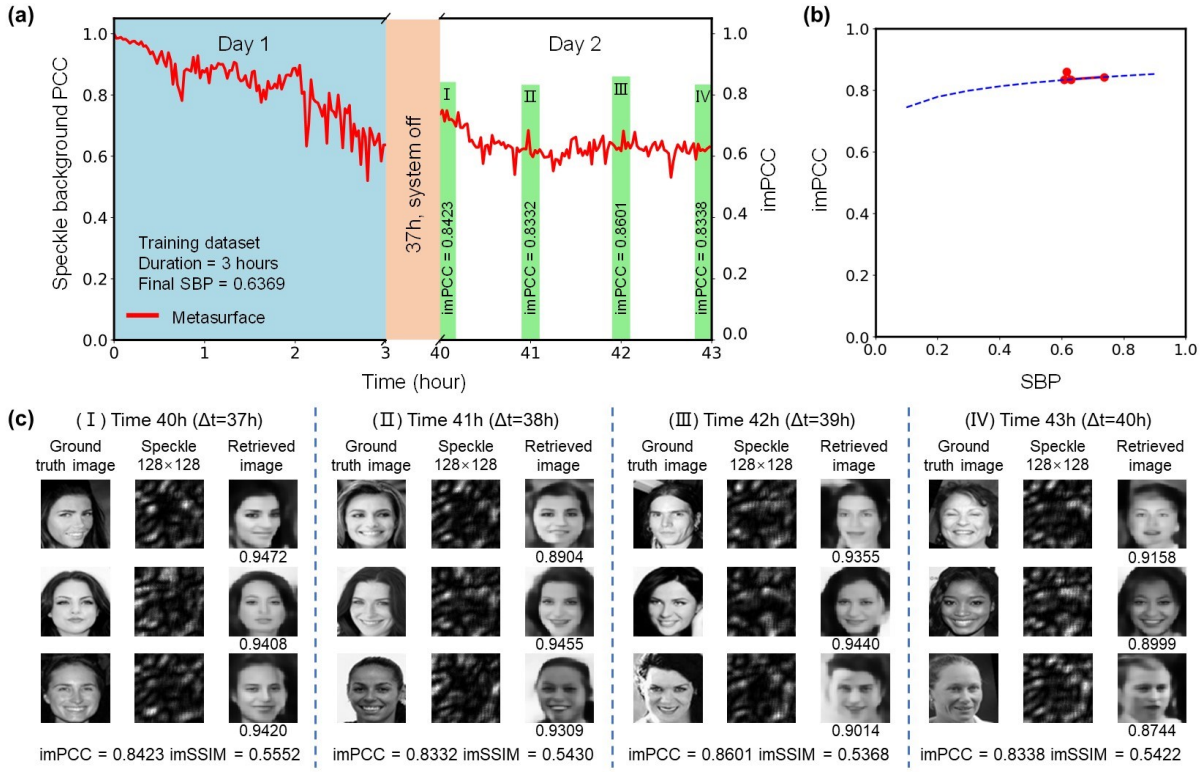

Figure 3-6 Metasurface experimental results. (a) SBP (the red solid curves) on Day 1 and Day 2, between which the optical system is turned off for 37 hours. The network is trained based on the data acquired during the first 3 hours on Day 1, with Final SBP = 0.6369 (marked in light blue) containing 60,000 speckle-image pairs. The testing dataset is acquired on Day 2, with the representative imPCCs at Δt = 37, 38, 39, and 40 hours being represented by the green bars. (b) The resultant imPCCs versus SBP in the testing dataset (the blue dashed curve is the fitted curve using the logarithmic function). (c) The ground

truths, speckles, and the retrieved images. The PCC between each retrieved image and the corresponding ground truth is marked under each retrieved image. The average PCC and structural similarity index measure (SSIM) between the retrieved image and the corresponding ground truth in the testing dataset are marked as imPCC and imSSIM, respectively. The ground truth images are selected from the Large-scale CelebFaces Attributes (CelebA) Dataset [33].

3.4 Discussions

Although the diffuser remains unchanged during each group of data collection, environmental perturbations vary across different days and moments, which is a common challenge in optical experiments. To address this, we have demonstrated that the proposed framework possesses sufficient spatiotemporal generalizability to adapt to scattering medium statuses that are not present in the training data. This exciting feature is enabled by training the network with a diverse set of statuses from the scattering medium, allowing it to extract the underlying speckle features.

We define the quantified perturbation of a non-stationary scattering medium as the speckle

background PCC (SBP). Here, SBP is derived from speckles resulting from loading a uniform phase pattern on the SLM and can be used to evaluate the influence of scattering medium instabilities. Then, complex and dynamic relations between speckles and images are learned by the proposed GAN, which has been compared with other neural networks in Figure 3-7. The generator in the proposed GAN extracts inherent features from speckles through complex-valued convolutions. The complex-valued convolutions mimic the random scattering process more accurately as modeled by the transmission matrix theory. Accordingly, the proposed GAN performs better than the real-valued GAN in Figure 3-7. Furthermore, the discriminator in the proposed GAN evaluates the retrieved images to help the generator further improve outputs, so that the retrieved images are continually refined to approach the corresponding ground truth image during network training. Accordingly, the proposed GAN performs better than the complex-valued U-Net without the discriminator in Figure 3-7.

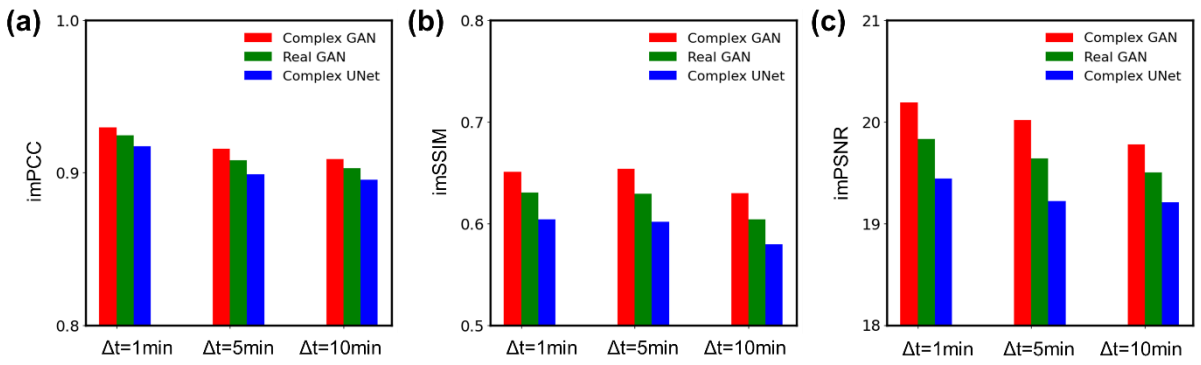

Figure 3-7 Comparisons of different neural networks, including complex-valued convolution-based generative adversarial network (Complex GAN), real-valued convolution-based generative adversarial network (Real GAN), and complex-valued convolution-based deep neural network (Complex UNet). The training datasets used here are from Group 1 (Final SBP = 0.8846) with training dataset durations T = 30 min. Time intervals between training and testing datasets Δt = 1 min, 5 min, and 10 min. The (a) imPCC / (b) imSSIM / (c) imPSNR are average similarities between the retrieved images and the corresponding ground truth images in testing datasets.

Notably, after the metasurface experiments (fluctuating SBP on Day 2 in Figure 3-6a), we can conclude that the proposed GAN successfully overcomes speckle decorrelation even for a time interval $\Delta t > 37$ hours (or longer, if needed). To the best of our knowledge, this is the first research in speckle imaging of complex objects (*e.g.*, grayscale human face images with

detailed features) whose network training and testing are performed on various days with significantly different or unpredicted system statuses. This capability opens new venues for applications where networks can be trained in advance but maintain their validity for data acquired later, regardless of the status of the medium. This is essential for extending the impact of the learning-based approaches in the field.

Additionally, due to its universality across different scattering media (ground glass and metasurface in experiments), the proposed GAN can be further extended to imaging or information retrieval through living biological tissues. This is long desired yet considered highly challenging, as the optical field decorrelates rapidly on the order of milliseconds due to physiological motions such as breathing and blood flow [15]. With the proposed framework, one can train the network with datasets that cover a relatively long temporal window. This allows the network to learn the different statuses of the sample. Consequently, high-fidelity information retrieval can be achieved instantaneously by inputting a single speckle captured later.

Furthermore, the performance of the proposed GAN can be improved by incorporating more encoded speckle information during the training stage or by adopting an optical neural network for faster processing [34]. Retrieved image quality is always challenging, especially when

speckles decorrelate significantly from the training data. Including more training speckle samples with various scattering medium statuses will generally yield better-retrieved images. Additionally, speckle data with a larger region of interest (ROI) includes more speckle grains and accesses more encoded information, potentially resulting in better-retrieved images, albeit at the cost of longer training time. From another point of view, if the target information is less complicated, such as binary digits and letters [35-36], less training data and a lighter network structure may reduce the time required for network training and computing resources, as shown in Figure 3-8, where digits can be retrieved from speckles with high fidelity, even for T = 10 min and Δt = 20 min. Moreover, since the distribution of target information encoded in the speckle field is spatially redundant, down-sampling [14] or using a partial field of view of the speckles [26] could significantly reduce the size of the training data, thereby accelerating the learning process of the network.

These enhancements could improve the prospects of the proposed framework for practical applications of speckle imaging, especially with an ultra-stable disordered metasurface as the scattering medium. For instance, speckles from different weather conditions can be used to train a GAN. The pre-trained GAN can then image through various scattering conditions with speckles acquired later, even if the statuses of the scattering medium (*i.e.*, weather conditions)

are unknown to the trained network. The spatiotemporal decorrelation generalizability of the network could help auto-driving cars perceive their surroundings more accurately in a variety of weather conditions.

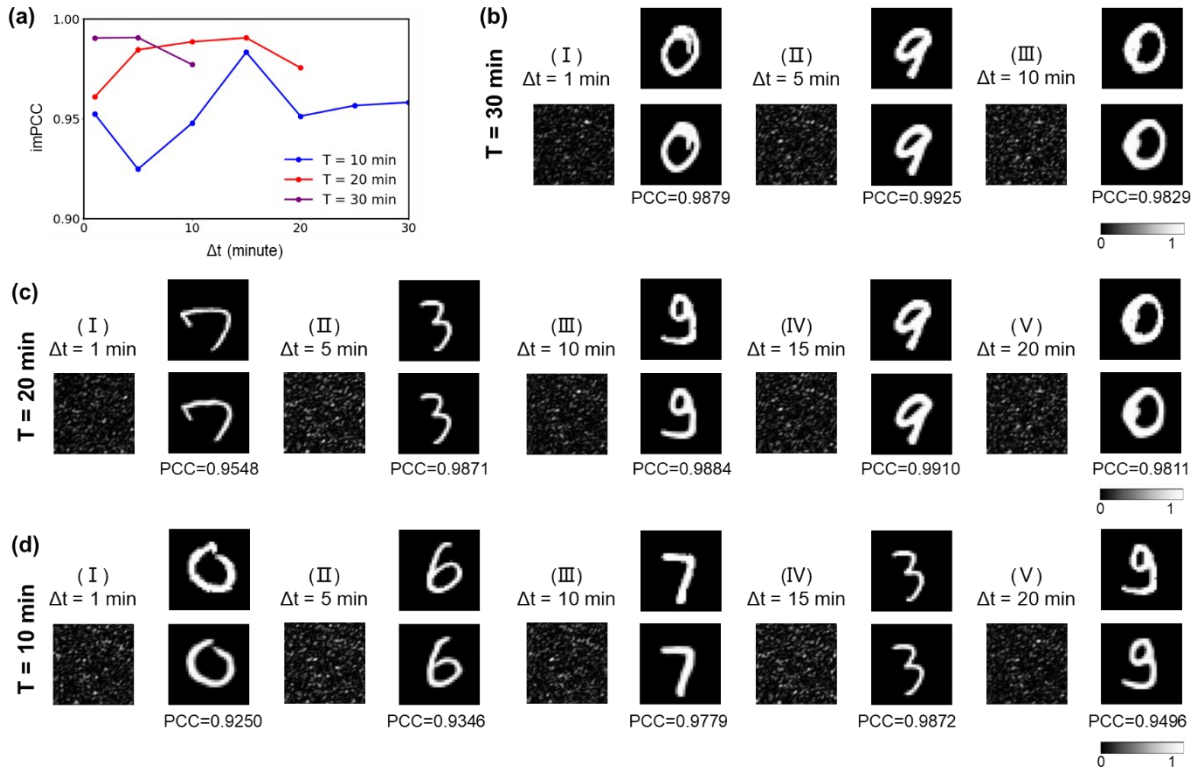

Figure 3-8 Experiments with speckles resulted from digits: (a) imPCC versus different time intervals for training dataset duration T = 30, 20, 10 min. (b-d) Retrieved images from speckles with different training durations (T) and different time intervals (Δt) between training and testing datasets: (b) T = 30 min, Δt = 1, 5, 10 min. (c) T = 20 min, Δt = 1, 5, 10, 15, 20 min. (d) T = 10 min, Δt = 1, 5, 10, 15, 20 min.

3.5 Conclusion

In this chapter, we propose a GAN-based learning framework to non-holographically retrieve images from spatiotemporally decorrelated speckles acquired from non-stationary scattering media. The extended generalizability of the proposed framework is first demonstrated through experiments using ground glass as the scattering medium. Despite the non-stationary nature of the scattering medium and the separation of data for network training and testing into different periods, the proposed framework successfully retrieves face images from speckles with high fidelity. Furthermore, using a disordered metasurface as the scattering medium, the framework effectively retrieves images from speckles acquired on the second day of network training, even after the optical system has been turned off for 37 hours and then restarted for testing data acquisition. These results demonstrate the excellent spatiotemporal decorrelation generalizability of the proposed GAN-based framework, enabling high-fidelity retrieval of complex information (*e.g.*, human face images) from speckles acquired under varying media or system conditions. This capability, allowing for a network to be trained in advance and tested with input data acquired later, highlights the impactful potential of learning-based approaches in various speckle imaging scenarios, such as non-holographic imaging through biological tissues and target sensing in autopilot systems under inclement weather.

References

1. Yun, S. H., & Kwok, S. J. (2017). Light in diagnosis, therapy and surgery. *Nature biomedical engineering, 1*(1), 0008.

2. Zhu, L., Soldevila, F., Moretti, C., d'Arco, A., Boniface, A., Shao, X., ... & Gigan, S. (2022). Large field-of-view non-invasive imaging through scattering layers using fluctuating random illumination. *Nature communications, 13*(1), 1447.

3. Luo, Y., Yan, S., Li, H., Lai, P., & Zheng, Y. (2021). Towards smart optical focusing: deep learning-empowered dynamic wavefront shaping through nonstationary scattering media. *Photonics Research*, *9*(8), B262-B278.

4. Turtaev, S., Leite, I. T., Altwegg-Boussac, T., Pakan, J. M., Rochefort, N. L., & Čižmár, T. (2018). High-fidelity multimode fibre-based endoscopy for deep brain in vivo imaging. *Light: Science & Applications*, *7*(1), 92.

5. Wang, L., Qi, T., Liu, Z., Meng, Y., Li, D., Yan, P., ... & Xiao, Q. (2022). Complex pattern transmission through multimode fiber under diverse light sources. APL Photonics, 7(10).

6. Wiersma, D. S. (2013). Disordered photonics. *Nature Photonics, 7*(3), 188-196.

7. Popoff, S. M., Lerosey, G., Carminati, R., Fink, M., Boccara, A. C., & Gigan, S. (2010).

Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media. *Physical review letters, 104*(10), 100601.

8. Lee, H., Yoon, S., Loohuis, P., Hong, J. H., Kang, S., & Choi, W. (2022). High-throughput volumetric adaptive optical imaging using compressed time-reversal matrix. *Light: Science & Applications, 11*(1), 16.

9. Tian, L., Hunt, B., Bell, M. A. L., Yi, J., Smith, J. T., Ochoa, M., ... & Durr, N. J. (2021). Deep learning in biomedical optics. *Lasers in surgery and medicine, 53*(6), 748-775.

10. Turpin, A., Vishniakou, I., & d Seelig, J. (2018). Light scattering control in transmission and reflection with neural networks. *Optics express*, *26*(23), 30911-30929.

11. Doronin, A., & Meglinski, I. (2011). Online object oriented Monte Carlo computational tool for the needs of biomedical optics. *Biomedical optics express*, *2*(9), 2461-2469.

12. Horisaki, R., Takagi, R., & Tanida, J. (2016). Learning-based imaging through scattering media. *Optics express*, *24*(13), 13738-13743.

13. Tang, P., Zheng, K., Yuan, W., Pan, T., Xu, Y., Fu, S., ... & Qin, Y. (2022). Learning to transmit images through optical speckle of a multimode fiber with high fidelity. *Applied physics letters*, *121*(8).

14. Li, H., Yu, Z., Zhao, Q., Luo, Y., Cheng, S., Zhong, T., ... & Lai, P. (2023). Learning-based

super-resolution interpolation for sub-Nyquist sampled laser speckles. *Photonics Research*, *11*(4), 631-642.

15. Cheng, S., Li, H., Luo, Y., Zheng, Y., & Lai, P. (2019). Artificial intelligence-assisted light control and computational imaging through scattering media. *Journal of innovative optical health sciences*, *12*(04), 1930006.

16. Yu, Z., Li, H., Zhong, T., Park, J. H., Cheng, S., Woo, C. M., ... & Lai, P. (2022). Wavefront shaping: A versatile tool to conquer multiple scattering in multidisciplinary fields. *The Innovation*, *3*(5).

17. d'Arco, A., Xia, F., Boniface, A., Dong, J., & Gigan, S. (2022). Physics-based neural network for non-invasive control of coherent light in scattering media. *Optics Express*, *30*(17), 30845-30856.

18. Li, Y., Xue, Y., & Tian, L. (2018). Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media. *Optica*, *5*(10), 1181-1190.

19. Zhao, Q., Li, H., Yu, Z., Woo, C. M., Zhong, T., Cheng, S., ... & Lai, P. (2022). Speckle-based Optical Cryptosystem and its Application for Human Face Recognition via Deep Learning. *Advanced Science*, *9*(25), 2202407.

20. Caravaca-Aguirre, A. M., & Piestun, R. (2017). Single multimode fiber endoscope. *Optics*

express, *25*(3), 1656-1665.

21. Vasquez-Lopez, S. A., Turcotte, R., Koren, V., Plöschner, M., Padamsey, Z., Booth, M. J., ... & Emptage, N. J. (2018). Subcellular spatial resolution achieved for deep-brain imaging in vivo using a minimally invasive multimode fiber. *Light: science & applications*, *7*(1), 110.

22. Li, H., Woo, C. M., Zhong, T., Yu, Z., Luo, Y., Zheng, Y., ... & Lai, P. (2021). Adaptive optical focusing through perturbed scattering media with a dynamic mutation algorithm. *Photonics Research*, *9*(2), 202-212.

23. Wu, G., Sun, Y., Yin, L., Song, Z., & Yu, W. (2023). High-definition image transmission through dynamically perturbed multimode fiber by a self-attention based neural network. *Optics Letters*, *48*(10), 2764-2767.

24. Resisi, S., Popoff, S. M., & Bromberg, Y. (2021). Image transmission through a dynamically perturbed multimode fiber by deep learning. *Laser & Photonics Reviews*, *15*(10), 2000553.

25. Zheng, S., Wang, H., Dong, S., Wang, F., & Situ, G. (2021). Incoherent imaging through highly nonstatic and optically thick turbid media based on neural network. *Photonics Research*, *9*(5), B220-B228.

26. Lyu, M., Wang, H., Li, G., Zheng, S., & Situ, G. (2019). Learning-based lensless imaging

through optically thick scattering media. *Advanced Photonics*, *1*(3), 036002.

27. Li, Z., Zhou, W., Zhou, Z., Zhang, S., Shi, J., Shen, C., ... & Dai, Q. (2024). Self-supervised dynamic learning for long-term high-fidelity image transmission through unstabilized diffusive media. *Nature Communications*, *15*(1), 1498.

28. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, *63*(11), 139-144.

29. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention-MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18* (pp. 234-241). Springer International Publishing.

30. Odena, A., Olah, C., & Shlens, J. (2017, July). Conditional image synthesis with auxiliary classifier GANs. In *International conference on machine learning* (pp. 2642-2651). PMLR.

31. Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4401-4410).

32. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... &

Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, *27*.

33. Liu, Z., Luo, P., Wang, X., & Tang, X. (2015). Deep Learning Face Attributes in the Wild. In *Proceedings of the IEEE international conference on computer vision* (pp. 3730-3738).

34. Li, H., Yu, Z., Zhao, Q., Zhong, T., & Lai, P. (2022). Accelerating deep learning with high energy efficiency: from microchip to physical systems. *The Innovation*, *3*(4).

35. Fan, P., Ruddlesden, M., Wang, Y., Zhao, L., Lu, C., & Su, L. (2021). Learning enabled continuous transmission of spatially distributed information through multimode fibers. *Laser & Photonics Reviews*, *15*(4), 2000348.

36. Zhu, C., Chan, E. A., Wang, Y., Peng, W., Guo, R., Zhang, B., ... & Chong, Y. (2021). Image reconstruction through a multimode fiber with a simple neural network architecture. *Scientific reports*, *11*(1), 896.

4 DELOCALIZED INFORMATION IN OPTICAL SPECKLES

This chapter has been prepared as the following paper and will be submitted for publication:

Zhao, Q.[†], Li, H.[†,#], Yu, Z.[†], Li, H.[†], Cheng, S., Huang, H., Zhong, T., Woo, C. M., Wang, Z., Zheng, Y., Liu, H.[#], & Lai, P.[#] (2024). Delocalized information in optical speckles: a learning-based study. *In preparation*.

Light travels in a straight line through a homogeneous and transparent medium, resulting in a one-to-one information mapping between the object and its image. However, in a scattering medium, incident light undergoes multiple scattering events due to wavelength-scale inhomogeneities of the refractive index in the medium. Consequently, photons from a single point on the object are significantly diffused and spatially delocalized across many different regions in the resultant optical field, and vice versa. This dispersal of information leads to a

multi-to-multi correspondence, with information being delocalized and presented as speckles. As discussed in Chapter 3, significant research efforts have focused on training neural networks to extract information from speckles. However, comparatively less attention has been given to understanding how information is encoded or delocalized in speckles, and the conditions under which high-fidelity information can be retrieved from speckles. This chapter comprehensively investigates the phenomenon of information delocalization in speckles. Experimental findings reveal that object information is uniformly delocalized among fully developed speckles, maintaining consistent information across different regions of interest (ROIs) of the same size and ensuring the equivalent fidelity of the retrieved information. Furthermore, it has been experimentally confirmed that delocalized information can be retrieved with high fidelity if the sampled speckle ROIs contain sufficient information.

4.1 Introduction

Optical techniques have revolutionized exploration and advancement across various fields [1]. These techniques have become indispensable tools for unveiling the intricacies of the physical world. Notable examples include confocal microscopes for deep tissue imaging and stimulated

emission depletion microscopes for super-resolution imaging [2-4]. In biomedical applications, optical methods are primarily categorized into ballistic and scattering schemes.

In typical optical imaging applications, ballistic imaging through homogeneous and transparent media is the most common scenario. The ballistic scheme involves light traveling in straight lines along ballistic paths from the object to the image plane [5]. Each point on the image corresponds directly to a point on the object, ensuring clarity in imaging, as shown in Figure 4-1a [6]. However, when some ballistic light paths are obstructed, crucial information from the object cannot reach the image plane or camera, resulting in information loss. In contrast, optical imaging through scattering media is dominated by multiply scattered light, as shown in Figure 4-1b [7-8]. The scattering medium has heterogeneous compositions, leading to variations in refractive indices and the propagation of light through processes like refraction, scattering, absorption, *etc*. [9]. As a result, the optical field at any point in a speckle is the superposition of fields from different points of the object, with contributions weighted by the transmission channels in the scattering medium. Despite obstructions in some transmission channels, photons from the same region can still reach the image plane through other channels. This principle in scattering underscores the multiple-to-multiple mapping between information and speckles in strong scattering schemes, differing significantly from ballistic schemes.

To address the challenges posed by scattering schemes, researchers have developed various techniques aimed at mitigating optical speckles or retrieving information from speckles, including wavefront shaping, transmission matrix, deep neural networks, *etc*. Recently, the integration of optical imaging and deep learning has led to remarkable advancements in the field, particularly in the context of speckle imaging [10-11]. Neural networks, with their advanced learning capabilities, have been employed to retrieve information from speckles. These models, including fully connected layer-based models [12], convolutional layer-based models [13], and transformer block-based models [14-16], have showcased their effectiveness in extracting features from multiple dimensions and building relations between information and speckles. Additionally, by applying physical models (such as transmission matrix and speckle autocorrelation) in neural networks, researchers have been able to improve image quality and successfully retrieve high-fidelity information from speckles [17-19]. High-level applications, such as non-line-of-sight imaging [20-21] and encrypted human face images for face recognition [22], have also been achieved, further validating the effectiveness of these neural network-based models in the field of speckle imaging.

Among studies of speckle imaging, researchers have proposed that information is not confined to a single ROI in speckles, but is spread across multiple speckle ROIs [23]. This has led to

novel research areas such as non-line-of-sight imaging [24] and speckle super-resolution imaging [25]. Despite advancements in retrieving information from speckles, the specifics of how information is delocalized in speckles and the conditions for ensuring high-fidelity information retrieval from speckles still remain unanswered. Understanding delocalization is crucial for developing effective methods to retrieve information from speckles, particularly in applications where clear, high-resolution images are essential for accurate diagnoses and treatments. To the best of our knowledge, this gap in understanding has not been explored, highlighting the importance of continued research in this field.

In this chapter, we focus on investigating how information is delocalized in speckles using neural networks. Theoretical analyses of information delocalized in speckles are conducted, and the speckle sampling conditions for ensuring high-fidelity information retrieval are explored. Experimental results demonstrate that information is uniformly delocalized among speckle ROIs with different sizes and locations. Additionally, we perform a quantitative analysis of the delocalized information using the concept of entropy, which is a widely adopted measure of information [26] and has been introduced to optical research about transmission invariants through scattering media [27]. Our findings indicate that if the entropy of speckle autocorrelation exceeds that of image autocorrelation, neural networks can be trained to retrieve

information from speckles with high fidelity, as evidenced by a Pearson correlation coefficient (PCC) greater than 0.9. This novel application of Shannon entropy to speckle analysis provides a new perspective on the relationship between speckle and image information.

Overall, this work contributes to a deeper understanding of information delocalization in speckles, which is crucial for advancing speckle imaging applications. By introducing the concept of entropy to speckles and discussing the relationship between the entropy of speckle autocorrelation and image autocorrelation, this research paves the way for future developments in speckle-related research. These findings have the potential to inspire new applications and methodologies for biomedical imaging.

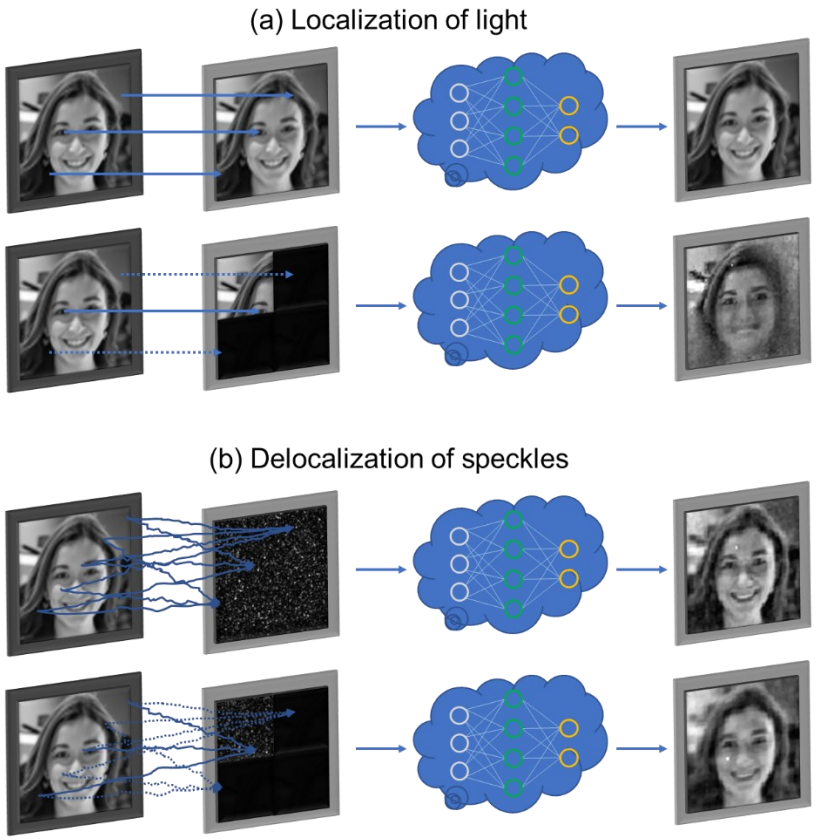

Figure 4-1 The concept of delocalized information in speckles: (a) Localization of light: one-to-one mapping between the object and the image in ballistic light imaging. Obstruction results in information loss and low-quality information retrieval. (b) Delocalization of speckles: multi-to-multi correspondence in speckle-based imaging. Delocalized information in speckles results in high-fidelity information retrieval even in the presence of obstruction. Face image: Copyright 2018, Deya at San Antonio Cocktail Conference, by Nan Palmero, Flickr (https://www.flickr.com/photos/nanpalmero/38756513965/); the original images are converted to grayscale, under terms of the CC-BY 2.0 license.

4.2 Theoretical Model

Before experiments, we first analyze the physical models that describe the intricate relations between images and the corresponding speckles. The analysis is conducted from two aspects: the concept of delocalization and the speckle sampling condition for information retrieval from speckles. The former refers to the phenomenon where information is spread out over a large area within the field of view of speckles, rather than being confined to a single small region; the latter focuses on the speckle sampling condition.

4.2.1 Concept of Delocalization

The concept of delocalization is first analyzed through the commonly used transmission matrix model [28]. In the spatial domain, the incident field E and the resultant speckle field U are modeled as vectors, and the scattering medium is modeled as a transmission matrix T. As a result, the relations between the image information and the corresponding speckles can be elaborated through matrix operations, as shown in Equation (4-1):

$$U = \begin{vmatrix} U_1 \\ U_2 \\ \vdots \\ \vdots \\ \vdots \\ U_M \end{vmatrix} = T \cdot E = \begin{bmatrix} t_{11} & \cdots & t_{1n} & \cdots & \cdots & t_{1N} \\ \vdots & \ddots & & & & \vdots \\ t_{m1} & \cdots & t_{mn} & \cdots & \cdots & t_{mN} \\ \vdots & & \vdots & \ddots & & \vdots \\ \vdots & & \vdots & & \ddots & \vdots \\ t_{M1} & \cdots & t_{Mn} & \cdots & \cdots & t_{MN} \end{bmatrix} \times \begin{vmatrix} e_1 \\ e_2 \\ \vdots \\ \vdots \\ \vdots \\ e_N \end{vmatrix}. \quad (4\text{-}1)$$

When E is multiplied by T, the resulting U is the result of the combined effect of transformations. One element in U is related to each individual element in E, since each element in U is a result of the transformation applied to each element in E. In the context of optical scattering, this implies that the transformation applied to the original image results in an output field where each element is a combination of all elements in the original image. On the other hand, this relationship is not just a one-way process. One element in E is also related to each element in U. Therefore, the elements in E can still be retrieved from U even if some elements in U are lost.

The mutual interaction between E and U exemplifies the core principles of information distribution in the optical scattering process. In this process, multiple elements of the image are encoded into multiple elements of the speckle, and vice versa. The inherent characteristics of matrix multiplication and the properties of linear transformations offer a theoretical basis for comprehending the multi-to-multi relationship between images and speckles. This relationship, referred to as delocalization in this chapter, describes how information is spread across multiple

elements rather than being confined to a single location.

4.2.2 Conditions for High-fidelity Information Retrieval from Speckles

Next, we explore the speckle sampling condition necessary for high-fidelity information retrieval from speckles. Information entropy is introduced as a measure to gauge the amount of information encoded within speckles, which is crucial for understanding the speckle sampling conditions under which images can be retrieved. Entropy (H) plays a critical role in assessing this information, as it quantifies the randomness and complexity of the information contained in the speckles [29]:

$$H = -\sum P(I) \times \log[P(I)], \tag{4-2}$$

where $P(I)$ is the probability of digital intensity level I (0-255 for 8 bits, or 0-65535 for 16 bits). Accordingly, higher entropy indicates more information, and lower entropy suggests less information.

From another point of view, the wavefronts are distorted by the scattering medium located at the Fourier plane, indicating that the information present at this plane is the Fourier transform

of the images, rather than the original images. This distinction is crucial because it means that the analysis of information retrieval conditions should not be directly based on the comparison of original images and speckles. Instead, the relationship between speckles and images should be examined through their autocorrelations [30]. Additionally, according to Wiener-Khinchine Theorem, the autocorrelation is the inverse Fourier transform of the power spectrum, indicating that the autocorrelation provides insights into the information about the Fourier transform [31]. Next, the analysis of information retrieval from speckles involves the relationship between the entropy of speckle autocorrelation and the entropy of image autocorrelation. Within the memory effect range, the autocorrelation of the speckle is the convolution of the autocorrelation of the image with the autocorrelation of the point spread function PSF, as shown in Equation (4-3) [30]:

$$Autocorr(Speckle) = Autocorr(Image) * Autocorr(PSF). \quad (4\text{-}3)$$

In other words, the speckle autocorrelation contains information about the image autocorrelation and PSF autocorrelation, which allows for the analysis of how information in speckles and images are related. Additionally, in common sense, lost information leads to reduced entropy and cannot be recovered [32]. In the context of neural networks, pre-training to retrieve information from speckles also implies that speckle autocorrelation should contain

enough information about the original images. Therefore, we expect that the entropy of speckle autocorrelation should exceed the entropy of image autocorrelation to ensure high-fidelity information retrieval, *i.e.,*

$$Entropy[Autocorr(Speckle)] > Entropy[Autocorr(Image)]. \quad (4\text{-}4)$$

Note that if the speckle autocorrelation contains less information than the image autocorrelation, the pre-trained neural network may still be capable of retrieving information due to the learned memory. However, the quality of the retrieved information may be low, and it is likely that the recovered information is generated from the training data rather than representing the actual information.

In brief, the entropy analysis involves understanding delocalized information in speckles. The proposed speckle sampling condition to ensure high-fidelity information retrieval highlights the complexity and intricacy of the relationship between images and speckles. Currently, the proposed speckle sampling condition is a necessary condition, rather than a necessary and sufficient condition. This means that retrieved information may still not be satisfactory even if the speckle sampling condition is met, due to various experimental factors, such as optical setups, environmental perturbations, and neural network parameters. In the following sections, we will validate the proposed speckle sampling condition.

4.3 Methods

4.3.1 Optical Setup for Dataset Generation and Acquisition

The optical experiments involve a meticulous process of capturing speckle data from images using the optical system shown in Figure 4-2a. The process begins with the use of a continuous-wave 532 nm laser source (EXLSR-532-300-CDRH, Spectra-Physics, Single mode, 300 mW, USA), which is expanded by a 4-f optical system (L1 and L2). The expanded beam fully illuminates the spatial light modulator (SLM, HOLOEYE PLUTO VIS056 1080p, German) to modulate the optical wavefront, where image intensities are converted into wavefront phase delays (0 to 2π). Following this, an objective lens (RMS20X, Olympus, Japan) focuses the modulated beam onto a scattering medium (220-grit ground glass, diameter of 1.0 inch, DG10-220-MD, Thorlabs, USA). On the other side of the medium, optical speckles are generated and recorded by a CMOS camera (FL3-U3-32S2M-CS, PointGrey, Canada). Corresponding image-speckle pairs are stored as data for analysis. During experiments, image data loaded on the SLM are 20,000 thumbnails from FFHQ dataset with permission [33].

4.3.2 Neural Networks

Following the preparation of the speckle dataset, neural networks are trained to explore and retrieve delocalized information from speckles. The neural networks, based on a complex number-based fully connected layer [25], are designed to effectively handle optical speckles, as shown in Figure 4-2b. The loss function (L) employed during the training of these neural networks is a combination of negative Pearson correlation coefficient (PCC), negative structural similarity index measure (SSIM), and mean square error (MSE). This combined loss function allows for a comprehensive evaluation of the neural network's performance, balancing overall similarity (PCC and SSIM) with pixel-wise similarity (MSE). The empirical parameters (a) and (b) are set to 0.8 to optimize the fidelity of image retrieval from speckles:

$$L = MSE(y, \hat{y}) - aPCC(y, \hat{y}) - bSSIM(y, \hat{y}), \tag{4-5}$$

where y is the ground truth, and $\hat{y}$ is the neural network output. During the network training process, 19,800 speckle-image pairs are used for training, and the remaining 200 image-speckle pairs are used for testing. The used optimizer is Adam with batch size = 16, and the training is conducted over 20 epochs. Initially, the combined loss function L in Equation (4-5) is used, and in the final 4 epochs, only the MSE loss is used to fine-tune the parameters. This approach ensures that the neural network is effectively trained to retrieve information from speckles with

high fidelity. During experiments, we use the same network structures and training parameters for fair comparisons between different groups. The framework used for network training and testing includes PyTorch 2.0.1, PyTorch Lighting 2.1.3, Python 3.11.4, and CUDA 11.7, running on an Nvidia GeForce RTX 3090 GPU.

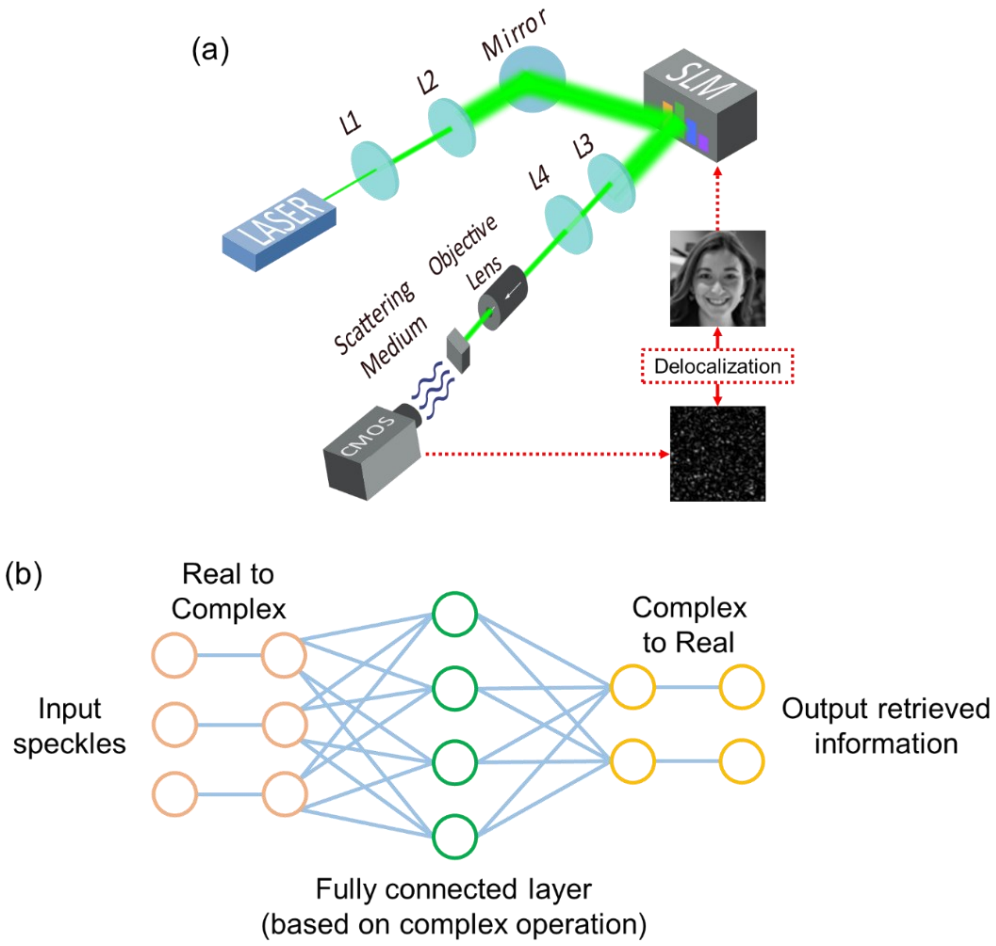

Figure 4-2 (a) The optical setup in experiments: images are loaded on the SLM to modulate the wavefront of the incident laser beam (λ = 532 nm), and the resultant speckles

are captured by a CMOS camera positioned after the scattering medium. (b) Complex fully connected neural network model: one fully connected layer is used to build the neural network, with speckles as the input and retrieved information as the output.

4.4 Results

4.4.1 Information Retrieval without and with Delocalization

The experiments aim to explore the concept of delocalization, specifically focusing on the differences in information retrieval with and without delocalization, which is crucial for understanding how delocalization influences the quality of the retrieved information from speckles. Here, information is retrieved from partial images (*i.e.*, without delocalization) and partial speckles (*i.e.*, with delocalization), respectively. We separately use partial ROIs from images (1/4 of full images) and speckles (1/4 of speckles with 256×256 resolution) to train neural networks to retrieve the encoded information. The results from two experiments are compared in Figure 4-3.

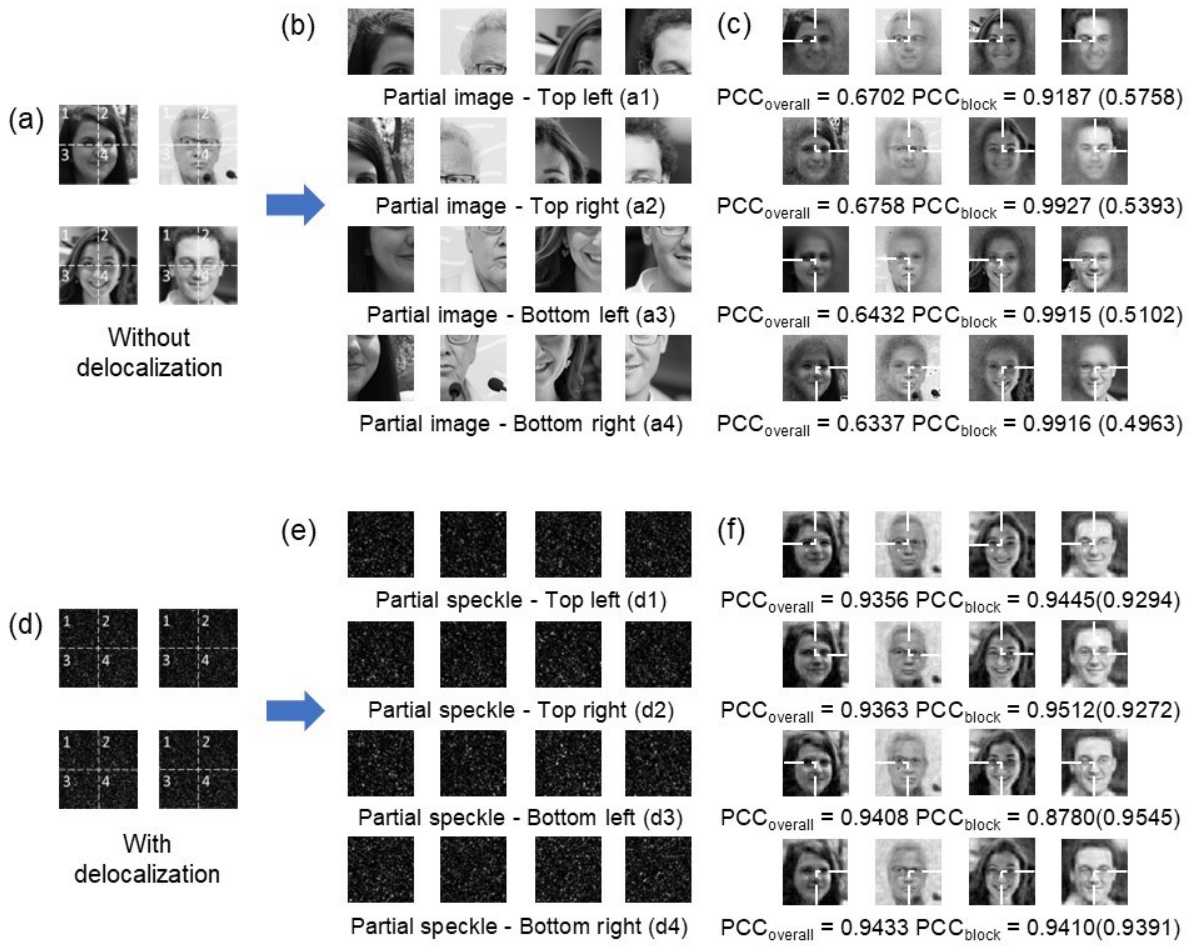

Figure 4-3 Comparisons of information retrieval without and with delocalization: (a-c) Information retrieval without delocalization: information retrieved from partial images and the corresponding information loss. (d-f) Information retrieval with delocalization: information retrieved from partial speckles with high fidelity. $PCC_{overall}$ is the PCC between the whole picture of the retrieved information and the corresponding ground truth. PCC_{block} is the PCC between the blocks marked with dotted lines and the corresponding ground truth blocks; the PCC between the remained regions and the corresponding remained ground truths are given in brackets. Ground truth image (top

left): Copyright 2018, DSC_0431, by Aireal Robbins, Flickr (https:// www.flickr.com/photos/148747390@N04/41350064164/); the original images are converted to grayscale, under terms of the CC-BY 2.0 license; Ground truth image (top right): Copyright 2018, 3ª Sessão Ordinária de 2018, by Câmara Municipal de Braganca Paulista, Flickr (https://www.flickr.com/photos/camarabraganca/39526445765/); the original images are converted to grayscale, under terms of the Public Domain Mark 1.0 license; Ground truth image (bottom left): Copyright 2018, Deya at San Antonio Cocktail Conference, by Nan Palmero, Flickr (https://www.flickr.com/photos/nanpalmero/ 38756513965/); the original images are converted to grayscale, under terms of the CC-BY 2.0 license; Ground truth image (bottom right): Copyright 2018, Laurea - Valerio 2, by Enrico, Flickr (https://www.flickr.com/photos/onefromrome/275113916/); the original images are converted to grayscale, under terms of the CC-BY 2.0 license.

As shown, in traditional imaging without delocalization (Figure 4-3 a-c), light propagates in straight lines and the transmission of information is localized. When partial information on the received field is lost, information retrieval is challenging, and the quality of the retrieved information is significantly lower. This is evident from the low PCCs (just between 0.63 and

0.67), indicating that the retrieved information is quite different from the original images. The vision perceptions in Figure 4-3c further highlight the limitations, where only the regions corresponding to the partial images can be retrieved clearly (PCC > 0.9), and other regions are obscure (PCC < 0.6). Additionally, neural networks just average over the ground truth images, rather than recover the lost information. As a result, face contours of the retrieved information all look similar to each other. For example, the mouth of the young man in Figure 4-3c is open although it should be closed, and the eyeglasses of the old and young men cannot be retrieved at all. This limitation is attributed to the localization nature of the traditional image transmission. In contrast, speckles with delocalization in Figure 4-3 d-f show different behaviors. Even when partial speckles are used for information retrieval, the important facial features can be retrieved with high fidelity (PCC between 0.94 to 0.95). It suggests that the partial speckles contain sufficient information of the entire face image, which allows for the retrieval of detailed and recognizable facial features, including cheeks, eyes, hair, expressions, *etc*. Accordingly, neural networks can be used in experiments to validate information delocalization in speckles.

Another difference between localization and delocalization is that different regions of the retrieved information in Figure 4-3f exhibit almost the same fidelity, unlike the various results seen in Figure 4-3c. The consistent quality of the retrieved information across different ROIs in

speckles further validates the concept of delocalized information. This indicates that the entire face image information is evenly distributed across different ROIs in speckles, making it possible to retrieve high-fidelity information.

4.4.2 Delocalized Information among Different ROIs

From the foregoing results, information is distributed across different speckle ROIs, regardless of their locations. Next, we will further explore how information is encoded in speckles. We crop different sizes of ROIs from speckles with 256×256 pixels and train a set of neural networks to retrieve information from these ROIs (*i.e.*, 128×128, 85×85, 64×64, 51×51, 42×42, 36×36, and 32×32).

Results from these experiments, presented in Figure 4-4a, show that even when ROIs of varying positions and sizes are used, delocalized information can be retrieved from different speckle ROIs. This is true despite the fact that PCC between the retrieved information and the corresponding images (*i.e.*, retrieved PCC) varies. Notably, the details (including eyes, hair, teeth, and facial expressions) in the retrieved information from 256×256, 128×128, and 85×85 speckle ROIs exhibit satisfying visual results, indicating that speckles contain sufficient human

face features. These findings support the proposed theory that information is delocalized among speckles, as evidenced by the satisfying similarities between the retrieved information and the ground truth.

Furthermore, for the same ROI sizes, the consistent retrieved PCC further confirms that information is evenly delocalized among speckles. By leveraging this delocalized nature, neural networks can effectively retrieve detailed and high-fidelity information from speckles. This capability is particularly useful in applications involving scattering and speckles, such as biomedical imaging and other scenarios where the retrieval of high-fidelity images is crucial.

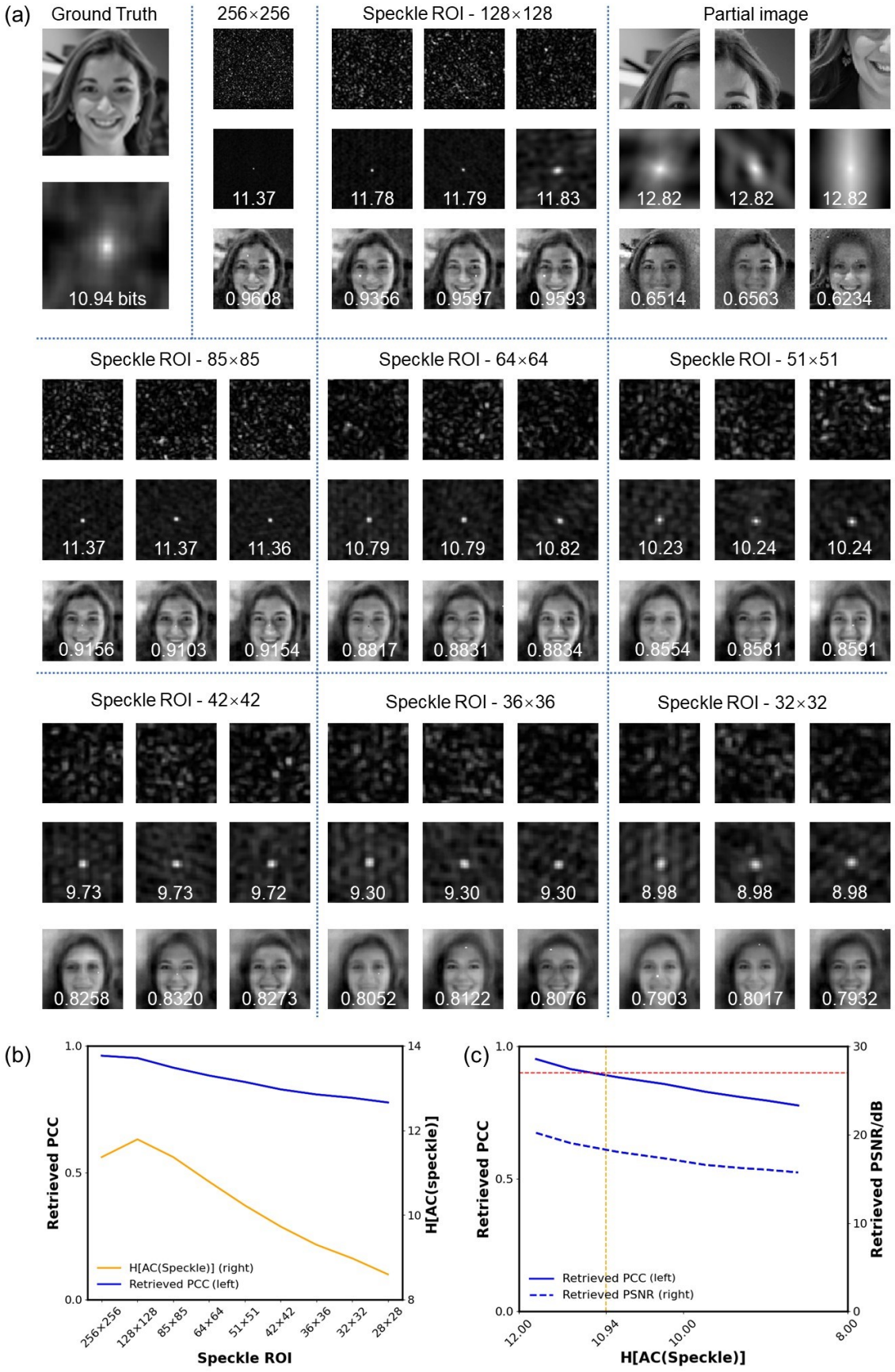

Figure 4-4 (a) Information retrieved from different ROIs within speckles: the first rows

are speckles from different ROIs, the second rows are the corresponding speckle autocorrelations, and the third rows are the corresponding information retrieved from speckles, with retrieved PCC (the PCC between retrieved information and the corresponding ground truth) and entropy of speckle autocorrelations marked on retrieved information and autocorrelations, respectively. (b) Entropy of speckle autocorrelation and corresponding retrieved PCC from different speckle ROI sizes (marked as Speckle ROI). (c) Relations between entropy of speckle autocorrelation and retrieved information quality (retrieved PCC/PSNR). Ground truth image: Copyright 2018, Deya at San Antonio Cocktail Conference, by Nan Palmero, Flickr (https://www.flickr.com/photos/nanpalmero/38756513965/); the original images are converted to grayscale, under terms of the CC-BY 2.0 license.

4.4.3 Entropy Relations in Delocalization

We have already verified the delocalized information in speckles, and will next explore the amount of information encoded in speckles. The use of entropy as a measure of uncertainty or variability associated with random variables provides a quantitative measure of the amount of

information delocalized in speckles. The quantitative analysis of the speckle autocorrelation entropy will further validate the proposed speckle sampling condition to retrieve information from speckles with high fidelity, which is evidenced by PCC greater than 0.9.

By comparing speckle autocorrelation and image autocorrelation in Figure 4-4a, the analysis reveals distinct differences in the distribution of information. Speckle autocorrelation exhibits higher amplitudes in the center regions, corresponding to short-distance correlation, indicating that speckle grains decorrelate fast in space. In contrast, image autocorrelation shows more information outside the center regions, reflecting the high correlation among different facial features.

Furthermore, the entropy of speckle autocorrelations and the entropy of the corresponding ground truth autocorrelations are summarized in Figure 4-4b, and the retrieved PCC is also summarized to verify the proposed speckle sampling condition. From the two curves (retrieved PCC in the blue solid line, and entropy of speckle in the yellow solid line), it's clear that speckles with larger ROIs include more information, as indicated by higher entropy values. Accordingly, speckles with larger ROIs lead to better fidelity in retrieved information, which is evidenced by higher retrieved PCC and more detailed structures in visual results. On the contrary, as the ROIs in speckles become smaller, the quality of the retrieved information

decreases.

As for speckle ROIs containing 256×256, 128×128 and 85×85 pixels, the entropies of speckle autocorrelations are 11.37, 11.79, and 11.37 bits, respectively. These are greater than the entropy of the ground truth image autocorrelation (*i.e.*, 10.94 bits), indicating that these speckles contain more information than the original images. As shown in Figure 4-4c, the retrieved information from speckle ROIs, where the entropy of speckle autocorrelation surpasses the entropy of image autocorrelation (yellow dashed line), exhibits high-fidelity information retrieval (PCC > 0.9, highlighted by the red dashed line). The corresponding visual results in Figure 4-4a appear satisfactory. This observation aligns with the hypothesis that neural networks can be trained to transform information distribution from speckles to that of ground truth images and retrieve information from speckles with high fidelity, as evidenced by PCC > 0.9.

Additionally, the analysis of speckle autocorrelation for smaller ROI sizes, such as 64×64 and 51×51, reveals slightly lower entropy values (10.79 bits and 10.23 bits) than the entropy of image autocorrelation, leading to slightly lower PCC for retrieved information. For even smaller ROI sizes (42×42, 36×36, 32×32, and 28×28), the entropy of speckle autocorrelation is significantly lower than the entropy of the ground truth image autocorrelation. As the obscure visual results shown in Figure 4-4a, only common facial features can be retrieved; most specific

and personal features are lost. The entropy disparity, along with the low retrieved information PCC in Figure 4-4c, indicates that some information is lost in these speckle ROIs and makes it challenging to retrieve information with high fidelity.

4.5 Discussions

Upon examining ROIs of varying sizes and positions, a pronounced distinction emerges between traditional imaging (*i.e.*, without delocalization) and speckles (*i.e.*, with delocalization). Traditional imaging, characterized by a straightforward mapping from the object to the image, encounters significant challenges in retrieving complete information from partial images, as it cannot compensate for blocked information. In contrast, speckles demonstrate remarkable results across ROIs of various sizes and locations, allowing for the retrieval of information with high fidelity. This attribute is primarily due to the inherent delocalization of information in speckles, where each point in speckles is correlated with many points in images, as evidenced by analyses conducted through the TM model.

Furthermore, we analyze the relationship between speckle autocorrelation and image autocorrelation. We propose that the entropy of speckle autocorrelation should be greater than

that of image autocorrelation, and neural networks can be trained to retrieve information from speckles with high fidelity, as evidenced by retrieved information PCC greater than 0.9. This suggests that the delocalized information in the captured ROI in the speckle is sufficient to encompass information for high-fidelity information retrieval. However, when the speckle autocorrelation entropy is much lower than that of image autocorrelation, the quality of the retrieved information diminishes significantly. Despite the ability of neural networks to retrieve information from speckles, the details are lost, and the retrieved information may not be suitable for further applications, such as face recognition. This scenario highlights the critical role of speckle autocorrelation entropy in determining the retrievability and the amount of information delocalized in speckles. As a result, using entropy as a measure to analyze information delocalization in speckles provides valuable insights into the information retrieval from speckles. The proposed speckle sampling condition underscores the potential of neural networks in speckle imaging for high-fidelity information retrieval. This not only advances the understanding of delocalized information in speckles, but also lays the groundwork for future research in this area.

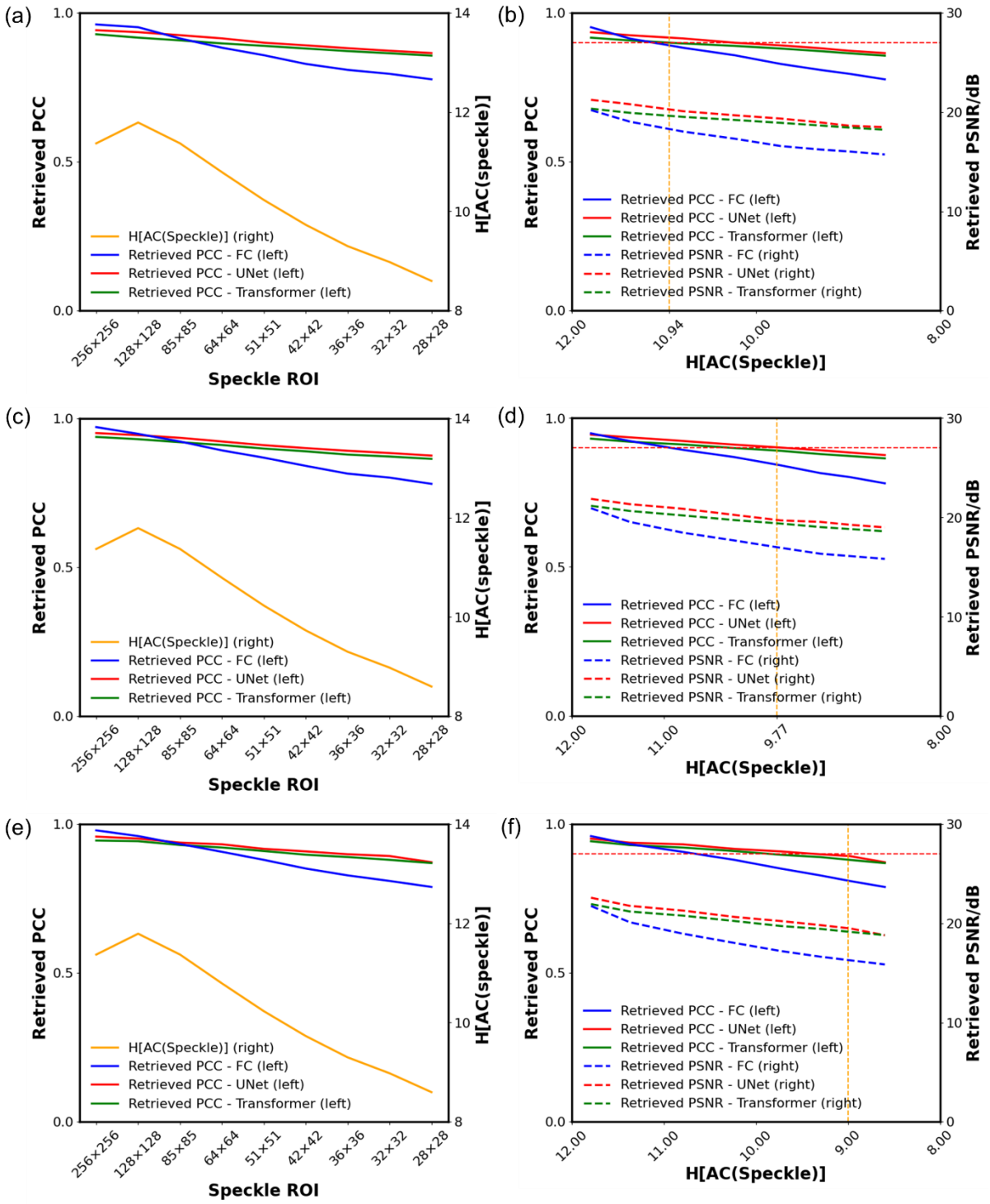

Figure 4-5 (a) Entropy of speckle autocorrelation and corresponding retrieved PCC from different speckle ROI sizes using different neural network models. (b) Relationship between entropy of speckle autocorrelation and retrieved information quality (retrieved

PCC / PSNR) using different neural network models. (c-d) Reproduction of the same experiments in (a-b) with 42×42 ground truth images. (e-f) Reproduction of the same experiments in (a-b) with 32×32 ground truth images.

To further verify the proposed speckle sampling condition under various situations, another two neural network models (the commonly used U-Net and Vision Transformer) are also tested, with results shown in Figure 4-5a and Figure 4-5b. For the U-Net models, the encoder-decoder structure is highly efficient in retrieving information from speckles, and the overall retrieved information PCC is greater than the results using complex fully connected models. The U-Net results (red solid line) almost coincide with the intersection of PCC = 0.9 and entropy of image autocorrelation in Figure 4-5b, indicating that the proposed speckle sampling condition fits well for the U-Net models. For Vision Transformer models, the overall results (green solid line) are near to the U-Net results, and the speckle sampling condition still holds well. Overall, despite the differences in the specific results obtained from different neural networks, the overall trends are consistent, reinforcing the speckle sampling condition for high-fidelity information retrieval from speckles.

Apart from 64×64 ground truth images with lots of detailed structures, the experiments in

Figure 4-5a and Figure 4-5b are further reproduced using 42×42 and 32×32 ground truth images, as demonstrated in Figure 4-5 c-f. As for the U-Net (red solid line) and Transformer (green solid line) models, the overall results in Figure 4-5 c-f still align with the intersection of PCC = 0.9 and the entropy of image autocorrelation. This validation supports the proposed speckle sampling condition for high-fidelity information retrieval from speckles.

However, when comparing these findings to those obtained using complex fully connected models (blue solid line), we observe a decline in performance for Figure 4-5 c-f. Despite the fact that the entropy of speckle autocorrelation exceeds that of image autocorrelation, the information retrieval results still fall short of PCC = 0.9 for complex fully connected models. The primary reason for this discrepancy lies in the reduced number of model parameters as the input and output dimensions decrease, leading to diminished performance in experiments involving 42×42 and 32×32 ground truth images. As a result, even when the proposed speckle sampling condition is met, the choice of models still significantly influences the retrieved information. These results further underscore that the proposed speckle sampling condition is necessary but not sufficient, as various factors can influence the retrieved information, such as optical setups, environmental disturbances, and neural network models.

From another point of view, experiments with speckles resulting from digits (*i.e.*, simple targets)

demonstrate the versatility of the proposed speckle sampling condition for retrieving various types of information from speckles. As shown in Figure 4-6, the Transformer-based networks accurately retrieve the shapes and details of the digits. Additionally, when the entropy of speckle autocorrelation is greater than that of image autocorrelation, digits retrieved from speckles can achieve high-fidelity (*i.e.*, PCC > 0.9). These cross-validations with different network models, ground truth resolutions, and targets further underscore the universality of the proposed theory. This universality is crucial for advancing the field of speckle imaging, as it suggests that the same principles can be applied across a wide range of applications, enhancing the versatility and robustness of speckle-based imaging techniques.

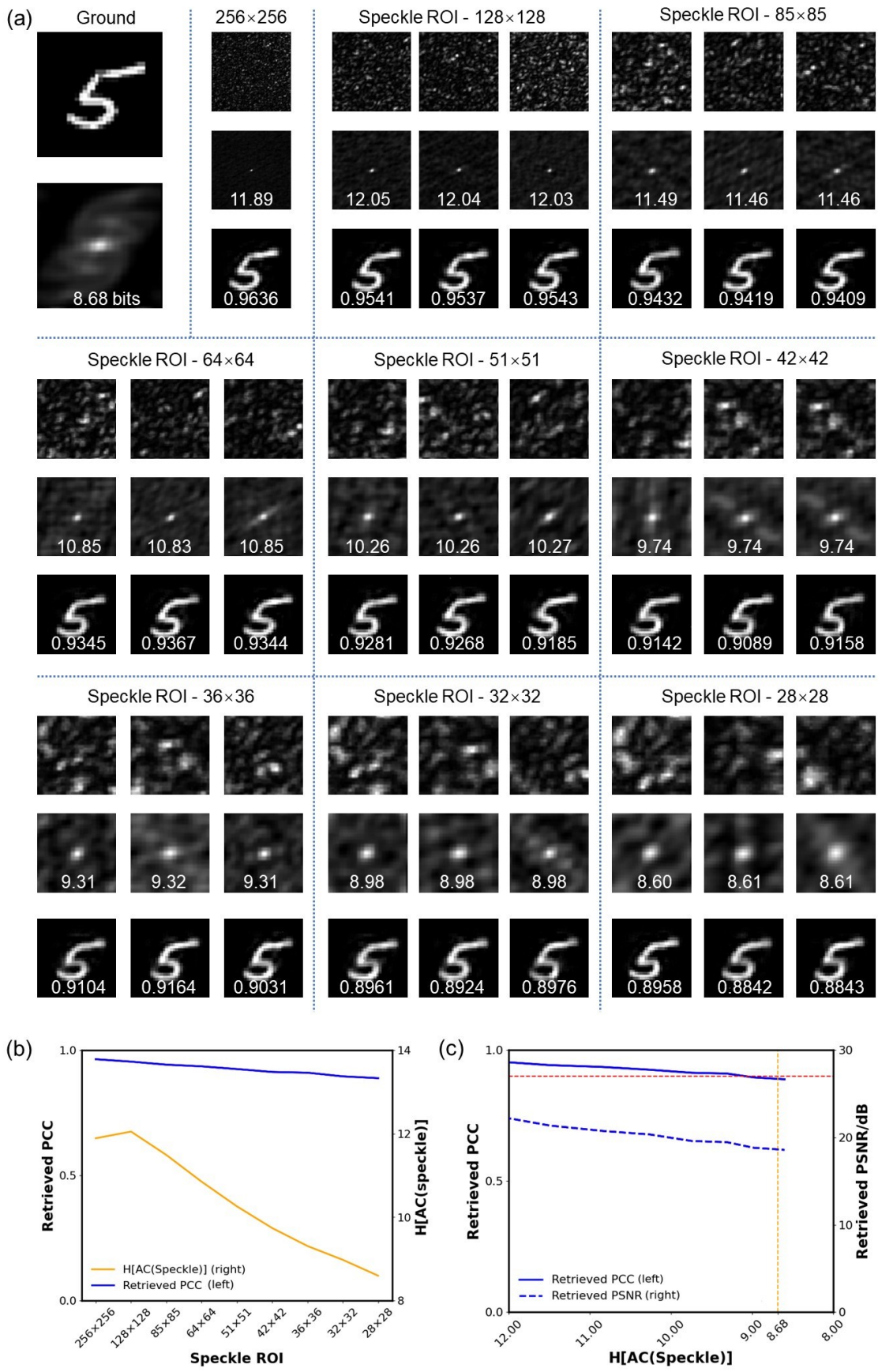

Figure 4-6 Results of information retrieval from speckles corresponding to digits. (a)

Information retrieved from delocalized information in different speckle ROIs. The first rows are speckles from different ROIs; the second rows are the corresponding speckle autocorrelations; the third rows are the corresponding information retrieved from speckles, with retrieved PCC (PCC between retrieved information and the corresponding ground truth) and entropy of speckle autocorrelations marked on retrieved information and autocorrelations, respectively. (b) Entropy of speckle autocorrelation and corresponding retrieved PCC from different Speckle ROIs. (c) Relations between entropy of speckle autocorrelation and retrieved information quality (retrieved PCC / PSNR).

Last but not least, the theoretical and experimental findings suggest potential applications for speckle imaging, such as non-line-of-sight imaging, where speckle ROIs can be reduced to utilize less data for information retrieval behind barriers [24]. Additionally, the findings could enable the development of low-cost cameras with high throughput, achieving super-resolution through compressive sampled speckles [25]. This approach could also contribute to high-level storage security, as information can still be retrieved even if some speckles are lost. Overall, the experimental results and the proposed theories provide a solid foundation for understanding how information is delocalized in speckles and under what conditions information can be

retrieved from speckles with high fidelity. These findings not only contribute to the field of speckle imaging but also open up new avenues for future research and application development.

4.6 Conclusion

The distribution of information in optical speckles has been explored through learning-based models. The theoretical models suggest that information is spatially delocalized among speckles. By analyzing both physical models and experimental results, the speckle sampling condition has been proposed. When the entropy of speckle autocorrelation exceeds that of image autocorrelation, neural networks can be trained to retrieve information from speckles with high fidelity, as evidenced by PCC greater than 0.9. The entropy relationship between speckle autocorrelation and image autocorrelation reveals the speckle sampling condition for information retrieval from speckles with high fidelity. Overall, the findings of this work not only contribute to the understanding of information delocalization in speckles, but also have the potential to inspire new research and applications in speckle-based imaging, computing, storage, *etc*.

References

1. Umar, A., & Atabo, S. (2019). A review of imaging techniques in scientific research/clinical diagnosis. *MOJ Anat Physiol*, *6*(5), 175-83.

2. Yun, S. H., & Kwok, S. J. (2017). Light in diagnosis, therapy and surgery. *Nature biomedical engineering*, *1*(1), 0008.

3. Sinclair, M. B., Haaland, D. M., Timlin, J. A., & Jones, H. D. (2006). Hyperspectral confocal microscope. *Applied optics*, *45*(24), 6283-6291.

4. König, K. (2000). Multiphoton microscopy in life sciences. *Journal of microscopy*, *200*(2), 83-104.

5. Bandres, M. A., Kaminer, I., Mills, M., Rodríguez-Lara, B. M., Greenfield, E., Segev, M., & Christodoulides, D. N. (2013). Accelerating optical beams. *Optics and Photonics News*, *24*(6), 30-37.

6. Fossum, E. R. (1997). CMOS image sensors: Electronic camera-on-a-chip. *IEEE transactions on electron devices*, *44*(10), 1689-1698.

7. Wiersma, D. S. (2013). Disordered photonics. *Nature Photonics*, *7*(3), 188-196.

8. Bertolotti, J., Van Putten, E. G., Blum, C., Lagendijk, A., Vos, W. L., & Mosk, A. P. (2012).

Non-invasive imaging through opaque scattering layers. *Nature*, *491*(7423), 232-234.

9. Gigan, S. (2017). Optical microscopy aims deep. *Nature Photonics*, *11*(1), 14-16.

10. Tian, L., Hunt, B., Bell, M. A. L., Yi, J., Smith, J. T., Ochoa, M., ... & Durr, N. J. (2021). Deep learning in biomedical optics. *Lasers in surgery and medicine*, *53*(6), 748-775.

11. Cheng, S., Li, H., Luo, Y., Zheng, Y., & Lai, P. (2019). Artificial intelligence-assisted light control and computational imaging through scattering media. *Journal of innovative optical health sciences*, *12*(04), 1930006.

12. Horisaki, R., Takagi, R., & Tanida, J. (2016). Learning-based imaging through scattering media. *Optics express*, *24*(13), 13738-13743.

13. Tang, P., Zheng, K., Yuan, W., Pan, T., Xu, Y., Fu, S., ... & Qin, Y. (2022). Learning to transmit images through optical speckle of a multimode fiber with high fidelity. *Applied physics letters*, *121*(8).

14. Wang, Y., Wang, H., & Gu, M. (2023). High performance "non-local" generic face reconstruction model using the lightweight Speckle-Transformer (SpT) UNet. *Opto-Electronic Advances*, *6*(2), 220049-1.

15. Turpin, A., Vishniakou, I., & d Seelig, J. (2018). Light scattering control in transmission and reflection with neural networks. *Optics express*, *26*(23), 30911-30929.

16. Y Li, Y., Xue, Y., & Tian, L. (2018). Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media. *Optica*, *5*(10), 1181-1190.

17. Caramazza, P., Moran, O., Murray-Smith, R., & Faccio, D. (2019). Transmission of natural scene images through a multimode fibre. *Nature communications*, *10*(1), 2029.

18. Horisaki, R., Takagi, R., & Tanida, J. (2016). Learning-based imaging through scattering media. *Optics express*, *24*(13), 13738-13743.

19. Liu, J., Yang, W., Song, G., & Gan, Q. (2023). Directly and instantly seeing through random diffusers by self-imaging in scattering speckles. *PhotoniX*, *4*(1), 1.

20. Wang, Z., Huang, H., Li, H., Chen, Z., Han, J., & Pu, J. (2023). Non-line-of-sight imaging and location determination using deep learning. *Optics and Lasers in Engineering*, *169*, 107701.

21. Cao, R., De Goumoëns, F., Blochet, B., Xu, J., & Yang, C. (2022). High-resolution non-line-of-sight imaging employing active focusing. *Nature Photonics*, *16*(6), 462-468.

22. Zhao, Q., Li, H., Yu, Z., Woo, C. M., Zhong, T., Cheng, S., ... & Lai, P. (2022). Speckle-based Optical Cryptosystem and its Application for Human Face Recognition via Deep Learning. *Advanced Science*, *9*(25), 2202407.

23. Lyu, M., Wang, H., Li, G., Zheng, S., & Situ, G. (2019). Learning-based lensless imaging

through optically thick scattering media. *Advanced Photonics, 1*(3), 036002.

24. Cao, R., De Goumoëns, F., Blochet, B., Xu, J., & Yang, C. (2022). High-resolution non-line-of-sight imaging employing active focusing. *Nature Photonics*, *16*(6), 462-468.

25. Li, H., Yu, Z., Zhao, Q., Luo, Y., Cheng, S., Zhong, T., ... & Lai, P. (2023). Learning-based super-resolution interpolation for sub-Nyquist sampled laser speckles. *Photonics Research*, *11*(4), 631-642.

26. Wu, Y., Zhou, Y., Saveriades, G., Agaian, S., Noonan, J. P., & Natarajan, P. (2013). Local Shannon entropy measure with statistical tests for image randomness. *Information Sciences*, *222*, 323-342.

27. Lu, X., Wang, Z., Zhan, Q., Cai, Y., & Zhao, C. (2024). Coherence entropy during propagation through complex media. *Advanced Photonics*, *6*(4), 046002-046002.

28. Kim, M., Choi, W., Choi, Y., Yoon, C., & Choi, W. (2015). Transmission matrix of a scattering medium and its applications in biophotonics. *Optics express*, *23*(10), 12648-12668.

29. Bromiley, P. A., Thacker, N. A., & Bouhova-Thacker, E. (2004). Shannon entropy, Renyi entropy, and information. *Statistics and Inf. Series (2004-004)*, *9*(2004), 2-8.

30. Zhu, S., Guo, E., Gu, J., Bai, L., & Han, J. (2021). Imaging through unknown scattering

media based on physics-informed learning. *Photonics Research*, *9*(5), B210-B219.

31. Leibovich, N., & Barkai, E. (2015). Aging wiener-khinchin theorem. *Physical review letters*, *115*(8), 080602.

32. Baez, J. C., Fritz, T., & Leinster, T. (2011). A characterization of entropy in terms of information loss. *Entropy*, *13*(11), 1945-1957.

33. Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4401-4410).

5 CLASSIFICATION BASED ON SPECKLES

This chapter has been prepared as the following paper and will be submitted for publication:

Zhao, Q.†, Li, H.†, Yu, Z., Li, H., Cheng, S., Huang, H., Zhong, T., Woo, C. M., Wang, Z., & Lai, P.# (2024). Speckle transformer: classification through scattering media with limited information. *In preparation*.

Speckle imaging has garnered significant research interest. However, retrieving images from speckles with limited information remains a substantial challenge. As discussed in Chapter 4, achieving high-fidelity information retrieval is particularly difficult when the information in speckles is insufficient. Accordingly, classification accuracies are expected to be low due to blurry retrieved images. To address this issue, we introduce Speckle Transformer, a vision transformer-based approach that leverages the limited information in speckles for high-accuracy classification. By directly extracting features from speckles for classification, this method circumvents the need to retrieve original image information before classification. This

enables classification based on the inherent features of speckles with limited information, achieving higher accuracy than classification after image retrieval. Furthermore, entropy analyses highlight the influences of delocalized speckle information on classification accuracies. Overall, Speckle Transformer not only overcomes the limitations of traditional methods, but also provides a new perspective on classification, opening up new avenues for research and application in speckle processing.

5.1 Introduction

Optical scattering leads to speckles rather than clear images, and it significantly hinders the application of optical technologies in deep tissues [1-3]. Recently, retrieving images from speckles has gained prominence for its potential in medical diagnoses and characterizing biological tissues [4-5]. Among the related research, transmission matrix-based models [6-8] and deep learning methods [9-12] have shown great potential for practical applications, due to their high efficiency and ease of implementation. However, to retrieve information from speckles with high fidelity, these retrieval methods require speckle sampling with sufficient information [13-14]. Additionally, speckle resolutions must be higher than those of ground truth

images to achieve near-unity fidelity [15]. Furthermore, the increased sampled information results in extended processing time for training neural networks or measuring transmission matrices, potentially impacting the efficiency of these methods in handling optical speckles. Meanwhile, some applications involving retrieved images from speckles, such as classification, recognition, and segmentation, may not necessitate very high resolution [16-17]. Additionally, achieving high speckle resolutions may be challenging due to hardware limitations, integration complexities, cost considerations, *etc*. [13]. For classification scenarios, only the target type information is required, rather than retrieving clear target images. This suggests that direct classification of information from low-resolution speckles with limited information is applicable [18-19]. Despite these advancements, a significant challenge remains in classifying information from speckles, particularly when insufficient sampling prevents the acquisition of complete speckle information [14].

To address these challenges, we introduce Speckle Transformer, a novel vision transformer-based model [20-21]. Speckle Transformer is originally designed to classify original images with high accuracies to achieve accurate classification based on the limited information available in speckles. By circumventing the need for complete speckle data to retrieve images, Speckle Transformer enables the classification of images based on the features extracted from

speckles, and achieves higher accuracies than traditional methods that rely on image retrieval before classification. Furthermore, the amount of information (*i.e.*, information entropy) in speckles is analyzed through the derivation of the entropy of speckle autocorrelation, and the influences of information in speckles on classification accuracies are discussed.

5.2 Methods

In this chapter, we integrate the widely recognized transformer model into speckle research, offering a novel approach for analyzing and classifying speckles. For the optical experiments, the optical setup in Figure 5-1a is utilized to generate speckles [22]. The experiments begin with the use of a continuous-wave 532 nm laser source (EXLSR-532-300-CDRH, Spectra-Physics, Single mode, 300 mW, USA), whose output is expanded by a 4-f system (L1 and L2) to illuminate the spatial light modulator (SLM, HOLOEYE PLUTO VIS056 1080p, German). Then the images containing digits from Modified National Institute of Standards and Technology (MNIST) dataset are displayed on the SLM to modulate the input beam. The modulated wavefronts are then focused by an objective lens (RMS20X, Olympus, Japan) and pass through a scattering medium (220-grit ground glass, diameter of 1.0 inch, DG10-220-MD,

Thorlabs, USA), generating optical speckles. Subsequently, these optical speckles are captured by a CMOS camera (FL3-U3-32S2M-CS, PointGrey, Canada), providing a tangible dataset for the following experiments. During optical experiments, we captured 20,000 image-speckle pairs, which were divided into two groups, with 19,500 pairs for network training and the other 500 pairs for network testing.

Following the optical experiments, a transformer-based neural network, *i.e.*, Speckle Transformer, as shown in Figure 5-1b, was trained to classify speckles according to the images displayed on the SLM. The input speckles are initially partitioned to fit the input requirements of Speckle Transformer [20-21]. Given the focus of classification tasks, Speckle Transformer primarily consists of a transformer encoder that incorporates multi-head attention (MHA) [23] and a multi-layer perceptron (MLP) [24]. The multi-head attention mechanism is instrumental in extracting features from speckles, effectively encoding the input speckles into extracted features. The extracted features are then utilized by the MLP for classification, yielding speckle classification outputs that align with the digits in the ground truth images. Additionally, the encoder includes normalization layers and skip connections, and the binary cross entropy function BCE is utilized as the loss function:

$$BCE = -\frac{1}{n}\sum[y_n \times \ln \hat{y}_n + (1 - y_n) \times (1 - \ln \hat{y}_n)], \tag{5-1}$$

where y is the ground truth, and $\hat{y}$ is the predicted output.

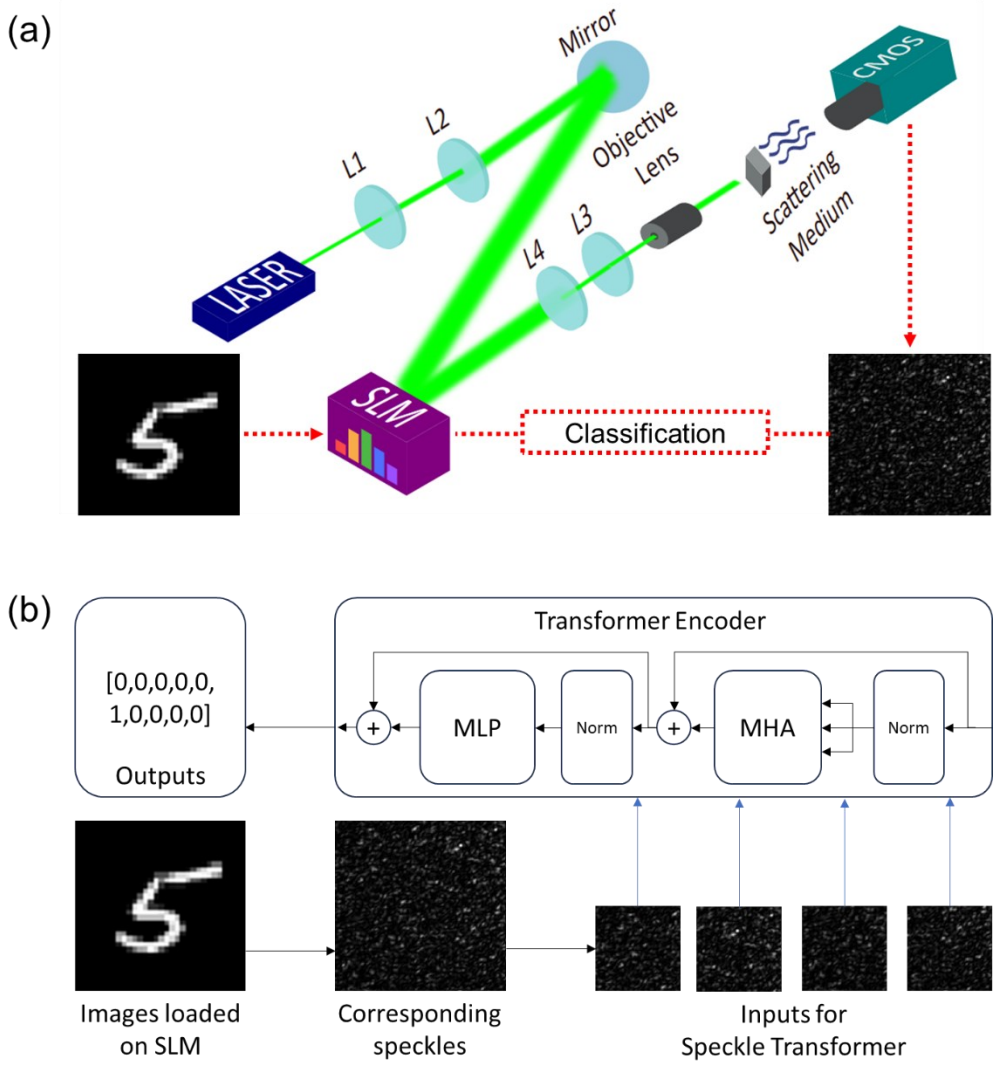

Figure 5-1 (a) Schematic of the optical setup: digits are loaded on the SLM to modulate the incident wavefront, and a CMOS camera captures the corresponding optical speckles after the scattering medium. (b) Speckle Transformer: The inputs are speckles corresponding to the images loaded on the SLM. The main block in Speckle Transformer

is Transformer Encoder, which includes a multi-head attention (MHA) and a multi-layer perceptron (MLP). The output of Speckle Transformer is the classification results corresponding to the images loaded on the SLM.

During neural network training and testing, Speckle Transformer was executed on a Dell Precision Tower 5810 workstation equipped with Intel Xeon E5-1650 V3 CPU and Nvidia GeForce RTX 3090 GPU. The employed software frameworks were PyTorch 2.0.1, CUDA 11.7, and Python 3.11.4. Additionally, during network optimization in training, the Adam optimizer with batch size = 256 was utilized [25], with a learning rate of 0.0001, complemented by cosine annealing to adjust the learning rate dynamically during training.

5.3 Results

5.3.1 Classification Based on Cropped Speckles and Images

Speckle Transformer is trained for classifying speckles through the scattering medium. During training, the speckle-label pairs were input into Speckle Transformer to tune the parameters. During testing, the classification accuracy was used to evaluate the performance of the trained

Speckle Transformer, which is defined as the ratio of correct predictions to the total number of input speckles [26-27].

First, by partitioning the speckles in Figure 5-2c into four distinct blocks, we trained Speckle Transformer to classify these speckles and output digits. The testing results, as shown in Figure 5-2d, reveal that Speckle Transformer can achieve accuracies approaching 90%. This indicates that Speckle Transformer can directly classify speckles while bypassing image retrieval, and the output results align well with the corresponding digits in the ground truth images. The potential of Speckle Transformer offers a promising alternative to traditional image retrieval methods by focusing on extracting features directly from speckles themselves, making it a valuable tool for speckle-related research and applications.

In contrast, images are also partitioned into four blocks in the same manner for comparison, as shown in Figure 5-2a and Figure 5-2b. When these partitioned images are used for training, the same Transformer model encounters difficulties in achieving satisfactory classification outcomes, with accuracies hovering around 80%. This is attributed to the loss of information in Figure 5-2b, which significantly impacts the model's ability to classify accurately. Moreover, the accuracies across different image ROIs in Figure 5-2b vary significantly. In comparison, the accuracies across different speckle ROIs in Figure 5-2d remain consistent, suggesting that

information is evenly delocalized among speckle ROIs of the same size.

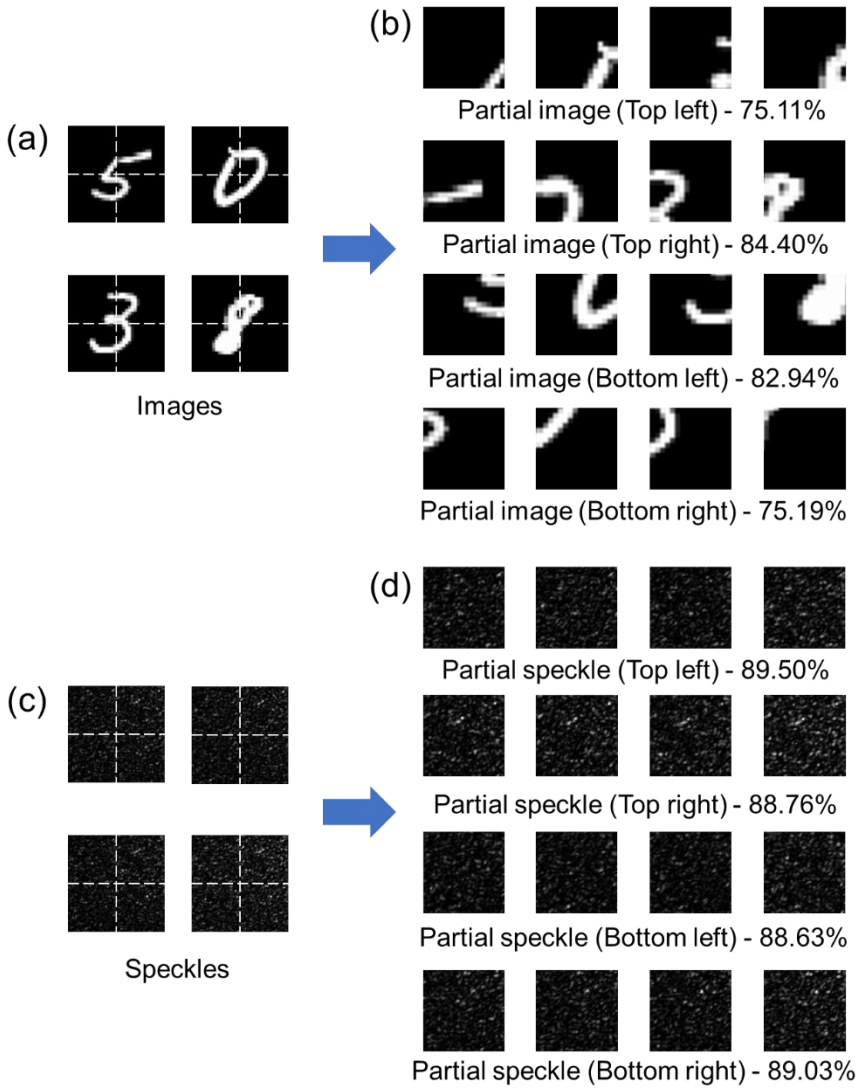

Figure 5-2 Speckle Transformer classification results. (a) Ground truth images are split into four sub-ROIs for classification. (b) Classification accuracies correspond to (a). (c) Optical speckles corresponding to (a) are also cropped into four sub-ROIs for classification. (d) Classification accuracies correspond to (c).

5.3.2 Classification Based on Retrieved Images from Speckles

In the previous experiments, the feasibility of Speckle Transformer for classification based on speckles and images has been demonstrated. To highlight the advantages of direct classification over classification after retrieval (from speckles to retrieved images), we compare the results of classifying speckles directly and classifying images retrieved from speckles (*i.e.*, classification after retrieval), as shown in Figure 5-3. The neural network employed for retrieving images from speckles is a fully connected layer based on complex algebra, while the network for classifying digits utilizes convolutional kernels and is trained on the MNIST dataset.

For full speckles and speckles divided into four ROIs (1/4 speckles), the results of direct classification and classification after retrieval exhibit minimal differences. Additionally, the images retrieved from these speckles are of high fidelity, with the PCC between the retrieved images and the corresponding ground truth images being greater than 0.9. These results are attributed to the fact that speckles contain sufficient information for retrieving digits, resulting in high-fidelity retrieved images that are crucial for achieving high accuracies in image-based classification.

However, for speckles with smaller ROIs, including 1/16, 1/64, 1/256, and 1/1024 of the full speckles, direct classification apparently outperforms classification after retrieval. Notably, for

1/1024 of full speckles, the accuracy of classification after retrieval is only slightly above 10%, suggesting that the classifier just produces some random results, and the retrieved images are actually not classifiable. These outcomes are due to the limited information available in smaller speckles for retrieving high-fidelity images. Furthermore, the loss function used for retrieving digits from speckles focuses on overall image similarity, which may not effectively capture key points in digits necessary for high-accuracy classification. Consequently, the retrieved images are of low quality and difficult to classify, leading to lower accuracies in subsequent image-based classification tasks.

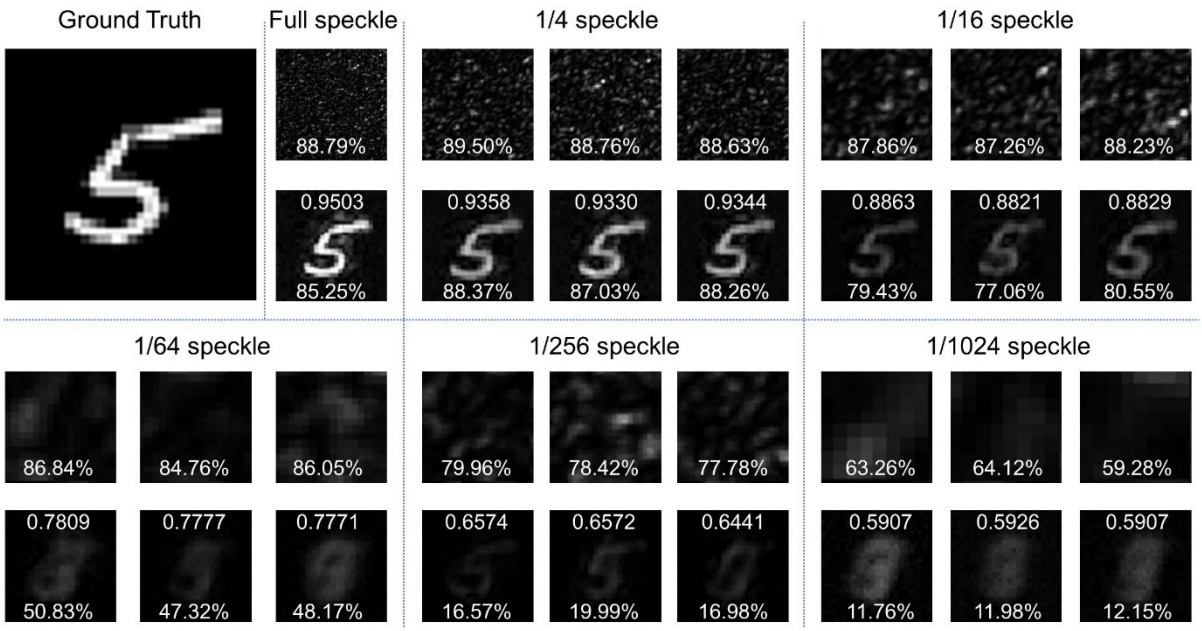

Figure 5-3 Direct speckle classification vs. classification after image retrieval: The first row is for direct classification based on speckles, including speckles from different ROIs.

The corresponding classification accuracies are marked on speckles. The second row is classification based on retrieved images from the corresponding speckles. The similarities between retrieved images and ground truths (*i.e.*, Pearson correlation coefficients, PCC) and the corresponding classification accuracies of the retrieved images (*i.e.*, classification after retrieval) are marked on each retrieved image.

5.3.3 Comparisons

The comparison results in Figure 5-3 are further elaborated in Figure 5-4. The curves in Figure 5-4a clearly illustrate that as the sizes of speckle ROIs decrease, Speckle Transformer (*i.e.*, direct classification) consistently outperforms its counterparts (*i.e.*, classification after retrieval). Moreover, when comparing different speckle ROIs of the same size, small variations (*i.e.*, shadows of the curves) in the results of Speckle Transformer are observed. The same minor variations are also observed in classification after retrieval. These findings underscore the uniform distribution of information across different speckle ROIs of the same size.

Conversely, when classifying based on digit images (no speckles are involved), the results in Figure 5-4b vary significantly across different ROIs of the same size, regardless of whether

direct classification (based on cropped images) or classification after retrieval (based on retrieved images) is employed. This discrepancy is attributed to the fact that each ROI contains only a portion of digit information, leading to low accuracies in the classification outcomes. These results highlight the distinct advantages of Speckle Transformer in classification, particularly in scenarios where limited information is encoded or recorded in speckles.

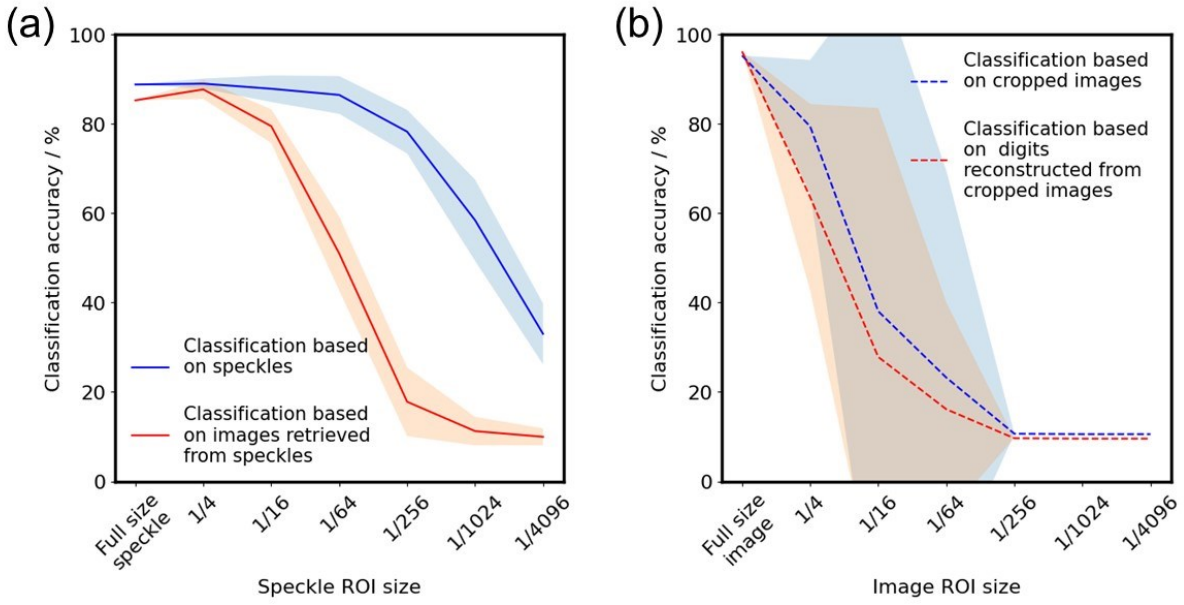

Figure 5-4 Experimental results of direct classification and classification after retrieval: (a) Classification based on cropped speckles of different ROIs (*i.e.*, direct classification) and classification based on images retrieved from speckles of different ROIs (*i.e.*, classification after retrieval). (b) Classification based on cropped images of different ROIs (*i.e.*, direct classification) and classification based on images retrieved from cropped images of different ROIs (*i.e.*, classification after retrieval).

5.4 Discussions

In this chapter, we introduce Speckle Transformer, an approach designed for direct classification of speckles without the need to retrieve images from speckles. Our experiments demonstrate that direct classification based on speckles, facilitated by Speckle Transformer, outperforms classification after information retrieval, especially when there is limited information in the speckles.

From a different perspective, we have previously mentioned that smaller speckle ROIs contain insufficient information for retrieving digits with high fidelity. To quantify the amount of information in different speckle ROIs, we employ information entropy to analyze speckle autocorrelation. The speckle autocorrelation is derived according to Wiener-Khinchin theorem, which states that autocorrelation is the inverse Fourier transform of the power spectrum [28]. We choose autocorrelation because the input wavefront in optical experiments is distorted in the Fourier plane, and autocorrelation contains information about the Fourier transform. We then compare the entropy of speckle autocorrelation with the entropy of image autocorrelation, as depicted in Figure 5-5.

The entropy of ground truth image autocorrelation is 8.84 bits. For full speckles, the entropy of speckle autocorrelation is 13.71 bits, surpassing the entropy of image autocorrelation. This

indicates that the information encoded in speckles is sufficient, which facilitates high-accuracy digit classification. Similarly, for 1/4 speckles and 1/16 speckles, the entropies of speckle autocorrelation are 12.64 bits and 10.95 bits, respectively, both exceeding the entropy of image autocorrelation. Consequently, the overall classification accuracies are acceptable due to the ample information encoded in speckles. For 1/64 speckles, the entropy of speckle autocorrelation (9.00 bits) is only slightly higher than the entropy of image autocorrelation (8.84 bits). The information encoded in 1/64 speckles is just sufficient, leading to classification accuracies slightly lower than 90%. However, for 1/256 speckles and 1/1024 speckles, the encoded information is significantly reduced, resulting in much lower classification accuracies (about 80% and 60%, respectively).

Overall, the entropy of speckle autocorrelation can serve as a critical parameter for evaluating information in speckles and has the potential for classification with high accuracy. These findings underscore the importance of considering the entropy of speckle autocorrelation in speckle-related scenarios. Furthermore, this knowledge is crucial for advancing speckle-related research and applications, ensuring that the full potential of information in speckles is harnessed.

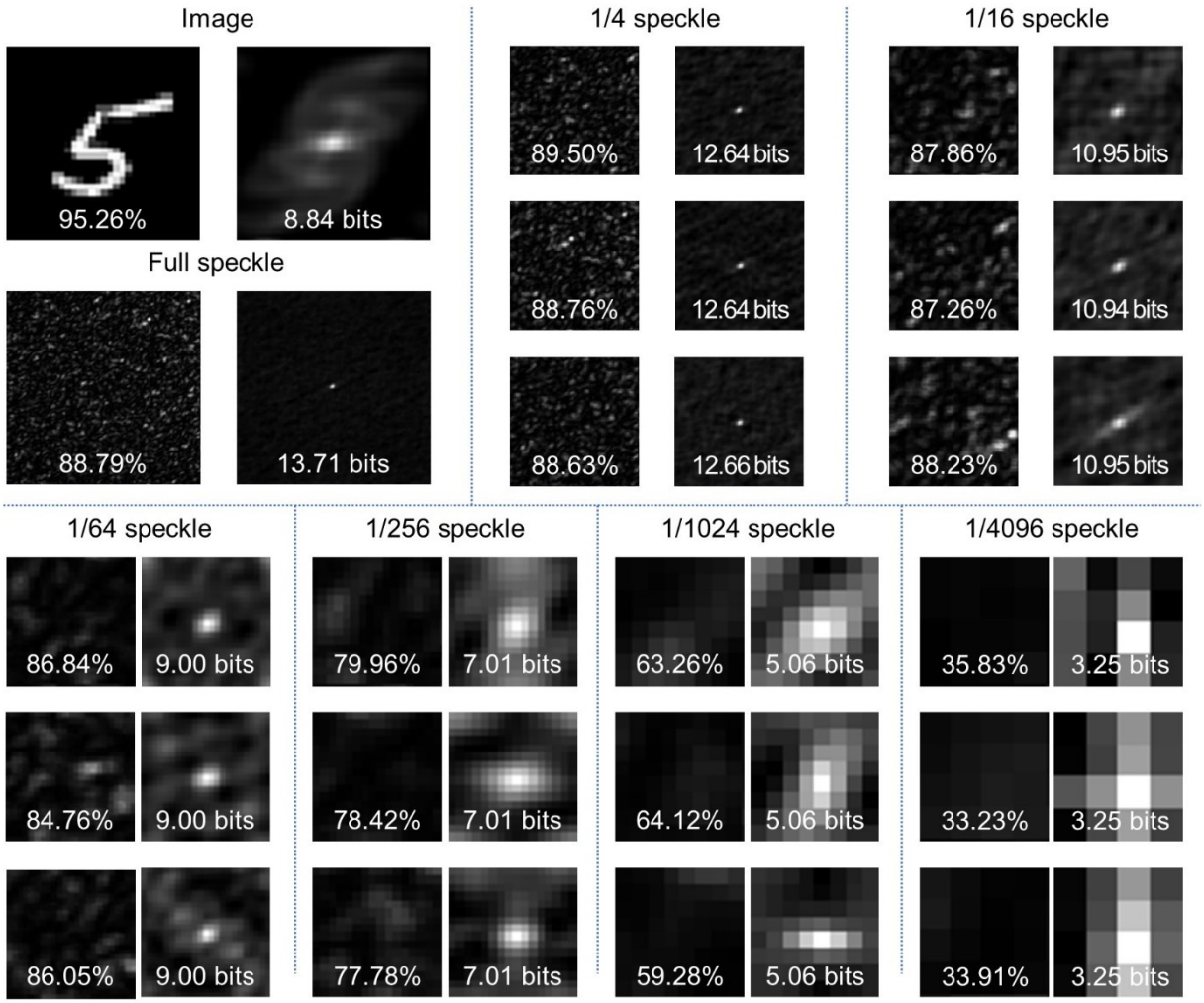

Figure 5-5 Entropy of different speckle ROIs: Speckles (including full speckle, 1/4 speckle, 1/16 speckle, 1/64 speckle, 1/256 speckle, 1/1024 speckle, and 1/4096 speckle) are shown in the left columns. The corresponding speckle autocorrelations are shown in the right columns. The speckle classification accuracies and entropies of speckle autocorrelation are marked on speckles and autocorrelations, respectively.

5.5 Conclusions

In this chapter, we propose Speckle Transformer to facilitate information classification based on speckles directly. The experimental findings indicate that direct classification, enabled by Speckle Transformer, can achieve higher accuracies compared to classification after image retrieval, particularly in scenarios where speckles contain insufficient information for high-fidelity image retrieval. Furthermore, entropy analyses highlight the influences of limited speckle information on classification accuracies. Overall, Speckle Transformer offers a promising alternative to traditional techniques, circumventing the need for complete speckle data to retrieve images and leveraging the inherent features of optical speckles for direct classification. In the future, Speckle Transformer holds significant potential for application to non-line-of-sight classifications and privacy-protected classifications.

References

1. Boas, D. A., & Dunn, A. K. (2010). Laser speckle contrast imaging in biomedical optics. *Journal of biomedical optics*, *15*(1), 011109-011109.

2. Yu, Z., Li, H., Zhong, T., Park, J. H., Cheng, S., Woo, C. M., ... & Lai, P. (2022). Wavefront

shaping: a versatile tool to conquer multiple scattering in multidisciplinary fields. *The Innovation*, *3*(5).

3. Cheng, S., Li, H., Luo, Y., Zheng, Y., & Lai, P. (2019). Artificial intelligence-assisted light control and computational imaging through scattering media. *Journal of innovative optical health sciences*, *12*(04), 1930006.

4. Park, J. H., Yu, Z., Lee, K., Lai, P., & Park, Y. (2018). Perspective: Wavefront shaping techniques for controlling multiple light scattering in biological tissues: Toward in vivo applications. *APL photonics*, *3*(10).

5. Gigan, S., Katz, O., De Aguiar, H. B., Andresen, E. R., Aubry, A., Bertolotti, J., ... & Yılmaz, H. (2022). Roadmap on wavefront shaping and deep imaging in complex media. *Journal of Physics: Photonics*, *4*(4), 042501.

6. Popoff, S. M., Lerosey, G., Carminati, R., Fink, M., Boccara, A. C., & Gigan, S. (2010). Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media. *Physical review letters*, *104*(10), 100601.

7. Wang, Z., Wu, D., Huang, G., Luo, J., Ye, B., Li, Z., & Shen, Y. (2021). Feedback-assisted transmission matrix measurement of a multimode fiber in a referenceless system. *Optics Letters*, *46*(22), 5542-5545.

8. Kim, M., Choi, Y., Yoon, C., Choi, W., Kim, J., Park, Q. H., & Choi, W. (2012). Maximal energy transport through disordered media with the implementation of transmission eigenchannels. *Nature photonics*, *6*(9), 581-585.

9. d'Arco, A., Xia, F., Boniface, A., Dong, J., & Gigan, S. (2022). Physics-based neural network for non-invasive control of coherent light in scattering media. *Optics Express*, *30*(17), 30845-30856.

10. Borhani, N., Kakkava, E., Moser, C., & Psaltis, D. (2018). Learning to see through multimode fibers. *Optica*, *5*(8), 960-966.

11. Caramazza, P., Moran, O., Murray-Smith, R., & Faccio, D. (2019). Transmission of natural scene images through a multimode fibre. *Nature communications*, *10*(1), 2029.

12. Luo, Y., Yan, S., Li, H., Lai, P., & Zheng, Y. (2021). Towards smart optical focusing: deep learning-empowered dynamic wavefront shaping through nonstationary scattering media. *Photonics Research*, *9*(8), B262-B278.

13. Zhao, Q., Li, H., Yu, Z., Woo, C. M., Zhong, T., Cheng, S., ... & Lai, P. (2022). Speckle-based Optical Cryptosystem and its Application for Human Face Recognition via Deep Learning. *Advanced Science*, *9*(25), 2202407.

14. Li, H., Yu, Z., Zhao, Q., Luo, Y., Cheng, S., Zhong, T., ... & Lai, P. (2023). Learning-based

super-resolution interpolation for sub-Nyquist sampled laser speckles. *Photonics Research*, *11*(4), 631-642.

15. Ye, J., Pan, T., Zheng, K., Luo, Z., Xu, Y., Fu, S., ... & Qin, Y. (2023). Light field information transmission through scattering media with high fidelity. *Chinese Optics Letters*, *21*(12), 121101.

16. Raskatla, V., Singh, B. P., Patil, S., Kumar, V., & Singh, R. P. (2022). Speckle-based deep learning approach for classification of orbital angular momentum modes. *JOSA A*, *39*(4), 759-765.

17. Castilho, V. M., Balthazar, W. F., da Silva, L., Penna, T. J. P., & Huguenin, J. A. O. (2023). Machine learning classification of speckle patterns for roughness measurements. *Physics Letters A*, *468*, 128736.

18. Li, Z., Zhang, L., Zhang, Z., Xu, R., & Zhang, D. (2022). Speckle classification of a multimode fiber based on Inception V3. *Applied Optics*, *61*(29), 8850-8858.

19. Wang, P., & Di, J. (2018). Deep learning-based object classification through multimode fiber via a CNN-architecture SpeckleNet. *Applied optics*, *57*(28), 8258-8263.

20. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image

recognition at scale. *arXiv preprint arXiv:2010.11929*.

21. Beyer, L., Zhai, X., & Kolesnikov, A. (2022). Better plain vit baselines for imagenet-1k. *arXiv preprint arXiv:2205.01580*.

22. Zhao, Q., Woo, C. M., Li, H., Zhong, T., Yu, Z., & Lai, P. (2021). Parameter-free optimization algorithm for iterative wavefront shaping. *Optics Letters*, *46*(12), 2880-2883.

23. Cordonnier, J. B., Loukas, A., & Jaggi, M. (2020). Multi-head attention: Collaborate instead of concatenate. *arXiv preprint arXiv:2006.16362*.

24. Kruse, R., Mostaghim, S., Borgelt, C., Braune, C., & Steinbrecher, M. (2022). Multi-layer perceptrons. In *Computational intelligence: a methodological introduction* (pp. 53-124). Cham: Springer International Publishing.

25. Zhang, Z. (2018, June). Improved adam optimizer for deep neural networks. In *2018 IEEE/ACM 26th international symposium on quality of service (IWQoS)* (pp. 1-2). Ieee.

26. Grover, D., & Toghi, B. (2020). MNIST dataset classification utilizing k-NN classifier with modified sliding-window metric. In *Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 2 1* (pp. 583-591). Springer International Publishing.

27. Yalug, B. B., Arslan, D. B., & Ozturk-Isik, E. (2021). Prospect of data science and artificial

intelligence for patient-specific neuroprostheses. In *Somatosensory Feedback for Neuroprosthetics* (pp. 589-629). Academic Press.

28. Leibovich, N., & Barkai, E. (2015). Aging wiener-khinchin theorem. *Physical review letters*, *115*(8), 080602.

6 SPECKLE-BASED OPTICAL CRYPTOSYSTEM

This chapter is modified based on a published paper in a peer-reviewed journal:

Zhao, Q.[†], Li, H.[†], Yu, Z.[†], Woo, C. M., Zhong, T., Cheng, S., Zheng, Y., Liu, H., Tian, J.[#], & Lai, P.[#] (2022). Speckle-based optical cryptosystem and its application for human face recognition via deep learning. *Advanced Science*, *9*(25), 2202407.

In the preceding chapters, we focused on overcoming, understanding, and interpreting optical speckles. However, from a different perspective, the inherent randomness of speckles makes them ideal candidates for use as ciphertexts in cryptosystems. Recently, face recognition has become ubiquitous for authentication or security purposes. Concurrently, concerns about the privacy of face images, which are sensitive biometric data, have increased. While software-based cryptosystems are widely adopted to encrypt face images, their security is often limited by insufficient digital secret key length or computing power. In contrast, hardware-based optical cryptosystems can generate enormously longer secret keys and enable encryption at the speed

of light. However, most reported optical methods, such as double random phase encryption, are less compatible with other systems due to system complexity.

In this chapter, we propose and implement a plain yet highly efficient speckle-based optical cryptosystem. We utilize a scattering ground glass to generate speckles that supports physical secret keys of 17.2 gigabit length, enabling the encryption of face images through seemingly random optical speckles at light speed. These encrypted face images can then be decrypted from the random speckles by a well-trained decryption neural network, achieving face recognition with up to 98% accuracy.

6.1 Introduction

The human face is a personal identifier, and an adult can hardly change the appearance. In modern society, numerous face recognition scenes have been set up for authentication or security purposes, due to the increasing concern for personal privacy and public safety [1]. The storage of human face data is therefore highly confidential. If the face database is leaked, hackers may use this information to attack key sectors, including bank accounts [2]. Therefore, effective protection of face image data is essential for privacy and security [3].

Various cryptosystems, including software-based and hardware-based, have been put forward to protect private data. For software-based cryptosystems, well-known encryption algorithms have been developed, such as Rivest-Shamir-Adleman encryption (RSA) [4], Advanced Encryption Standards (AES) [5], Message Digest Algorithm (MD5) [6], *etc*. These algorithms are all based on mathematical theories, whose digital secret key lengths range from tens to hundreds of bits. The selection of the secret key lengths involves a trade-off or balance between security level and processing speed. Such a limited key length seems to be sufficiently secure for conventional attacks by general computers, but is vulnerable to attacks by the rapidly evolving quantum computers, whose computing power is 10^8 times that of the general ones [7]. As a result, researchers keep exploiting novel cryptosystems to achieve higher security, and hardware-based solutions are therefore in demand.

Amongst current hardware-based solutions, speckle-based cryptosystems are of extensive interest with the development of optical computing and computational imaging [8-9], due to their superior performance, such as fast speed, high security, low cost, *etc*. [10] In speckle-based cryptosystems, optical speckles are utilized as ciphertext to encrypt plaintext. The random feature of the speckles seems meaningless and is usually annoying, but speckles constitute nearly infinite information channels, leading to the tremendously long physical secret key

length in a cryptosystem [11]. Accordingly, speckles can be exploited to yield high-level security and information protection. Thus far, a few methods, such as transmission matrix, support vector regression, deep neural networks, *etc.*, have been developed to retrieve images from speckles [11-13]. Among these approaches, neural networks can automatically learn the complex relations between the plaintext and the ciphertext, resulting in image retrieval with higher fidelity than other methods [14-19]. Since the physical models in speckle-based optical cryptosystems are similar to those for imaging through scattering media, neural networks can also be applied in speckle-based optical cryptosystems to decrypt speckles for high-level applications, such as face recognition. The main challenge here is to decrypt images from rapidly changing optical speckles and to recognize faces in the decrypted images. Moreover, to achieve high accuracy in face recognition, decryption with high fidelity in key features and detailed structures is required. In this chapter, we propose a scheme that utilizes optical speckles for face image encryption and a deep neural network for speckle decryption, and the decrypted images are then used for face recognition.

The concept, as illustrated in Figure 6-1, can be decomposed into three stages: first, face images are optically scrambled into speckles for encryption, which protects the data during transmission and storage; then, a neural network is trained to decrypt the face images with high

fidelity from the ciphertext (*i.e.*, speckles); last, the decrypted images are compared with the known face encodings for recognition. In the cryptosystem, face images are encrypted into seemingly random speckles that are nearly impossible to be decrypted without the knowledge of the physical key (*i.e.*, the scattering medium) or the learned digital key (*i.e.*, the trained neural network). Moreover, only speckles but no face images are stored in the database to avoid any potential private information leakage. To the best of our knowledge, this is the first demonstration of a speckle-based optical cryptosystem for face recognition, and the accuracy has reached more than 98%, which is applicable in a wide range of applications.

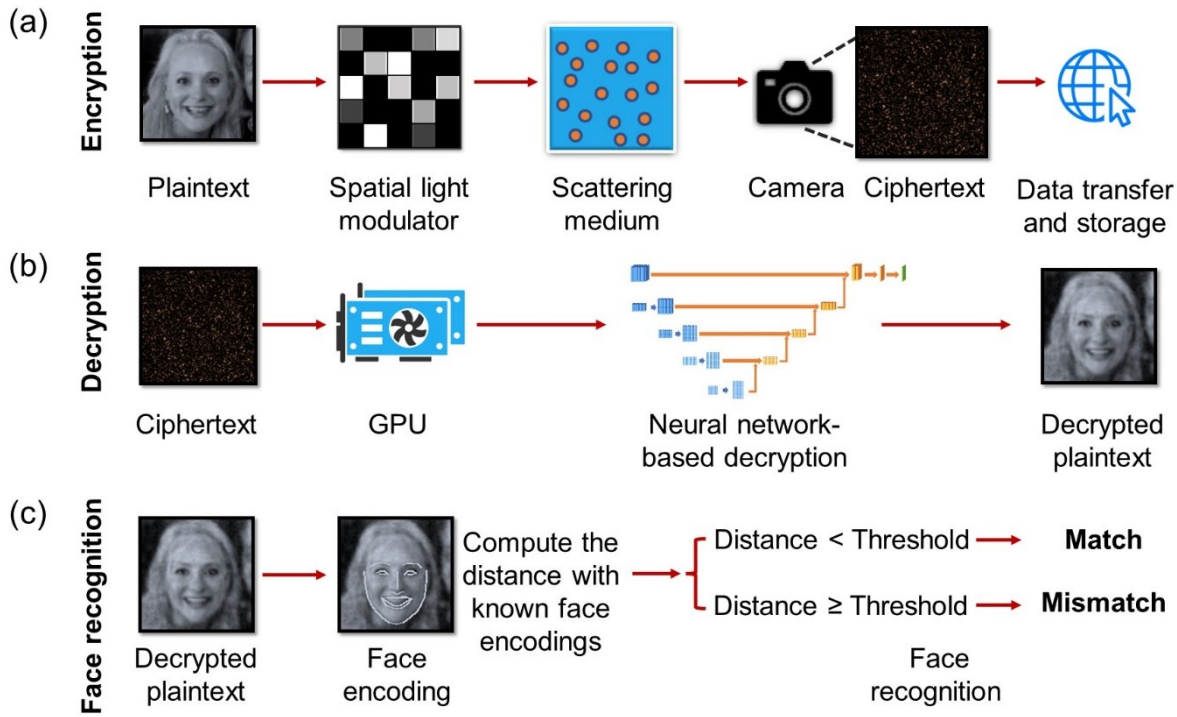

Figure 6-1 The flowchart of the proposed cryptosystem for face recognition. (a) Speckle

encryption: face images (plaintext) are loaded on a spatial light modulator (SLM) to generate the corresponding speckles (ciphertext) when coherent light is reflected by the SLM and transmits through a scattering medium. The ciphertext is safely transferred and stored via the cloud, and no face images need to be kept in the database after encryption. (b) Learning-based decryption: a neural network is trained in advance to link the plaintext with the ciphertext. After training, new random speckles (ciphertext) are directly fed into the neural network for decryption, and the decrypted face images are then utilized for face recognition. (c) Face recognition: the camera-recorded face images are encoded to unique 128-dimensional vectors of each known face image. After decryption, the face encoding distances between the decrypted images and the known face encodings are computed. If the encoding distance is less than a pre-set threshold, the face recognition result is Match (the same person), otherwise, it is Mismatch (different people). Plaintext image: Reproduced under terms of the CC-BY 2.0 license. Copyright 2015, Lawrence Lessig at Second Home London, by Innotech Summit, Flickr (https://www.flickr.com/photos/115363358@N03/18260388752/). The original image is cropped and converted to grayscale.

6.2 Results

6.2.1 Speckle-based Encryption

Figure 6-2 shows the experimental optical setup for information encryption. Face images with a resolution of 128×128 from the thumbnails of the Flickr Faces High Quality (FFHQ) database [21] are displayed on a phase-modulating spatial light modulator (SLM, HOLOEYE PLUTO VIS056 1080p, German) to modulate the incident coherent light from a 532 nm single mode laser source (300 mW, EXLSR-532-300-CDRH, Spectra-Physics, USA). Thus, the information of the face images (*i.e.*, plaintext) is carried by the wavefront modulated laser beam. Then, the modulated wavefront is focused by an objective lens (RMS20X, Olympus, Japan) onto and passes through a scattering medium (220-grit ground glass, diameter of 1.0 inch, DG10-220-MD, Thorlabs, USA). Accordingly, the wavefront is multiply scattered to form random speckles (*i.e.*, ciphertext), which are captured by a digital camera (FL3-U3-32S2M-CS, PointGrey, Canada) with a resolution of 256×256. During encryption, which is the process of generating speckles, a MATLAB program synchronizes all devices to ensure each captured speckle (*i.e.*, ciphertext) is paired with one exclusive face image (*i.e.*, plaintext) displayed on the SLM. As seen, the ciphertext appears random and exhibits no direct relationship with the plaintext, and the mean Pearson correlation coefficient (PCC) between them is as low as 0.02.

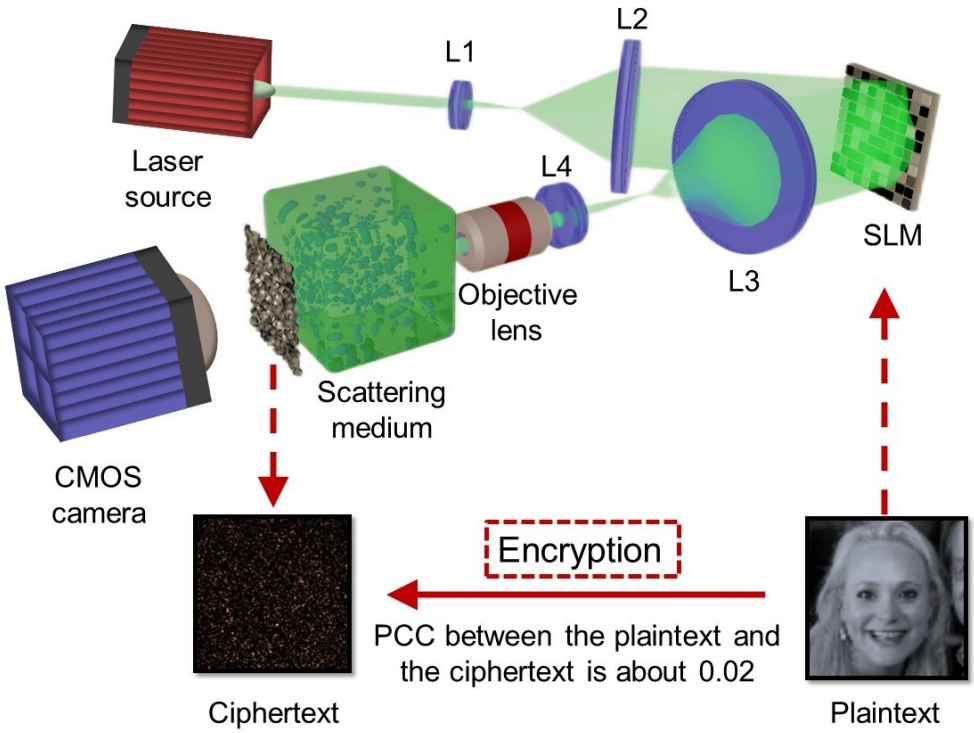

Figure 6-2 The optical setup for encryption. Face images (plaintext) are displayed on the SLM, generating speckles (ciphertext) through a scattering medium. The speckles are recorded by a CMOS camera, which is synchronized by a Matlab program to ensure one-to-one mapping with the displayed face image for network training. Plaintext image: Reproduced under terms of the CC-BY 2.0 license. Copyright 2015, Lawrence Lessig at Second Home London, by Innotech Summit, Flickr (https://www.flickr.com/photos/115363358@N03/18260388752/). The original image is cropped and converted to grayscale.

6.2.2 Learning-based Decryption

For information decryption from speckles, a neural network is constructed first. The structure of the neural network is shown in Figure 6-3a, which is a U-Net [22] concatenated with a complex fully connected layer [15] and a normalization layer. The encoders in the U-Net contained 4 down-sampling blocks, and the decoders in the U-Net contained 4 up-sampling blocks. In addition, the fully connected layer was based on complex numbers. In Figure 6-3a, the blue arrows and filters represented the encoders in the U-Net, and the orange arrows and filters represented the decoders in the U-Net. The encoder tended to extract low-dimensional features from the speckles and encode them, and the decoder then tended to extract high-dimensional features and decode them [22]. As a result, the encoder and decoder neural network could extract features of different dimensions. The fully connected layer was used as the last layer to transform the extracted features into images. The normalization layer limited the output range to [0,1]. At last, the final output was the face images decrypted from random speckles.

During network training, the speckles used as the network input were 256×256 speckle images captured by the CMOS camera, and the images used as the network output were 64×64 images that were down-sampled from the FFHQ dataset (128×128) to avoid using up the GPU memory [21]. These resolutions were chosen to make full use of the experimental setup and achieve

high-fidelity image decryption. Before the speckle data was input to the neural network, the input data were linearly normalized to [0,1] for better neural network performance [23]. Then the neural network is trained with 19,800 pairs of face images and their corresponding speckles. Additionally, the loss function L used for training the neural network is:

$$L = MSE(y, \hat{y}) - PCC(y, \hat{y}) \tag{6-1}$$

$$PCC = \frac{mean[(y - mean\ (y)) \times (\hat{y} - mean\ (\hat{y}))]}{std\ (y) \times std\ (\hat{y})} \tag{6-2}$$

$$MSE = mean\ [(\hat{y} - y)^2] \tag{6-3}$$

where y is the ground truth, and $\hat{y}$ is the predicted output from the neural network. Here, we adopt PCC to measure the overall similarity and mean square error (MSE) to measure the pixel-wise error. The optimizer used in training the neural network was stochastic gradient descent (SGD) [24] with batch size = 3, and the learning rate was 0.15 with cosine annealing. During the experiments, the neural network was trained for 30 epochs, and the neural network was then tested. The software framework used was Pytorch 1.8.0 with Python 3.7.6 and Compute Unified Device Architecture (CUDA) 10.1 for GPU acceleration. The hardware used was Dell Precision Tower 5810 with Intel Xeon E5-1650 V3 CPU, 64 GB RAM, and Nvidia GeForce RTX 2080Ti 11GB GPU. During the training, one epoch took about 30 min, and the whole training process

takes about 15 h.

The experimental results of the neural network are shown in Figure 6-3b. Apart from PCC, we also measure other commonly used criteria, including the peak signal to noise ratio (PSNR) and the structural similarity index measure (SSIM), respectively. In Figure 6-3b, four groups of exampled plaintexts, ciphertexts, and decrypted images during network testing are shown. The PCC, MSE, SSIM, and PSNR between the decrypted images and the original plaintexts are marked under the decrypted images. Overall, the average PCC, MSE, SSIM, and PSNR among all the testing data with 100 image-speckle pairs are 0.9422, 0.0083, 0.6884, and 21.25, respectively, demonstrating high accuracy of information decryption, which is essential for face recognition in the next stage. After network training, the plaintexts can be deleted from the cryptosystem to avoid privacy data leakage.

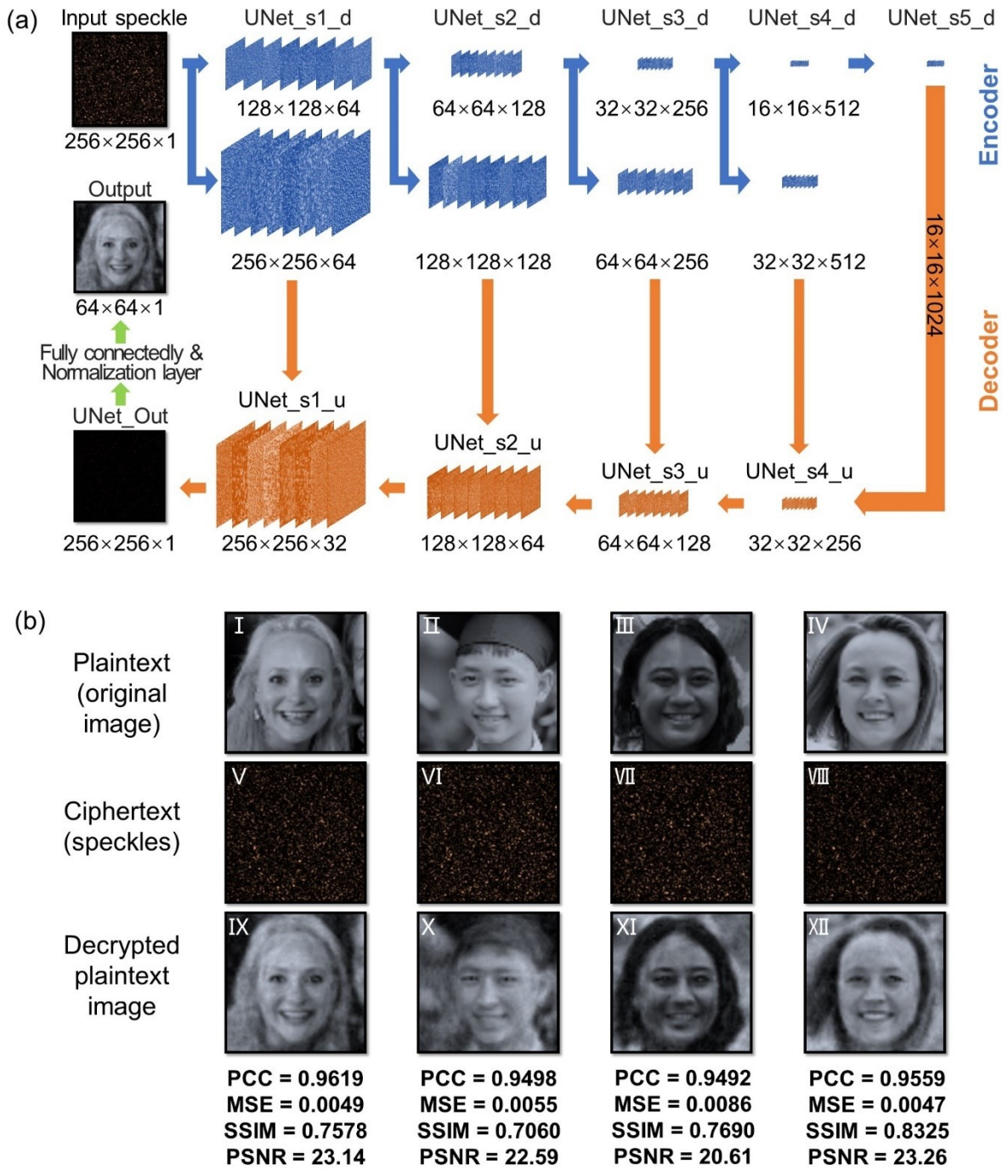

Figure 6-3 Neural network structure and the decryption performance. (a) Architectures of the neural network based on U-Net with an additional layer of a complex fully connected layer and normalization layer. The U-Net mainly contains 4 layers, with 4 down-sampling blocks for encoders (marked in blue) and 4 up-sampling blocks for

decoders (marked in orange) [22]. The final outputs are face images decrypted from speckles, which are then used for face recognition. The dimensions of the filters are described as length × height × amount, and the filters shown here are visualized by inputting one speckle into the neural network. (b) Four groups of exampled plaintexts, ciphertexts, and decrypted plaintext images during network testing. The ciphertexts are all from the same scattering medium, and the decrypted plaintext images are the results of inputting ciphertexts to the pre-trained neural network for decryption. The PCC, MSE, SSIM, and PSNR between the decrypted and original images are marked under the corresponding decrypted images. Plaintext image I: Reproduced under terms of the CC-BY 2.0 license. Copyright 2015, Lawrence Lessig at Second Home London, by Innotech Summit, Flickr (https://www.flickr.com/photos/115363358@N03/18260388752/). Plaintext image II: Reproduced under terms of the Public Domain Mark 1.0 license. Copyright 2018, kỉ yếu 12c, by khanhkhokhao201, Flickr (https://www.flickr.com/photos/154663983@N08/28538465128/). Plaintext image III: Reproduced under terms of the Public Domain Mark 1.0 license. Copyright 2016, Future Leaders of the Pacific 2016 by US Embassy, Flickr (https://www.flickr.com/photos/us_embassy_newzealand/29355772191/). Plaintext image IV: Reproduced under terms of the CC-BY 2.0 license.

Besides, the noise-resisting ability of the network is examined since noise always exists in experiments due to environmental disturbances, vibration, airflow, *et al*. [16]. Here, some computer-generated Gaussian noises with different standard deviations [25] are added to the speckles for testing, and the decryption performance is updated with the pre-trained neural network. The results are given in Figure 6-4a, where the quality of the decrypted images deteriorates considerably when the standard deviation of the noise is greater than or equal to 0.5 (*i.e.*, the noise amplitude is half of the mean of the signal amplitude), and the face outline becomes indistinct. These results suggest that the neural network can handle low and moderate noise conditions to the testing data, which is meaningful to the applicability of the method.

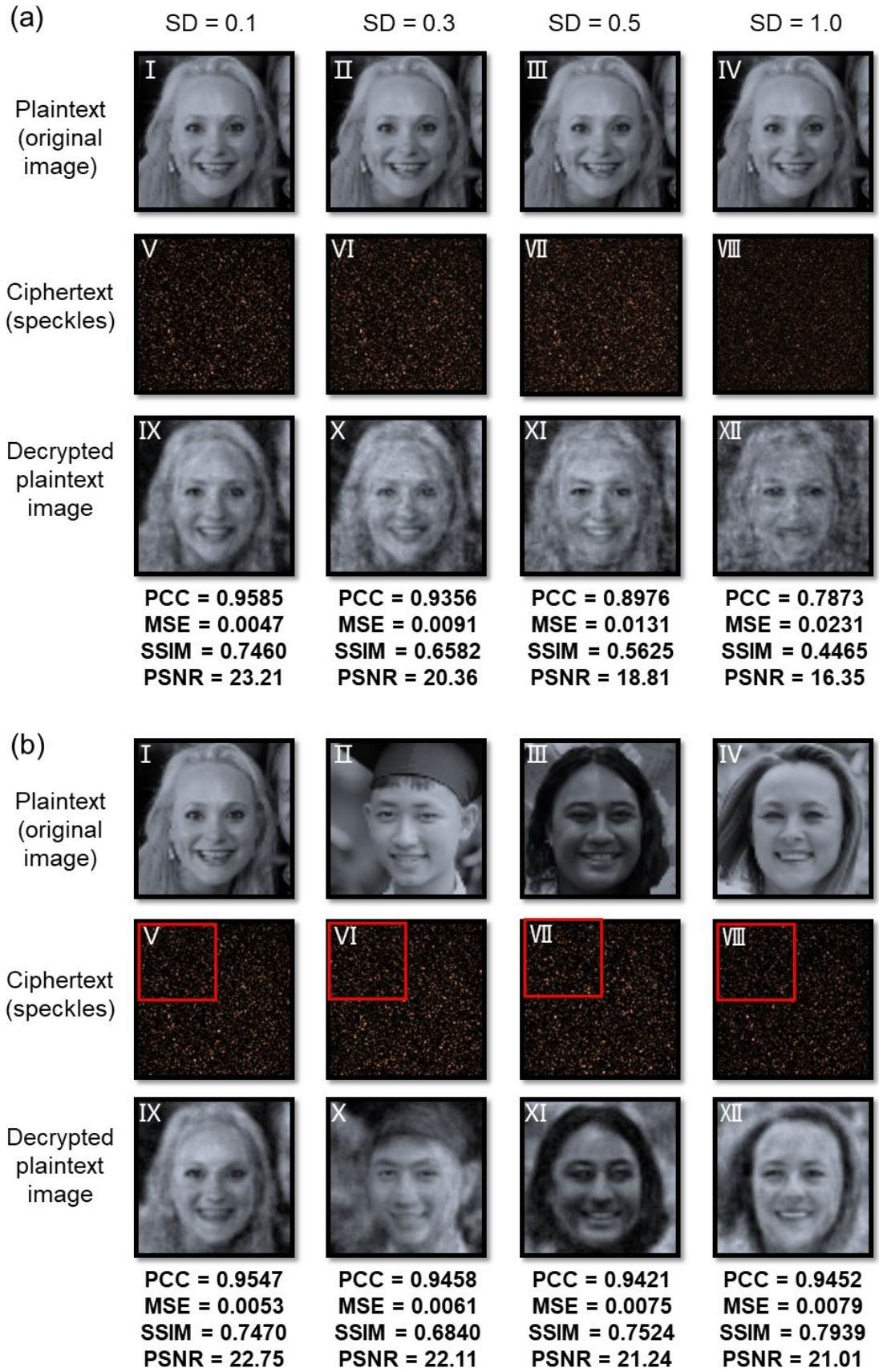

Figure 6-4 (a) Decryption performance with noisy speckles: the speckles with computer-

generated random noise are fed into the pre-trained neural network for decryption. The noisy speckles and the corresponding decrypted images are marked with the corresponding noise standard deviation (SD) and similarity criteria, respectively. (b) Decryption performance with partial speckles: only the top left corners (*i.e.*, quarter field of view, marked in red box) of the speckles are used to train and test the neural network. The plaintext image I-IV: Reproduced under terms of the CC-BY 2.0 license. Copyright 2015, Lawrence Lessig at Second Home London, by Innotech Summit, Flickr (https://www.flickr.com/photos/115363358@N03/18260388752/). The plaintext image I: Reproduced under terms of the CC-BY 2.0 license. Copyright 2015, Lawrence Lessig at Second Home London, by Innotech Summit, Flickr (https://www.flickr.com/photos/115363358@N03/18260388752/). The plaintext image II: Reproduced under terms of the Public Domain Mark 1.0 license. Copyright 2018, kỉ yếu 12c, by khanhkhokhao201, Flickr (https://www.flickr.com/photos/154663983@N08/28538465128/). The plaintext image III: Reproduced under terms of the Public Domain Mark 1.0 license. Copyright 2016, Future Leaders of the Pacific 2016 by US Embassy, Flickr (https://www.flickr.com/photos/us_embassy_newzealand/29355772191/). The plaintext image IV: Reproduced under terms of the CC-BY 2.0 license. Copyright 2018, Ekaterina by Wonder Woman, Flickr

(https://www.flickr.com/photos/zamerzla/ 28685633938/). The original images were cropped and converted to grayscale.

Furthermore, due to multiple light scattering and the conceptualized infinite information channels [26] within the scattering medium, it is hypothesized that the information in the plaintext is scrambled and distributed to the whole speckle. Spatially, this speckle could be large in practice, especially if the incident light is focused on the front sample surface or the detection plane is far away from the sample. It is thus possible that only part of the speckle is captured by the detection camera in experiments [27]. To study whether this factor may affect the performance, an additional group of experiments is conducted by using a quarter region of interest (ROI) of the speckles for network training and testing. That is, the dimension of the speckles is reduced from 256×256 to 128×128 under the same spatial sampling condition. The experimental results are shown in Figure 6-4b. As seen, partial ROI leads to decryption results (Figure 6-4b) that are very comparable to those obtained with a larger ROI (Figure 6-3b), confirming the hypothesis above. Such a non-point-to-point information mapping between the plaintext and the ciphertext is distinctive to most existing cryptosystems. It allows smaller speckle ROIs to be adopted in network training and testing, which can relieve the burdens of

data collection, storage, and processing.

6.2.3 Face Recognition

During decryption, we utilize PCC and other criteria to test similarities. However, these criteria are not suitable for face recognition as they may be affected by many factors other than face features, such as backgrounds, orientations, and expressions of faces [28]. Therefore, the original and decrypted face images are further processed with an open-source Python face-recognition library [29]. The neural network used for face recognition is based on ResNet [30], which is well-trained based on 3 million faces and has 99.38% accuracy on the Labeled Faces in the Wild benchmark [31-32]. The face recognition network encodes each face image with a unique 128-dimensional vector, which extracts the specific features of human faces, including eyebrows, eyes, noses, mouths, and cheeks (as illustrated in Figure 6-5 a-b). If the Euclidean distance [33] between two face vectors is lower than a pre-set threshold, two corresponding faces are defined as Match with each other; otherwise, they are defined as Mismatch, as exampled in Figure 6-5c. The commonly used pre-set threshold is 0.6 (for general situations) or 0.5 (for higher security scenes).

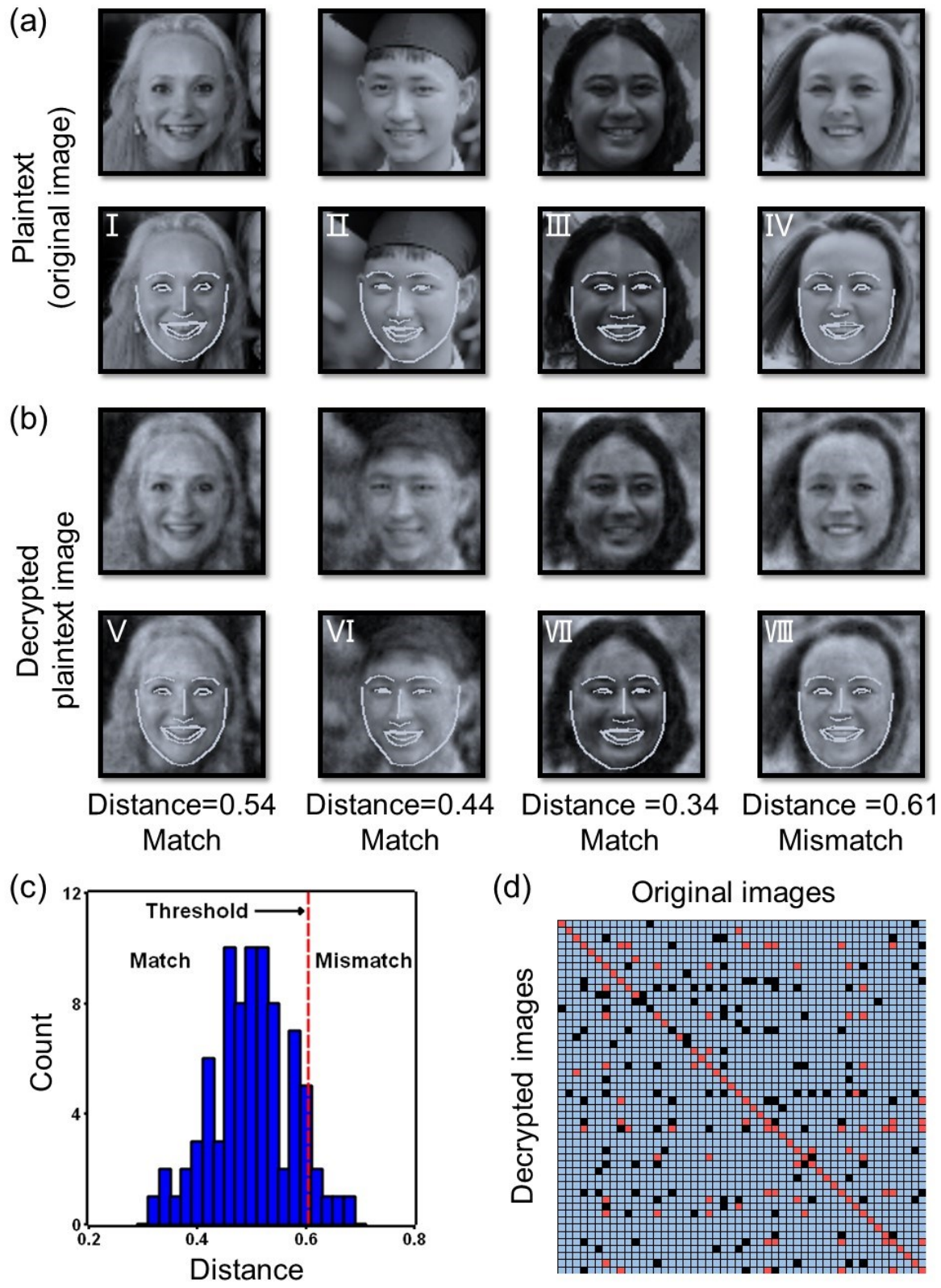

Figure 6-5 Face recognition results based on face images from FFHQ and the corresponding decrypted images from speckles. (a) The original face images (*i.e.*, plaintext)

and their key features for face recognition. (b) The decrypted face images by feeding speckles into the trained neural network and their key features. The face encoding distances between the decrypted and original face images with a threshold = 0.6 are marked under the decrypted images. (c) Face encoding distances between the decrypted and original images in the testing dataset. If the distance is less than or equal to the threshold = 0.6, the recognition result is Match; otherwise, it is Mismatch. (d) The face recognition results of the decrypted images. True positives are marked in red, true negatives are marked in blue, while false positives and false negatives are marked in black. The first-row plaintext image I: Reproduced under terms of the CC-BY 2.0 license. Copyright 2015, Lawrence Lessig at Second Home London, by Innotech Summit, Flickr (https://www.flickr.com/photos/115363358@N03/18260388752/). The first-row plaintext image II: Reproduced under terms of the Public Domain Mark 1.0 license. Copyright 2018, kỉ yếu 12c, by khanhkhokhao201, Flickr (https://www.flickr.com/photos/154663983@N08/28538465128/). The first-row plaintext image III: Reproduced under terms of the Public Domain Mark 1.0 license. Copyright 2016, Future Leaders of the Pacific 2016 by US Embassy, Flickr (https://www.flickr.com/photos/us_embassy_newzealand/29355772191/). First-row plaintext image IV: Reproduced under

terms of the CC-BY 2.0 license. Copyright 2018, Ekaterina by Wonder Woman, Flickr (https://www.flickr.com/photos/zamerzla/28685633938/). The original images were cropped and converted to grayscale.

Before recognizing the decrypted face images from the trained neural network, some images with sunglasses and babies were excluded since some of their facial key points were ambiguous. Then, the target is that if the Euclidean distances between the encoding vectors of two original images are smaller than the preset threshold (indicating that they are the same person), the distances between the two corresponding decrypted images are also expected to be smaller than the preset threshold, indicating that the people in the decrypted images and the original images match. The encodings of the decrypted images were also compared with each encoding of the original images. If the two original images' encoding distances were smaller than the preset threshold, the two samples were treated as positive samples. If the corresponding two decrypted images' encoding distances are also smaller than the preset threshold, the results are true positives, otherwise, they are false negatives. On the contrary, if the two original images' encoding distances are larger than the preset threshold, the two samples are treated as negative samples. If the corresponding two decrypted images' encoding distances are also larger than the

preset threshold, the results are true negatives; otherwise, they are false positives. During network testing, recall, precision, accuracy, and F1-score were used to test the performance, as defined in Equation (6-4), Equation (6-5), Equation (6-6), and Equation (6-7), respectively.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{6-4}$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{6-5}$$

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Negative + True\ Positive + False\ Negative + False\ Positive} \tag{6-6}$$

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{6-7}$$

As one person might be recognized as two different people, while two different people should not be recognized as the same person, accuracy is more meaningful than the other three criteria. Accordingly, in the experiments to test the proposed cryptosystem, various thresholds between 0.5 and 0.6 are tested with decrypted face images. As an example, the results of face recognition with a threshold distance of 0.6 are shown in Figure 6-5. The key features of the original and decrypted face images are extracted by the face recognition neural network and marked in the second row of Figure 6-5a and Figure 6-5b, respectively [29]. As seen, most of these decrypted images appear akin to their corresponding original plaintext images (*e.g.*, image pairs I-V, II-VI, and III-VII, whose PCC are all more than 0.94) and hence are recognized as Match. Note

that, however, some image pairs seem visually alike, such as IV-VIII whose PCC ≈ 0.96, but are still recognized as Mismatch since the distance is 0.61, being above the threshold of 0.6. Nevertheless, it shows that the face recognition library can extract key features and scale the differences between decrypted and original face images.

Furthermore, we test the accuracy of face recognition. The 128-dimension face encodings from the decrypted images are compared with the corresponding encodings from the original face images, as shown in Figure 6-5d. The results with different distance thresholds are shown in Table 6-1 and compared with other face recognition algorithms [34-38]. It is not surprising that different thresholds result in different recalls, precisions, and accuracies. It can be observed that our accuracy reaches greater than 98% when the threshold is below 0.58. Compared with FaceNet and VGGFace, the method proposed in this chapter has higher accuracy and is therefore more suitable for practical applications [35-36]. Moreover, the precision is 100% when the threshold is set at 0.5, indicating high confidence during face recognition. However, the recall and the F1 score are not as good as those from FaceNet and VGGFace, which can be attributed to the fact that there are more negative samples than positive samples in the data we use. The performance can be further improved by adjusting the threshold in face recognition according to the sample distribution in the dataset, or tuning the structure or the parameters of

the neural network.

Table 6-1 Face recognition results by our method and other algorithms with optimal thresholds.

	Threshold	Recall	Precision	Accuracy	F1 score
This work	0.60	66.18%	64.02%	97.87%	65.08%
	0.58	62.73%	69.66%	98.49%	66.01%
	0.56	61.65%	78.10%	98.93%	68.91%
	0.54	61.34%	87.95%	99.19%	72.28%
	0.52	56.07%	92.31%	99.25%	69.77%
	0.50	46.53%	100.00%	99.22%	63.51%
FaceNet[39]	0.90	96.42%	100.00%	98.21%	98.18%
VGGFace[40]	0.79	80.71%	97.41%	89.28%	88.28%
OpenFace[41]	0.47	16.42%	95.83%	57.85%	28.04%
DeepFace[42]	0.51	9.28%	100.00%	54.64%	16.99%

6.3 Discussions

In this chapter, an optical speckle-based optical cryptosystem has been proposed, implemented, and demonstrated in experiments, where a ground glass scattering medium has been exploited

as the physical secret key to generate speckles that uniquely encrypt information. As for a cryptosystem, security is the topmost concern, and we will discuss the security of the proposed method from three aspects, including the physical secret key, the optical setup, and cracking attacks.

6.3.1 Length of the Secret Key

The equivalent key length of the scattering medium can be modeled by the transmission matrix, whose dimension is $(256 \times 256) \times (64 \times 64)$, and each element is 64 bits (for complex float numbers) in the computer. Thus, the digital key of this cryptosystem is of length $64 \times [(256 \times 256) \times (64 \times 64)] = 1.72 \times 10^{10}$ bits, that is, 17.2 gigabits, which is enormous for brute force attacks even with a quantum computer. In comparison, for pure software-based encryption approaches, such as Advanced Encryption Standard (AES) [5] and Compression Friendly Encryption Scheme (CFES) [39], the digital cryptosystems are all based on matrix manipulations. As the size of the matrix (*i.e.*, digital secret key length) increases, more multiplicative manipulations are needed, and the computational complexity grows exponentially. Therefore, to balance computational efficiency and security, the digital secret

key lengths in digital cryptosystems are usually limited to hundreds of bits. However, in our speckle-based physical encryption process, no mathematical algorithms are involved, so the computational burden can be ruled out during encryption, and users can achieve high security without compromising encryption speed. Note that, when it comes to decryption, our optical cryptosystem involves a large amount of computation. Fortunately, these decryption processes can be accelerated by using a high-performance GPU.

6.3.2 Unclonable Feature of the Secret Key

As for the optical setup, it is nearly impossible to generate the same speckles with a different scattering medium (*i.e.*, the physical secret key), in which the scatterers are randomly distributed. The light-medium interactions are very complicated, and the resultant optical propagation involves intricate multipath scattering; minor variations in the scattering medium can influence the optical field considerably, resulting in a totally different transmission matrix of the scattering medium. Therefore, compared with existing digital encryption matrix-based approaches (*i.e.*, relays only on digital secret keys) [40], it is nearly impossible to duplicate the inhomogeneous refractive index distribution of the scattering medium to crack the cryptosystem,

except for a self-defined medium such as a metasurface [41-42]. Therefore, the speckles can be viewed as nearly unclonable, and the decryption process is exclusive to the quantification of the scattering medium, that is, a deep neural network (DNN) trained with ciphertext (*i.e.*, speckles) as the input and plaintext as the output. If speckles generated from another scattering medium (*i.e.*, wrong physical secret keys) are input to the pre-trained neural network for decryption, as exampled in Figure 6-6, the decrypted results are obscure and very different from the plaintext. Consequentially, the decrypted images cannot be used for face recognition, and thus the security of the proposed system can be guaranteed.

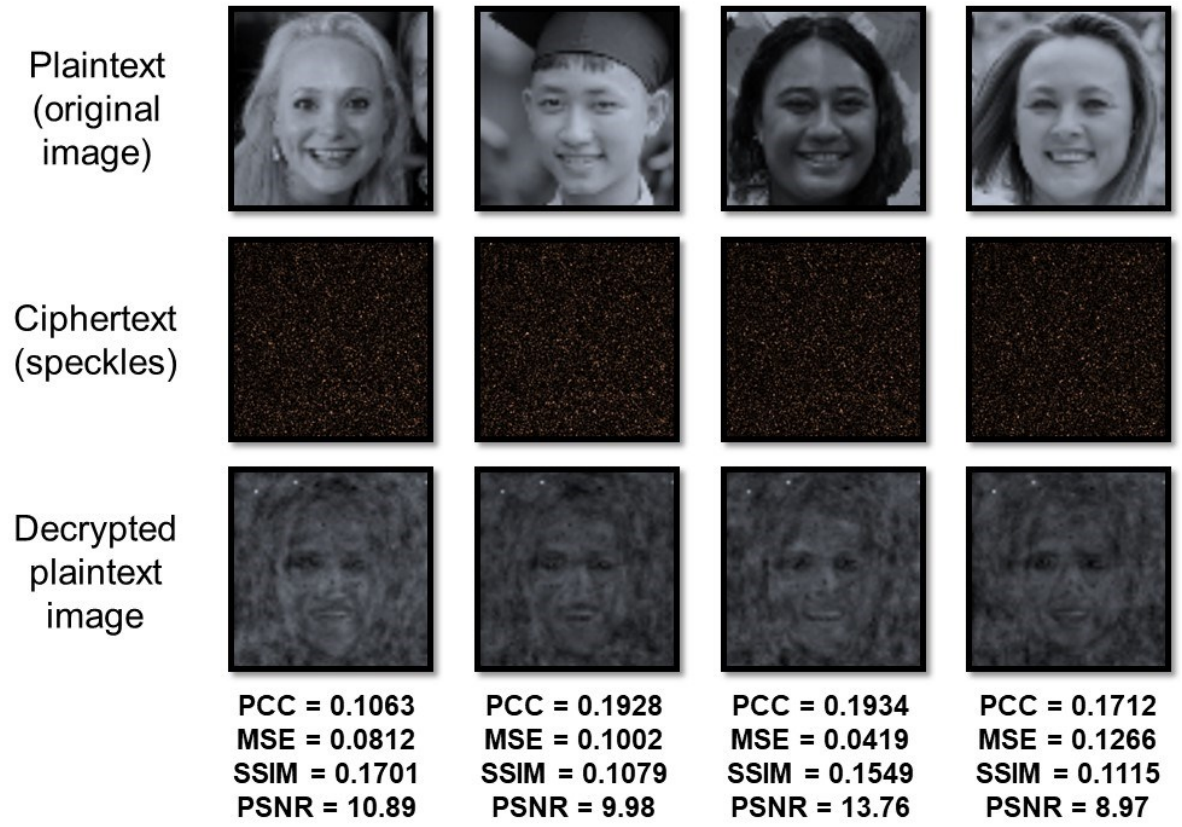

Figure 6-6 Wrong physical secret key attack: the same plaintext images are used, but another scattering medium is utilized to generate the speckles (*i.e.*, ciphertext), which are input to the pre-trained neural network to yield the decrypted plaintext images. The PCC, MSE, SSIM, and PSNR between the decrypted and the corresponding original face images are marked. The transmission matrix similarity, as measured by PCC, between the correct and wrong physical secret keys is 0.00012. Plaintext image I: Reproduced under terms of the CC-BY 2.0 license. Copyright 2015, Lawrence Lessig at Second Home London, by Innotech Summit, Flickr (https://www.flickr.com/photos/115363358@N03/18260388752/). Plaintext image II: Reproduced under terms of the Public Domain Mark 1.0 license.

Copyright 2018, kỉ yếu 12c, by khanhkhokhao201Flickr (https://www.flickr.com/photos/154663983@N08/28538465128/). Plaintext image III: Reproduced under terms of the Public Domain Mark 1.0 license. Copyright 2016, Future Leaders of the Pacific 2016 by US Embassy, Flickr (https://www.flickr.com/photos/us_embassy_newzealand/29355772191/). Plaintext image IV: Reproduced under terms of the CC-BY 2.0 license. Copyright 2018, Ekaterina by Wonder Woman, Flickr (https://www.flickr.com/photos/zamerzla/28685633938/). Plaintext image V: Reproduced under terms of the Public Domain Mark 1.0 license. 2015, Resiliency Day, Sept. 11, Copyright 2015 by Presidio of Monterey, Flickr (https://www.flickr.com/photos/presidioofmonterey/21442846325/). Plaintext image VI: Reproduced under terms of the CC-BY 2.0 license. Copyright 2008, P1020227 by Kyle Peyton, Copyright 2008, Flickr (https://www.flickr.com/photos/kylepeyton/2779218214/). The original images were cropped and converted to grayscale.

6.3.3 Uniqueness of Optical Setups

Under extreme situations when hackers have obtained the scattering medium (*i.e.*, the physical secret key), in order to produce the same speckles, the error in duplicating the optical system

alignment and the light-medium interaction should be within the optical wavelength scale [43]. That is, the optical setup ensures that the interaction between the light and medium is hard to reproduce due to the narrow range (approximately milliradians for tilt and submicrons for shift) of the memory effect. What's more, within the memory effect, neural networks can be built to retrieve images from speckle autocorrelations, and the trained neural networks can be generalized to unknown scattering media, that is, the trained neural networks based on speckle autocorrelations can be used for a ciphertext-only attack. However, beyond the memory effect, it is theoretically impossible to build and train neural networks based on speckle autocorrelations to decrypt complex-structured face images from an unknown scattering medium, due to weak relations between speckle autocorrelations and image autocorrelations [43-44]. In experiments, the memory effect range is less than a quarter of the face image size, thus the cryptosystem is safe under ciphertext-only attacks.

Furthermore, chosen-plaintext and known-plaintext attacks are possible only when attackers can get access to at least 10,000 image-speckle sets, as discussed in Figure 6-7. As seen, to achieve satisfactory performance, for example, PCC and face recognition accuracy greater than 0.9, the training datasets need to be larger than 10,000 and 15,000, respectively. In the proposed cryptosystem, obtaining such a large number of image-speckle sets is possible only when

attackers have access to the optical setup and the unique physical secret key simultaneously, which, however, is very demanding and already beyond the scope of the topic. Even in that situation, if the unique physical secret key is stolen, it can be replaced with a new secret key to protect data.

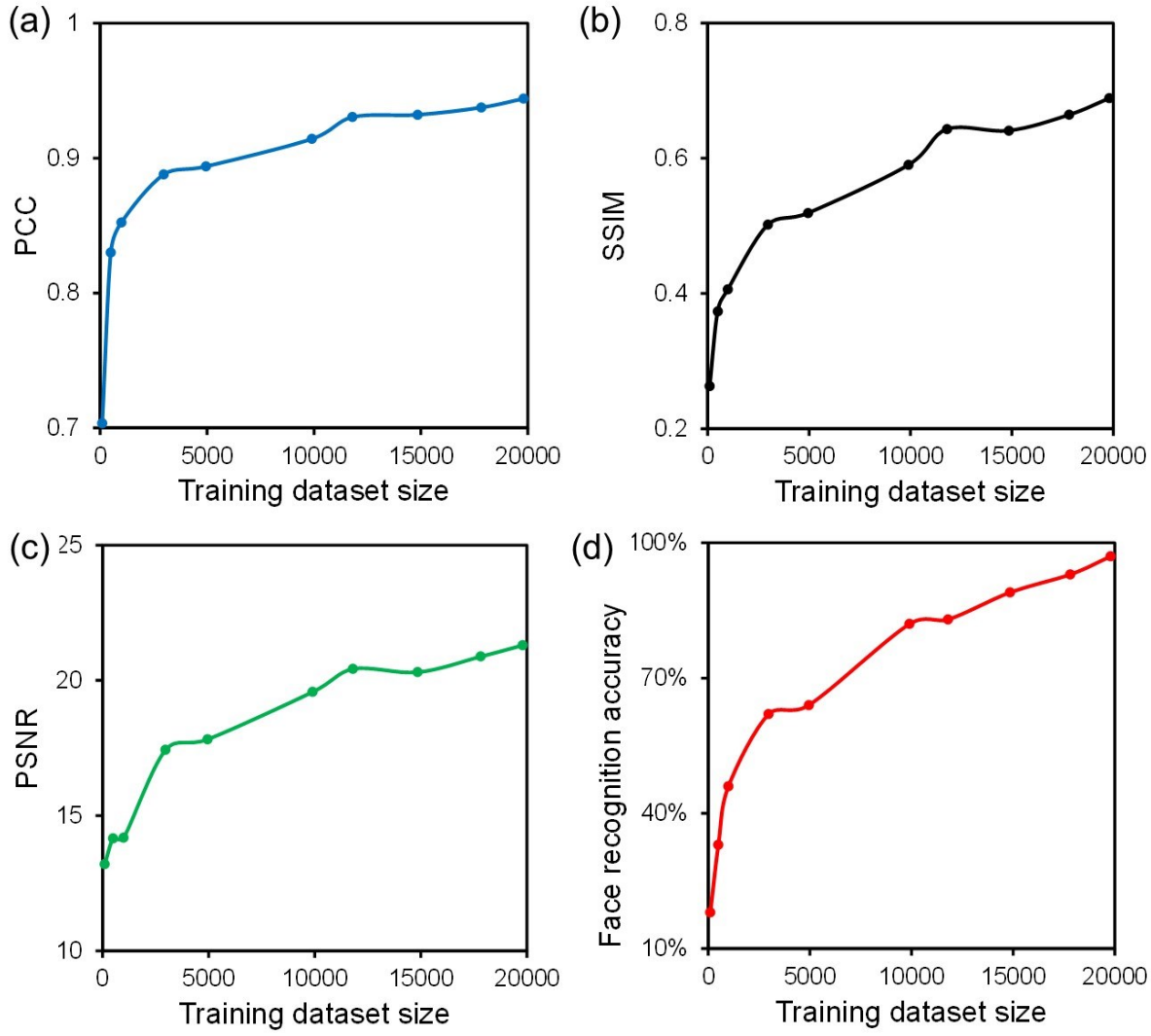

Figure 6-7 Experimental results with different dataset sizes: similarities between the decrypted and original images as measured by (a) PCC, (b) SSIM, and (c) PSNR, as well

as (d) face recognition accuracy, as a function of training dataset size.

6.3.4 Others

The intervention of optics further boosts the efficiency of encryption (*i.e.*, at the speed of light), which overwhelms the software-based cryptosystems. Optical solutions, including the proposed speckle-based method and the DRPE method, can enable highly efficient encryption and generate high-dimensional secret keys [8]. Notably, compared with DRPE, the proposed method is advantageous due to its simpler optical design. DRPE requires two SLMs in the optical setup since the information is encrypted by two random phase masks [50]. In our cryptosystem, encryption can be performed with a scattering medium. This not only facilitates integration with other systems, but also reduces the cost of the cryptosystem. The most expensive component in the current system is the SLM, which is only responsible for loading the images and is indeed replaceable in practice since direct illumination of human faces can be used as input images for the cryptosystem. As a result, the cost of the proposed cryptosystems becomes comparable to the software-based cryptosystems, which only require computers for encryption and decryption.

When it comes to system latency, although well-known edge computing can help to recognize face images and protect privacy through computing in cloudlets, its scalability is refrained by the computing power, leading to applications of limited databases [45]. In comparison, the proposed light-based system can achieve fast encryption speed and high scalability. Moreover, with the development of high throughput communication networks, such as 5G, the latency of the proposed system can also be comparable to edge computing-based face recognition [46].

When it comes to the quality of decrypted images, the proposed neural network delivers high similarity between decrypted and original images, resulting in accurate face recognition (*i.e.*, 98%) that is comparable to other state-of-the-art methods [34-38]. That said, some high-frequency information (*i.e.*, detailed structures, such as hair) in images may still be lost after the speckle-based encryption in experiments, due to non-ideal experimental setups such as aberrations from the SLM curvature, optical lens, and camera. The lost high-frequency information is therefore difficult to be retrieved by neural networks during decryption. Furthermore, to simplify the optical setup, we have just recorded speckle intensity during experiments. The missing phase of the speckle field also results in information loss. These all lead to moderate PSNR of decrypted images, which will be improved in the next phase of the study by optimizing the optical setup or the neural network structures.

Last but not least, the novelty of the proposed speckle-based optical cryptosystem is highlighted from three aspects. First, although some literature has mentioned speckle-based encryption recently [51-52], they have mainly focused on the encryption of simple digits and characters, rather than complex-structured images such as face images. The cryptosystem for face reconstruction and recognition is considerably more complicated than that for digits and characters. Second, although learning-based decryption has also been demonstrated in [13] and [51], our efforts have gone beyond that. After decryption with high fidelity, face recognition is demonstrated with 98% accuracy, which is comparable to the state-of-the-art algorithms in the field. Third and most importantly, the proposed speckle-based optical cryptosystem has a very high level of security. The length of the physical security key is more than 17 gigabits, being many magnitudes longer than that of pure software-based encryption approaches and sufficiently secure for brute force attacks. Due to the nature of the speckle-based mechanism, there is no computational burden or compromised speed during encryption. Meanwhile, the complicated light-medium interaction assures that every physical secret key (*i.e.*, the scattering medium) is unique and nearly unclonable. Furthermore, the narrow memory effect range of the optical system determines that the interaction between the light and the medium is hard to reproduce, protecting the cryptosystem from ciphertext-only attacks, chosen-plaintext attacks,

and known-plaintext attacks. The only exception is when the optical setup and the physical secret key are both leaked, which, however, is beyond what a cryptosystem can handle.

6.4 Conclusion

In this chapter, we demonstrate a speckle-based optical cryptosystem for face recognition, and the accuracy has reached more than 98%, which is comparable to that of other state-of-the-art methods. With the proposed speckle-based optical cryptosystem, the encrypted private data (*e.g.*, face images) is difficult to crack and reduces the risk of information leakage. The speckle-based optical cryptosystem is suitable for practical applications due to its high security, fast speed, low cost, insensitivity to the ROI, as well as immunity to low and moderate noise in the ciphertexts. That said, the accuracy of face recognition can still be further improved by constructing more complex neural networks that lead to an all-speckle-based optical cryptosystem for decryption and face recognition [47-48], where there is no need to decrypt optical speckles to face images. Moreover, to further enhance the security of the encryption processes, multi-channel laser diffraction by high-dimensional scattering media can be adopted to increase the speckle randomness. On the other hand, binary speckles can be used to reduce

data storage space and increase data transmission speed [49]. Collectively, although the experiments contain only a proof-of-principle demonstration for face encryption and recognition, we believe that with further optimization, the proposed speckle-based optical cryptosystem may find or inspire wide applications in high-security encryption and decryption.

References

1. Senior, A. W., & Pankanti, S. (2011). Privacy protection and face recognition. In *Handbook of Face Recognition* (pp. 671-691). London: Springer London.

2. Berghel, H. (2017). Equifax and the latest round of identity theft roulette. *Computer, 50*(12), 72-76.

3. Chavis, K. (2021). Transformative Policing Technologies: Balancing Public Safety, Privacy, and Community Consent in Vulnerable Communities in the United States. *Policing: A Journal of Policy and Practice, 15*(1), 425-439.

4. Gidney, C., & Ekerå, M. (2021). How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits. *Quantum, 5*, 433.

5. F D'souza, F. J., & Panchal, D. (2017, May). Advanced encryption standard (AES) security enhancement using hybrid approach. In *2017 International Conference on Computing, Communication and Automation (ICCCA)* (pp. 647-652). IEEE.

6. Rachmawati, D., Tarigan, J. T., & Ginting, A. B. C. (2018, March). A comparative study of Message Digest 5 (MD5) and SHA256 algorithm. In *Journal of Physics: Conference Series* (Vol. 978, p. 012116). IOP Publishing.

7. Denchev, V. S., Boixo, S., Isakov, S. V., Ding, N., Babbush, R., Smelyanskiy, V., ... & Neven, H. (2016). What is the computational value of finite-range tunneling. *Physical Review X, 6*(3), 031015.

8. Liu, S., Guo, C., & Sheridan, J. T. (2014). A review of optical image encryption techniques. *Optics & Laser Technology, 57*, 327-342.

9. Javidi, B., Carnicer, A., Yamaguchi, M., Nomura, T., Pérez-Cabré, E., Millán, M. S., ... & Markman, A. (2016). Roadmap on optical security. *Journal of Optics, 18*(8), 083001.

10. Zhao, A., Jiang, N., Liu, S., Zhang, Y., & Qiu, K. (2021). Physical layer encryption for WDM optical communication systems using private chaotic phase scrambling. *Journal of Lightwave Technology, 39*(8), 2288-2295.

11. Qu, G., Yang, W., Song, Q., Liu, Y., Qiu, C. W., Han, J., ... & Xiao, S. (2020).

Reprogrammable meta-hologram for optical encryption. *Nature communications, 11*(1), 5484.

12. Caramazza, P., Moran, O., Murray-Smith, R., & Faccio, D. (2019). Transmission of natural scene images through a multimode fibre. *Nature communications, 10*(1), 2029.

13. Horisaki, R., Takagi, R., & Tanida, J. (2016). Learning-based imaging through scattering media. *Optics express, 24*(13), 13738-13743.

14. Cheng, S., Li, H., Luo, Y., Zheng, Y., & Lai, P. (2019). Artificial intelligence-assisted light control and computational imaging through scattering media. *Journal of innovative optical health sciences, 12*(04), 1930006.

15. Li, S., Deng, M., Lee, J., Sinha, A., & Barbastathis, G. (2018). Imaging through glass diffusers using densely connected convolutional networks. *Optica, 5*(7), 803-813.

16. Barbastathis, G., Ozcan, A., & Situ, G. (2019). On the use of deep learning for computational imaging. *Optica, 6*(8), 921-943.

17. Wu, J., Cao, L., & Barbastathis, G. (2021). DNN-FZA camera: a deep learning approach toward broadband FZA lensless imaging. *Optics Letters, 46*(1), 130-133.

18. Luo, Y., Yan, S., Li, H., Lai, P., & Zheng, Y. (2020). Focusing light through scattering media by reinforced hybrid algorithms. *APL photonics, 5*(1).

19. Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., ... & Zhang, J. (2021). Artificial intelligence: A powerful paradigm for scientific research. *The Innovation, 2*(4).

20. Clemente, P., Durán, V., Tajahuerce, E., & Lancis, J. (2010). Optical encryption based on computational ghost imaging. *Optics letters, 35*(14), 2391-2393.

21. Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4401-4410).

22. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional Networks for Biomedical Image Segmentation. In *Medical image computing and computer-assisted intervention-MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18* (pp. 234-241). Springer International Publishing.

23. Salimans, T., & Kingma, D. P. (2016). Weight normalization: A simple reparameterization to accelerate training of deep neural networks. *Advances in neural information processing systems, 29*.

24. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

25. M. Donadio, How to generate white gaussian noise, https://qualityassignmenthelp.com/wp-content/uploads/2017/03/Gaussian-noise_How-to-Generate-.pdf (accessed: January 2022).

26. Hou, J., & Situ, G. (2022). Image encryption using spatial nonlinear optics. elight, 2(1), 3.

27. Lyu, M., Wang, H., Li, G., Zheng, S., & Situ, G. (2019). Learning-based lensless imaging through optically thick scattering media. *Advanced Photonics, 1*(3), 036002.

28. Dagher, I., Dahdah, E., & Al Shakik, M. (2019). Facial expression recognition using three-stage support vector machines. *Visual Computing for Industry, Biomedicine, and Art, 2*, 1-9.

29. A. Geitgey, face recognition, https://github.com/ageitgey/face_recognition (accessed: October 2021).

30. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

31. Learned-Miller, E., Huang, G. B., RoyChowdhury, A., Li, H., & Hua, G. (2016). Labeled faces in the wild: A survey. *Advances in face detection and facial image analysis*, 189-248.

32. D. King, High Quality Face Recognition with Deep Metric Learning, http://blog.dlib.net/2017/02/high-quality-face-recognition-with-deep.html (accessed: October 2021).

33. Wang, L., Zhang, Y., & Feng, J. (2005). On the Euclidean distance of images. *IEEE*

transactions on pattern analysis and machine intelligence, 27(8), 1334-1339.

34. Serengil, S. I., & Ozpinar, A. (2020, October). Lightface: A hybrid deep face recognition framework. In *2020 innovations in intelligent systems and applications conference (ASYU)* (pp. 1-5). IEEE.

35. Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A Unified Embedding for Face Recognition and Clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 815-823).

36. Parkhi, O., Vedaldi, A., & Zisserman, A. (2015). Deep face recognition. In *BMVC 2015-Proceedings of the British Machine Vision Conference 2015*. British Machine Vision Association.

37. Baltrušaitis, T., Robinson, P., & Morency, L. P. (2016, March). Openface: an open source facial behavior analysis toolkit. In *2016 IEEE winter conference on applications of computer vision* (WACV) (pp. 1-10). IEEE.

38. Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). DeepFace: Closing the Gap to Human-Level Performance in Face Verification. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1701-1708).

39. Ahmed, F., Siyal, M. Y., & Abbas, V. U. (2010, November). A perceptually scalable and

jpeg compression tolerant image encryption scheme. In *2010 Fourth Pacific-Rim Symposium on Image and Video Technology* (pp. 232-238). IEEE.

40. Tao, C., Diene, A., Tang, S., & Ding, J. (2013). Simple matrix scheme for encryption. In *Post-Quantum Cryptography: 5th International Workshop, PQCrypto 2013, Limoges, France, June 4-7, 2013. Proceedings 5* (pp. 231-242). Springer Berlin Heidelberg.

41. Chen, H. T., Taylor, A. J., & Yu, N. (2016). A review of metasurfaces: physics and applications. *Reports on progress in physics, 79*(7), 076401.

42. Bukhari, S. S., Vardaxoglou, J., & Whittow, W. (2019). A metasurfaces review: Definitions and applications. *Applied Sciences, 9*(13), 2727.

43. Osnabrugge, G., Horstmeyer, R., Papadopoulos, I. N., Judkewitz, B., & Vellekoop, I. M. (2017). Generalized optical memory effect. *Optica, 4*(8), 886-892.

44. Zhu, S., Guo, E., Gu, J., Bai, L., & Han, J. (2021). Imaging through unknown scattering media based on physics-informed learning. *Photonics Research, 9*(5), B210-B219.

45. Satyanarayanan, M. (2017). The emergence of edge computing. *Computer, 50*(1), 30-39.

46. Osseiran, A., Boccardi, F., Braun, V., Kusume, K., Marsch, P., Maternia, M., ... & Fallgren, M. (2014). Scenarios for 5G mobile and wireless communications: the vision of the METIS project. *IEEE communications magazine, 52*(5), 26-35.

47. Yuan, X., & Han, S. (2021). Single-pixel neutron imaging with artificial intelligence: Breaking the barrier in multi-parameter imaging, sensitivity, and spatial resolution. *The Innovation, 2*(2).

48. Lin, X., Rivenson, Y., Yardimci, N. T., Veli, M., Luo, Y., Jarrahi, M., & Ozcan, A. (2018). All-optical machine learning using diffractive deep neural networks. *Science, 361*(6406), 1004-1008.

49. Zhou, Z., Xia, J., Wu, J., Chang, C., Ye, X., Li, S., ... & Tong, G. (2020). Learning-based phase imaging using a low-bit-depth pattern. *Photonics Research, 8*(10), 1624-1633.

50. Chen, W., Javidi, B., & Chen, X. (2014). Advances in optical security systems. Advances in optical security systems. *Advances in Optics and Photonics, 6*(2), 120-155.

51. Zhou, L., Xiao, Y., & Chen, W. (2020). Learning complex scattering media for optical encryption. Optics Letters, 45(18), 5279-5282.

52. Liu, Y., Yu, P., Li, Y., & Gong, L. (2020). Exploiting light field imaging through scattering media for optical encryption. OSA Continuum, 3(11), 2968-2975.

7 SUMMARY

A significant obstacle to optical technologies is strong scattering, which distorts desired information and results in optical speckles rather than clear images. Despite the complexity of speckles, retrieving delocalized information from speckles has attracted considerable research interests. This thesis has explored overcoming, understanding, and utilizing optical speckles, aiming to enhance optical imaging capabilities at substantial penetration depths and exceptional resolutions. The thesis is structured into six chapters, each contributing uniquely to the overarching theme.

Chapter 1 introduces the research fields of the thesis, including backgrounds, state-of-the-art research, and the motivations of the studies in the following chapters.

Chapter 2 proposes a parameter-free algorithm for iterative wavefront shaping to overcome optical speckles. Iterative wavefront shaping has been approved to be an effective way to overcome scattering and has seen many exciting developments, such as focusing light and lossless image transmission through or inside scattering media. While encouraging, lots of efforts might be needed to tune parameters towards robust and optimum optimization. Moreover, optimal parameters might differ for different scattering samples and experimental conditions.

This chapter provides a robust method to enhance optical focusing through various scattering media, and the time-consuming and experience-dependent parameter tuning process, which is inevitable for existing iterative algorithms, is no longer needed with the proposed method. The integration of genetic, bat, and dynamic mutation algorithms to optimize parameters automatically is a significant step towards practical applications, reducing the dependency on manual parameter tuning.

Chapter 3 addresses spatiotemporal decorrelation in optical speckles using a GAN-based framework. This chapter extends the focusing capabilities developed in Chapter 2 to imaging from decorrelated speckles by tackling the dynamic nature of scattering media. For long, researchers have made strides in retrieving target information from speckles, primarily through calibrating the transmission matrix of the scattering medium or employing neural networks. That said, most of these approaches are designed for stationary scattering media, and the fidelity of the retrieved images is significantly compromised when the scattering medium's status changes due to motions, perturbations, or vibrations. Additionally, time intervals between acquiring the training and test datasets were neglected, and data from the training and test sets are highly correlated. In practical applications, time intervals between training and testing data are usually inevitable due to the lengthy training and the need for repeated usage. Hence, testing

data are often collected under different statuses of the medium/system, leading to decorrelation from the training data. The decorrelation has hindered these approaches so far from seeing wide applications in practice. In this chapter, we have proposed a GAN-based framework with extended generalizability, aiming to address the spatiotemporal instabilities of scattering media and the resultant decorrelation between training and testing data. Experiments demonstrate that the proposed GAN framework can be trained to retrieve face images from speckles with high fidelity, even when the scattering medium has undergone random decorrelation to unknown statuses after network training. Compared with existing learning-based implementations, the proposed GAN can non-holographically retrieve images from unstable scattering media and effectively address speckle decorrelation, even after the optical system has been inactive for an extended period (up to 37 hours in experiments) and subsequently reactivated. This capability paves the way for broad applications where networks can be pre-trained and maintain their effectiveness for data acquired at a later time. Such resilience is pivotal for extending the applications of learning-based methodologies in speckle imaging, encompassing applications like non-holographic imaging through scattering media.

Chapter 4 delves deeper into the nature of speckles and explores how information is delocalized within optical speckles through the prism of information entropy. This chapter complements

the research in Chapters 2 and 3, providing a deeper theoretical understanding of speckle-related studies and applications. The concept of information delocalization in speckles has been introduced, proposing that object information is uniformly delocalized among optical speckles. Experimental findings reveal that object information is uniformly delocalized among speckles, maintaining consistent information across different ROIs of the same size and ensuring the equivalent fidelity of the retrieved information. This is the first work to systematically summarize the concept of delocalization. Furthermore, the concept of entropy is utilized to provide a quantitative understanding of delocalized information in speckles. Then, we propose the speckle sampling condition for high-fidelity information retrieval: the entropy of speckle autocorrelation should exceed that of image autocorrelation. That said, various factors can influence the information retrieval from speckles, including optical setups, environmental perturbations, neural network parameters, *etc*. At this moment, the speckle sampling condition is a necessary condition, with ongoing exploration into establishing a sufficient and necessary condition. Collectively, this chapter contributes significantly to the understanding of information delocalization in speckles and has the potential to inspire new research and speckle-related applications, including high-throughput speckle imaging, non-line-of-sight imaging, optical speckle storage, *etc*.

Chapter 5 further explores delocalized information within speckles for direct classification through scattering media rather than solely for imaging, which is a synthesis of the practical algorithmic advancements and theoretical insights from the previous chapters. In Chapter 4, achieving high-fidelity information retrieval from optical speckles proves challenging, often requiring extensive speckle sampling that might be difficult in practical applications. From another point of view, certain applications of these retrieved speckle images do not necessarily demand high resolutions. For instance, in classification tasks, what matters most is identifying the type, rather than obtaining clear images. This implies that directly classifying the limited information in speckles could be possible and potentially yield better accuracy compared to classifying retrieved images from speckle retrieval. The advantage of direct classification stems from the fact that the quality of retrieved images from limited speckle information is often blurry, resulting in decreased accuracy in subsequent classification. Accordingly, this chapter introduces Speckle Transformer, a novel vision transformer-based model designed to classify original images with high accuracy using the limited information available in speckles with small ROIs. Due to bypassing the need for complete speckle data to retrieve images, Speckle Transformer enables the classification of images based on the features extracted from speckles, thereby achieving higher accuracy than traditional methods that rely on image retrieval.

Collectively, the experimental findings indicate that direct classification, enabled by Speckle Transformer, can achieve higher accuracies compared to classification after image retrieval, particularly in scenarios where speckles contain insufficient information for high-fidelity image retrieval. In the future, Speckle Transformer holds significant potential for application to non-line-of-sight imaging and privacy-protected classifications.

Chapter 6 shifts focus from overcoming speckles to utilizing speckles for encryption, which signifies a departure from the primary focus on imaging through scattering media to embracing the natural randomness of speckles. It is known that face recognition has recently become ubiquitous in many scenes for authentication or security purposes. In the meantime, there are increasing concerns about the privacy of face image data, which should be carefully protected. Software-based cryptosystems are widely adopted nowadays to encrypt face images, but security is limited by the insufficient digital secret key length. Hardware-based optical cryptosystems can generate enormously longer secret keys and enable encryption at the speed of light, but most reported optical encryption methods, such as double random phase encryption, have not yet been widely adopted as the optical design is complicated to be integrated with other systems. From another point of view, when coherent light propagates within and through scattering media, optical speckles are formed. The random features of speckles appear

meaningless and usually annoying, but constitute infinitely information channels, which may contribute to the tremendously long physical secret key length in a cryptosystem. However, as far as we know, speckle-based optical cryptosystems for complex tasks, such as encrypted face recognition, have rarely been explored. In this chapter, a plain yet high-efficient speckle-based optical cryptosystem is proposed, implemented, and demonstrated, where a ground glass is exploited to generate optical speckles that serve as an unclonable physical secret key with gigabit length and encrypt face images at light speed. The concept is decomposed into three steps: first, face images are optically scrambled into speckles for encryption, which protects the data during transmission and storage; then, a neural network is trained to decrypt the face images of high fidelity from the ciphertext (*i.e.*, speckles); last, the decrypted images are compared with the known face encodings and recognized. In this cryptosystem, face images are encrypted into random speckles that are nearly impossible to decrypt without the knowledge of the physical key (*i.e.*, the scattering medium) or the learned digital key (*i.e.*, the neural network). To the best of our knowledge, this is the first demonstration of a speckle-based optical cryptosystem for face recognition. Although the current study contains only a proof-of-principle demonstration for encrypting face images, the proposed method may find or inspire speckle-based applications in high-security information encryption and decryption.

Overall, this thesis contributes to overcoming the challenges posed by optical scattering, understanding delocalized information in optical speckles, and utilizing the properties of speckles for diverse applications. Each chapter builds upon the previous ones, creating a cohesive narrative that advances the field of speckle-related research and expands speckle-related applications. Among these studies, the theory of delocalized information within optical speckles not only deepens our comprehension of this phenomenon but also paves the way for advancements in speckle imaging and related applications.

To move forward, future work will explore the necessary and sufficient conditions for retrieving information from speckles with high fidelity. This includes investigating the limits of information retrieval in increasingly complex scattering environments and developing more sophisticated algorithms for information retrieval. Additionally, further research will focus on enhancing the robustness and generalizability of neural network-based frameworks to handle a wider range of dynamic conditions in biological applications. Another promising direction is the integration of speckle-based techniques with other imaging modalities, such as fluorescence and photoacoustic imaging, to improve the overall imaging performance in deep tissue applications. Exploring the potential of speckle-based cryptosystems for other types of biometric data and expanding their applications in data storage are also important areas for

future investigation. Ultimately, these studies will provide a promising paradigm for applying optical imaging in biomedical optics with greater penetration depths and higher resolutions, benefiting researchers within and beyond deep tissue imaging and diagnosis.