國立臺灣大學理學院物理學系
碩士論文

Department of Physics
College of Science
National Taiwan University
Master's Thesis

基於雙光子內視鏡與深度學習增強的高速深腦神經體積成像

Rapid Volumetric Imaging of Deep-Brain Neural Activity Using Two-Photon Endoscopy and Post-Processing Contrast Enhancement

廖品淳

Pin-Chun Liao

指導教授：朱士維 博士

Advisor: Shi-Wei Chu, Ph.D.

中華民國 114 年 10 月
October, 2025

doi:10.6342/NTU202504634

口試委員審定書

國立臺灣大學碩士學位論文
口試委員會審定書
MASTER'S THESIS ACCEPTANCE CERTIFICATE
NATIONAL TAIWAN UNIVERSITY

基於雙光子內視鏡與深度學習增強的高速深腦神經體積成像

（論文中文題目）(Chinese title of Master's thesis)

Rapid Volumetric Imaging of Deep-Brain Neural Activity Using Two-Photon Endoscopy and Post-Processing Contrast Enhancement

（論文英文題目）(English title of Master's thesis)

本論文係 廖品淳 (姓名) R12222016 (學號) 在國立臺灣大學 一般物理所 (系/所/學位學程) 完成之碩士學位論文，於民國 114 年 10 月 13 日承下列考試委員審查通過及口試及格，特此證明。

The undersigned, appointed by the Department / Institute of Physics on 13 (date) 10 (month) 2025 (year) have examined a Master's thesis entitled above presented by Pin-Chun, Liao (name) R12222016 (student ID) candidate and hereby certify that it is worthy of acceptance.

口試委員 Oral examination committee:

朱士維 吳順吉 陳永昭

（指導教授 Advisor）

doi:10.6342/NTU202504634

致謝

記得第一次與士維老師面談時，老師對我說：「我覺得你是一個很有勇氣的人。」對於一個從大學時期做理論模擬、如今踏入截然不同領域的我而言，當時的心情既期待又害怕。感謝一路上協助與支持我的每一位夥伴，讓我有信心在研究的道路上堅持下去。在發展光學系統、合作精進影像處理技術、分析數據與觀察生物現象的過程中，我獲得了充實而珍貴的經驗。這段歷程讓我以更熱忱的心擁抱研究。

我衷心感謝我的指導教授朱士維老師。在一次次的討論中，老師讓我學會如何有效吸收新知。無論是在團體會議或書面討論中，老師總能提出新穎的想法，並以認真教學、自律不懈的態度帶領我們。從老師身上，我學到的不僅是光學知識，更是重要的思考模式及與人分享的樂趣。

能夠參與不同的研究專案，是難得的機會。透過團隊合作，我吸收了豐富的知識，也結識了許多志同道合的研究夥伴。感謝示國老師和丞涵協助準備可愛的老鼠夥伴，讓實驗得以順利進行；多才多藝的柏廷親手製作實驗器材，令人敬佩；謝謝順吉老師和文琪用心與我分享模型架構，並高效率地討論；Risa 在課業繁忙之餘仍努力推進 endoscopy 技術，很開心能與妳一起玩、一起研究。謝謝志昌學長教我老鼠相關知識，並慷慨分享化學品的準備方法；感謝 AO 夥伴 Antonio 與我一同優化光路表現；也謝謝建昇像大哥般照顧我們，並一同整理 GRIN lens 主題、發表 review paper。

同時，也要感謝研究室的每一位夥伴。于碩總是熱情分享各式電子「新玩具」，也是讓我對自組系統充滿信心的夥伴；謝謝尹慈在初期帶我熟悉環境，庭禎協助處理控制程式問題；東翰語出驚人、又有奇妙默契，很開心和你一起與老鼠培養感情。和家總是為研究室增添活

doi:10.6342/NTU202504634

發氣氛，謝謝你與我一起在立陶宛與拉脫維亞的化學品苦樂與共並發表論文；奕蓁總是為我打氣、陪我學習光學；致嘉在研究之餘與我打桌球放鬆；Kentaro 邏輯清晰、幫我整理思緒，出遊時也貼心照顧大家；德心幽默又細心，樂於分享見解；品壬與俞安是我們的小幫手，主動協助解決各種問題。能與你們共同面對挑戰、交流想法、解決難題，是我研究生活中最精彩的部分。謝謝你們的陪伴，這些日子將是我難忘的回憶。

最後，深深感謝一路默默支持我的父母。你們給予我充分的空間與自由，讓我能專注追求所愛的事物。家人是我最好的聆聽者，無論開心或難過，我都能自在地與你們分享。

再次衷心感謝所有支持與幫助我的人。這是一段充滿挑戰與成就的旅程，有了你們的鼓勵與陪伴，我才能克服困難、堅持到底。希望自己能在未來的學習路上持續成長，豐富人生。

doi:10.6342/NTU202504634

摘要

理解大腦的複雜結構與功能仍然是神經科學中的一項重大挑戰。雙光子激發顯微術克服了傳統光學顯微術的限制，實現了更深層的組織穿透、更低的光毒性，以及具備體積解析度的活體成像。近期的技術進展，例如可調式聲學梯度（TAG）透鏡，透過沿 z 軸快速調制焦點，進一步將雙光子激發拓展至高速體積成像。當 TAG 透鏡與梯度折射率（GRIN）透鏡結合時，可進行活體深層神經迴路成像，並達到近乎影像速率的體積獲取。我們開發了一套自製的雙光子內窺系統，結合 TAG 與 GRIN 透鏡，以捕捉位於小鼠大腦深層 6 毫米的視交叉上核（SCN）的神經動態。然而，這些技術進步的代價是在深層組織成像時影像對比度降低。為了提升影像品質，我們應用了基於深度學習的後處理方法，利用自監督模型挖掘時空冗餘以達到有效的降噪。此外，我們還開發了一套客製化演算法用於像差校正。我們的系統成功地獲取了活體中三維的神經元群體活動，並提升了影像的對比與清晰度，從而捕捉到 SCN 的神經動態。這個方法能幫助我們更清楚地研究晝夜節律和大腦深部的神經連結，也提供了一個實用的工具來進一步探索深腦中複雜的神經活動。

關鍵詞：雙光子內視鏡、高速體積成像、視交叉上核、深度學習降噪、像差校正、神經連結

doi:10.6342/NTU202504634

ABSTRACT

Understanding the intricate structure and function of the brain remains a critical challenge in neuroscience. Two-photon excitation microscopy has addressed the limitations of traditional optical microscopy, enabling deep tissue penetration, reduced phototoxicity, and volumetric resolution for *in vivo* imaging. Recent advancements, such as tunable acoustic gradient (TAG) lenses, have extended two-photon excitation capabilities to high-speed volumetric imaging by rapidly modulating focus along the z-axis. When integrated with gradient-index (GRIN) lenses, TAG lenses enable *in vivo* imaging of deep neural circuits, achieving nearly video-rate volumetric acquisition. We have developed a home-built two-photon endoscopy system incorporating TAG and GRIN lenses to capture neural dynamics in the suprachiasmatic nucleus (SCN), located 6 mm deep in the mouse brain. However, these advancements come at the cost of reduced image contrast, particularly during deep tissue imaging. To enhance image quality, we applied deep-learning-based post-processing techniques, leveraging self-supervised models that exploit spatiotemporal redundancies for effective denoising. Moreover, a customized algorithm is developed for aberration calibration. Our system successfully derived three-dimensional neuronal population activity *in vivo* with enhanced contrast and clarity, capturing neuronal dynamics in the SCN. This innovative approach provides a powerful tool for advancing our understanding of circadian rhythms and deep brain functional connectivity, paving the way for further exploration of complex neural processes.

Keywords: Two-photon endoscopy, High-speed volumetric imaging, Suprachiasmatic nucleus, **Deep-learning denoising, Aberration calibration, Neural connection**

doi:10.6342/NTU202504634

CONTENTS

doi:10.6342/NTU202504634

doi:10.6342/NTU202504634

LIST OF FIGURES

doi:10.6342/NTU202504634

doi:10.6342/NTU202504634

LIST OF TABLES

doi:10.6342/NTU202504634

Chapter 1 Introduction

Understanding brain function demands deep, high-speed imaging of neuronal circuits. Two-photon excitation microscopy enables deep-tissue imaging with reduced phototoxicity, but its speed is limited by point scanning. To observe the functional response of the suprachiasmatic nucleus (SCN) *in vivo* in mice, we developed dual-GRIN two-photon endoscopy with a tunable acoustic GRIN lens for video-rate volumetric imaging. However, deep imaging and fast scanning degrade image contrast, which is further worsened by GRIN lens aberrations. Here, we apply a self-supervised deep-learning method that leverages spatial and temporal redundancies, along with a deconvolution algorithm to correct aberrated PSFs. This approach improves image contrast and preserves neural activity signals without requiring high-quality ground truth.

1.1. The challenge of understanding the deep brain region

Unveiling how the brain works is a significant challenge in scientific research [1]. An essential goal is to understand neuron activities and connections *in vivo*. A neuron is an electrically excitable cell that communicates with other cells through specialized connections called synapses, as shown in **Figure 1-1**. It consists of message-receiving dendrites, a nucleus-containing cell body, and a signal-sending axon. The size of a cell body is about 10 μm. Neural signal transmission occurs via electrical impulses known as action potentials, which propagate along the axon and trigger neurotransmitter release at synapses [2, 3]. Various technologies, including electrophysiology, fMRI, and optical microscopy, have emerged for studying and understanding the complexities of the brain [4]. Since action potentials result from changes in membrane potential, the most direct way to monitor them is by measuring these

doi:10.6342/NTU202504634

voltage shifts. This principle forms the basis of electrophysiology, where conductive microelectrodes are inserted into or placed on neurons to detect electrical signals such as voltage and current. Microelectrodes provide excellent temporal resolution, making them well suited for capturing rapid neural activity. Although miniaturized electrode arrays [5] have been developed to record signals from multiple neurons, this lack of spatial information continues to impose significant limitations. Microelectrodes only capture extracellular voltage fluctuations from neurons in the immediate vicinity, and precise placement of an electrode [6] onto a single target neuron is technically challenging, particularly in regions with dense neuronal populations. Unlike optical imaging, which directly reports activity through changes in ion channel dynamics and intracellular calcium concentrations, electrodes lack molecular-level specificity.

To monitor brain activity without invasive methods, functional magnetic resonance imaging (fMRI) is a widely adopted technique [7]. It leverages the fact that oxyhemoglobin and deoxyhemoglobin produce different magnetic resonance signals, allowing the detection of the blood-oxygen-level-dependent (BOLD) signal using strong magnetic fields. When neurons become active, they consume more oxygen, leading to a local decrease in blood oxygen levels. This change enables fMRI to indirectly track brain activity by observing variations in the BOLD signal. With its non-invasive nature and whole-brain imaging capability, fMRI has become one of the preferred tools for studying brain function at the organ level. However, its spatial and temporal resolutions, typically in the sub-millimeter and sub-second ranges, are insufficient for single-cell activity recording.

doi:10.6342/NTU202504634

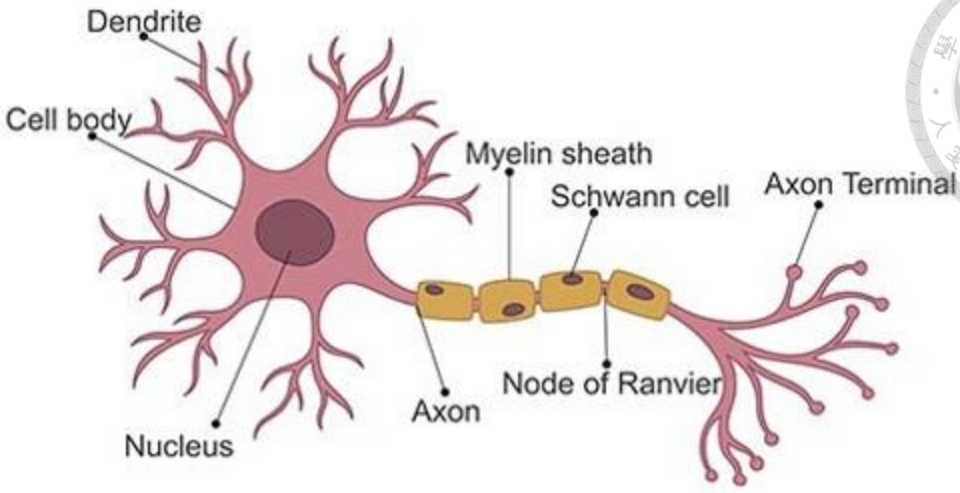


Figure 1-1. Schematic of a neuron structure.

A neuron consists of message-receiving dendrites, a nucleus-containing cell body, and a signal-sending axon.

To study neuronal connectivity, optical imaging has emerged as a powerful tool, offering sub-micrometer spatial resolution and sub-millisecond temporal resolution [8, 9, 10], particularly when paired with genetically encoded probes such as calcium indicators [11, 12]. These indicators detect neural activity by sensing changes in intracellular calcium concentration during action potentials. Among them, green fluorescent protein (GFP)-based GCaMP sensors have been engineered for this purpose [13]. For example, the jGCaMP7f indicator shows a response duration of approximately 265 ± 20 ms following an action potential, as illustrated in **Figure 1-2**. To investigate deep and functionally critical regions, advanced optical imaging methods are required. This leads directly to two-photon fluorescence microscopy, which has become indispensable for *in vivo* studies of neuronal dynamics.

doi:10.6342/NTU202504634

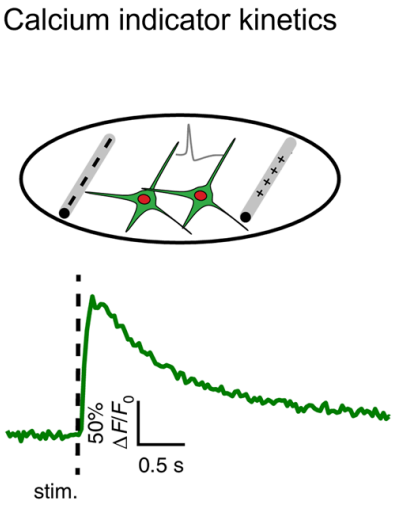


Figure 1-2. Calcium indicator kinetics.

Top, schematic cultured neurons with stimulation electrodes (gray). Cultured neurons expressed a cytosolic jGCaMP7 variant and nuclear mCherry. Bottom, change in fluorescence of a single well with cells expressing jGCaMP7 after firing one action potential [14].

1.2. Two-photon Fluorescence Microscopy *in vivo*

Optical imaging of deep tissue is hindered by scattering from complex tissue components, which historically restricted imaging to thin slices or superficial layers. The advent of two-photon microscopy transformed *in vivo* imaging by enabling high-resolution measurements hundreds of micrometers below the surface. Two-photon microscopy uses NIR for deep penetration and less scattering [15, 16]. Thanks to rapid improvements in fluorescent markers and two-photon microscopy unique ability to capture dynamic biological processes with fine spatial (<1 μm) and temporal (millisecond) precision at depth, it has transformed *in vivo* imaging by allowing direct observation of cellular activity deep within living tissues (for

doi:10.6342/NTU202504634

example, up to about 1 mm in the mouse brain) [17]. Nowadays, two-photon microscopy is an essential technique across many research fields, including neuroscience, immunology, and cancer biology. It is estimated that there are between 1,000 and 10,000 multiphoton microscopes in use globally.

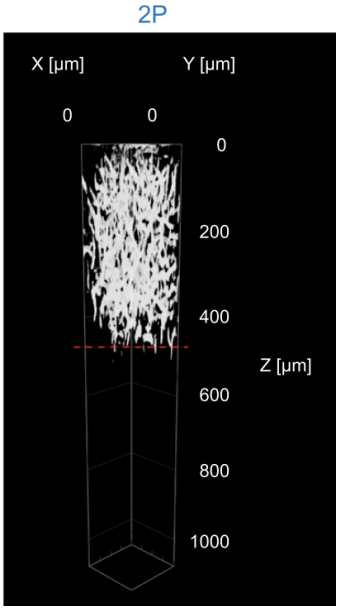


Figure 1-3. 3D renderings showing two-photon microscopy imaging of dendrites in *in vivo* mice down to approximately 500 µm [18].

Two-photon microscopy achieves optical sectioning via nonlinear excitation of fluorescence through the coincident absorption of two illumination photons. This process relies on two photons that interact nearly simultaneously with a fluorescent molecule [19]. Two-photon microscopy offers two significant advantages. First, for commonly used fluorescent reporters, two-photon absorption occurs in the near-infrared wavelength range (680 − 1,030 nm), resulting in a deep penetration depth of tissue and less phototoxic to biological tissues than visible light used for one-photon fluorescence excitation [20]. Second, the localization of excitation not only

doi:10.6342/NTU202504634

reduces photodamage but also provides volumetric resolution without any spatially resolved detection through a confocal pinhole. These properties are crucial for deep imaging, as all fluorescence photons are known to originate from near the focus and thus increase the efficiency of fluorescence collection [21, 22]. Two-photon microscopy is achieved by ultra-short pulses of 100 fs duration at a repetition rate of 80 MHz to generate high intensity within the objective focus. Despite these successes, the penetration depth of *in vivo* mice is still limited to 500 μm (**Figure 1-3**) [23] and the average laser power within the sample is typically 1–100 mW [24], [25]. Overcoming this depth barrier has motivated the development of new optical and hardware strategies, as discussed next.

1.3. Extending Depth and Speed in Optical Imaging

To observe live brain functions beyond 500 μm depth using light microscopy, researchers have developed several approaches [26]. One common strategy involves increasing the excitation power and enhancing the number of ballistic photons by incorporating a regenerative amplifier. This allows imaging depths of up to approximately 1 mm in the neocortex [27]. Another approach uses three-photon excitation at longer wavelengths (1300 nm or 1700 nm), which experience less scattering in tissue, enabling imaging depths of around 1.2 mm [28]. More recently, powerful long-wavelength lasers have pushed this limit further to about 2 mm [29]. Despite these advancements, reaching the full depth of a mouse brain remains out of reach with these noninvasive optical methods [30].

To reach deep areas, researchers often implant miniature optical elements such as GRIN lenses and optical fibers [31]. Optical fibers constitute the foundation of contemporary communication technologies, providing low-loss light guidance and high transmission capacity.

doi:10.6342/NTU202504634

Developments in multimode and few-mode fibers have further expanded bandwidth through spatial multiplexing, yet the resulting complexity in mode management underscores the importance of optical elements such as GRIN lenses [32].

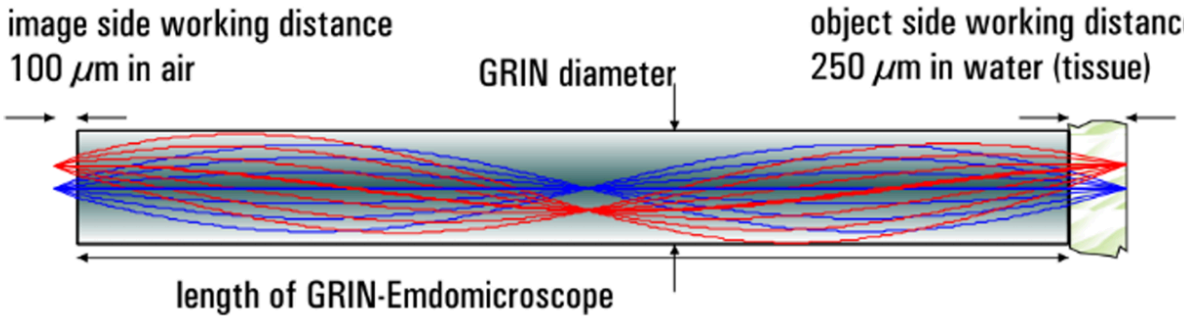


Figure 1-4. Schematic illustration of a GRIN lens.

Schematic of a GRIN endomicroscope. Light rays (red and blue) propagate through the GRIN lens, which relays the image from the tissue to the image plane. The system has an image-side working distance of 100 μm in air and an object-side working distance of 250 μm in water (tissue). The GRIN diameter and length define the physical dimensions of the probe.

GRIN lenses, in particular, have become popular due to their compact size which is typically around 0.5 mm in diameter and 7 mm in length [33, 34, 35]. These lenses work by varying the refractive index radially, following a hyperbolic secant function controlled by dopant concentrations.This design allows light to be focused and relayed along the optical axis through the lens (**Figure 1-4**). Owing to the use of GRIN lenses, the penetration depth of two-photon endoscopy is no longer limited to 2 mm. This technique enables researchers to look into deep brain functions.

Building on this foundation, it is highly desirable to study the intricate neuronal networks in specific brain regions through volumetric functional imaging within the mouse brain [36, 37, 38], for which several strategies have been developed for two-photon endoscopy (**Figure**

doi:10.6342/NTU202504634

1-5). The most straightforward approach involves mechanically moving the objective lens **(Figure 1-5A**); however, this method is inherently slow and limited to partial sampling of neural responses [39]. Alternatively, spatial light modulators (SLMs) [40] can be employed to generate patterned illumination or Bessel beams, enabling simultaneous observation of a volume beneath a GRIN lens (**Figure 1-5B**). While effective in principle, this approach is primarily restricted to sparse samples due to cross-talk between planes and typically requires additional post-processing to separate signals from different depths. Electrically tunable lenses (ETLs) have also been incorporated to accelerate axial scanning (**Figure 1-5C**), but their response speed is generally limited to multi-plane imaging rather than continuous volumetric acquisition [41]. In practice, these techniques either suffer from slow axial scanning speeds or compromised axial resolution, making it difficult to resolve fast, simultaneous activity from neurons located at different depths. To address these limitations, we aim to integrate a tunable acoustic gradient (TAG) lens (**Figure 1-5D**), which enables high-speed axial scanning at hundreds of kilohertz (**Table 1-1**), thereby allowing true volumetric functional imaging in real time. Yet, while such strategies extend imaging depth and speed, they also introduce emerging difficulties, including tissue scattering, reduced contrast, and optical aberrations from GRIN lenses. These issues highlight the need for effective image-quality correction, which we address in the next section.

doi:10.6342/NTU202504634

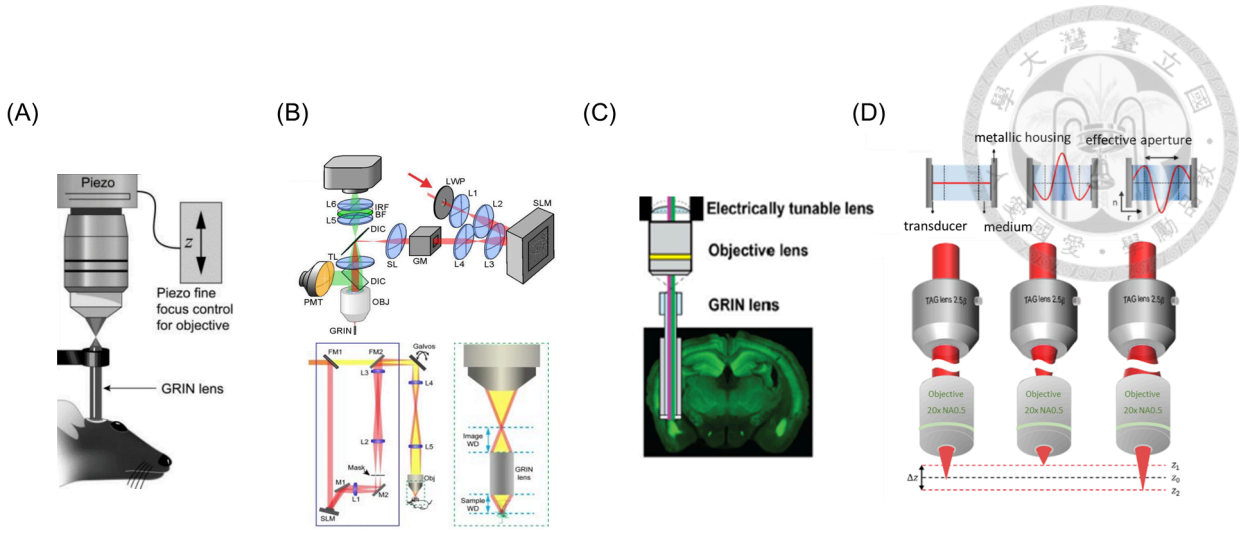


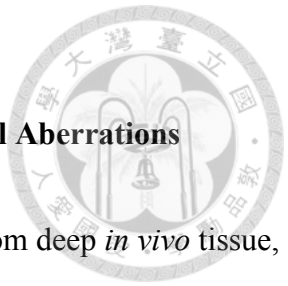
Figure 1-5. Strategies for fast volumetric imaging in two-photon endoscopy.

Corresponding techniques: (A) Moving objective. (B) Spatial light modulator. (C) Electrically tunable lens. (D) Tunable acoustic gradient (TAG) lens**.**

	Technique	Speed	Feature
A	Moving objective	~100 Hz	Too slow/ Partial sampling
B	Spatial light modulator (Patterned illumination or Bessel beam)	Simultaneous	Z Projected image Sparse sample
C	Electrically tunable lens	~1 kHz	Few-plane images
D	TAG lens	~100 kHz to 1MHz	Whole volume image

Table 1-1. Comparison of axial scanning strategies for volumetric two-photon endoscopy.

doi:10.6342/NTU202504634



1.4. Handling Image Degradation Caused by Deep Imaging and Optical Aberrations

Although our 3D imaging system is capable of capturing signals from deep *in vivo* tissue, the combination of tissue observation and high-speed scanning inevitably compromises image contrast [42]. This limitation is further aggravated by the intrinsic fourth-order aberrations inherent to GRIN lenses. To address such issues, image processing plays a pivotal role in enabling accurate data analysis. Conventional approaches typically rely on handcrafted algorithms and follow a sequence of predefined steps, including histogram equalization, image projection (averaging), smoothing filters, and deconvolution [43, 44]. While effective to some extent, these methods fail under conditions of low signal-to-noise ratio (SNR) and may introduce artifacts that distort functional imaging outcomes [45, 46].

In recent years, deep learning [47] has transformed image processing tasks such as noise reduction, segmentation, recognition, and classification [48]. As a subset of machine learning inspired by neural processing in the human brain, it has become an essential tool for handling complex imaging problems [49]. The introduction of convolutional neural networks (CNNs) in the late 1990s marked a breakthrough in this field, and continuous development has since advanced supervised, unsupervised, and self-supervised learning paradigms. Supervised learning relies on labeled datasets to achieve high prediction accuracy, unsupervised learning extracts hidden patterns without labels, and self-supervised learning [50] generates supervisory signals directly from the internal structure of the data [51], thereby reducing dependence on manual annotation.

doi:10.6342/NTU202504634

In parallel, optical correction strategies have been extensively investigated to mitigate aberrations in GRIN lenses. Early efforts employed low-order electrostatic membrane mirrors for optical compensation, while pupil segmentation techniques enabled more efficient aberration correction, thereby enhancing fluorescence signal intensity and extending the field of view (FOV) [52]. Subsequent work introduced geometric transformation adaptive optics, which provided improved correction but required substantial modifications to the optical path, additional software control, and often compromised temporal resolution. Alternative approaches, such as the integration of millimeter-scale plano-convex lenses with GRIN optics, demonstrated on-axis correction and increased numerical aperture (NA) [53]. However, these methods remain impractical for GRIN lenses with cross-sectional diameters below 1 mm due to manufacturing challenges in producing precise small-scale optics. More recently, 3D micro-printing of polymer aspherical corrective lenses directly at the GRIN back end has been proposed as a built-in solution [54], offering compactness and ease of use. Nevertheless, demonstrations of this method have been limited to GRIN lenses with a radius of FOV around 150 μm and remain constrained in imaging depth.

Given these limitations, we developed two post-processing approaches. One is a self-learning model designed to denoise the image series, and the other is a deconvolution-based algorithm to compensate for GRIN-induced aberrations. Unlike adaptive optics or hardware-based correction methods [55], our approach requires no modification of the optical system, preserves temporal resolution, and provides a practical framework for enhancing image quality in deep tissue imaging

doi:10.6342/NTU202504634

1.5. Deep brain target of interest – the Suprachiasmatic Nucleus

By combining our home-built two-photon endoscopy with two post-processing methods, we are able to investigate the biological mechanisms of specific deep-brain regions in live animals. Among these, particular attention has been devoted to the SCN, a small structure in the hypothalamus located ~6 mm beneath the surface of the mouse brain (**Figure 1-6A**) [56]. Serving as the central circadian pacemaker, the SCN synchronizes peripheral clocks across diverse tissues to maintain temporal homeostasis [57]. The presence of an endogenous circadian rhythm, which can be entrained by environmental cues, is essential for physiological and behavioral regulation. A critical component of this system is the light pulse signal transmitted from intrinsically photosensitive retinal ganglion cells (ipRGCs), which aligns the internal circadian clock with the external light–dark cycle, a process known as circadian photoentrainment [58]. Across species, light stimulation induces distinct phase-shift responses, exposure during the early subjective night typically produces a phase delay, while exposure during the late subjective night induces a phase advance. In contrast, exposure during the so-called “dead zone” of the subjective day yields minimal phase shifts. Despite decades of research, the fundamental properties of the SCN that underlie these phase-shifting behaviors remain a subject of ongoing debate [59, 60]. In this context, our two-photon endoscopy system, employing a 7 mm implanted GRIN lens in the mouse brain, provides an experimental model that enables investigation of neuronal responses within the SCN region (**Figure 1-6B**).

doi:10.6342/NTU202504634

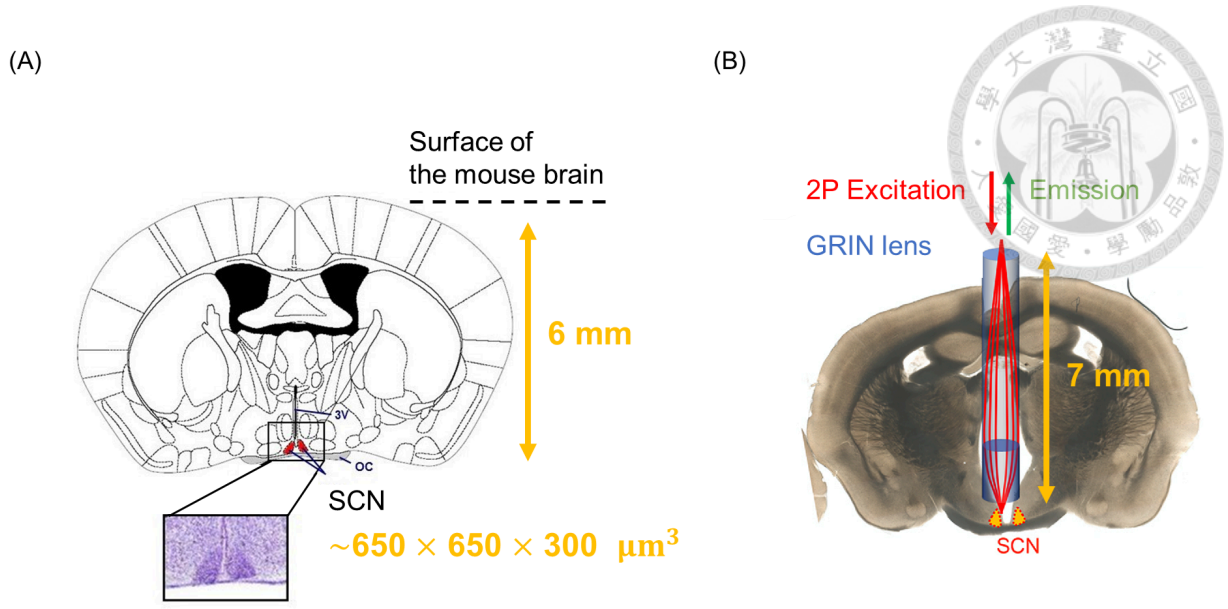


Figure 1-6. Optical access to the mouse SCN via implanted GRIN lens.

(A) Schematic coronal brain atlas showing the SCN, located ~6 mm beneath the cortical surface. The implanted GRIN lens provides a volumetric imaging field of ~650 × 650 × 300 μm³ within the SCN. (B) A 7-mm gradient-index (GRIN) lens is stereotaxically implanted into the mouse brain, allowing two-photon (2P) excitation light (red) to reach the SCN and enabling fluorescence emission (green) to be collected from the same pathway.

doi:10.6342/NTU202504634

1.6. Aim of this Study

In this study, we aim to develop and demonstrate a custom-built two-photon endoscopy system capable of capturing dynamic three-dimensional neural activity in deep brain regions with enhanced image quality by deep-learning denoising and aberration correction. Specifically, we target the SCN, located approximately 6 mm beneath the cortical surface of the mouse brain. To achieve this, we integrate TAG and GRIN lenses into our optical system. To overcome the reduced image contrast commonly encountered in deep tissue imaging, we apply deep-learning-based post-processing techniques. These methods rely on self-supervised learning models that exploit spatial and temporal redundancies to perform effective denoising. Furthermore, we implement a customized algorithm to correct optical aberrations introduced by the imaging system.

Chapter 2 introduces the underlying principles of the technologies used, including the two-photon laser scanning system, TAG lens, GRIN lenses, and the self-attention (SA) deep-learning model. Chapter 3 details the experimental procedures, including performance evaluation using fluorescent beads and mouse brain slices, as well as the development of an algorithm for calibrating dark image patterns.

By combining advanced optical hardware with intelligent computational processing, our system successfully reconstructs high-contrast, high-resolution volumetric images of neuronal activity *in vivo*. Chapter 4 presents the experimental results and corresponding post-processed images, along with quantitative analyses. This approach enables precise visualization of the SCN and provides a promising platform for investigating circadian rhythm regulation and deep brain functional connectivity.

doi:10.6342/NTU202504634

Chapter 2 Principle

In this chapter, we present the optical and computational foundations underlying our imaging system, beginning with the principle and implementation of two-photon laser scanning microscopy. We first discuss GRIN lenses, which serve as compact optical relays for deep-brain access, followed by the introduction of the TAG lens, a high-speed adaptive optical element that enables rapid axial scanning for volumetric imaging. Building upon these hardware components, we next describe our development of TAG-SPARK 2.0, a self-supervised deep learning framework that enhances imaging quality by incorporating attention mechanisms into a 3D U-Net architecture. Finally, we address deconvolution techniques, which compensate for optical aberrations and scattering to further restore fine structural details. Together, these elements establish the basis of our methodology, combining advanced microscopy, adaptive optics, and computational correction to achieve high-resolution volumetric imaging of neural activity *in vivo*.

2.1 Two-photon laser scanning microscopy

To understand two-photon laser scanning microscopy, we begin by introducing its theoretical foundation, which is the concept of two-photon absorption. We then explain how this principle is applied in laser scanning microscopy. The idea of two-photon absorption was first proposed by Maria Göppert-Mayer in 1931 [61]. As the name suggests, this phenomenon occurs when two photons, usually of equal energy, are absorbed by a single molecule at the same time. The combined energy of these photons excites the molecule as if it had absorbed one photon with twice the energy (**Figure 2-1**). If the molecule is fluorescent, it subsequently returns to its ground state by emitting a single fluorescent photon. Because this process requires both photons to

doi:10.6342/NTU202504634

arrive nearly simultaneously, the probability of two-photon absorption depends on the square of the excitation light intensity ($S \propto I^2$) [62]. The efficiency of two-photon absorption is measured using the unit Göppert-Mayer (GM), which defines the two-photon absorption cross-section [63]. Most commonly used fluorescent dyes and proteins have cross-section values in the range of 1 to 300 GM . Due to the extremely small likelihood of two-photon absorption, high-intensity light is required to observe this effect.

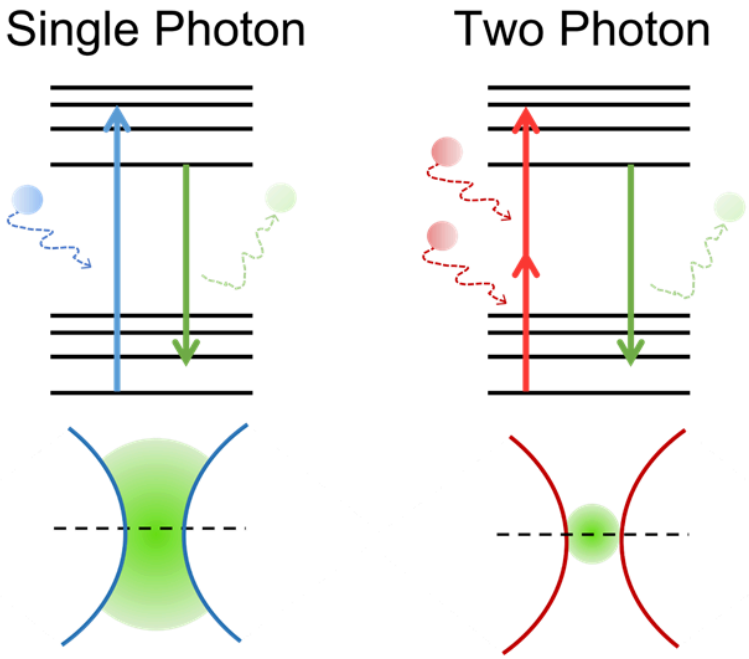


Figure 2-1. Fluorescence emission in one- and two-photon microscopy.

To capture a two-dimensional image, the focal spot of the laser must be scanned across the sample. This process is similar to how images were formed on old cathode-ray tube (CRT) televisions, where electron beams were deflected to scan across a phosphorescent screen and build up the image line by line. In two-photon laser scanning microscopy, a similar principle is applied. The excitation laser is deflected using a pair of galvanometric mirrors, which perform a

doi:10.6342/NTU202504634

raster scanning motion to sequentially illuminate each point on the focal plane. As the laser excites the sample point by point, the resulting fluorescence emitted from each location is collected by a point detector, typically a photomultiplier tube (PMT) [64].The detected signals are then sent to a computer, which assigns the intensity values to the corresponding positions in a two-dimensional matrix, effectively reconstructing the image pixel by pixel. This forms a 2D image of the focal plane. By shifting the focal plane along the axial (z) direction and repeating the scanning process, a stack of 2D images can be acquired at different depths. These image layers can then be combined to generate a full three-dimensional representation of the sample [65]. With its ability to perform optical sectioning and to penetrate deeper into tissue than traditional fluorescence microscopy, two-photon laser scanning microscopy allows researchers to visualize volumetric structures in living tissues with high resolution and clarity at the cellular level.

Two-Photon scattering principle

Only photons that travel straight without scattering, which is known as ballistic photons, are responsible for generating two-photon (2P) excited fluorescence within the small focal region [66]. Since the focal spot is very small (around 1 μm or less) compared to the imaging depth (which can range from hundreds to thousands of micrometers), photons that have been scattered and deviated from this direct path are highly unlikely to reach the focus. Because the number of ballistic photons decreases exponentially as imaging depth increases, the fluorescence signal detected in two-photon fluorescence microscopy (2PFM) also diminishes exponentially with depth, a trend confirmed by many experimental studies. The characteristic length over which this signal decays, called the effective attenuation length (EAL), mainly depends on how much the tissue absorbs and scatters light. Although increasing the excitation power can partly compensate

doi:10.6342/NTU202504634

for signal loss at greater depths up to the point where tissue damage occurs, doing so reduces the three-dimensional confinement of excitation. This effect is measured by the signal-to-background ratio (SBR), where the signal (S) is the fluorescence generated at the focal point, and the background (B) is fluorescence coming from other areas. When considering only ballistic photons, the SBR can be approximated by the formula $SBR \approx z^2 e^{-\frac{2z}{EAL}}$, where z represents imaging depth. Studies using simulations and tissue phantoms have shown that scattered photons do not contribute to the focused signal but significantly add to background fluorescence. As imaging goes deeper, fluorescence generated outside the focal volume, which lacks high-resolution information, increases and eventually overwhelms the signal, causing the SBR to approach zero and resulting in a complete loss of image contrast.

2.2 Gradient-index lenses (GRIN) lens

The concept of the GRIN lens can be traced back to James Clerk Maxwell in 1854 and was later further developed by Robert William Wood in the early twentieth century [67]. Fundamentally, GRIN lenses are designed to achieve ideal on-axis optical focal relays. The refractive index distribution of such lenses typically follows a hyperbolic secant profile, expressed as

$$n(r) = n_0 sech(gr). \qquad \text{Eq. 2-12}$$

where n_0 denotes the refractive index along the optical axis, g represents the gradient constant that characterizes the radial variation of refractive index, and r is the radial distance from the axis. This profile allows the focal point to be periodically relayed along the optical axis with minimal aberration.

doi:10.6342/NTU202504634

The gradient constant g is directly related to the physical length of a full-pitch GRIN lens, given by

$$L_{pitch} = \frac{2\pi}{g}. \quad \text{Eq. 2-13}$$

, and simultaneously determines the maximum numerical aperture (NA) that the GRIN lens can support. The relationship between the NA and lens geometry is described by

$$NA = n_0\sqrt{1 - sech^2(\frac{gd}{2})}. \quad \text{Eq. 2-14}$$

, where d is the lens diameter. These characteristics highlight the ability of GRIN lenses to serve as efficient optical relays while maintaining compact geometries and reduced aberration compared to conventional refractive optics [68].

Aberrations of GRIN lens

Since the GRIN lens is inherently designed to achieve perfect on-axis focal relay, light rays deviating from the optical axis experience aberrations. Two dominant off-axis aberrations are observed, **field curvature** and **astigmatism**, both classified as fourth-order aberrations with quadratic dependence on the field position . Two types of aberrations, astigmatism and field curvature, arise in cylindrical GRIN structures and result in fourth-order aberrations in the imaging system.

doi:10.6342/NTU202504634

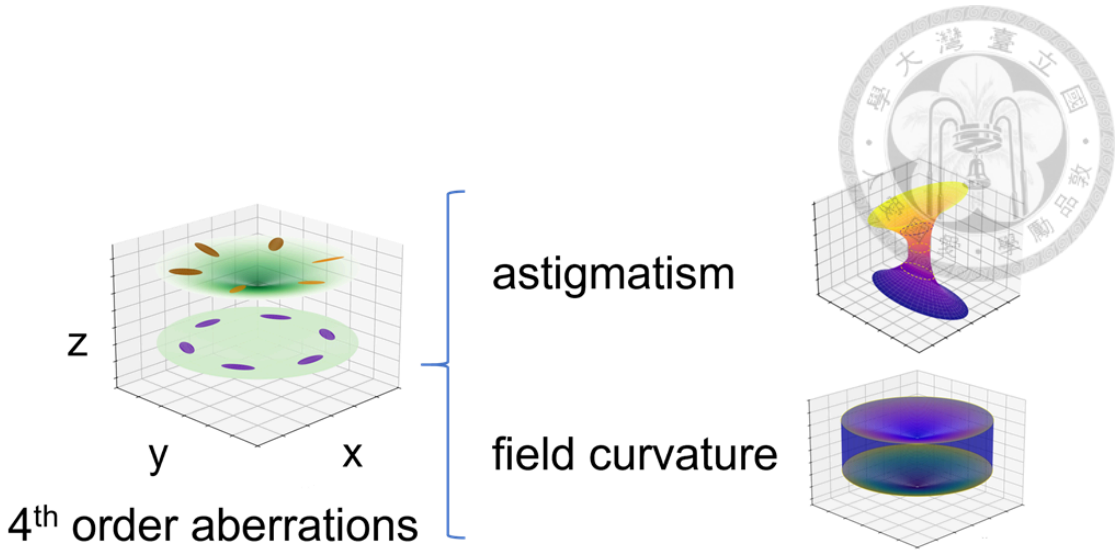


Figure 2-2. Aberrations caused by a GRIN lens.

Astigmatism is particularly detrimental to image quality, and its effects become pronounced with increasing off-axis distance. As illustrated in upper part of **Figure 2-4**, this multiplicity of foci leads to artifacts in volumetric imaging, where the same structural feature may appear more than once within the reconstructed volume. Such distortions significantly compromise both spatial accuracy and interpretability of the acquired data.

Field curvature, on the other hand, results in a non-planar imaging surface (the bottom part of **Figure 2-4**). While the tangential foci remain approximately on a flat plane, the sagittal foci instead reside on a quadratic surface described by

$$x^2 + y^2 = az. \qquad \text{Eq. 2-15}$$

The accompanying 3D visualization illustrates this curvature of the focal surface, with colored ellipses highlighting the separation between sagittal and tangential focal planes caused by

doi:10.6342/NTU202504634

astigmatism across different radial regions [69]. In the context of volumetric deep-brain imaging with GRIN lenses, field curvature directly impacts both image intensity uniformity and effective imaging throughput [70]. Specifically, the central region of the volume of view (VoV) appears at greater depth within the tissue and thus suffers from reduced focal intensity due to an increased scattering path length. This non-uniformity may be partially mitigated through modulation of the excitation laser intensity as a function of field position. Furthermore, field curvature produces characteristic artifacts in volumetric recordings: in the upper portion of the imaging volume, the periphery of the VoV becomes dark because the true focal positions fall outside the tissue boundaries; conversely, in the lower portion of the volume, the central region darkens as the focal depth exceeds the maximum penetration limit of the system [71].

2.3 Tunable Acoustic Gradient Lens (TAG lens)

A TAG lens consists of a cylindrical cavity filled with silicone oil, enclosed by a hollow piezoelectric tube with two flat glass windows providing optical access (**Figure 2-2A**) [72]. When an AC voltage is applied to the piezoelectric transducer, it excites radial vibrations in the tube wall. These vibrations launch an acoustic standing wave in the fluid, periodically modulating its refractive index (**Figure 2-2B**). The focal length of the lens changes continuously during each oscillation cycle, and because the acoustic resonance frequency is typically in the hundreds of kHz range. Its tuning can be achieved with sub-microsecond temporal resolution [73].

doi:10.6342/NTU202504634

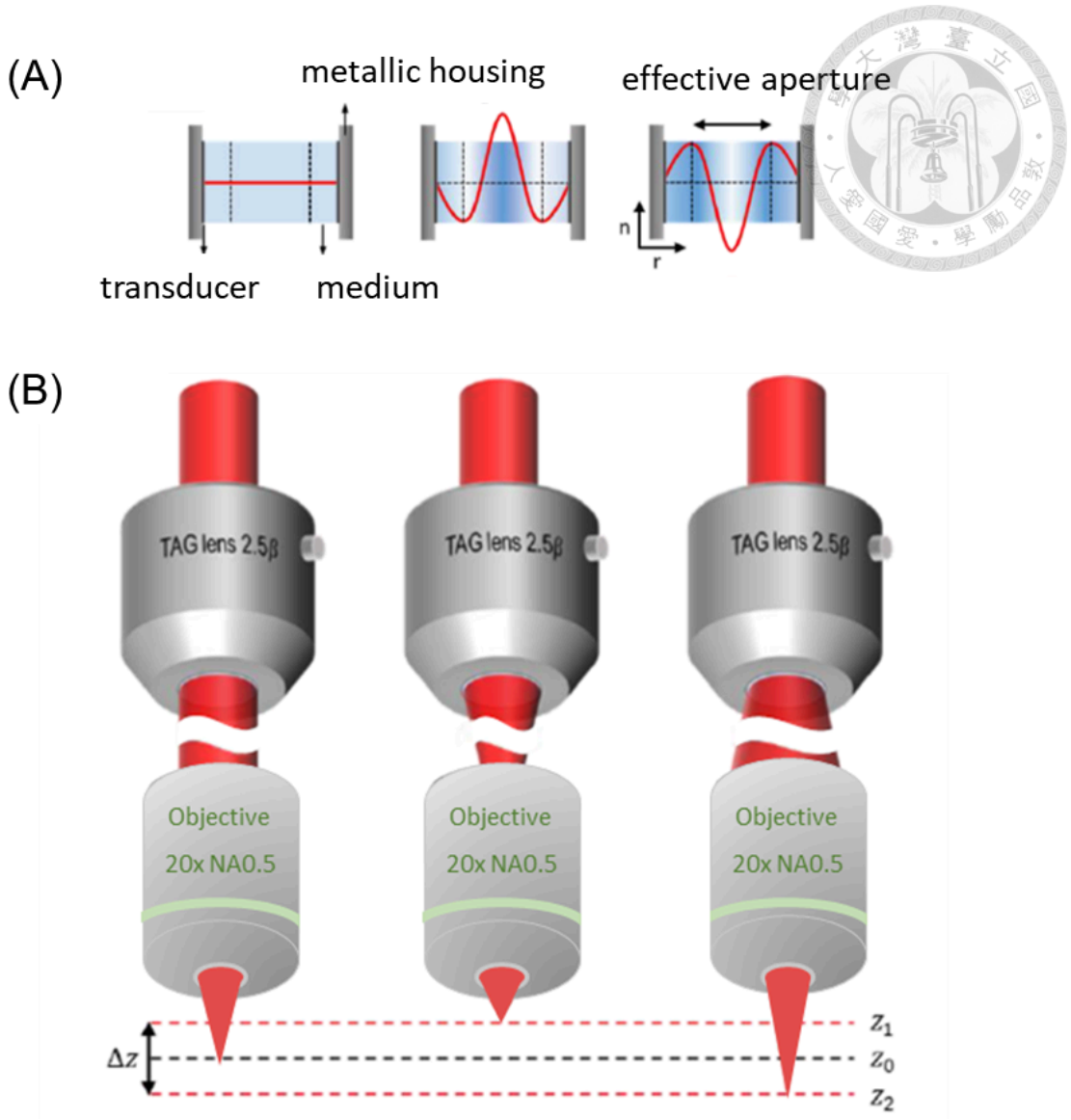


Figure 2-3. Axial focal shift of objective induced by TAG lens resonance.

(A) Three conditions of the TAG lens are shown, namely no effect, convergent effect, and divergent effect. The red curve in each condition represents the refractive index profile along the radial direction inside the TAG lens. (B) Black dotted lines indicate the effective aperture of the TAG lens. d: distance between the objective and the TAG lens. z_0: inherent focal plane. z_1: nearest focal plane. z_2: farthest focal plane. Δz: axial shift between z_1 and z_2.

doi:10.6342/NTU202504634

Piezoelectric Excitation and Acoustic Field

Electrodes are deposited on the inner and outer surfaces of the piezoelectric cylinder. The radial vibration velocity V_A of the inner surface is proportional to the driving voltage amplitude

$$V = V_A sin(\omega t). \quad \text{Eq. 2-1}$$

where ω is the driving frequency of the TAG lens.

The acoustic field inside the fluid forms concentric standing-wave rings in the radial direction [74]. The spatial distribution follows a Bessel function, with geometry and spacing determined by the drive conditions and refractive index modulation. Under the linearized fluid mechanics approximation, the refractive index distribution is given by:

$$n(\rho, t) = n_0 + n_A J_0(\frac{\omega\rho}{c_s}) sin(\omega t). \quad \text{Eq. 2-2}$$

where n_0 is the static refractive index of the fluid, n_A is related to modulation amplitude, ω is TAG lens driving frequency, c_s is acoustic velocity in the fluid, and ρ is the radial coordinate in the lens plane. Eq. 2-2 shows that the refractive index varies radially as a Bessel function, which defines the strength and location of the focusing rings.

Optical Phase Modulation

At the lens plane, the transverse coordinates are defined such that the optical phase imparted to the light is:

doi:10.6342/NTU202504634

$$t_{lens}(\xi, \eta) = e^{ik_0(nL_0)}. \quad \text{Eq. 2-3}$$

, where k_0 is the free-space wavenumber and L_0 is the lens thickness.

Since the ray deflection (<50 μm) is much smaller than the acoustic wavelength, the thin lens approximation holds that light exits the lens at the same transverse location it enters.

If the wavefront propagates an additional distance [75], z, beyond the exit surface (Eq. 2-4), the constant-phase surface satisfies

$$t_{lens}(\xi, \eta) = e^{ik_0(n(L_0+(\rho)))} = constant = C. \quad \text{Eq. 2-4}$$

$$n(\rho)(L_0 + z(\rho)) = C \text{ and } z(\rho) = \frac{C}{n(\rho)} - L_0. \quad \text{Eq. 2-5}$$

, where z<0 denotes the distance from the exit window to the wavefront.

When $|z| \ll L_0$, the beam deflection angle inside the fluid is

$$tan(\theta(\rho)) = \frac{\partial z}{\partial \rho} = \frac{-C}{n^2}\frac{\partial n}{\partial \rho} = -\frac{L_0 \partial z}{n \partial \rho}. \quad \text{Eq. 2-6}$$

At the interface, Snell's law gives

$$sin(\theta'(\rho)) = n(\rho)sin(\theta(\rho)). \quad \text{Eq. 2-7}$$

where θ' is the angle of light after leaving the lens.

Assuming $n_{air} = 1$ and applying the small-angle approximation to Eqs. 2-6 and 2-7 yields

doi:10.6342/NTU202504634

$$\theta' = n\theta \text{ and } \theta'(\rho) = -\frac{L_0 \partial z}{\partial \rho}. \quad \text{Eq. 2-8}$$

The exit angle of the beam is proportional to the radial gradient of the refractive index, meaning stronger modulation produces larger angular deviations. During a single oscillation period, three representative states occur in the TAG lens.

Uniform index, $sin(\omega t) = 0$ and $n = n_0$, no lensing effect.

Positive index, $sin(\omega t) > 0$ and $n = n_{max}$, convex lens (converging beam).

Negative index, $sin(\omega t) < 0$ and $n = n_{min}$concave lens (diverging beam).

For sin (t) = 0, the refractive index is uniform and the beam's wavefront is preserved. A positive sin (t) corresponds to a local refractive index maximum, where the TAG lens behaves as a convex lens, converging the beam. A negative sin (t) corresponds to a local refractive index minimum, where the TAG lens behaves as a concave lens, diverging the beam.

Increasing the driving frequency compresses the radial of the major rings, and the first major ring, is relative to the effective numerical aperture (NA) of the TAG lens [76], which is approximately given by,

$$\rho' = \frac{3.832 c_s}{\omega}. \quad \text{Eq. 2-9}$$

where 3.832 is the first minimum of $J_0(\rho)$.

doi:10.6342/NTU202504634

Depth of Field and Axial Scanning

When the TAG lens is placed in the conjugate plane of an objective, it shifts the focal position along the optical axis [77]. The axial range or **depth of field (DOF)** is given by:

$$DOF = n \times f^2 \times \delta_{total} \times \left(\frac{f_2}{f_1}\right)^2 \text{ and } \delta_{total} = \frac{1}{f_{TAG}}. \qquad \text{Eq. 2-10}$$

, where n is refractive index of the sample, f is the focal length of the objective, δ_{total} is the optical power of TAG lens, f_1 and f_2 are the focal lengths of relay lenses (beam magnification) between TAG lens and objective lens respectively. A larger f or smaller beam magnification increases the DOF.

The optical power scales quadratically with the driving frequency:

$$\delta(t) = \frac{L_0 n \omega^2 sin(\omega t)}{2v^2}. \qquad \text{Eq. 2-11}$$

Thus, increasing ω extends DOF but reduces effective NA, introducing a trade-off between axial range and lateral resolution.TAG lenses offer high-speed tuning, with their performance limited only by the acoustic resonance frequency, which can reach hundreds of kilohertz. They also possess a high damage threshold, making them more robust than liquid-crystal tunable lenses. In addition, their analog operation enables smooth, continuous focal variation without pixelation, ensuring precise and seamless optical adjustments [78].

doi:10.6342/NTU202504634

Signal to noise ratio (SNR) of TAG lens

Owing to the rapid axial focal shift characteristic of the TAG lens, typically tuned to 188 kHz, the number of photons captured by the photomultiplier tube (PMT) is limited. In other words, the pixel dwell time in TAG lens imaging is approximately 1000 times shorter than that in conventional low-speed scanning mode.

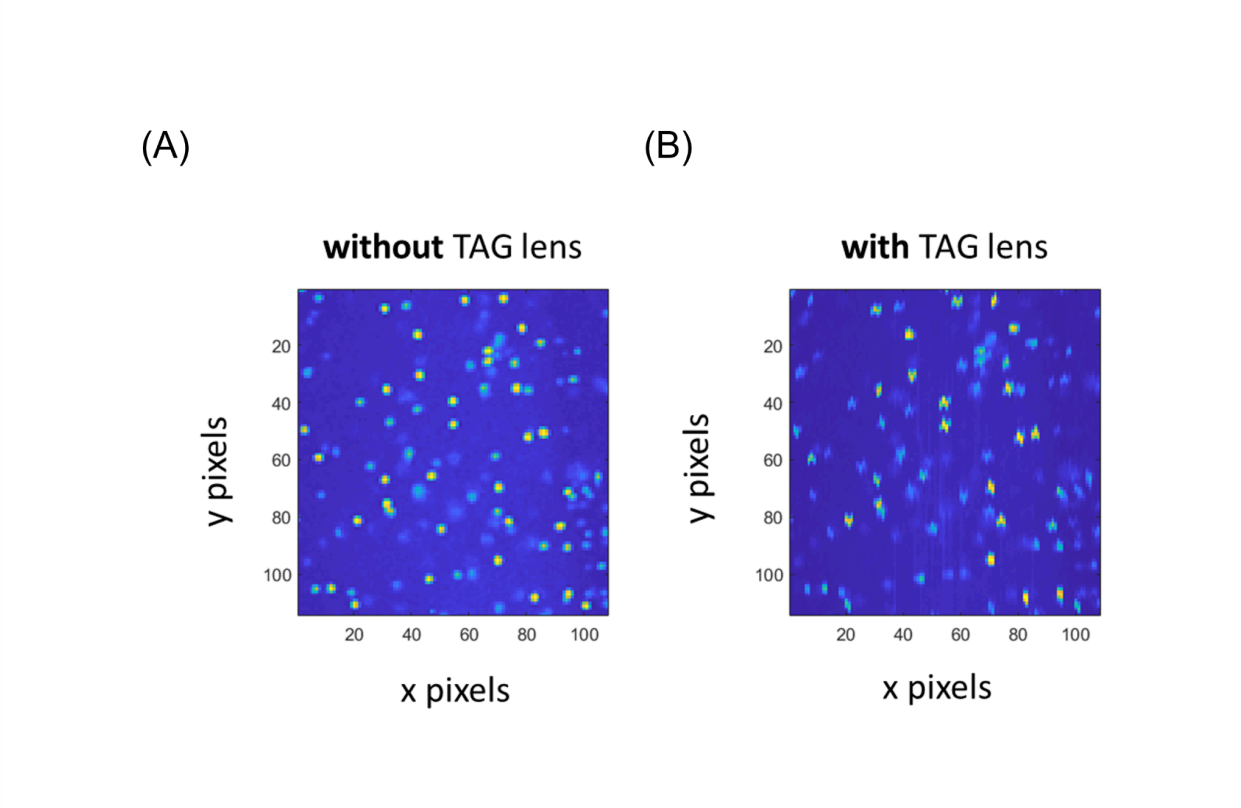


Figure 2-4. Comparison of images acquired with and without the TAG lens.

(A) Image of a layer of 10-μm fluorescent beads (FLBs) obtained with low-speed scanning, where the pixel dwell time is approximately 4000 μs. Image size: 128 × 128 pixels. (B) Image of a layer of 10-μm fluorescent beads (FLBs) obtained with high-speed scanning, where the pixel dwell time is approximately 5 μs and the TAG lens amplitude is 15% (corresponding to 56 μm along the axial direction). Image size: 128 × 128 pixels.

doi:10.6342/NTU202504634

Figure 2-3 presents a comparison of two images of the same 10-μm FLB sample with different scanning speed in our system. The image acquired with the TAG lens exhibits poorer contrast than the image obtained without it (**Figure 2-3A**). To quantify the signal-to-noise ratio (SNR), regions of interest (ROIs) were selected around each FLB shown in the images. The SNR of the left image is 25.3 dB, whereas that of the right image is 21.1 dB. Thus, with the TAG lens (**Figure 2-3B**), the SNR decreases and the image quality is reduced by approximately a factor of 2.63 compared to low-speed scanning. Intuitively, although the TAG lens enhances volumetric imaging speed, this improvement comes at the expense of image quality. Balancing temporal resolution and spatial resolution without significant compromise is therefore one of the central objectives of this research. This trade-off highlights the necessity of developing strategies to improve image quality in high-speed TAG-lens microscopy, which forms a key motivation of this thesis.

2.4 TAG-SPARK 2.0 model

When imaging the brain, one of the most important challenges is seeing fine details clearly. Contrast, or the difference in brightness between regions, allows us to distinguish the shapes and boundaries of cells. But capturing high-quality images is not always easy. Traditionally, researchers have relied on step-by-step image processing methods such as histogram equalization, smoothing filters, and deconvolution [79]. These techniques are based on predefined rules and often perform well when the signal is clean. However, under noisy conditions, such as during live imaging where light levels must be kept low, these methods can struggle and may even introduce unwanted artifacts into the data [80]. This is where deep learning becomes valuable. Inspired by how the human brain processes information, deep learning uses neural networks to learn patterns directly from data [81]. Over the past few

doi:10.6342/NTU202504634

decades, especially with the development of convolutional neural networks (CNNs), deep learning has transformed the field of image analysis. Unlike traditional methods, these models can learn how to reduce noise, identify structures, and interpret complex data. Depending on the training method, deep learning can be classified as supervised, unsupervised, or self-supervised. In supervised learning, models are trained using labeled data provided by humans. In unsupervised learning, the model looks for structure in data without labels. Self-supervised learning stands somewhere in between, where the model generates its own training signals based on the hidden structure of the input data.

For researchers studying brain activity in live animals, self-supervised learning has proven especially useful [82]. One major reason is that clean reference images, which are essential for supervised training, are extremely difficult to obtain. Neural signals change rapidly, and capturing a perfect reference image is often not possible. On the other hand, fully unsupervised methods may compromise the accuracy of the signal. Self-supervised learning avoids these limitations by relying on the data itself. It often uses redundancy in time or space, such as the similarity between consecutive image frames, to guide the learning process [83].

A well-known example is DeepCAD-RT, a self-supervised model developed for calcium imaging. Rather than depending on ground truth data, it learns from temporal similarities across neighboring frames within the same dataset [45]. This approach effectively reduces noise while preserving the structure and dynamics of neural activity. With the help of such models, researchers can gain clearer insights into how the brain functions in real time, opening up new possibilities for understanding complex neural processes.

doi:10.6342/NTU202504634

In this study, we propose an enhanced denoising algorithm developed in collaboration with Wen-Chi Chen, called TAG-SPARK 2.0 model, which is based on the TAG-SPARK framework by incorporating **attention mechanisms** into the 3D U-Net architecture [84]. The goal is to improve overall denoising performance across datasets with diverse structural characteristics and noise types. To accommodate these variations, the algorithm of TAG-SPARK 2.0 (self-attention and spatial attention) model is further developed into two tailored strategies optimized for different application scenarios.

For image sequences with structurally intact content and high spatial redundancy across adjacent slices, spatial attention blocks are inserted before the first-layer convolution in the encoder to emphasize salient regions in the input data. This enhancement facilitates more effective feature extraction and improves subsequent model performance. In addition, self-attention blocks are appended after the convolutional layers in the second to fourth encoding stages to capture long-range dependencies among features. These refined features are then forwarded either to the next encoder layer or fused via skip connections, with the self-attention mechanism enhancing representational capacity in critical regions and improving the model's ability to capture relevant information. Considering the scale difference between Mean Absolute Error (MAE) and Mean Squared Error (MSE), which can lead to imbalance during optimization, the loss function is formulated as a weighted combination to harmonize their contributions (**Figure 2-5**).

$$\text{Loss} = 0.5 \times \text{MAE} + 0.5 \times \sqrt{\text{MSE}}. \qquad \text{Eq. 2-16}$$

For datasets containing structural distortions or lacking sufficient spatial redundancy, a two-stage training scheme is adopted to enhance denoising capability. In the first stage, two

doi:10.6342/NTU202504634

parallel models are trained using different target image pairs to extract both spatial structures and temporal dynamics. At the time of designing the first-stage model, the goal was to capture the dominant noise pattern in the time series, not necessarily the full calcium event. Using 9 consecutive frames (t – 4 … t … t + 4) gives the network a short temporal window over a duration of about 0.72 s, i.e. 8 time intervals that is long enough to learn what is consistent over time (structure) and what is inconsistent (noise), but still short enough that the underlying signal does not change too much. If the SCN calcium response occurs on a temporal scale of approximately 1 s, the temporal parameter in the denoised model should be adjusted from "t – 4 to t + 4" to "t – 5 to t + 5". The averaged outputs of the two models serve as the input for the second stage, in which the model is further trained to complete the denoising task. This progressive refinement leverages enhanced structural features and redundancy established during the first stage (**Figure 2-5)**.

doi:10.6342/NTU202504634

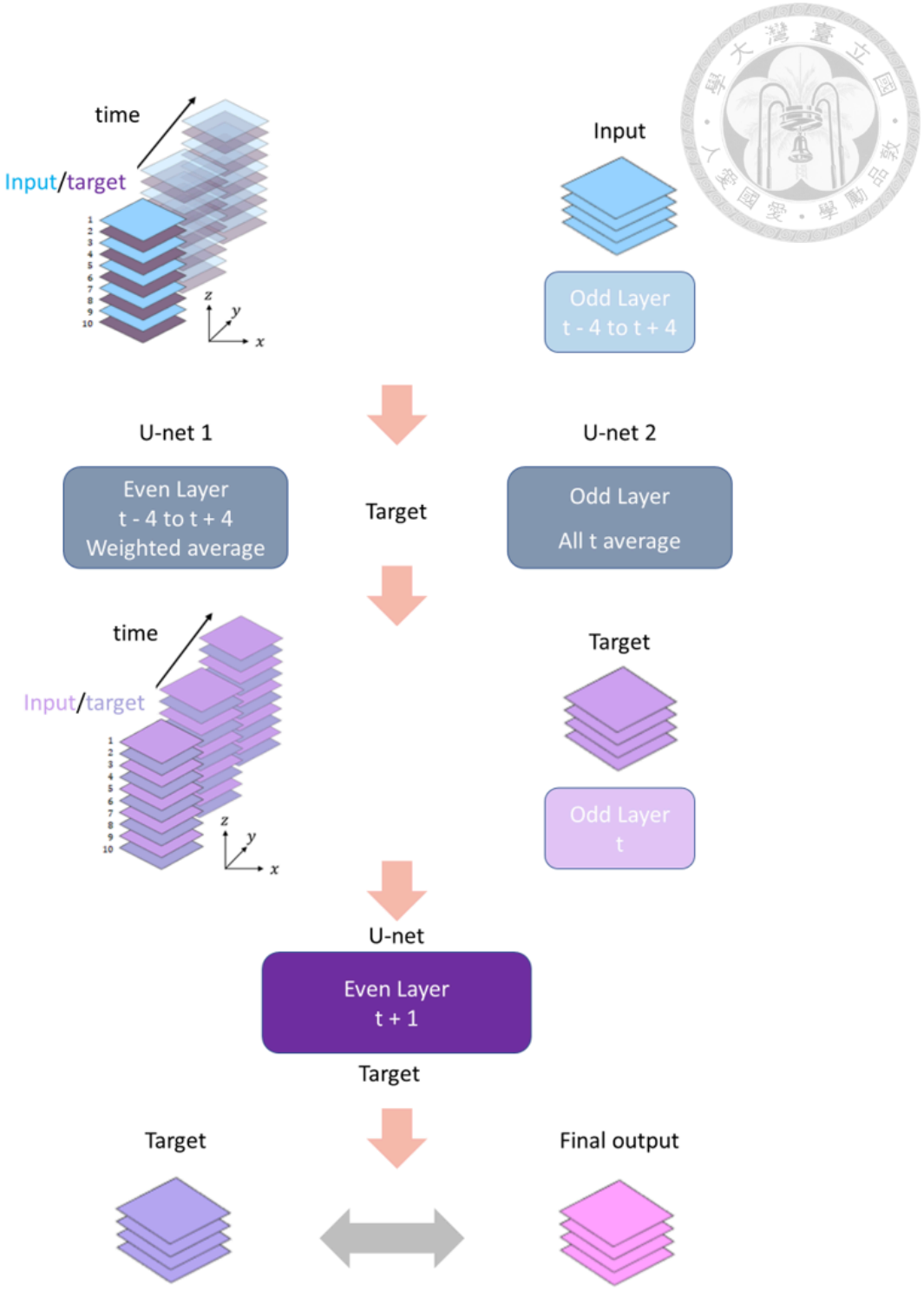


Figure 2-5. Illustration of TAG-SPARK 2.0 model.

doi:10.6342/NTU202504634

2.5 Deconvolution

In optical imaging systems, the point spread function (PSF) characterizes how a single point source of light is blurred by the system [85]. Due to optical aberrations, scattering, and diffraction, the recorded image is often a degraded version of the true object, with the PSF acting as a blurring kernel [86].

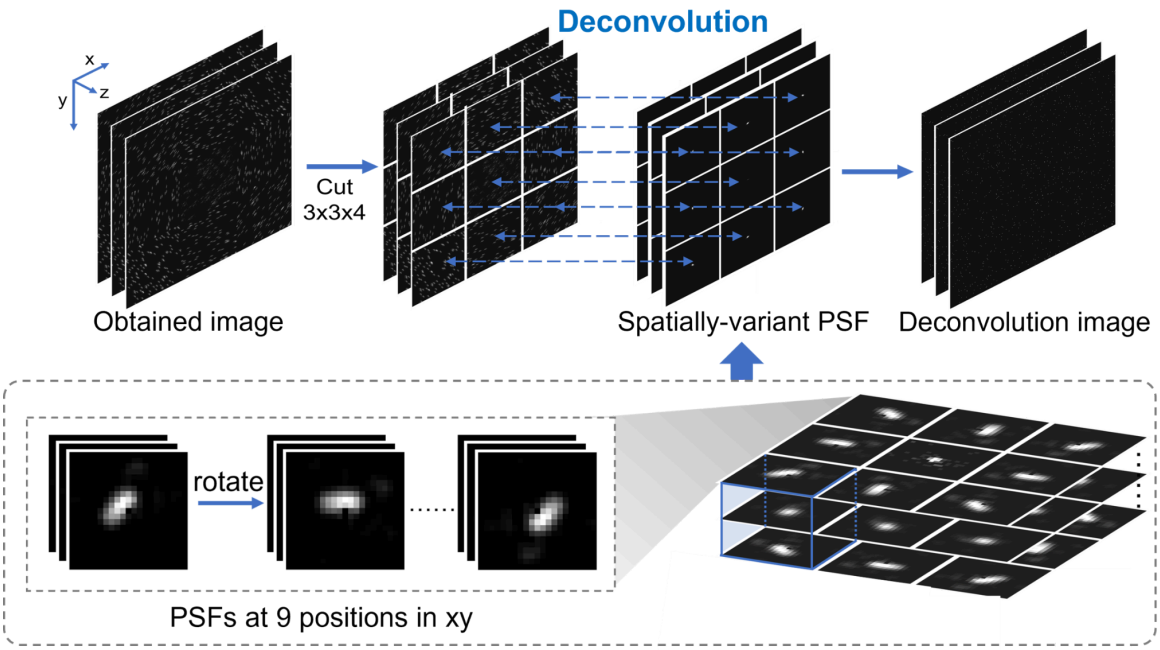


Figure 2-6. Illustration of aberration correction.

The upper schematic shows the flowchart of deconvolution processing. The lower schematic shows the PSF of 1 μm in a 3D space.

Deconvolution is a computational technique that aims to reverse this effect by mathematically estimating and removing the influence of the PSF from the observed image [87], [88, 89]. By applying PSF-based deconvolution, fine structural details such as cell boundaries

doi:10.6342/NTU202504634

and subcellular features can be restored with greater clarity. In this study, we incorporated a deconvolution step following denoising, using either a theoretical or experimentally measured PSF.

Optical aberrations were calibrated using a spatially variant deconvolution algorithm developed in collaboration with Risa Kitamura. Two versions of the algorithm were implemented to address aberrations across the imaging field.

In the first version, each acquired image was partitioned into nine spatial subregions to compensate for local variations in aberration across the FOV. Deconvolution was performed independently for each subregion using a PSF derived from 1 μm fluorescent beads (**Figure 2-6**). This correction process effectively restored distorted cellular structures, improving both their morphological accuracy and contrast.

The second version of the algorithm exploits the cylindrical symmetry of the PSF in three-dimensional space, which arises from field-dependent astigmatism (**Figure 2-7A**). In this approach, the FOV was divided into 4 × 4 subregions during deconvolution. At each imaging depth, the PSF varies with radial distance from the FOV center but retains a consistent shape along any given radius which is different only by rotational orientation. This astigmatism generates radial and tangential line above and below the nominal focal plane, thereby reducing spatial resolution and contrast in peripheral regions. To address this, the spatially variant deconvolution framework leverages the PSF's radial symmetry by aligning the radial axis with one dimension of the image. Each radial ring and its corresponding PSF are transformed into the polar coordinate domain, where the PSF can be assumed to be approximately shift-invariant within each local subregion. This allows the application of standard deconvolution techniques in

doi:10.6342/NTU202504634

a locally adaptive manner. For each radial segment *i*, the PSF is padded and circularly shifted to ensure proper spatial alignment, then transformed into the polar domain with angular constraints to minimize artifacts (**Figure 2-7B**). The full FOV is reconstructed by aggregating the deconvolved results from all radial bands.

This cylindrical deconvolution strategy offers a physically grounded balance between modeling accuracy and computational efficiency. However, it should be noted that the spatially variant deconvolution using 4 × 4 cylindrical subregions requires approximately **4.5 times** longer computation time compared to the 3 × 3 subregion approach.

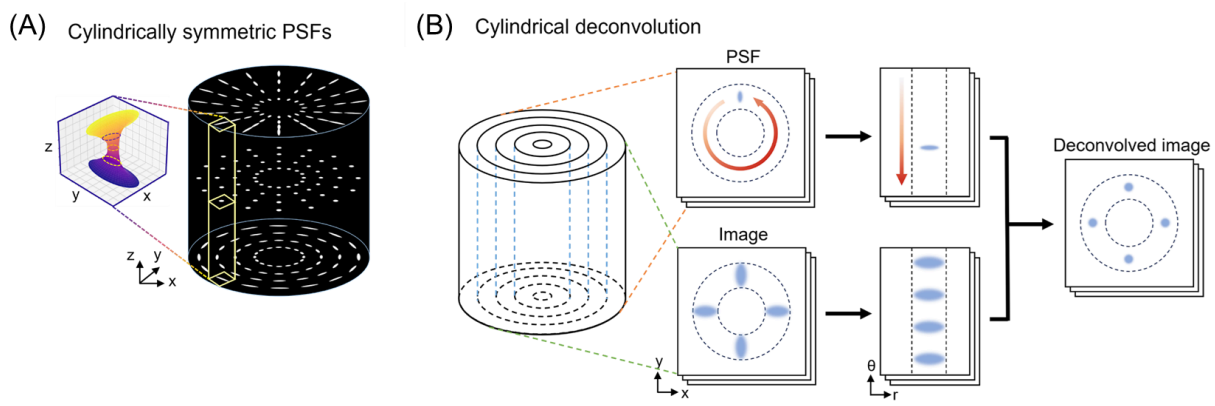


Figure 2-7. Spatially-variant aberration correction for cylindrically symmetric PSFs.

(A) Cylindrically symmetric PSFs caused by GRIN-induced astigmatism. (B) Spatially-variant deconvolution: the image stack is divided into concentric cylindrical subregions by radial distance, each deconvolved in the polar domain with its corresponding PSF.

doi:10.6342/NTU202504634

Chapter 3 Method

This chapter describes our optical setup, the expected depth of field, the actual system performance, and the procedures for sample preparation. Building upon the previous chapter, which outlines the principles of the techniques used, we now focus on the experimental methods and key considerations to be addressed before conducting the experiments.

3.1. Optical setup

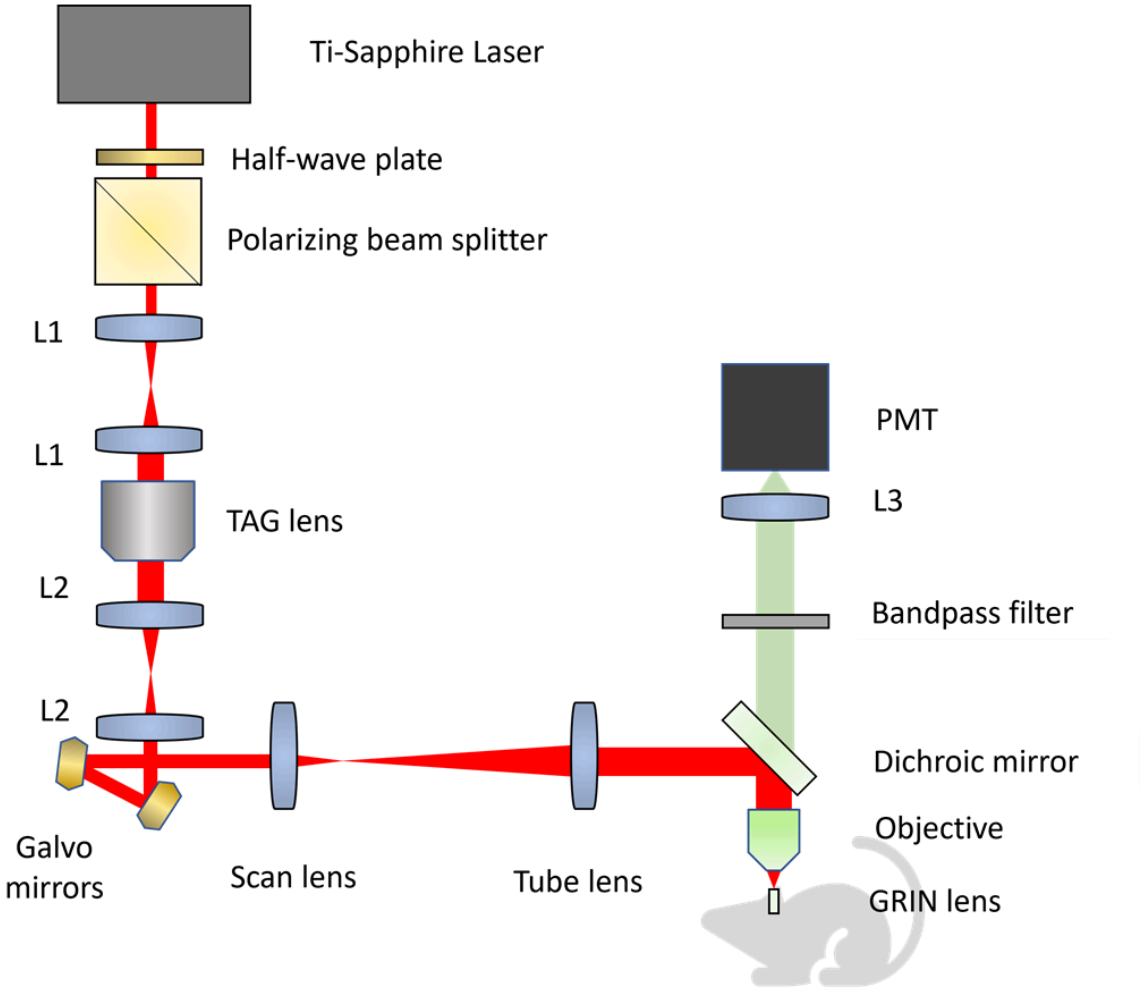


Figure 3-1. Optical setup of the high-speed two-photon volumetric imaging system.

doi:10.6342/NTU202504634

The home-built high-speed two-photon volumetric imaging system is driven by a tunable Ti: Sapphire oscillator (Chameleon Ultra II, Coherent). The excitation wavelength was tuned to 940 nm for two-photon excited fluorescence of green fluorescent protein GCaMP7. With a lens pair (LA1608-B and AC254-150-A-ML, Thorlabs), the laser beam size was adjusted to match the effective aperture of the tunable acoustic gradient lens (TAG lens 2.5β, TAG Optics), which is 4 mm. The laser beam was guided through a lens pair (AC254-060-B-M and LA1422-B, Thorlabs) and then relayed to a pair of galvanometric mirrors (6215H, Cambridge Technology) for 2D raster scanning. Through a telecentric scan lens and tube lens pair (SL50-2P2 & TL200-2P2, Thorlabs), the scanning pattern was reflected by a dichroic beam splitter and directed onto the back aperture of a 20X objective (UPLanFLN20X0.5NA, Olympus). The emitted two-photon fluorescence was epi-collected by the same objective, passing through the dichroic beam splitter, and then focused by an imaging lens onto a photomultiplier tube (PMT) module. Two bandpass filters (FF01-520/15-25, Semrock) were placed in front of the PMT to filter out the reflected excitation laser and undesired background signals. The main difference from our previous publication is that to observe neuronal dynamics in the SCN, a GRIN lens (NEM-060-50-00-920-S-1.5p, Grintech) was inserted into the mouse brain, enabling high-resolution imaging of deep-brain regions using two-photon endoscopy (**Figure 3-1** and **Table 3-1**). During image acquisition sessions, awake mice, which expressed GCaMP7f in SCN, were mounted on a custom-designed two-dimensional tilting stage, with a free-rotating wheel provided to alleviate stress from head constraints.

doi:10.6342/NTU202504634

Instrument	Manufacturer	Model number
Ti: Sapphire Laser	Coherent	Chameleon Ultra II
Half-wave plate	Thorlabs	AHWP05M-980
Polarizing beam splitter	Thorlabs	CM1-PBS252
Lens1	Thorlabs	LA1608-B
Lens 2	Thorlabs	AC254-150-A-ML
TAG lens	TAG Optics	TAG lens 2.5β
4f system, lens 1	Thorlabs	AC254-60-B
4f system, lens 2	Thorlabs	LA1422-B
Galvanometric mirrors	Cambridge Technology	6215HM40B
Scan lens	Thorlabs	SL50-2P2
Tube lens	Thorlabs	TL200-2P2
Dichroic	Semrock	FF749-SDi01-25×36×3.0
Objective lens	Olympus	UPLanFLN20X0.5NA
Bandpass filter	Semrock	FF01-520/15-25
PMT	Hamamatsu	H14119-40
Gradient-index (GRIN) lens	Grintech	NEM-060-50-00-920-S-1.5p

Table 3-1. List of elements with manufacturer and model number.

3.2. Sample preparation

Before conducting the experiments, we evaluated the system's performance using several samples. To assess the system's resolution, we measured the spot size of the image obtained from 2D fluorescent bead samples. To evaluate the DOF for volumetric imaging, we used 3D

doi:10.6342/NTU202504634

fluorescent bead samples and mouse brain slices. Additionally, we describe the surgical procedure for inserting a GRIN lens into the SCN region of an *in vivo* mouse.

2D Fluorescent beads

Fluorescent beads with a diameter of 200 nm are diluted in deionized (DD) water at a 1:100 ratio. A small volume of the diluted solution, which is typically 5 µL in size, is then dispensed onto a slide and covered with a coverslip. To let the sample air dry, the edges of the coverslip are sealed using nail polish (**Figure 3-2**).

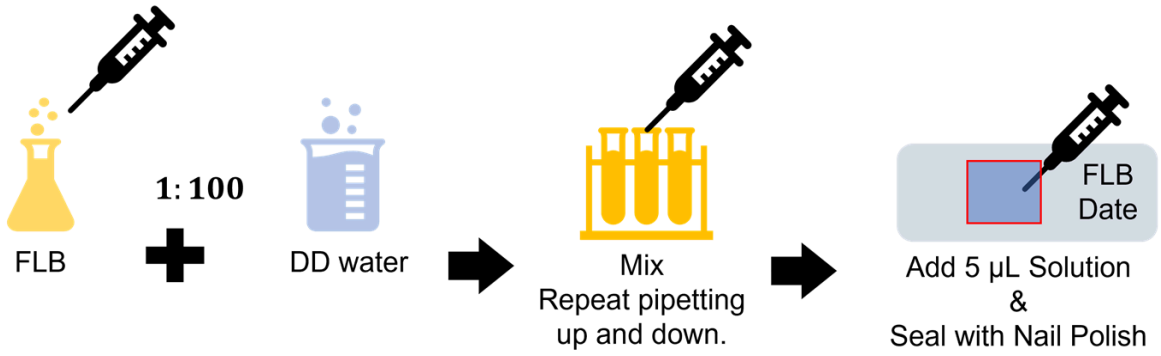


Figure 3-2. 2D Fluorescent Bead Sample Preparation.

3D Fluorescent beads

To evaluate the DOF for volumetric imaging, 10 µm fluorescent microspheres are used to simulate mouse neurons. The agarose hydrogel concentration is 0.7%. Agarose is heated and stirred until it reaches 90 °C and becomes fully transparent. Once cooled to approximately 50 °C, 10 µL of diluted fluorescent beads are added to the upper and central regions of the gel. The mixture is then refrigerated to allow solidification, forming a jelly-textured sample. To preserve

doi:10.6342/NTU202504634

the sample, Parafilm is used to seal the edges of the container, which is then stored in a refrigerator at 7 °C (**Figure 3-3**).

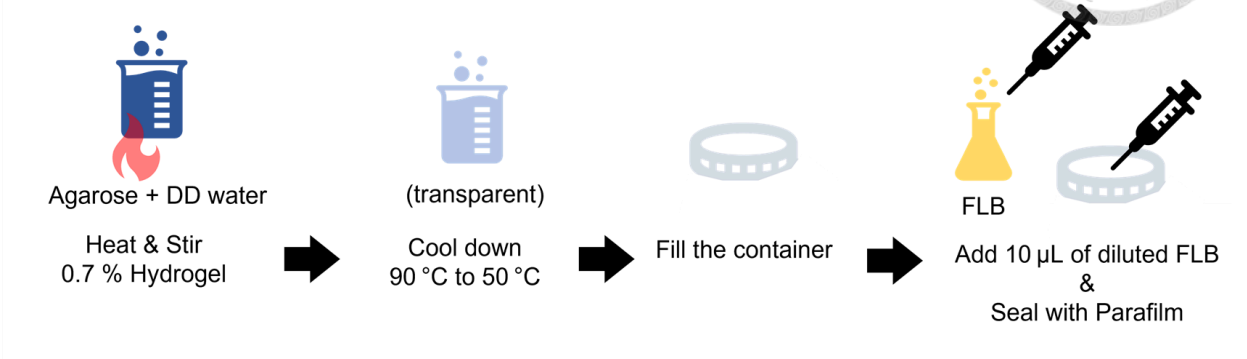


Figure 3-3. 3D Fluorescent Bead Sample Preparation.

Mouse brain slice

The Thy1-eGFP (Jackson Laboratory, #007788) mice were used as one of the samples to check the resolution performance. The mice were housed under a 12-hour light/12-hour dark cycle with food and water provided ad libitum. All experimental procedures were approved by the Institutional Animal Care and Use Committee of National Taiwan University (approval number: NTU110-EL-00023). The mice were placed in a chamber and exposed to 5% isoflurane via an anesthetic vaporizer for 5 minutes to euthanize them. After breathing had stopped, the mice were transcardially perfused with 30 ml phosphate-buffered saline followed by 30 ml 4% paraformaldehyde. When the perfusion was completed, the brain was dissected and postfixed in 4% paraformaldehyde for 2 days at 4° C. A vibratome was used to cut the brain into 500 µm sagittal or coronal sections. The sections were immersed in RapiClear solution (SunJin Lab Co., Taiwan) overnight at room temperature. The clarified brain sections were mounted on a chamber

doi:10.6342/NTU202504634

composed of a 0.5-mm plastic spacer and cover glass for imaging. (The mouse brain tissue sample was prepared by Dr. Jye-Chang Lee.)

3.3. SCN Mouse preparation

All surgeries were performed with a stereotaxic platform and Hamilton syringes controlled by a motorized micromanipulator to ensure precise virus injection and lens implantation into the SCN. Mice were anesthetized by isoflurane inhalation (5% vapor for induction and 1-1.5% vapor for maintenance). For *in vivo* imaging, the GRIN endoscopes (0.6 mm in diameter and 7.4 mm in length, GRINTECH NEM-060-50-00920-S-1.5P) were implanted after 300 nL AAV9-GCaMP7f (Addgene #104488) injection. The injection and implantation trajectories for the GRIN endoscope were rolled 6° clockwise on the coronal plane. Injection needle tips and endoscopes were aimed at anterior-posterior (AP) -0.46 mm and medial-lateral (ML) -0.725 mm from the bregma, and DV -5.2 to -5.4 from the surface of the cerebral cortex. Endoscopes were protected by custom-designed head plates, and all materials were secured using dental cement and superbond. For TRAP2 mice in the behavioral tests, 150 nL of AAV9-hSyn-DIO161 rM3D(Gs)-mCherry or AAV9-hSyn-DIO-EGFP (Addgene #50458 and #50457) were bilaterally injected, aimed at AP -0.46 mm and ML ±0.2 mm from the bregma and DV 163 -5.76 from the surface of the cerebral cortex. (The SCN mouse brain tissue sample was prepared by Cheng-Han Wang.)

3.4. Performance evaluation of the optical system

To evaluate the performance of the optical system, we first calculate the theoretical values based on the parameters of the optical elements. We then measure the spot size to assess the practical resolution and DOF.

doi:10.6342/NTU202504634

By measuring the spot size, we evaluate the resolution of our two-photon microscope. After calculating the theoretical spot size, we imaged 200 nm fluorophores, which are more than three times smaller than the theoretical resolution limit. As a result, the measured spot size closely matches the PSF.

Theoretical value

The theoretical two-photon focal volume corresponds to

$$FWHM\ (X,Y) = \begin{cases} 2\sqrt{ln2}\times 0.320\times\frac{\lambda}{\sqrt{2}NA} & NA < 0.7 \\ 2\sqrt{ln2}\times 0.325\times\frac{\lambda}{\sqrt{2}NA^{0.91}} & NA > 0.7. \end{cases}$$

Eq. 3-1

$$FWHM\ (Z) = \frac{0.532\lambda}{\sqrt{2}} \times \frac{1}{n-\sqrt{n^2-NA^2}}\ .$$

, where $FWHM(X,Y)$ and $FWHM(Z)$ are the full width at half maximum (FWHM) of the PSF in the lateral and axial direction, λ is the wavelength of the laser, NA is the effective numerical aperture of the objective, and n is the refractive index of the sample.

However, since the laser beam size does not fill the aperture of the objective, we consider the effective NA by measuring the laser beam size with a camera and calculating the effective NA by the following formula.

$$\text{FWHM (laser beam)} = D = 6.3 \text{ mm}.$$

$$NA = n \times sin\theta.$$

$$2 \times sin\theta = \frac{D}{f}.$$ Eq. 3-2

$$NA_{eff} = n \times sin(D/2f) = 1 \times sin(6.3/2 \times 9.0) = 0.34.$$

, where n is a refractive index and f is the focal length of an objective lens.

Substituting the parameters of our two-photon microscopy system into Eq. 3-1,

doi:10.6342/NTU202504634

$$theoretical\ FWHM(X,Y) = 2\sqrt{ln2}\times 0.320\times\frac{940}{\sqrt{2}(0.34)} \cong 1032\ \text{nm}.$$

Eq. 3-3

$$theoretical\ FWHM(Z) = 2\sqrt{ln2}\times\frac{0.532*940}{\sqrt{2}}\times\frac{1}{1-\sqrt{1^2-(0.34)^2}} \cong 9711\ \text{nm}.$$

, where $\lambda = 940$ nm, effective NA = 0.34 of the objective (Olympus, UPLFLN 20X, NA0.5), and $n = 1$ which is the same as gas for 2D samples.

This measurement evaluates the resolution of two-photon microscopy by measuring the spot size using a theoretical model and 200 nm fluorophores, achieving a spot size close to the Point Spread Function (PSF). Fluorescent bead intensity profiles fitted with a Gaussian curve determined a lateral size of 1046.4 nm ± 78.3 nm and an axial size of 5972.4 nm ± 1156.2 nm, with minimal errors. DOF measurements for 10 μm fluorescent beads and mouse slices, correlated between slow and fast scan images, revealed variations with different Tunable Acoustic Gradient (TAG) lens amplitudes. DOF for 15% amplitude was about 10% smaller than expected, 25% amplitude was slightly larger, and 40% amplitude showed 11.4% smaller DOF for beads and 9.6% larger for mouse slices. Additionally, the GRIN lens produced a circular field of view, with axial spot size varying with depth.

doi:10.6342/NTU202504634

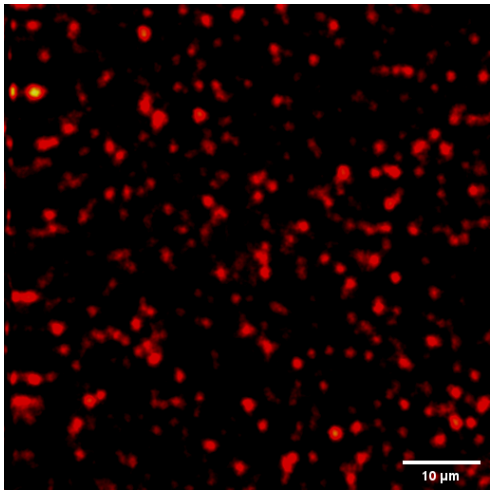


Figure 3-4. The z-projection image of a sample of the 200 nm fluorophore.

According to the sample preparation, the concentration of the hydrous solution is 2%. Scale bar: 10 µm.

By fitting the intensity profile of several Fluorescent beads with a Gaussian curve, we derive the measured FWHM. Table 3-1. shows the statistical result of the spatial resolution. The measured lateral size is 1046.4 nm ± 78.3 nm, and the error compared to the theoretical size is 1.4% while the measured axial size is 5972.4 nm ± 1156.2 nm, and the error compared to the theoretical size is 2.4% (**Table 3-2**).

FWHM	**X-Y (nm)**	**Z (nm)**
Theoretical size	1032	5832
Measured size	1046.4	5972.4
Standard Deviation	78.3	1156.2
Error	1.4%	2.4%

Table 3-2. The statistical result of the lateral and axial FWHM of 2D 200 nm.

doi:10.6342/NTU202504634

Analysis of DOF corresponds to different amplitudes of TAG lens with an objective (Olympus, UPLFLN 20X, NA 0.5)

To analyze the DOF of the system, the amplitude of the TAG lens is set to 15%, 25%, and 40%. Two types of samples are used: 10 μm fluorescent beads (FLB) and a mouse brain slice. Both samples are scanned using low-speed and high-speed scanning modes. In the following figures **(Figures 3-5, 3-6, and 3-7)**, the first image shows the correlation analysis between the low-speed and high-speed scans. Since the actual layer thickness is known from the low-speed scanning mode, the measured DOF is determined by identifying the layers in the high-speed scan that yield the highest correlation with the top and bottom layers from the low-speed scan.

1. Theoretical DOF with TAG lens 188k, Amplitude = 15% ,

$$n = 1.33, \quad f_{obj} = 9.0mm, \quad M = \left(\frac{40}{60}\frac{200}{50}\right) = \frac{8}{3}, \quad \frac{1}{f_{TAG}} = \delta = 3.75.$$

$$\Delta z = n\frac{f_{obj}^2}{f_{TAG}}\left(\frac{f_2}{f_1}\right)^2 = n\frac{f_{obj}^2}{f_{TAG}}\left(\frac{1}{M}\right)^2 = 1.33 * 81.0 * 3.75 * 9/64 \cong 56.8\ \mu m.$$

2. Theoretical DOF with TAG lens 188k, Amplitude = 25%

$$n = 1.33, \quad f_{obj} = 9.0mm, \quad M = \left(\frac{40}{60}\frac{200}{50}\right) = \frac{8}{3}, \quad \frac{1}{f_{TAG}} = \delta = 6.25.$$

$$\Delta z = n\frac{f_{obj}^2}{f_{TAG}}\left(\frac{f_2}{f_1}\right)^2 = n\frac{f_{obj}^2}{f_{TAG}}\left(\frac{1}{M}\right)^2 = 1.33 * 81.0 * 6.25 * 9/64 \cong 94.6\ \mu m.$$

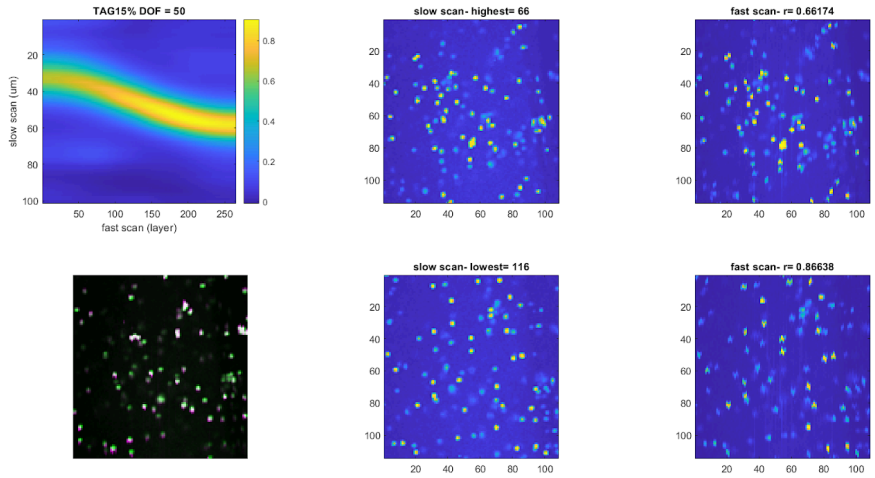
3. Theoretical DOF with TAG lens 188k, Amplitude = 40%

doi:10.6342/NTU202504634

$$n = 1.33, \quad f_{obj} = 9.0mm, \quad M = \left(\frac{40}{60}\frac{200}{50}\right) = \frac{8}{3}, \quad \frac{1}{f_{TAG}} = \delta = 10.$$

$$\Delta z = n\frac{f_{obj}^2}{f_{TAG}}\left(\frac{f_2}{f_1}\right)^2 = n\frac{f_{obj}^2}{f_{TAG}}\left(\frac{1}{M}\right)^2 \quad = 1.33 * 81.0 * 10 * 9/64 \cong 151.4\ \mu m.$$

(A)



(B)

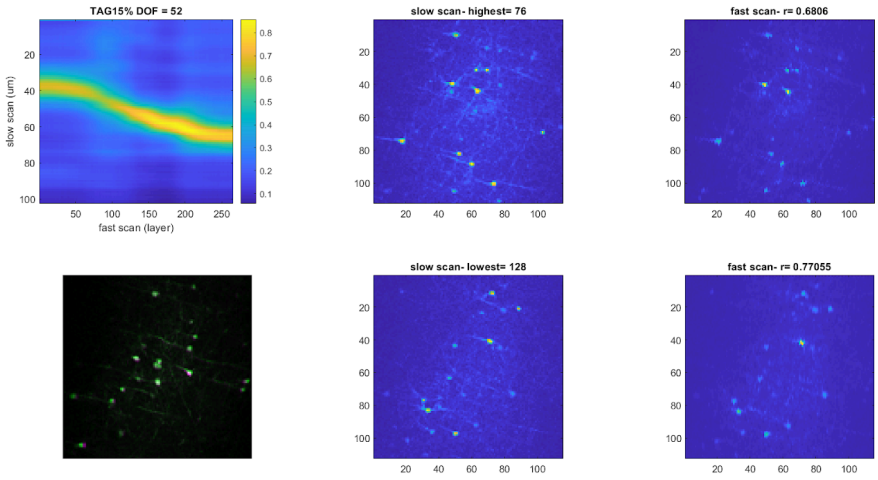


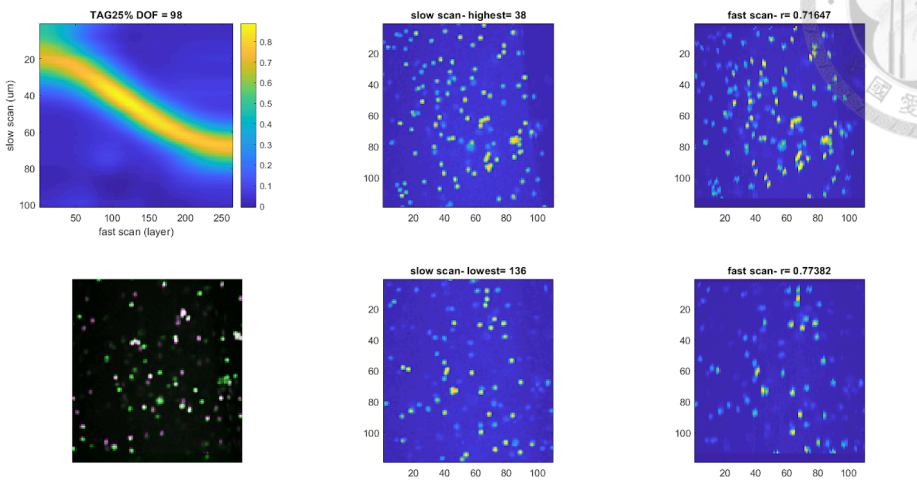
Figure 3-5. Measured DOF analysis with TAG amplitude set as 15%.

(A) The measured DOF for the 10 μm fluorescent beads (FLB) is 50 μm. Compared to the theoretical value, this corresponds to a relative error of –11.9%.

(B) The measured DOF for the mouse brain slice is 52 μm, with a relative error of –8.4%.

doi:10.6342/NTU202504634

(A)



(B)

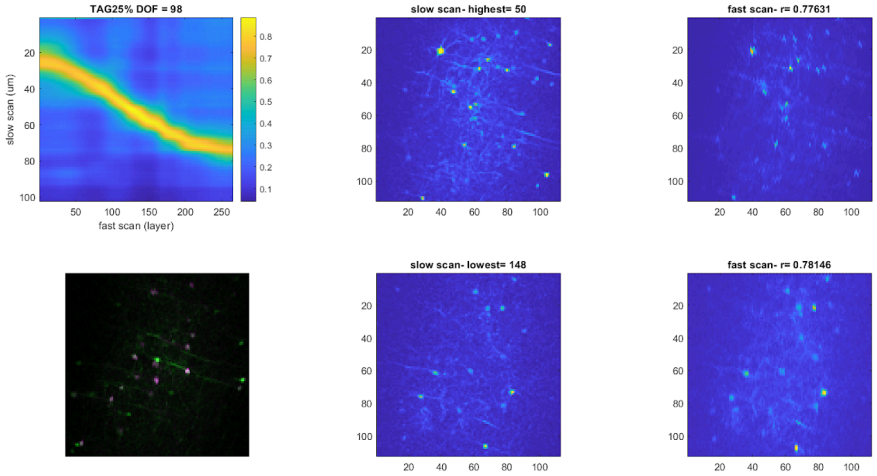


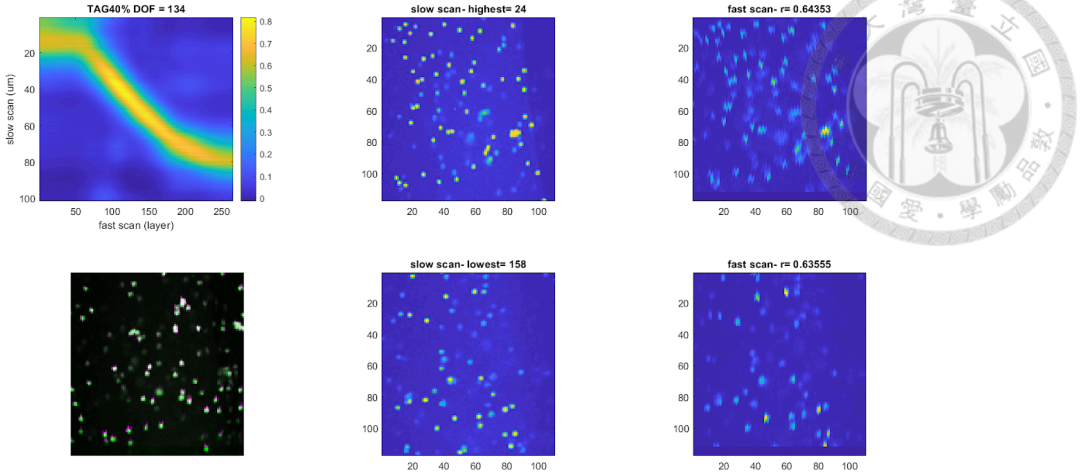
Figure 3-6. Measured DOF analysis with TAG amplitude set as 25%.

(A) The measured DOF for the 10 μm fluorescent beads (FLB) is 98 μm. Compared to the theoretical value, this corresponds to a relative error of +3.5%.

(B) The measured DOF for the mouse brain slice is 98 μm, with a relative error of +3.5%.

(A)

doi:10.6342/NTU202504634



(B)

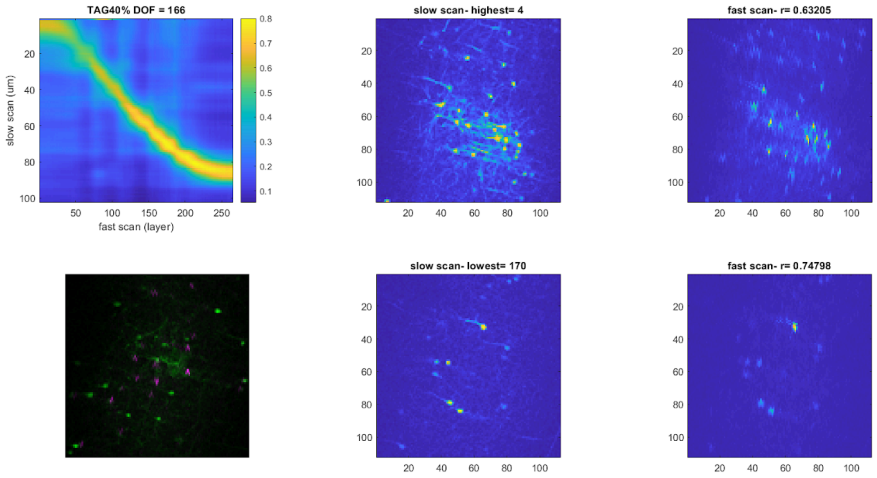


Figure 3-7. Measured DOF analysis with TAG mplitude set as 40%.

(A) The measured DOF for the 10 μm fluorescent beads (FLB) is 134 μm. Compared to the theoretical value, this corresponds to a relative error of -11.4%.

(B) The measured DOF for the mouse brain slice is 166 μm, with a relative error of +9.6%.

Examining the error values, we observe that the measured DOF of the TAG lens at 15% amplitude is approximately 10% smaller than the theoretical value for both samples. At 25% amplitude, the measured DOF is slightly larger than expected. At 40% amplitude, the measured DOF for the 10 μm FLB is 11.4% smaller than the theoretical value, while the measured DOF for

doi:10.6342/NTU202504634

the mouse brain slice is 9.6% larger. These results (**Table 3-3**) indicate that the system's DOF closely matches the theoretical predictions across different TAG lens amplitudes.

Amplitude of TAG lens		
15%	25%	40%
Theoretical DOF (um)		
56.8	94.6	151.4
3D 10 um FLB by high-speed scanning with TAG lens Measured DOF (um)		
50	98	134
Error		
-11.9%	+3.5%	-11.4%
Mouse brain slice by high-speed scanning with TAG lens Measured DOF (um)		
52	98	166
Error		
-8.4%	+3.5%	+9.6%

Table 3-3. The measured DOF corresponds to different amplitudes of the TAG lens by high-speed and slow-speed scanning.

doi:10.6342/NTU202504634

3.5. Image processing

Dark pattern calibration

After deriving the high-speed scanning images, a dark pattern randomly appears **(Figure 3-8A**). To fix this problem, we developed an image processing method to calibrate the dark pattern. This method relies on homemade Python programming.

The algorithm is designed to detect the appearance of a dark pattern in the image series and then analyze its intensity. Accordingly, it recovers the intensity of the dark pattern to its original value. The details are as follows: there are three loops during the calibration process. In the first loop, the input data is denoised from high-speed imaging. The threshold is defined as a parameter, 0.8 times the average intensity of all pixels.

Next, to detect the appearance of a dark pattern in the image series, the inspection is conducted line by line. If the average intensity of a line is smaller than the threshold, the intensity of that line is replaced by the average of nearby lines. Therefore, the output images are saved in Folder 1.

In the second loop, the input data is from Folder 1, and the parameter is set to 0.9. The main process is the same as the previous loop. The output images are saved in Folder 2. Similarly, in the last loop, the input data is from Folder 2, and the parameter is set to 1. The output images are well-calibrated.

By comparing the images before and after calibration, we selected one ROI to analyze its intensity. In the uncorrected image series, the ROI exhibits a drastic drop in intensity. However, after calibration, the intensity curve becomes smooth and consistent. This demonstrates that our homemade calibration program effectively corrects the dark pattern without compromising the true intensity values in the image **(Figure 3-8B**).

doi:10.6342/NTU202504634

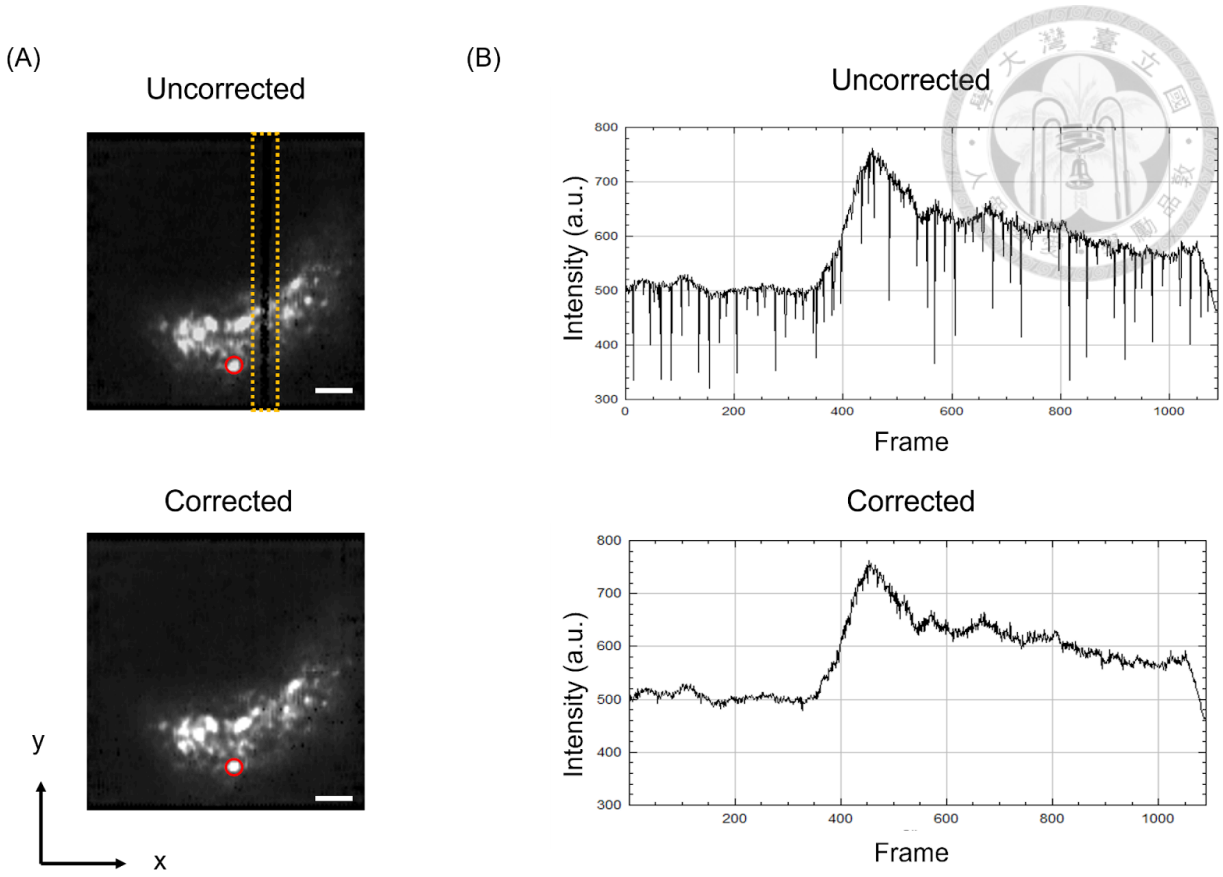


Figure 3-8. Dark pattern calibration and ROI analysis.

(A) Uncorrected and corrected images of a single layer at one time point. The dashed line indicates the dark pattern. The red circle marks a region of interest (ROI). Scale bar: 100 μm.

(B) Intensity analysis within the ROI. (Top): Before correction. (Bottom): After correction.

doi:10.6342/NTU202504634

Deconvolution algorithm with 3D 1 μm fluorescent beads (FLB) using an inserted GRIN lens

After addressing the random appearance of dark patterns, the reconstructed images were substantially improved. However, distortion and aberrations introduced by the GRIN lens remained evident. To mitigate these effects, a dedicated deconvolution algorithm was developed by our collaborator, Risa Kitamura. To evaluate whether this algorithm is appropriate for correcting GRIN lens–induced aberrations, 3D image stacks of 1 μm FLB were acquired under two conditions: without GRIN lens insertion and with GRIN lens insertion.

To evaluate the effectiveness of our deconvolution algorithm in correcting GRIN lens–induced aberrations, we imaged 1 μm fluorescent beads (FLB) embedded in a three-dimensional matrix. **Figure 3-9A** shows a representative raw image acquired at a depth of 162 μm with an inserted GRIN lens. For comparison, two deconvolution methods were applied to the raw image. The theoretical PSF, obtained without the GRIN lens, is spatially invariant, whereas the experimental deconvolved image was reconstructed by applying the algorithm to the raw dataset, where the PSF is spatially variant. With the spatially variant PSF, the deconvolved image shows a more precise correction than the one produced by the traditional spatially invariant deconvolution method. Highlighted in **Figure 3-9B**, the raw images display characteristic aberrations, including lateral distortions (yellow arrows) and elongation of bead profiles (green arrows). Following deconvolution, bead morphology and symmetry were largely restored, indicating effective correction of optical artifacts.

Quantitative analysis further confirmed these improvements. As shown in **Figure 3-9C**, lateral FWHM measurements along tangential and radial directions revealed strong anisotropy in the raw images, consistent with GRIN lens–induced aberrations. After deconvolution, isotropy

doi:10.6342/NTU202504634

was restored, and line profiles illustrated reduced elongation of bead shapes. Similarly, axial resolution analysis (**Figure 3-10D**) demonstrated a marked reduction of aberration-induced broadening. Both depth-dependent and radial-dependent FWHM values showed significant improvements after deconvolution, yielding more uniform resolution across imaging depth and lateral positions. Together, these results validate the capability of the algorithm to correct GRIN lens–induced aberrations and enable accurate three-dimensional imaging in deep-brain applications.

doi:10.6342/NTU202504634

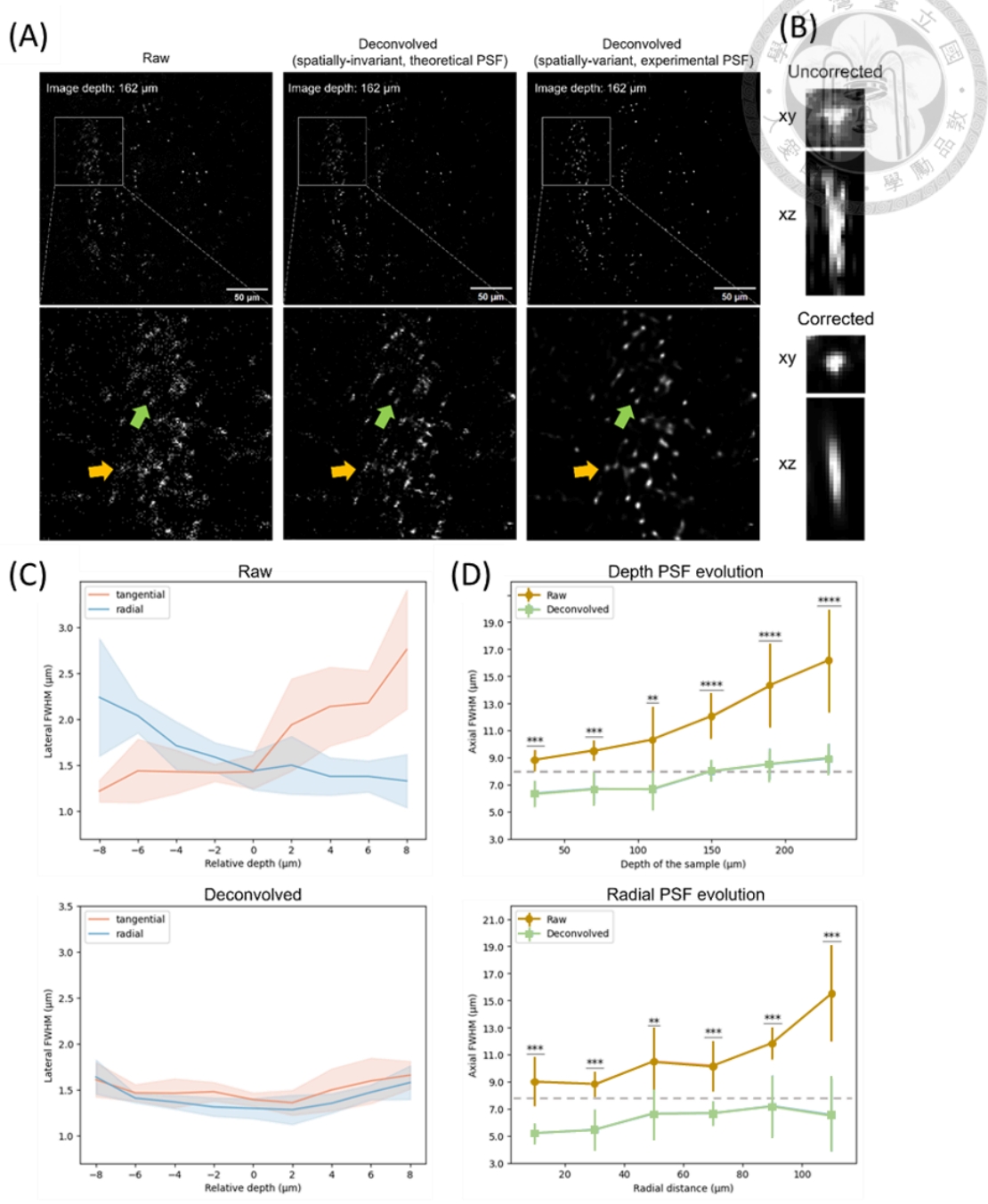


Figure 3-9. Correction of GRIN lens induced aberrations using a deconvolution algorithm.

(A) Representative raw image of 1 μm fluorescent beads (FLB) acquired at 162 μm depth with an inserted GRIN lens. For reference, the theoretical point spread function (PSF) was obtained without the GRIN lens, while the experimental deconvolved image was generated by applying the deconvolution algorithm to the raw dataset. Scale

doi:10.6342/NTU202504634

bar: 50 μm. (B) Enlarged views of bead images showing characteristic aberrationsin the raw data, including lateral distortion (yellow arrows) and elongation (green arrows). Application of the deconvolution algorithm recovers bead morphology and symmetry. (C) Lateral FWHM along tangential and radial directions. Raw images exhibit anisotropy introduced by the GRIN lens, while deconvolution restores isotropy. Line profiles further illustrate the reduction of elongation. (D) Axial FWHM analysis as a function of imaging depth (left) and radial distance (right), comparing raw and deconvolved conditions. Deconvolution significantly reduces aberration-induced broadening, yielding more uniform resolution across depth and lateral positions.

doi:10.6342/NTU202504634

Chapter 4 Results and Discussion

Based on the foundation laid in Chapter 3 regarding the optical system setup, we now examine the evidence that demonstrates the system's performance. In this chapter, we focus on applying our techniques to capture neural dynamics in the SCN, located 6 mm deep in the mouse brain. Our goal is to observe how neural signals are transmitted under different light-induced phases and examine the impact on the three-dimensional neuronal population activity *in vivo*.

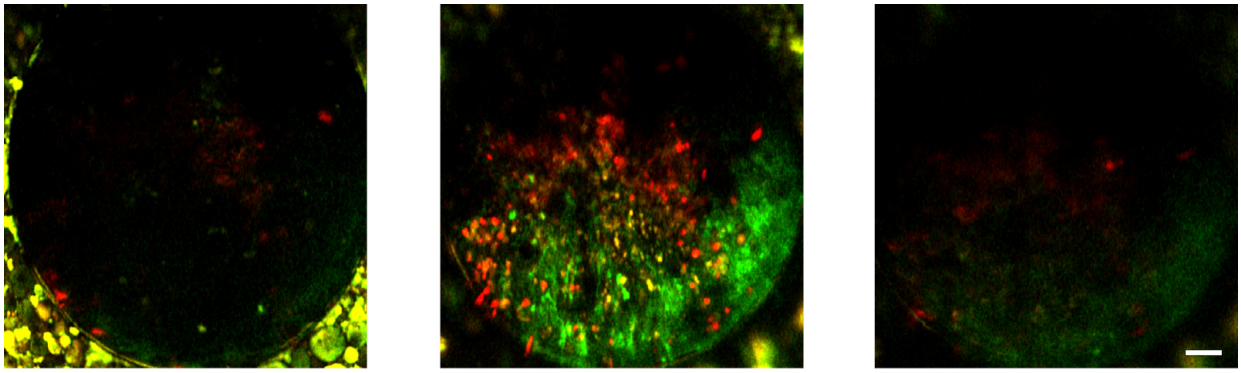
4.1. SCN structure in an *in vivo* mouse

Before conducting the experiments, we examined the structure of SCN in an alive mouse using a two-photon endoscopy system. Using a laser wavelength of 920 nm, cells expressing the calcium indicator GCaMP8s were imaged through low-speed scanning, with the translation stage advanced in 5 μm intervals. A 5-second average projection image was generated to visualize the fine structure of this deep brain region. Each frame was acquired with a pixel dwell time of 1.7 μs.

Approximately 120 cells were observed within the SCN volume, with an imaging area of $1000 \times 1000 \times 285\ \mu m^3$. By gradually moving the objective toward the GRIN lens, we captured images from the lens surface down to its bottom. At the surface, the image revealed a circular field of view surrounded by numerous dental cement bridges, which were used to secure the GRIN lens to the mouse skull. At the bottom of the GRIN lens, signal intensity was minimal. In the middle portion of the volume, two distinct types of signals were detected: green fluorescence from GCaMP8s-expressing cells and red fluorescence from tdTomato-expressing cells.

doi:10.6342/NTU202504634

Because tdTomato was expressed specifically in VIP (vasoactive intestinal peptide) neurons, which are localized exclusively within the SCN, we confirmed that the imaged region corresponded precisely to our target area (**Figure 4-1**).



$z = 0$ μm $z = 150$ μm $z = 285$ μm

Figure 4-1. Average projections of *in vivo* SCN structure via GRIN lens imaging.

Five-second average projection images of *in vivo* SCN cells labeled with GCaMP8s (green) and tdTomato (red), simultaneously acquired using a two-photon microscope.Left: Image at the surface of the GRIN lens. Middle: Image focused on the SCN region. Right: Image at the bottom of the GRIN lens. Scale bar: 100 μm.

4.2. High-speed volumetric image of SCN neural activity in an *in vivo* mouse

To study the functional activity of the deep brain region, SCN, we derived the volumetric image using a high-speed scanning system. Given that intrinsically photosensitive retinal ganglion cells (ipRGCs) are crucial for conveying light information to the SCN, which controls the circadian rhythm, the experiment is set to a specific zeitgeber time (ZT). To stimulate retinal photosensitive cells, a 488 nm LED light is diffused onto a 10 x 26 mm screen in front of the eyes of mice. The average power of light on the screen was 1.76 $\mu W/mm^2$. For the experimental design, each trial consists of a 90-second recording using a 940 nm wavelength laser to stimulate SCN cells in an alive mouse. The trial begins with a 30-second dark phase, during which no light

doi:10.6342/NTU202504634

is applied to simulate a dark environment. This is followed by a 60-second phase in which an LED light is exposed to the mouse's eyes. During this phase, we expect certain cells to exhibit a functional response. Afterward, both the laser and LED light are turned off to allow the mouse to rest and avoid interference with subsequent trials (**Figure 4-2**).

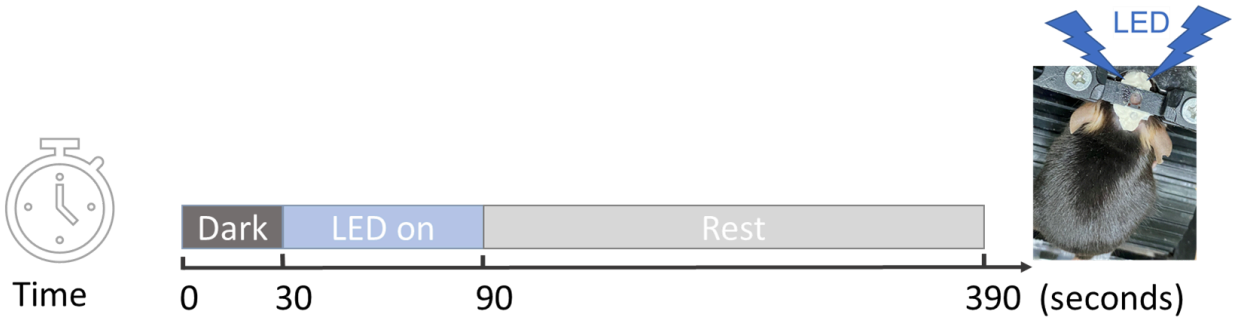


Figure 4-2. Schematic of the experimental design.

Each trial begins with a 30-second dark phase (Dark), followed by blue LED illumination from t = 30 s to t = 90 s (LED on). Afterward, the mouse enters a 300-second resting period (Rest).

Figure 4-3 presents *in vivo* images of the SCN at different depths. The left portion of the figure highlights the structural layers of the SCN, acquired through low-speed scanning during initial imaging (**Figure 4-3A**). With this low-speed mode, individual cells are clearly resolved within a single frame. Once the structure was confirmed, the stage was fixed at the middle layer of the volume for subsequent high-speed functional imaging.

High-speed scanning was performed using a tunable acoustic gradient (TAG) lens, which dynamically shifts the focal plane along the z-axis at its resonant frequency of 188 kHz, enabling an ultra-fast voxel dwell time of 0.02 μs. The volume acquisition rate depended on the lateral pixel resolution, allowing for flexible image capture that integrates both structural and functional

doi:10.6342/NTU202504634

information. In our setup, the TAG lens was driven at 15% amplitude, corresponding to an optical power of approximately 3.75 diopters. This setting extended the depth of field to around 56 μm. However, under high-speed scanning, structural details could no longer be discerned from a single frame due to the rapid acquisition.

Instead, high-speed imaging generated a cascade of overlapping layers and frames during each experimental trial. Owing to the fast axial scanning, neighboring frames showed high similarity (**Figure 4-3B**). The average time projection (0–6 s) was generated from the first 6 seconds of each trial, while the average time projection (6–12 s) was generated from the subsequent 6 seconds. By comparing these two projections, we observed that signal locations remained consistent, while noise appeared randomly distributed. These spatial and temporal redundancies allowed us to apply a TAG-SPARK 2.0 model for denoising, preserving both spatial detail and temporal dynamics. Additionally, a deconvolution algorithm was used to correct aberrations introduced by the GRIN lens.

doi:10.6342/NTU202504634

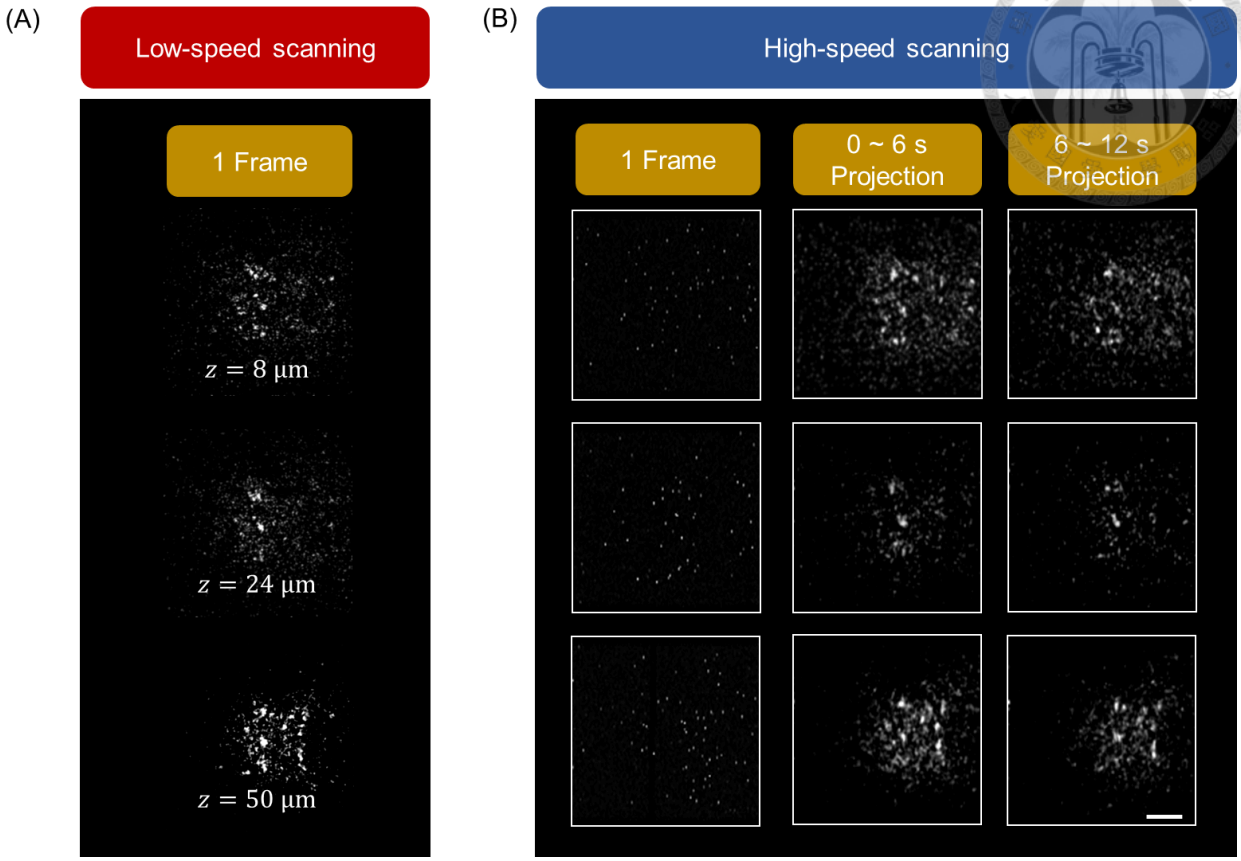


Figure 4-3. Adjacent frames of the image of the *in vivo* SCN at different depths.

(A) Representative single-frame image at a specific depth acquired via low-speed scanning. (B) Single-frame image, average time projection (0–6 s), and average time projection (6–12 s) at the same depth obtained using high-speed axial scanning.

4.3. Evaluation and comparison of functional images pre- and post-processing

In Chapter 3, we mentioned that there are two methods used in post-processing. The first method involves denoising the raw image using a deep learning model, while the second method corrects aberrations using a deconvolution algorithm. To evaluate the denoising performance of the TAG-SPARK 2.0 model, we compared the raw and denoised images shown in **Figure 4-4A**.

doi:10.6342/NTU202504634

By selecting a region of interest (ROI), the calcium signal responses over time are presented in **Figure 4-4B**.

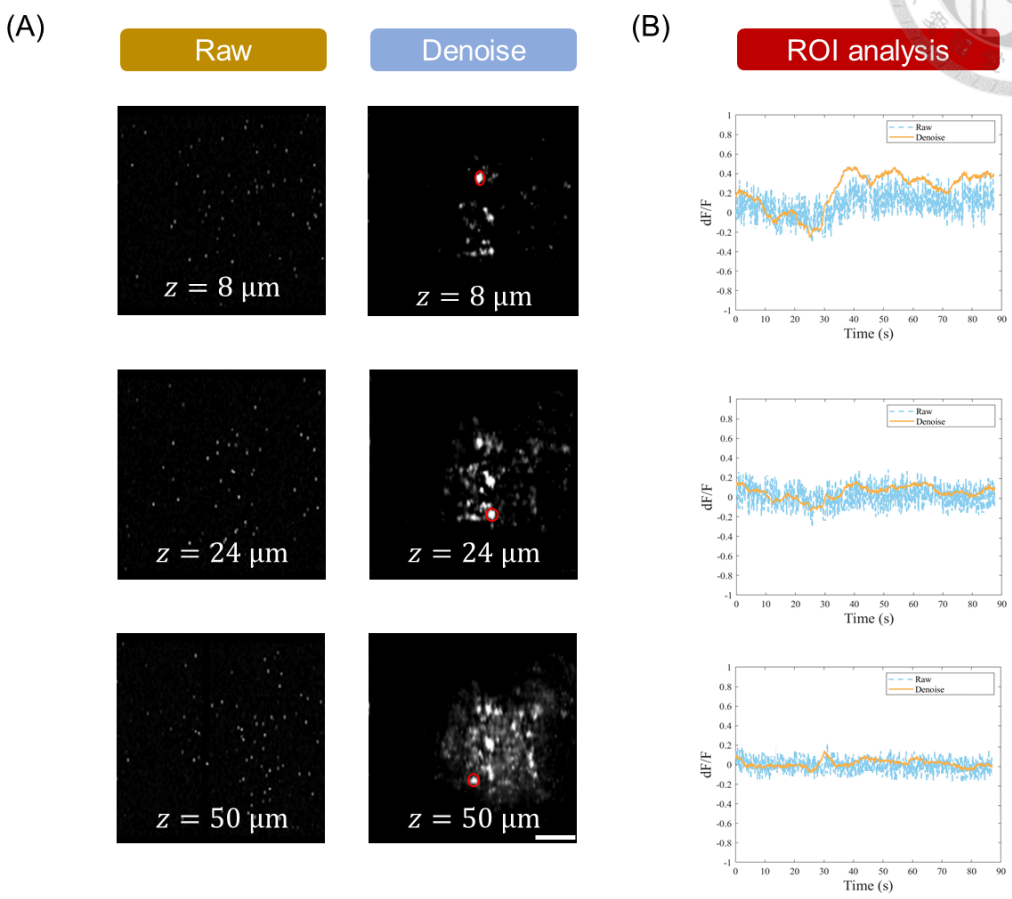


Figure 4-4. Improved signal clarity in SCN calcium imaging via TAG-SPARK 2.0 model denoising.

(A) In vivo calcium imaging of SCN neurons at different depths, shown as single-frame snapshots. Left: Raw images acquired using high-speed scanning. Right: Denoised images obtained by applying the TAG-SPARK 2.0 model to high-speed data. Scale bars: 100 μm. (B) Calcium traces from SCN neurons located at different depths were recorded over a 90-second period.

The raw trace shows a slight increase in signal intensity from t = 30 s to t = 90 s, but it contains a significant amount of noise, which makes the trace appear rigid. After applying the dark pattern correction and our TAG-SPARK 2.0-based denoising method, the denoised trace

doi:10.6342/NTU202504634

displays a much clearer trend during the same time period. These results indicate that the TAG-SPARK 2.0 model successfully reduces noise while preserving both the signal peaks and overall dynamics.

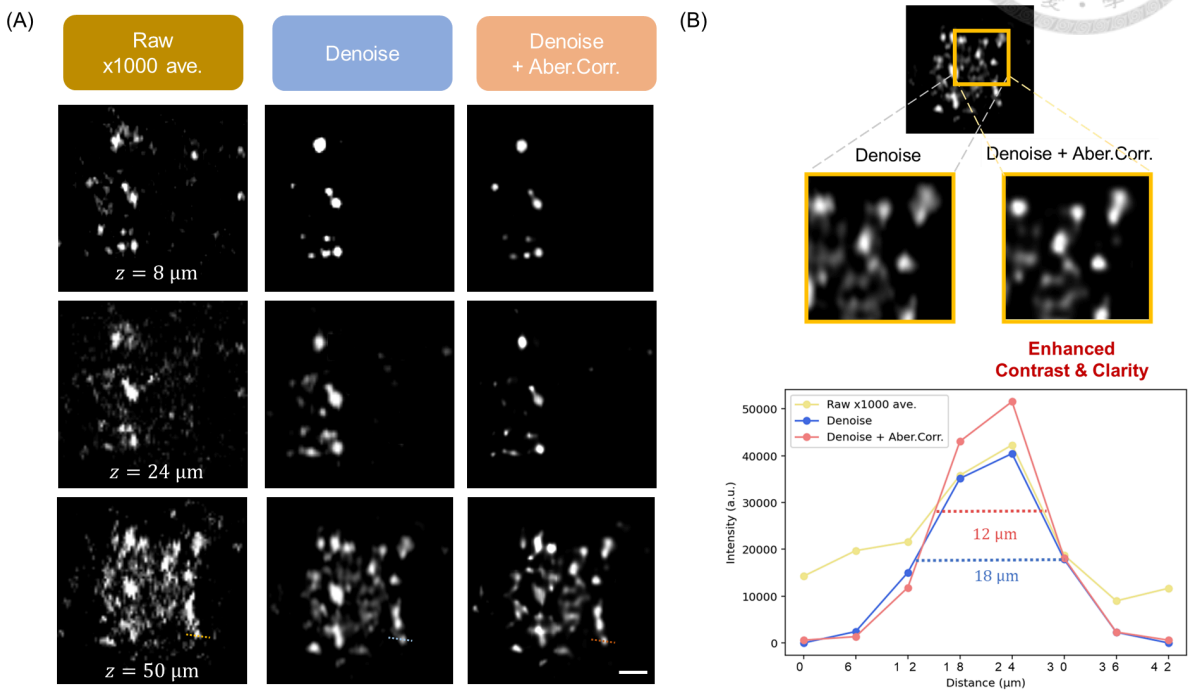


Figure 4-5. Cell shape calibration in SCN calcium imaging via a deconvolution algorithm.

(A) In vivo calcium imaging of SCN neurons at different depths. Left: Average projection of 1000 time points from raw high-speed scanning images. Middle: Denoised images generated by applying the TAG-SPARK 2.0 model to individual frames of high-speed data. Right: Denoised and aberration-corrected images obtained by applying the TAG-SPARK 2.0 model followed by deconvolution. Scale bars: 50 µm. (B) Quantification of cell size at different depths. Yellow curve: Average of 1000 time points from raw images. Blue curve: Denoised images. Pink curve: Denoised images with aberration correction.

The image quality is significantly enhanced by the denoising model. However, the cells appear distorted due to intrinsic aberrations introduced by the GRIN lens. Fortunately, our deconvolution algorithm corrects both the shape and size of the cells (**Figure 4-5A**). Compared

doi:10.6342/NTU202504634

to the average projection of 1000 time points from raw high-speed scanning images, the post-processed images after denoising and aberration correction show cells that closely approximate the theoretical size. The FWHM of the denoised image is about 18 μm, whereas the FWHM of the two post-processed images is approximately 12 μm **(Figure 4-5B**).

To ensure that the post-processing methods do not introduce any artifacts, we perform correlation analysis to compare the pre- and post-processed images. Moreover, peak signal-to-noise ratio (PSNR) analysis is used to evaluate image quality and demonstrate the effects of noise reduction.

During the correlation analysis, the correlation between denoised images from each frame and layer and a 1000-time-point moving average is over 0.8. After the aberration correction, the correlation increases to approximately 0.9. That is, combining both post-processing methods, we derive the images from high-speed scanning that have a high similarity with those from low-speed scanning. According to the PSNR analysis, the PSNR of TAG-SPARK 2.0 denoised images increases from 19 dB to 21.5 dB. To convert a change in PSNR (Peak Signal-to-Noise Ratio) from 19 dB to 21.5 dB into a percentage improvement in signal quality, you can use the fact that PSNR is a logarithmic measure of the ratio between signal power and noise power.

The formula for PSNR is

$$\text{PSNR (dB)} = 10 \log_{10} \left(\frac{\text{MAX}^2}{\text{MSE}} \right) \qquad \text{Eq. 4-1}$$

, where MAX is the maximum possible pixel value and MSE is the mean squared error.

$$\frac{\text{MSE}_2}{\text{MSE}_1} = 10^{-(21.5-19)/10} = 10^{-2.5/10} \approx 10^{-0.25} \approx 0.562 \qquad \text{Eq. 4-2}$$

doi:10.6342/NTU202504634

This means MSE decreased by approximately 43.8% since $1 - 0.562 \approx 0.438$. Therefore, the signal quality improved by about 43.8%, in terms of reduced noise/error. After the aberration correction, it increases 1 dB in PSNR, corresponding to a ~20.6% reduction in noise (**Figure 4-6)**.

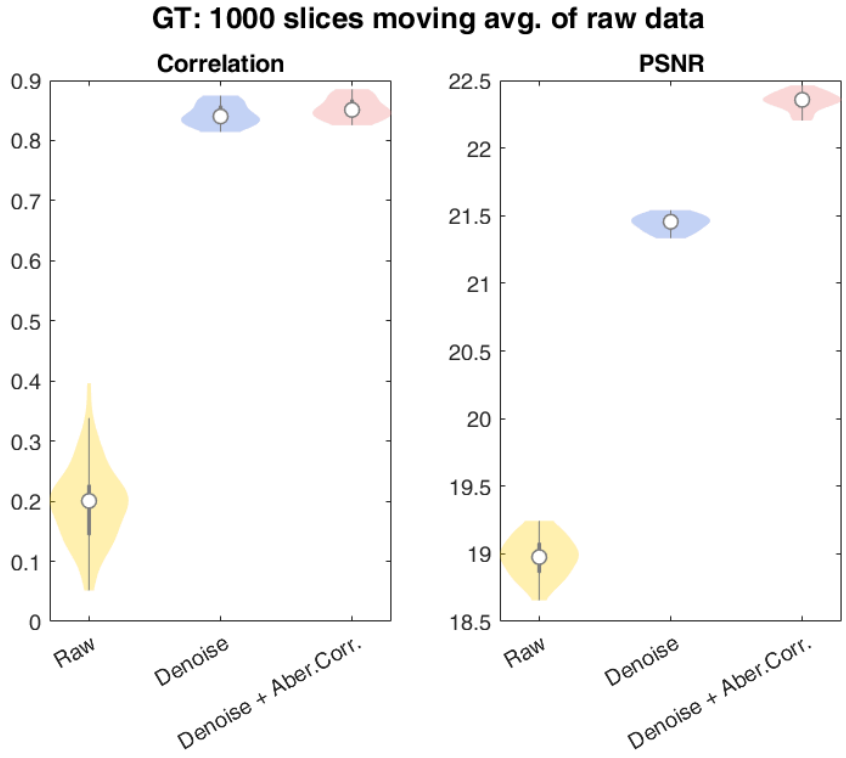


Figure 4-6. Correlation and PSNR analysis with three kinds of datasets.

Yellow dots: Average of 1000 time points from raw images. Blue dots: Denoised images. Pink dots: Denoised images with aberration correction.

doi:10.6342/NTU202504634

Chapter 5 Conclusion and Outlooks

Although previous studies have explored circadian rhythms using brain slices or two-dimensional *in vivo* recordings, these approaches are limited in their ability to reveal dynamic biological processes occurring deep within the brain. Our system addresses this challenge by enabling real-time, three-dimensional imaging of neuronal population in the SCN *in vivo* (**Figure 5-1)**.

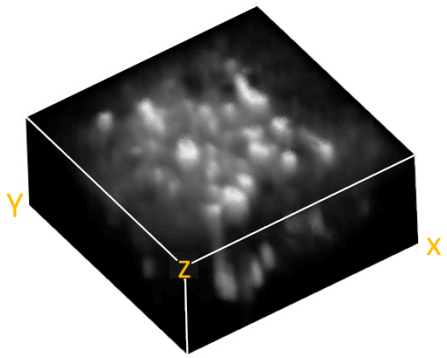


Figure 5-1. Three-dimensional SCN reconstruction obtained by high-speed scanning endoscopy and image post-processing. Image size: 500 × 500 × 56 μm³.

To enhance the quality and interpretability of the captured images, we implemented two post-processing methods: deep learning-based denoising and deconvolution-based aberration correction. The TAG-SPARK 2.0 denoising model significantly reduced noise while preserving signal fidelity, as confirmed by clearer calcium dynamics and a ~43.8% reduction in mean squared error (MSE). Additionally, the deconvolution algorithm corrected shape distortions

doi:10.6342/NTU202504634

introduced by the GRIN lens, resulting in cellular structures that more closely matched theoretical expectations. Post-processed images showed a full width at half maximum (FWHM) of 10 – 15 μm and a 1 dB increase in PSNR, corresponding to a ~20.6% noise reduction.

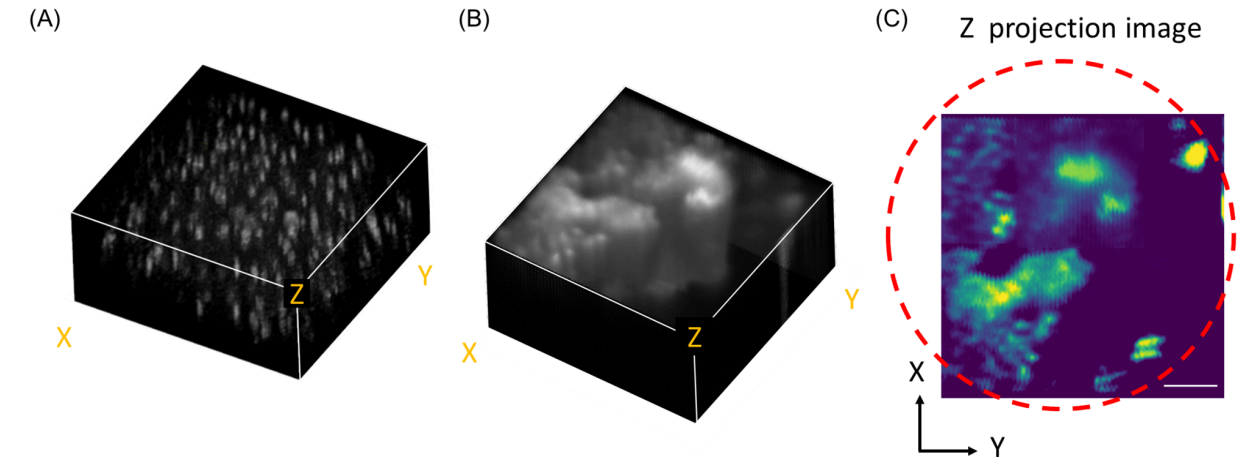


Figure 5-2. Maximum FOV of the three-dimensional sample obtained by high-speed scanning endoscopy. (A)Three dimensional reconstruction of a standard sample 10 μm FLB. (B) Three dimensional reconstruction of SCN. (C) Depth-projection image of SCN. Image size: 500 × 500 × 152 μm³. Scale bar:100 μm. Dashed curve: FOV of the GRIN lens

Together, these results demonstrate that the proposed post-processing pipeline not only improves the overall image quality but also enhances the reliability of functional interpretations. Using a standard sample of 10 μm FLB, the maximum FOV for 3D imaging is measured to be **5**00 × 500 × 152 μm³ (**Figure 5-2A**). This figure clearly illustrates the spatial distribution of the FLB population in three dimensions.

doi:10.6342/NTU202504634

To further explore the structural complexities of the SCN, the maximum *in vivo* FOV was obtained using our high-speed volumetric imaging system (**Figure 5-2B**). The axial resolution in the *in vivo* condition; however, is less well-defined compared to that of the standard sample. This observation might suggest mechanical instability caused by the insertion of the GRIN lens into the living mouse brain. Despite these limitations, the developed imaging system successfully captures multiple neurons within the maximum FOV (**Figure 5-2C**), demonstrating its capability for functional *in vivo* neuroimaging. The integrated imaging and analysis framework thus provides a powerful platform for uncovering the neural mechanisms underlying circadian rhythms and paves the way for future investigations of deep brain circuits *in vivo*.

In the context of the SCN, this approach can be leveraged to probe the complexity of neural networks across distinct circadian phases, as well as assess their responses to varying light exposure intensities. Such applications highlight the potential of the system to advance our understanding of how environmental cues shape circadian regulation at the network level.

doi:10.6342/NTU202504634

Reference list

1 Kennedy, D. & Norman, C. What don't we know? *Science* **309**, 75 (2005). https://doi.org/10.1126/science.309.5731.75

2 Lin, M. Z. & Schnitzer, M. J. Genetically encoded indicators of neuronal activity. *Nature Neuroscience* **19**, 1142-1153 (2016). https://doi.org/10.1038/nn.4359

3 Weisenburger, S. & Vaziri, A. A Guide to Emerging Technologies for Large-Scale and Whole-Brain Optical Imaging of Neuronal Activity. *Annual Review of Neuroscience* **41**, 431-452 (2018). https://doi.org/10.1146/annurev-neuro-072116-031458

4 Inácio, A.R., Lam, K.C., Zhao, Y. Pereira, F. Lee, G. Lee, S. Brain-wide presynaptic networks of functionally distinct cortical neurons. *Nature* **641**, 162–172 (2025). https://doi.org/10.1038/s41586-025-08631-w

5 Choi, J. R., Kim, S. M., Ryu, R. H., Kim, S. P. & Sohn, J. W. Implantable Neural Probes for Brain-Machine Interfaces - Current Developments and Future Prospects. *Experimental Neurobiology* **27**, 453-471 (2018). https://doi.org/10.5607/en.2018.27.6.453

6 Goense, J., Bohraus, Y. & Logothetis, N. K. fMRI at High Spatial Resolution: Implications for BOLD-Models. *Frontiers in Computational Neuroscience* **10**, 66 (2016). https://doi.org/10.3389/fncom.2016.00066

7 Dombeck, D. A., Harvey, C. D., Tian, L., Looger, L. L. & Tank, D. W. Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. *Nat Neurosci* **13**, 1433-1440 (2010). https://doi.org/10.1038/nn.2648

8 Kim, W. & Kim, S. K. Neural circuit remodeling and structural plasticity in the cortex

doi:10.6342/NTU202504634

during chronic pain. *The Korean Journal of Physiology & Pharmacology* **20**, 1-8 (2016). https://doi.org/10.4196/kjpp.2016.20.1.1

9 Dipalo, M. Amin, H. Lovato, L. Moia, F. Caprettini, V. Messina, Gabriele C. Tantussi, F. Berdondini, L. De Angelis, F. Intracellular and Extracellular Recording of Spontaneous Action Potentials in Mammalian Neurons and Cardiac Cells with 3D Plasmonic Nanoelectrodes. *Nano Letters* **17**, 3932-3939 (2017). https://doi.org/10.1021/acs.nanolett.7b01523

10 Jung, J. C., Mehta, A. D., Aksay, E., Stepnoski, R. & Schnitzer, M. J. In vivo mammalian brain imaging using one- and two-photon fluorescence microendoscopy. *Journal Neurophysiol* **92**, 3121-3133 (2004). https://doi.org/10.1152/jn.00234.2004

11 Wu, Z., Lin, D. & Li, Y. Pushing the frontiers: tools for monitoring neurotransmitters and neuromodulators. *Nature Reviews Neuroscience* **23**, 257-274 (2022). https://doi.org/10.1038/s41583-022-00577-6

12 Kim, T. H. & Schnitzer, M. J. Fluorescence imaging of large-scale neural ensemble dynamics. *Cell* **185**, 9-41 (2022). https://doi.org/10.1016/j.cell.2021.12.007

13 Dana, H. Sun, Y. Mohar, B. Hulse, B.K. Kerlin, A.M. Hasseman, J.P. Tsegaye, G. Tsang, A. Wong, A. Patel, R. Macklin, J.J. Chen, Y. Konnerth, A. Jayaraman, V. Looger, L.L. Schreiter, E.R. Svoboda, K. Kim, D.S. High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nature Methods* **16**, 649-657 (2019). https://doi.org/10.1038/s41592-019-0435-6

14 Helmchen, F. & Denk, W. Deep tissue two-photon microscopy. *Nature Methods* **2**, 932-940 (2005). https://doi.org/10.1038/nmeth818

15 Vasquez-Lopez, S.A., Turcotte, R., Koren, V. Plöschner, M. Padamsey, Z. Booth, M.J. Čižmár, T. Emptage, N.J. Subcellular spatial resolution achieved for deep-brain imaging in vivo using a minimally invasive multimode fiber. *Light: Science & Applications* **7**, 110 (2018). https://doi.org/10.1038/s41377-018-0111-0

doi:10.6342/NTU202504634

16 Luu, P., Fraser, S. E. & Schneider, F. More than double the fun with two-photon excitation microscopy. *Communications Biology* **7**, 364 (2024). https://doi.org/10.1038/s42003-024-06057-0

17 Schubert, M.C., Soyka, S.J., Tamimi, A. Maus , E. Schroers, J. Wißmann, N. Reyhan, E. Tetzlaff, S.K. Yang, Y. Denninger, R. Peretzke, R. Beretta, C. Drumm, M. Heuer , A. Buchert, V. Steffens, A. Walshon, J. McCortney, K. Heiland, S. Bendszus, M. Neher, P. Golebiewska, A. Wick , W. Winkler , F. Venkataramani, V. Deep intravital brain tumor imaging enabled by tailored three-photon microscopy and analysis. *Nature Communications* **15**, 7383 (2024). https://doi.org/10.1038/s41467-024-51432-4

18 Levene, M. J., Dombeck, D. A., Kasischke, K. A., Molloy, R. P. & Webb, W. W. In vivo multiphoton microscopy of deep brain tissue. *J Neurophysiol* **91**, 1908-1912 (2004). https://doi.org/10.1152/jn.01007.2003

19 Zhao, C., Chen, S., Zhang, L. Zhang, D. Wu, R. Hu, Y. Zeng, E. Li, Y. Wu, D. Yu, F. Zhang, Y. Zhang, J. Chen, L. Wang, a. Cheng, H. Miniature three-photon microscopy maximized for scattered fluorescence collection. *Nature methods* **20**, 617-622 (2023). https://doi.org/10.1038/s41592-023-01777-3

20 Li, Y. Xu, X. Zhang, C. Sun, X. Zhou, S. Li, X. Guo, J. Hu, R. Qu, J. Liu, L. In Vivo Neurodynamics Mapping via High-Speed Two-Photon Fluorescence Lifetime Volumetric Projection Microscopy. *Advanced Science* **12,** 7 (2024). https://doi.org/10.1002/advs.202410605

21 Chen, Y. Luo, C. Wang, S. Li, Y. Shen, B. Hu, R. Qu, J. Liu, L. Rapid, high-contrast, and steady volumetric imaging with Bessel-beam-based two-photon fluorescence microscopy. *Journal of Biomedical Optics* **29,** 1 (2024). https://doi.org/10.1117/1.Jbo.29.1.016501

22 Wang, S. W., Liu, J., Goh, C. C., Ng, L. G. R. & Liu, B. NIR-II-Excited Intravital Two-Photon Microscopy Distinguishes Deep Cerebral and Tumor Vasculatures with an Ultrabright NIR-I AIE Luminogen. *Advanced Materials* **31,** 44 (2019). https://doi.org/10.1002/adma.201904447

doi:10.6342/NTU202504634

23 Meng, G. Liang, Y. Sarsfield, S. Jiang, W. Lu, R. Dudman, J. Aponte, Y. Ji, N. High-throughput synapse-resolving two-photon fluorescence microendoscopy for deep-brain volumetric imaging in vivo. *eLife* **8**, e40805 (2019). https://doi.org/10.7554/eLife.40805

24 Chen, Q. Cichon, J. Wang, W. Qiu, L. Lee, S. Campbell, N. Destefino, N. Goard, M. Fu, Z. Yasuda, R. Looger, L. Arenkiel, B. Gan, W. Feng , G. Imaging neural activity using Thy1-GCaMP transgenic mice. *Neuron* **76**, 297-308 (2012). https://doi.org/10.1016/j.neuron.2012.07.011

25 Zhang, Y. Rózsa, M. Liang, Y. Bushey, D. Wei, Z. Zheng, J. Reep, D. Broussard, G. Tsang, A. Tsegaye, G. Narayan, S. Obara, C. Lim, J. Patel, Zhang, R. Ahrens, M. Turner, G. Wang, S. Korff, W. Schreiter, E. Svoboda, K. Hasseman, J. Kolb, I. Looger, L.L. Fast and sensitive GCaMP calcium indicators for imaging neural populations. *Nature* **615**, 884-891 (2023). https://doi.org/10.1038/s41586-023-05828-9

26 Xu, C., Nedergaard, M., Fowell, D. J., Friedl, P. & Ji, N. Multiphoton fluorescence microscopy for in vivo imaging. *Cell* **187**, 4458-4487 (2024). https://doi.org/10.1016/j.cell.2024.07.036

27 Prevedel, R. Ortas, J. Kerr, J. Waters, J. Breckwoldt, M. Deneen, B. Monje, M. Soyka, S. Venkataramani, V. Three-photon microscopy: an emerging technique for deep intravital brain imaging. *Nature Reviews Neuroscience* **26**, 521-537 (2025). https://doi.org/10.1038/s41583-025-00937-y

28 Liu, H. Deng, X. Tong, S. He, C. Cheng, H. Zhuang, Z. Gan, M. Li, J. Xie, W. Qiu, P. Wang, K. In Vivo Deep-Brain Structural and Hemodynamic Multiphoton Microscopy Enabled by Quantum Dots. *Nano Letters* **19**, 5260-5265 (2019). https://doi.org/10.1021/acs.nanolett.9b01708

29 Wang, C. & Ji, N. Characterization and improvement of three-dimensional imaging performance of GRIN-lens-based two-photon fluorescence endomicroscopes with adaptive optics. *Optics Express* **21**, 27142-27154 (2013).

doi:10.6342/NTU202504634

https://doi.org/10.1364/OE.21.027142

30 Sancataldo, G., Silvestri, L., Allegra Mascaro, A. L., Sacconi, L. & Pavone, F. S. Advanced fluorescence microscopy for in vivo imaging of neuronal activity. *Optica* **6**, 758-765 (2019). https://doi.org/10.1364/OPTICA.6.000758

31 Lu, K. Chen, Z. Chen, H. Zhou, W. Zhang, Z. Tsang, H. Tong, Y. Empowering high-dimensional optical fiber communications with integrated photonic processors. *Nature Communications* **15**, 3515 (2024). https://doi.org/10.1038/s41467-024-47907-z

32 Bortoletto, F., Bonoli, C., Panizzolo, P., Ciubotaru, C. D. & Mammano, F. Multiphoton Fluorescence Microscopy with GRIN Objective Aberration Correction by Low Order Adaptive Optics. *PLOS ONE* **6**, e22321 (2011). https://doi.org/10.1371/journal.pone.0022321

33 Zhang, L. Liang, B. Barbera, G. Hawes, S. Zhang, Y. Stump, K. Baum, I. Yang, Y. Li, Y. Lin, D. Miniscope GRIN Lens System for Calcium Imaging of Neuronal Activity from Deep Brain Structures in Behaving Animals. *Current Protocols in Neuroscience* **86**, e56 (2019). https://doi.org/10.1002/cpns.56

34 Liu, G. Ahn, S. Kang, J. Bhagavatula, S. Matthew, D. Martin, S. Marlin, C. So, P. Tearney, G. Jonas, O. Two-site microendoscopic imaging probe for simultaneous three-dimensional imaging at two anatomic locations in tissues. *Optics Letters* **49**, 3312-3315 (2024). https://doi.org/10.1364/OL.525945

35 Helmchen, F., Fee, M. S., Tank, D. W. & Denk, W. A Miniature Head-Mounted Two-Photon Microscope: High-Resolution Brain Imaging in Freely Moving Animals. *Neuron* **31**, 903-912 (2001). https://doi.org/https://doi.org/10.1016/S0896-6273(01)00421-4

36 Uhlířová, H., Čižmár, T., Pakan, J. M. P. & Gomes, A. Special Section Guest Editorial: Exploiting Complex Media Photonics to Illuminate Brain's Hidden Depth. *Neurophotonics* **11**, S11501 (2024). https://doi.org/10.1117/1.NPh.11.S1.S11501

doi:10.6342/NTU202504634

37 Zhou, Z. Gordon-Fennell, A. Piantadosi, S. Ji, N. Smith, L. Bruchas, M. Stuber, G. Deep-brain optical recording of neural dynamics during behavior. *Neuron* **111**, 3716-3738 (2023). https://doi.org/https://doi.org/10.1016/j.neuron.2023.09.006

38 Weber, T. D., Moya, M. V., Kılıç, K., Mertz, J. & Economo, M. N. High-speed multiplane confocal microscopy for voltage imaging in densely labeled neuronal populations. *Nature Neuroscience* **26**, 1642-1650 (2023). https://doi.org/10.1038/s41593-023-01408-2

39 Yang, Y., Forbes, A. & Cao, L. A review of liquid crystal spatial light modulators: devices and applications. *Opto-Electronic Science* **2**, 230026-230029 (2023). https://doi.org/10.29026/oes.2023.230026

40 Wang, D., Meng, Y., Chen, D., Yam, Y. & Chen, S.-C. High-speed 3D imaging based on structured illumination and electrically tunable lens. *Chinese Optics Letters* **15**, 090004 (2017).

41 Chien, Y. Lin, J. Yeh, P. Hsu, K. Tsai, Y. Chen, S. Chu, S. Dual GRIN lens two-photon endoscopy for high-speed volumetric and deep brain imaging. *Biomedical Optics Express* **12**, 162-172 (2021). https://doi.org/10.1364/BOE.405738

42 Gillies, R. J., Kinahan, P. E. & Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **278**, 563-577 (2016). https://doi.org/10.1148/radiol.2015151169

43 Lambin, P. Leijenaar, R. Deist, T. Peerlings, J. de Jong, E. Timmeren, J. Sanduleanu, S. Larue, R. Even, A. Jochems, A. Wijk, Y. Woodruff, H. Soest, J. Lustberg, T. Roelofs, E. Elmpt, W. Dekker, A. Mottaghy, F. Wildberger, J. Walsh, S. Radiomics: the bridge between medical imaging and personalized medicine. *Nature Reviews Clinical Oncology* **14**, 749-762 (2017). https://doi.org/10.1038/nrclinonc.2017.141

44 Eom, M. Han, S. Park, P. Kim, G. Cho, E. Sim, J. Lee, K. Kim, S. Tian, H. Böhm, U. Lowet, E. Tseng, H. Choi, J. Lucia, S. Ryu, S. Rózsa, M. Chang, S. Kim, P. Han, X. Piatkevich, Choi, M. Kim, C. Cohen, A. Chang, J. Yoon, Y. Statistically unbiased

doi:10.6342/NTU202504634

prediction enables accurate denoising of voltage imaging data. *Nature Methods* **20**, 1581-1592 (2023). https://doi.org/10.1038/s41592-023-02005-8

45 Choi, W. Park, B. Choi, S. Oh, D. Kim, J. Kim, C. Recent Advances in Contrast-Enhanced Photoacoustic Imaging: Overcoming the Physical and Practical Challenges. *Chemical Reviews Journal - ACS Publications* **123**, 7379-7419 (2023). https://doi.org/10.1021/acs.chemrev.2c00627

46 LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436-444 (2015). https://doi.org/10.1038/nature14539

47 Qu, L. Zhao, S. Huang, Y. Ye, X. Wang, K. Liu, Y. Liu, X. Mao, H. Hu, G. Chen, W. Guo, C. He, J. Tan, J. Li, H. Chen, L. Zhao, W. Self-inspired learning for denoising live-cell super-resolution microscopy. *Nature Methods* **21**, 1895-1908 (2024). https://doi.org/10.1038/s41592-024-02400-9

48 van den Boom, B. E, A. Rasmussen, P. Beest, E. Parthasarathy , A. Denys, D. Willuhn, I. Unraveling the mechanisms of deep-brain stimulation of the internal capsule in a mouse model. *Nature Communications* **14**, 5385 (2023). https://doi.org/10.1038/s41467-023-41026-x

49 Lehtinen, J. Munkberg, J. Hasselgren, J. Laine, S. Karras, T. Aittala, M. Aila, T.Noise2Noise: Learning Image Restoration without Clean Data. *ArXiv,* 10.48550.1803.04189 (2018). https://doi.org/10.48550/arXiv.1803.04189

50 Özgün, Ç., Ahmed, A., S., S. L., Thomas, B. & Olaf, R.3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. *arXiv,* 10.48550.1606.0665 (2016). https://arxiv.org/abs/1606.06650

51 Antonini, A. Sattin, A. Moroni, M. Bovetti, S. Moretti, C. Succol, F. Forli, A. Vecchia, D. Rajamanickam, V. Bertoncini, A. Panzeri, S. Liberale, C. Fellin, T.Extended field-of-view ultrathin microendoscopes for high-resolution two-photon imaging with minimal invasiveness. *Elife* **9**, e58882 (2020). https://doi.org/10.7554/eLife.58882

doi:10.6342/NTU202504634

52 Zhu, J., Natan, R., Zhong, J., Kang, I. & Ji, N. Aberration measurement and correction for ultrafast two-photon fluorescence imaging. *SPIE*, 13328 (2025).

53 Sattin, A. Nardin, C. Daste, S. Moroni, M. Reddy, I. Liberale, C. Panzeri, S. Fleischmann, A. Fellin, T. Aberration correction in long GRIN lens-based microendoscopes for extended field-of-view two-photon imaging in deep brain regions. *eLife* **13**, 101420 (2024).https://doi.org/10.7554/eLife.101420.3

54 Adie, S., Graf, B., Ahmad, A., Carney, P. & Boppart, S. Computational adaptive optics for broadband optical interferometric tomography of biological tissue. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 7175-7180 (2012). https://doi.org/10.1073/pnas.1121193109

55 Meijer, J. H., Michel, S., Vanderleest, H. T. & Rohling, J. H. Daily and seasonal adaptation of the circadian clock requires plasticity of the SCN neuronal network. *Eur J Neurosci* **32**, 2143-2151 (2010). https://doi.org/10.1111/j.1460-9568.2010.07522.x

56 Tsuno, Y. Peng, Y. Horike, S. Wang, M. Matsui, A. Yamagata, K. Sugiyama, M. Nakamura, T. Daikoku, T. Maejima, T. Mieda, M. In vivo recording of suprachiasmatic nucleus dynamics reveals a dominant role of arginine vasopressin neurons in circadian pacesetting. *PLOS Biology* **21**, e3002281 (2023). https://doi.org/10.1371/journal.pbio.3002281

57 Yeh, P. Jhan, K. Chua, E. Chen, W. Wu, S. Chen, S. Discrete photoentrainment of mammalian central clock is regulated by bi-stable dynamic network in the suprachiasmatic nucleus. *Nature Communications* **16**, 3331 (2025). https://doi.org/10.1038/s41467-025-58661-1

58 Ness, N. & Brancaccio, M. Network-level time computations in the suprachiasmatic nucleus. *Cell Research* **34**, 471-472 (2024). https://doi.org/10.1038/s41422-024-00969-6

59 Wang, Z. Yu, J. Zhai, M. Wang, Z. Sheng, K. Zhu, Y. Wang, T. Liu, M. Wang, L. Yan, M. Zhang, M. Xu, Y. Wang, X. Ma, L. Hu, W. Cheng, H. System-level time computation and

doi:10.6342/NTU202504634

representation in the suprachiasmatic nucleus revealed by large-scale calcium imaging and machine learning. *Cell Research* **34**, 493-503 (2024). https://doi.org/10.1038/s41422-024-00956-x

60 Xu, C. & Webb, W. W. Measurement of two-photon excitation cross sections of molecular fluorophores with data from 690 to 1050 nm. *Journal of the Optical Society of America B* **13**, 481-491 (1996). https://doi.org/10.1364/JOSAB.13.000481

61 Zhao, Z. Zhou, Y. Liu, B. He, J. Zhao, J. Cai, Y. Fan, J. Li, X. Wang, Z. Lu, Z. Wu, J. Qi, H. Dai, Q. Two-photon synthetic aperture microscopy for minimally invasive fast 3D imaging of native subcellular behaviors in deep tissue. *Cell* **186**, 2475-2491 (2023). https://doi.org/10.1016/j.cell.2023.04.016

62 Takasaki, K., Abbasi-Asl, R. & Waters, J. Superficial Bound of the Depth Limit of Two-Photon Imaging in Mouse Brain. *eneuro* **7**, 0255-0219 (2020). https://doi.org/10.1523/eneuro.0255-19.2019

63 Drobizhev, M., Makarov, N. S., Tillo, S. E., Hughes, T. E. & Rebane, A. Two-photon absorption properties of fluorescent proteins. *Nature Methods* **8**, 393-399 (2011). https://doi.org/10.1038/nmeth.1596

64 Ma, L. Zhang, X. Xu, Z. Späth, A. Xing, Z. Sun, T. Tai, R. Three-dimensional focal stack imaging in scanning transmission X-ray microscopy with an improved reconstruction algorithm. *Optics Express* **27**, 7787-7802 (2019). https://doi.org/10.1364/OE.27.007787

65 Durr, N. J., Weisspfennig, C. T., Holfeld, B. A. & Ben-Yakar, A. Maximum imaging depth of two-photon autofluorescence microscopy in epithelial tissues. *J Biomed Opt* **16**, 026008 (2011). https://doi.org/10.1117/1.3548646

66 He, C. Chang, J. Hu, Q. Wang, J. Antonello, J. He, H. Liu, S. Lin, J. Dai, B. Elson, D. Xi, P. Ma, H. Booth, M. Complex vectorial optics through gradient index lens cascades. *Nature Communications* **10**, 4264 (2019). https://doi.org/10.1038/s41467-019-12286-3

doi:10.6342/NTU202504634

67 Jung, J. C. & Schnitzer, M. J. Multiphoton endoscopy. *Optics Letters* **28**, 902-904 (2003). https://doi.org/10.1364/OL.28.000902

68 Göbel, W., Kerr, J. N., Nimmerjahn, A. & Helmchen, F. Miniaturized two-photon microscope based on a flexible coherent fiber bundle and a gradient-index lens objective. *Opt Lett* **29**, 2521-2523 (2004). https://doi.org/10.1364/ol.29.002521

69 Moretti, C., Antonini, A., Bovetti, S., Liberale, C. & Fellin, T. Scanless functional imaging of hippocampal networks using patterned two-photon illumination through GRIN lenses. *Biomedical Optics Express* **7**, 3958-3967 (2016). https://doi.org/10.1364/BOE.7.003958

70 Sato, M. Motegi, Y. Yagi, S. Gengyo-Ando, K. Ohkura, M. Nakai, J. Fast varifocal two-photon microendoscope for imaging neuronal activity in the deep brain. *Biomed Opt Express* **8**, 4049-4060 (2017). https://doi.org/10.1364/boe.8.004049

71 Mermillod-Blondin, A., McLeod, E. & Arnold, C. B. High-speed varifocal imaging with a tunable acoustic gradient index of refraction lens. *Optics Letters* **33**, 2146-2148 (2008). https://doi.org/10.1364/OL.33.002146

72 Yamato, K., Tanaka, Y., Oku, H., Yasutomi, K. & Kawahito, S. Quasi-simultaneous multi-focus imaging using a lock-in pixel image sensor and TAG lens. *Optics Express* **28**, 19152-19162 (2020). https://doi.org/10.1364/OE.394760

73 Theriault, C., Guttenfelder, J. & Arnold, C. B. TAG Lens: Revolutionizing Optical Microscopy With Ultra-High Speed Variable Focus. *Microscopy and Microanalysis* **20**, 1092-1093 (2014). https://doi.org/10.1017/S1431927614007181

74 Goodman, J. W. & Cox, M. E. Introduction to Fourier Optics. *Physics Today* **22**, 97-101 (1969). https://doi.org/10.1063/1.3035549

75 Olivier, N., Mermillod-Blondin, A., Arnold, C. B. & Beaurepaire, E. Two-photon microscopy with simultaneous standard and extended depth of field using a tunable

doi:10.6342/NTU202504634

acoustic gradient-index lens. *Opt Lett* **34**, 1684-1686 (2009). https://doi.org/10.1364/ol.34.001684

76 Kang, S., Duocastella, M. & Arnold, C. B. Variable optical elements for fast focus control. *Nature Photonics* **14**, 533-542 (2020). https://doi.org/10.1038/s41566-020-0684-z

77 Du, X., Florian, C. & Arnold, C. B. Single-lens dynamic z-scanning for simultaneous in situ position detection and laser processing focus control. *Light: Science & Applications* **12**, 274 (2023). https://doi.org/10.1038/s41377-023-01303-2

78 Li, Y. Su, Y. Guo, M. Han, X. Liu, J. Vishwasrao, H. Li, X. Christensen, R. Sengupta , T. Moyle, M. Rey-Suarez, I. Chen, J. Upadhyaya, A. Usdin, T. Colón-Ramos, D. Liu, H. Wu, Y. Shroff, H. Incorporating the image formation process into deep learning improves network performance. *Nature Methods* **19**, 1427-1437 (2022). https://doi.org/10.1038/s41592-022-01652-7

79 Li, X. Li, Y., Zhou, Y. Wu, J. Zhao, Z. Fan, J. Deng, F. Wu, Z. Xiao, G. He, J. Zhang, Y. Zhang, G. Hu, X. Chen, X. Zhang, Y. Qiao, H. Xie, H. Li, Y. Wang, H. Fang, L. Dai, Q. Real-time denoising enables high-sensitivity fluorescence time-lapse imaging beyond the shot-noise limit. *Nature Biotechnology* **41**, 282-292 (2023). https://doi.org/10.1038/s41587-022-01450-8

80 Jung, W. Lee, H. Seo, M. Nam, Y. Choi, Y. Shin, N. Ahn, K. Kim, B. Jang, J. MR-self Noise2Noise: self-supervised deep learning-based image quality improvement of submillimeter resolution 3D MR images. *European Radiology* **33**, 2686-2698 (2023). https://doi.org/10.1007/s00330-022-09243-y

81 Li, X. Zhang, G. Wu, J. Zhang,Y. Zhao,Z. Lin, X. Qiao, H. Xie, H. Wang, H. Fang, L. Dai, Q. Reinforcing neuron extraction and spike inference in calcium imaging using deep self-supervised denoising. *Nature Methods* **18**, 1395-1400 (2021). https://doi.org/10.1038/s41592-021-01225-0

doi:10.6342/NTU202504634

82 Deng, Q. Wen, Z. Dong, Z. Tang, J. Chen, W. Liu, X. Yang, Q. Spatially variant deblur and image enhancement in a single multimode fiber imaged by deep learning. *Optics Letters* **47**, 5040-5043 (2022). https://doi.org/10.1364/OL.469034

83 Hsieh, Y. Jhan, K. Lee, J. Huang, G. Chung, C. Chen, W. Chang, T. Chen, B. Pan, M. Wu, S. Chu, S. TAG-SPARK: Empowering High-Speed Volumetric Imaging With Deep Learning and Spatial Redundancy. *Advanced Science* **11**, e2405293 (2024). https://doi.org/10.1002/advs.202405293

84 Turcotte, R., Sutu, E., Schmidt, C. C., Emptage, N. J. & Booth, M. J. Deconvolution for multimode fiber imaging: modeling of spatially variant PSF. *Biomedical Optics Express* **11**, 4759-4771 (2020). https://doi.org/10.1364/BOE.399983

85 Wang, X., Park, J., Susztak, K., Zhang, N. R. & Li, M. Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. *Nature Communications* **10**, 380 (2019). https://doi.org/10.1038/s41467-018-08023-x

86 Bertero, M. & Boccacci, P. From Cells to Proteins: Imaging Nature across Dimensions. *Springer Netherlands,* 349-370 (2005). https://DOI:10.1007/1-4020-3616-7

87 Sage, D. Donati, L. Soulez, F. Fortun, D. Schmit, G. Seitz, G. Seitz, A. Guiet, R. Vonesch , C. Unser, M. DeconvolutionLab2: An open-source software for deconvolution microscopy. *Methods* **115**, 28-41 (2017). https://doi.org/https://doi.org/10.1016/j.ymeth.2016.12.015

88 Yanny, K., Monakhova, K., Shuai, R. W. & Waller, L. Deep learning for fast spatially varying deconvolution. *Optica* **9**, 96-99 (2022). https://doi.org/10.1364/OPTICA.442438

doi:10.6342/NTU202504634