

Wavefront shaping and deep learning in fiber endoscopy

Présentée le 2 octobre 2020

à la Faculté des sciences et techniques de l'ingénieur
Laboratoire d'optique
Programme doctoral en photonique

pour l'obtention du grade de Docteur ès Sciences

par

Eirini KAKKAVA

Acceptée sur proposition du jury

Prof. L. Thévenaz, président du jury
Prof. D. Psaltis, directeur de thèse
Prof. R. Piestun, rapporteur
Prof. L. Tian, rapporteur
Prof. A. Radenovic, rapporteuse

*Happiness can be found
even in the darkest of times,
when one only remembers to turn on
the light...*

J.K. Rowling

Acknowledgements

First and foremost, I would like to thank my thesis supervisor Prof. Demetri Psaltis for giving me the opportunity to join the group of the Optics Laboratory at EPFL and obtain this remarkable experience. His guidance and insight were always a reference point for the advancement of my research during this PhD. Furthermore, I would like to express the appreciation for his trust on me from the very beginning for several tasks such as assisting in his courses and representing our work in conferences worldwide as well as for his overall support in every step for these past years.

Special thanks to Prof. Moser with whom I collaborated closely on several projects, for all his ideas and feedback during my PhD studies and for being always ready to help and answer my questions. I would also like to thank him for always being welcoming to me participating in activities of his group and introducing me to new members.

Importantly, I would like to thank the members of my PhD thesis committee, Prof. Thévenaz, Prof. Piestun, Prof. Tian and Prof. Radenovic for reviewing my work and giving me insightful comments. The time they devoted on reading this thesis led to not only improving the manuscript but also my understanding of the various topics.

Every beginning can be difficult but for me it could not be smoother, because of all my colleagues in the group being always helpful and supportive and they deserve many thanks for that. Especially, I would like to thank Don Conkey and Nico Stasio for their guidance through the practical aspects of my PhD, their patience in explaining new concepts and their unlimited willingness to transfer experimental and theoretical knowhow. Many thanks to Marilisa Romito and Giulia Panusa with whom I basically lived together in the lab all these years and they have been a constant support at a professional but also a very personal level thus becoming from colleagues to valuable friends through this process. Furthermore, I would like to thank Navid Borhani for collaborating always in the most efficient way from the very beginning of his participation in the Optics Laboratory and of course for being there for me to advice and support with his uniquely logical and structured thinking. I would probably need many pages to separately thank every person that I had the pleasure and honor to work with during my PhD in my lab but nevertheless I want to warmly thank: Thomas Lanvin, Ye Pu, Elizabeth Antoine, Mohammad Hashemi, Joowon Lim, Amir Sabba, Pooria Hadikhani, Ahmed Bassam, Ugur Tegin, Morteza Shoreh, Alex Goy, Ulas Dinc, Alexa Guglielmelli, Miguel Modestino. Additionally, I would like to thank our secretaries in the lab that made every process easy for all of us, Carole Berthet, Anne De Witte, Silke Jan and Sabrina Martone.

Having one great lab to work in is by itself very important; in my case I was doubly lucky to have the opportunity to collaborate closely and almost feel like part of the Laboratory of Applied Photonic Devices. Therefore, I would like to thank all the LAPD group for not only their research feedback but also their friendship. Particularly, I would like to thank Damien Loterie that he has always found the time to answer my questions and give me suggestions no matter how many times I show up at his office door and I appreciate it greatly. Moreover, special thanks to Enrico Chinello, Chiara Bonati, Georgia Konstantinou and Babak Rahmani for their support and friendship as well as for constantly hosting me in their office for coffee breaks.

Alongside, I would also like to thank the people who made my moving in Lausanne easy and pleasant, Marina, Denis, Georgios and Milena Kiriakopoulos and Evita Papadopoulou. Many thanks to Iliana Spartali that was always there to advise, support and discuss on a personal and professional level, firmly believing in me in every step on the way. Furthermore, I would like to thank all of my people in Greece that supported my decision towards the PhD and they hold a special place in my heart for that, among which are Elina Petala, Dionisis Potamianos, Maria Kotzagianni, Irene Papagiannouli and particularly Kostas Papanikolaou who is there for me for more than eleven years.

Last but not least, I owe a special thanks to the people that made sure that I had all I needed to be here for this PhD, namely my family. I cannot thank them enough and appreciate all their efforts in my life. Especially, I would like to thank my brother for being my best friend and support for as long as I remember myself and quote him saying that we are the best present our parents gave to one another.

Abstract

Fiber endoscopy plays an important role in clinical diagnosis and treatment processes involved in modern medicine. Thin fiber probes can relay information from confined places in the human body that are inaccessible for conventional bulky microscopes. Therefore, they can provide a minimally invasive platform for optical diagnosis assisting the fast recovery of patients.

Aiming for more advanced and compact fiber endoscopic devices, we investigate different types of fibers that can integrate not only superior imaging modalities but also microsurgery capabilities while maintaining an ultrathin size (less than 400 μm). In our experiments, the imaging potential of multimode and multicore fibers is assessed. On the one hand, multimode fibers provide high information capacity for small core sizes but scramble the input field after its propagation through the fiber length, while on the other hand, multicore fibers are able to directly relay images from the inspection location to the observation side and for this reason they are the most common choice for endoscopic probes. However, the final images produced by multicore fibers are of poor quality because of the discretization effect induced by the individual core sampling and the final resolution is dependent on the core spacing. We show that each fiber type presents certain advantages and the type selection can be made based on the desired application aiming to improve the quality of imaging results.

Two approaches for imaging through the different fibers are investigated. In the first part of the present thesis, wavefront shaping using the transmission matrix approach to generate a focus spot at the distal fiber side is presented. The limitations concerning the maximum peak intensity guided through the different endoscopes is investigated for high power femtosecond pulses where nonlinear optical phenomena can hinder the overall performance of the system. Femtosecond laser ablation is demonstrated through multimode fibers for a range of materials. Furthermore, laser ablation is combined with two-photon fluorescence imaging in the same multimode fiber endoscope showing for the first time selective tissue modifications at a cellular level.

In the second part of this thesis, deep learning of the light propagation through the two fiber types is studied. Datasets of known input images and their respective fiber output images are generated to train deep neural network algorithms to map the fiber output to the fiber input for either classification or image reconstruction purposes. The deep neural networks show impressive performance to recover the information from intensity-only images of the speckle patterns emerging from multimode fibers, removing the need to record the full field information, which usually implies experimental complexity. Moreover, a trained deep neural network model proved to recover successfully the information delivered by multimode fiber for imaging even in the

presence of noise in the datasets related to mechanical, thermal and wavelength perturbations. Such measurement drifts caused by perturbations are usually catastrophic for calibration-based techniques based on wavefront shaping. In the case of multicore fibers, deep neural networks are trained to remove the discretization artefact from the final image, which is generated by the sampling using many individual cores and resolve features with an improved resolution. Finally, deep learning is employed to integrate phase imaging capabilities to a bright field imaging-based commercial endoscope by training deep neural networks to map the intensity-only image delivered by the multicore fiber to its corresponding phase map.

Overall, we demonstrate that multimode and multicore fibers can be more than just imaging devices and underline their potential for microsurgical applications using wavefront shaping. For more customized information recovery, deep learning is a robust method, which can be easily applied for different purposes such as classification and multimodal microscopy.

Keywords: fiber endoscopy, wavefront shaping, transmission matrix, deep neural networks, image classification, image reconstruction

Résumé

L'endoscopie par fibres optiques joue un rôle important dans les diagnostics et les traitements impliqués dans la médecine moderne. Les fibres optiques peuvent relayer des informations à partir d'endroits confinés dans le corps humain qui sont inaccessibles des microscopes conventionnels. Par conséquent, elles peuvent permettre la réalisation d'une plateforme peu invasive pour le diagnostic optique favorisant la récupération rapide des patients.

Dans le but de contribuer au développement d'appareils d'endoscopie par fibres optiques plus compacts, nous étudions dans cette thèse différents types de fibres optiques qui peuvent intégrer non seulement des modalités d'imagerie supérieures, mais également des capacités de microchirurgie tout en conservant une taille ultra-mince (moins de 400 μm). Dans nos expériences, les fibres multimodes et multicœurs sont testées pour évaluer leur potentiel pour l'imagerie. D'une part, les fibres multimodes présentent une capacité d'information élevée pour les petites tailles de cœurs mais brouillent le champ d'entrée après propagation, tandis que, d'autre part, les fibres multicœurs sont le choix le plus courant pour les endoscopes, mais elles génèrent des images pixellisées de mauvaise qualité. Nous montrons que le choix de fibre peut être optimisé en fonction de l'application souhaitée.

Deux voies sont suivies pour l'imagerie à travers les différentes fibres optiques. Dans la première partie de cette thèse, la matrice de transmission est utilisée pour générer un point focale à l'extrémité distale de la fibre. Les limites concernant l'intensité maximale guidée à travers les différents endoscopes sont étudiées pour les impulsions femtosecondes de haute puissance où les phénomènes optiques non linéaires peuvent entraver la performance globale du système. L'ablation laser femtoseconde est démontrée à travers des fibres multimodes pour une gamme de matériaux. De plus, l'ablation au laser est combinée à l'imagerie de fluorescence par excitation à deux photons dans le même endoscope à fibre multimode, montrant ainsi pour la première fois des modifications tissulaires sélectives au niveau cellulaire.

Dans la deuxième partie de cette thèse, nous étudions l'utilisation de l'intelligence artificielle et plus précisément de l'apprentissage profond de la propagation de la lumière à travers les deux types de fibres. Des ensembles de données d'images d'entrée connues ainsi que de leurs images de sortie de fibre respectives sont générés pour entraîner des algorithmes de réseaux de neurones profonds pour mapper la sortie à l'entrée de fibre à des fins de classification ou de reconstruction d'image. Les réseaux de neurones profonds affichent des performances impressionnantes pour récupérer les informations des images d'intensité des tavelures des fibres multimodes sans besoin d'enregistrer l'ensemble des informations de front d'onde. De plus, l'apprentissage profond pour

L'imagerie par fibre multimode s'est avérée résiliente aux perturbations liées aux dérives mécaniques, thermiques et même de longueurs d'onde introduites pendant les mesures. Dans le cas des fibres multicœurs, les réseaux de neurones profonds sont entraînés à supprimer les artefacts de pixellisation de l'image finale et à reproduire les détails avec une résolution améliorée. Enfin, nous utilisons l'apprentissage profond pour ajouter un module de contraste de phase sur un endoscope commercialement disponible avec lequel nous enregistrons seulement des images à fond clair.

Dans l'ensemble, nous démontrons que les fibres multimodes et multicœurs peuvent être plus que des dispositifs d'imagerie et soulignons leur potentiel pour des applications microchirurgicales utilisant l'adaptation de front d'onde. Pour une récupération d'informations plus adaptée aux besoins particuliers, l'apprentissage profond est une méthode robuste qui peut être facilement appliquée à la classification et la microscopie multimodale à travers des fibres optiques.

Mots clés: endoscopie par fibres optiques, matrice de transmission, adaptation de front d'onde, réseaux de neurones profonds, classification d'image, reconstruction d'image.

Contents

Acknowledgements.....	i
Abstract	iii
Résumé.....	v
Contents	vii
List of Figures	xi
List of Tables	xvii
List of Equations.....	xix
Chapter 1 Introduction	1
1.1 Basic concepts in optical fibers	2
1.1.1 Light propagation through optical fibers.....	2
1.1.2 Optical fiber types.....	4
1.1.3 Fiber imaging	6
1.2 State-of-the-art in fiber endoscopy	11
1.3 Motivation.....	14
PART A: Wavefront-shaping in fiber endoscopy.....	17
Chapter 2 Digital holography	19
2.1 Principles of holography.....	19
2.2 Digital holography	20
2.2.1 Minimum reference angle	21
2.2.2 Digital sampling	23
2.2.3 Reconstruction of the complex field	23
2.3 Spatial light modulators	24
2.4 Summary	25
Chapter 3 Transmission Matrix	27
3.1 Introduction	27

3.1.1 Theoretical aspects	27
3.1.2 Measurement basis	29
3.1.3 Practical requirements	30
3.2 Methods	31
3.2.1 Optical apparatus	31
3.2.2 Biological sample preparation	33
Chapter 4 High-intensity focusing through optical fibers	35
4.1 Multicore fiber probes for endoscopy	36
4.1.1 Wavefront control through MCFs	36
4.1.2 Nonlinear degradation of focusing efficiency	38
4.1.3 Comparison of two different MCFs	43
4.1.4 Ultrashort pulse ablation through an MCF	45
4.1.5 Conclusions	46
4.2 Multimode fiber probes for endoscopy	47
4.2.1 High peak intensity delivery through GRIN fibers	48
4.2.2 TPF imaging feedback-based selective FLA	53
4.2.3 FLA of a various materials through GRIN fibers	55
4.3 Conclusions	56
PART B: Deep learning in fiber endoscopy	57
Chapter 5 Deep Neural Networks	59
5.1 Basic concepts in Artificial Neural Networks	60
5.1.1 The perceptron	60
5.1.2 Activation functions	61
5.1.3 Error functions	62
5.2 Convolutional neural networks	63
5.3 Computational imaging	64
Chapter 6 Seeing through multimode fibers using deep learning	67
6.1 Information recovery through multimode fibers using Deep Neural Networks ..	68
6.1.1 Methods	69
6.1.2 Image reconstruction	70
6.1.3 Image classification	72

6.2	Intensity-only versus holographic data recording	75
6.2.1	Methods.....	76
6.2.2	Speckle image classification	77
6.2.3	Dataset processing effect on Deep Neural Networks performance	78
6.3	Effect of wavelength drifting on the deep learning performance.....	81
6.3.1	Methods.....	82
6.3.2	Results.....	82
6.4	Conclusions	87
Chapter 7	Deep learning-enhanced imaging through fiber bundles.....	89
7.1	Introduction	89
7.2	Image reconstruction of different datasets through a fiber bundle	90
7.2.1	Methods.....	91
7.2.2	Amplitude modulated inputs.....	92
7.2.3	Phase modulated inputs	98
7.3	Bending sensitivity of the fiber bundle probe	99
7.3.1	Methods.....	100
7.3.2	Results.....	101
7.4	Deep learning assisted phase contrast endoscope.....	104
7.4.1	Methods.....	106
7.4.2	Phase endoscopy of liver tissue samples using deep learning.....	111
7.5	Conclusions	119
Summary	121
Outlook	123
Appendix	125
References	147
Curriculum Vitae	157

List of Figures

Figure 1. Light guiding in step-index optical fibers (left) and the refractive index profile along the fiber core (right).	2
Figure 2. Intensity (first row) and phase (second row) profiles of fiber modes, LP01 mode (a,d), LP11 (b,e) and LP44 (c,f).	4
Figure 3. Light guiding in GRIN optical fibers (left) and the refractive index profile along the fiber core (right).	4
Figure 4. Fiber bundle scheme (left) and the refractive index profile along the bundle core (right).	5
Figure 5. Other fiber types: a) Silica fiber with air-hole dopants, b) solid-core PCF and c) hollow-core PCF	6
Figure 6. Image decomposition in plane waves via a 2D Fourier transform.	7
Figure 7. Image band-passing. a) Image result when limiting some of its high frequency components as shown in b)	8
Figure 8. Image propagation through a MMF. Snapshots at different distances from the input facet.	9
Figure 9. Image at the MCF a) input and b) output.	10
Figure 10. Holographic imaging. a) Hologram recording and b) hologram readout for image generation by illuminating the hologram with the conjugate reference beam.	20
Figure 11. Fourier spectral components of the hologram.	22
Figure 12. Extraction of phase and amplitude of a wavefront using digital holography.	23
Figure 13. Visualization of the generation of patterns through the GRIN fiber using the measured TM of the optical system.	28
Figure 14. Transmission matrix experimental setup for focusing and scanning the light through the optical fibers.	32
Figure 15. Visualization of wavefront shaping through an MCF.	38
Figure 16. The focusing efficiency when focusing ultrashort pulses of varying pulse widths through a MCF compared to the input pulse energy.	39
Figure 17. Spectral and temporal measurements in the focus (blue curve) and background (orange, yellow and purple curves) for different input pulse energies.	41

Figure 18. The estimated peak intensity in focus plotted for increasing input pulse energy and pulse width [51].	42
Figure 19. Characterization of the 10,000 core fiber system.	44
Figure 20. Widefield transmission optical images of ablated samples on thin gold films deposited on glass.	46
Figure 21. Laser ablation spots created using different pulse energies through the MCF.....	46
Figure 22. Spectral measurements for the lowest and highest values of input pulse energy in the case of the 200 μm GRIN fiber.	48
Figure 23. Focus delivery characterization of the 200 μm (a and b) and 400 μm (c and d) core diameter fibers..	50
Figure 24. Fiber damage of the 200 μm core GRIN when the input pulse energy increases above a certain threshold.....	51
Figure 25. Focus spot diameter as a function of the focusing distance from the distal fiber facet.....	52
Figure 26. TPF imaging of the Cochlear hair cell sample stained with PI....	53
Figure 27. BF images that show the time evolution of the bubble generated after FLA of a hair cell within the organ of Corti.	55
Figure 28. a) EPFL logo and b) a square of spots ablated on a glass slide...	56
Figure 29. Schematic description of a) a single perceptron, b) a multilayer perceptron neural network.	60
Figure 30. Example of a convolutional block. One input image is convolved with two different kernels, processed by a ReLU activation function and then downsampled to half the size using maxpooling.	64
Figure 31. The experimental setup for pattern transmission through the MMF.	68
Figure 32. Details of the implemented (a) VGG type image classifier and (b) U-net type image reconstruction convolutional neural networks [60].....	69
Figure 33. Images of the digits 0 and 4: a-b) input pattern on the SLM, c-d) amplitude modulated output from the SLM, e-f) phase modulated output from the SLM, g-h) speckle patterns of each digit respectively for amplitude inputs and i) the difference between the speckle patterns g and h. [60]	70
Figure 34. Examples and accuracies of the reconstructed SLM input images from the recorded distal speckle intensity patterns for amplitude modulated proximal inputs [60].	71

Figure 35. Speckle pattern decorrelation. a) Frame at 0s, b) frame at 2s and c) the difference between the two frames [60].	72
Figure 36. Normalized confusion matrices for the classification of the reconstructed SLM input images for the 1 km GRIN fiber for a phase modulated proximal input a) from the speckle pattern images and b) from the reconstructed by U-net images. c) Classification accuracy for different fiber lengths (Solid line: The inputs to the VGG-CNN are the recorded speckles at the distal fiber end, Dotted line: The inputs to the VGG-CNN are the reconstructed images obtained by the U-net-CNN, Circles: Amplitude input images Squares: Phase input images) [60].	74
Figure 37. Training and validation classification accuracies as a function of epoch for the (a) 10m fiber distal speckle intensity pattern, (b) 10m fiber SLM reconstructed input, (c) 1km fiber distal speckle intensity pattern, and (d) 1km fiber SLM reconstructed input [60].	74
Figure 38. Optical set for the collection of the datasets consisting of fiber input-output image pairs.	75
Figure 39. Image of the SLM output on CCD2 for the digit 6 for a) amplitude and b) phase modulation. Image of the speckle pattern generated at the fiber output for input (b) when the reference beam is c) blocked and d) interfering with the output resulting in a hologram.	76
Figure 40. a) Amplitude modulated SLM output for the digit 8, b) hologram intensity of the fiber output for the input (a), c) amplitude and d) phase, e) real and f) imaginary part of the output field obtained after processing the hologram intensity (b) [66].	77
Figure 41. Cropped images of speckle patterns recorded on the camera CCD1 after downsampling to 32x32 pixels. The upper row of images corresponds to intensity-only recording while the lower row to holographic recording of the speckle output.	79
Figure 42. Image of a speckle pattern recorded at the fiber output.	80
Figure 43. Classification accuracy calculated for datasets collected at different laser wavelengths for phase modulated inputs.	83
Figure 44. a) Speckle intensity correlation as a function of wavelength drifting bandwidth, b) classification accuracy for datasets collected in for the two cases discussed in the text: DNN trained at 800nm wavelength and tested at data collected at different wavelengths (red circles) and DNN trained with data that include the perturbation (blue circles), c)	

classification accuracy for a DNN trained in the presence of 12nm bandwidth drift and tested at data collected outside this range.....	84
Figure 45. Normalized confusion matrices showing the classification results for the case that the drifting wavelength is a) not part of the training set and be) is included in the drifting wavelengths for which the VGG classifier is trained with [161,162].....	85
Figure 46. Classification accuracy of the VGG-type DNN used in the experiment for increasing the number of available samples in the training dataset.	86
Figure 47. U-net type DNN architecture used to reconstruct the input images in the fiber bundle from the intensity images at the fiber output.	91
Figure 48. MNIST database. a) Input image on the SLM and b) the image (a) relayed through the Fujikura fiber bundle. Scale bar: 50 μm	93
Figure 49. DNN reconstruction of MNIST dataset propagated through a fiber bundle for a-c) 32x32 and d-f) 64x64 image size.	94
Figure 50. Fiber bundle imaging of the distal image of the digit 6.....	94
Figure 51. Example of an image from the CIFAR-10 database.	96
Figure 52. Fiber bundle deconvolution of the distal image from the CIFAR-10 database.....	96
Figure 53. DNN reconstruction example from the CIFAR-10 dataset propagated through a fiber bundle.	97
Figure 54. DNN reconstruction example from the MNIST dataset propagated through a fiber bundle.	98
Figure 55. U-net DNN architecture for input images of 256x256 pixels size.....	100
Figure 56. Fiber bundle output images digit 0 for input wavelengths of a) 700 nm, b) 800 nm and c) 900 nm. (d-f) Corresponding 1D waveguide simulation results of the mode coupling. The inset shows magnified the area of overlap. Scale bar: 25 μm	101
Figure 57. U-net DNN reconstruction results for the three light sources for data acquired after bending the fiber bundle.	104
Figure 58. Experimental setup for collecting bright field images of tissue samples through the SCHOTT fiber bundle (red light path) and the digital holographic microscope used to obtain the phase information from the sample (green light path).....	106
Figure 59. Description of the physical and digital alignment between the proximal and distal images of the dataset.	108

Figure 60. GAN-type DNN used for the reconstruction of phase images from the intensity output of the fiber bundle.	110
Figure 61. a) Intensity-only image of a liver tissue sample, b) intensity image of the hologram of the same liver tissue (the red square shows a magnified a region of the hologram to visualize the interference fringes) and c) the extracted phase from (b).	111
Figure 62. DNN training image pair, a) fiber bundle output and b) its proximal phase extracted by the holographic setup at the proximal side. Scale bar: 10 μm	112
Figure 63. The reconstruction results for different learning parameters are presented. In all cases we consider i) the DNN input, ii) the ground truth phase image, iii) the reconstructed phase from (i), iv) and v) the maximum intensity gradient map after applying the Sobel operator at the phase images (ii) and (iii) respectively.	114
Figure 64. Generator and discriminator network losses over the epochs.	117
Figure 65. GAN reconstruction results in the presence of bending.	118

List of Tables

Table 1. A comparison of the optical parameters for the two MCFs [51]...	43
Table 2. Classification accuracy for the four different fiber lengths using amplitude or phase input patterns. Classification was carried out on either the intensity image of the distal speckle patterns or on the reconstructed SLM inputs [60].	73
Table 3. Classification accuracies obtained after the DNN is trained with different input images [59].	78
Table 4. Classification accuracies obtained after the DNN is trained with differently sized cropped portions of the initial speckle image recorded either using intensity-only or holographic measurement [59].	80
Table 5. Classification accuracies obtained after the DNN is trained by cropping different locations on the initial speckle image recorded either using intensity-only or holographic measurement [59].	81
Table 6. MNIST and CIFAR-10 reconstruction metrics for amplitude inputs.	95
Table 7. Image reconstruction results for inputs obtained from MNIST and MNIST-Fashion databases in the case of three light sources. The results are presented for test data obtained without moving the fiber and after the fiber is set in a new configuration to study the effects of bending in the performance of the U-net DNN.	103
Table 8. Phase reconstruction results through the fiber bundle endoscope for different inputs and DNN learning parameters.	116

List of Equations

Equation 1. Optical fiber numerical aperture.....	2
Equation 2. Electric field of the 1m fiber mode.....	3
Equation 3. V-number of an optical fiber	3
Equation 4. Number of propagation modes for a step-index optical fiber ...	3
Equation 5. Plane wave function	7
Equation 6. 2D Fourier transform formula	7
Equation 7. Coupling coefficients of an image to the fiber modes	9
Equation 8. The field of the input image at a distance z along the fiber	9
Equation 9. Sample field	19
Equation 10. Reference field	19
Equation 11. Interference pattern intensity.....	20
Equation 12. Off-axis reference field	21
Equation 13. Off-axis hologram intensity	21
Equation 14. Fourier transform of the hologram intensity	22
Equation 15. Theoretical minimum reference angle.....	22
Equation 16. Minimum reference angle considering Nyquist theorem.....	23
Equation 17: Transmission matrix	27
Equation 18. Intensity of the focused spot after the MCF at the output point m	37
Equation 19. Nonlinear polarization.....	38
Equation 20. Nonlinear wave equation	39
Equation 21. Phase-shift per core because of SPM.....	40
Equation 22. Intensity of the focused spot after the MCF at the output point m in the presence of SPM	40
Equation 23. Output of one node of a multilayer perceptron.	61
Equation 24. Perceptron weights evolution per iteration step.....	63
Equation 25. Mean squared error between two images	92
Equation 26. Formula for the calculation of the structural similarity index	92

Equation 27. Definition of the Sobel filter operator..... 113

Chapter 1 Introduction

Optical fibers are nowadays implemented in a wide range of applications because of their unique properties in guiding light and thus information between far spaced places. The most well-known field where optical fibers play a key role is telecommunications [1,2]. Apart from the revolutionary contribution of fiber cables in the information transmission, optical fibers also find applications in pressure and temperature sensing [3,4], fiber-based laser sources [5–8] and medical imaging [9–17].

The current work mainly investigates different kinds of fibers in the context of endoscopy and material manipulation. Optical fibers are integrated in medical procedures for many decades now, starting as illumination tools together with camera chips for visual inspection of areas inside the human body [18–22]. Many famous companies have contributed in the advances of commercial endoscopes such as Olympus, Pentax, Fujifilm, Karl Storz etc. In most of these endoscopes, the optical fiber delivers the light in the area of interest and a camera chip placed at the fiber end records the image. With the evolution of the fiber drawing techniques, new kinds of fibers were manufactured (e.g. fiber bundles, graded-index multimode fibers, photonics-crystal fibers), that upgraded the endoscopic role of the fibers further from simple illumination tools [23–26]. In particular, fiber bundles opened the path for all-fiber imaging. Thousands of single-mode fiber cores inside the fiber bundle core act as individual pixels, which collect and transfer the local information from the sample to a camera on the user’s side forming an image. Multimode fibers have also shown remarkable potential in fiber imaging using new optical technologies that correct for the scrambling of the information along the propagation [14,15,27–31]. Other types of fibers such as photonic-crystal fibers (PCFs) empowered the delivery of ultrashort pulses of high peak intensities which are sufficient for not only multiphoton imaging but also material modifications (laser ablation) [32–34]. More details on each fiber technology will follow in the next chapters of the present work.

This first introductory chapter (*Introduction*) aims to present the basic concepts of light propagation and imaging through optical fibers as well as our motivation for further research on the topic. The rest of thesis is divided in two parts, which present two different ways for optimal recovery of information through fiber bundles and multimode fibers. The Part A (*Wavefront shaping through fibers*) describes the use of wavefront-shaping for controlling the light propagation through the aforementioned types of optical fibers and particularly the transmission matrix method employed to perform endoscopic nonlinear imaging and laser ablation. The Part A consists of three chapters, which go through the principles of *Digital Holography* (Chapter 2), the *Transmission Matrix* (Chapter 3) and finishes with the experimental results on *High-intensity*

focusing through fibers (Chapter 4). In the Part B (*Deep learning in fiber endoscopy*), the scrambling of the information through the various optical fibers is reversed using a deep learning approach. The second part starts with an introductory chapter on *Deep Neural Networks* (Chapter 5) and continues with the demonstration of imaging through multimode fibers using deep learning (Chapter 6: *Seeing through multimode fibers using deep learning*). The last chapter of this part presents the realization of *Deep learning-enhanced imaging through fiber bundles* (Chapter 7). The thesis concludes with the *Summary* of the presented research work and the *Outlook*, which includes ideas for future studies.

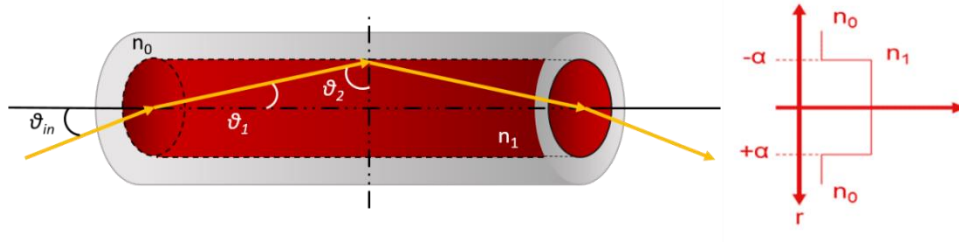


Figure 1. Light guiding in step-index optical fibers (left) and the refractive index profile along the fiber core (right).

1.1 Basic concepts in optical fibers

1.1.1 Light propagation through optical fibers

Optical fibers are dielectric waveguides characterized by cylindrical symmetry. Optical fibers consist of a high refractive index central core n_1 , in which the light is guided, and the cladding that surrounds the core made by a material with lower refractive index n_0 (Figure 1). The fiber types for which the radial distribution of the refractive index profile is constant in the core as shown in Figure 1 are called step-index fibers. As in slab waveguides, the light traps in the fiber core and propagates along its length as described by the principle of total internal reflection, which is caused by the refractive index difference between the fiber core and the cladding. However, only light waves with certain incoming angles θ_{in} for which the angle θ_2 is greater than the critical angle for total internal reflection of the fiber will be guided as depicted in Figure 1. The maximum input angle θ_{in} at which light is guided defines the numerical aperture (NA) of the fiber and it can be calculated based on the Snell's law of refraction. In order to have total internal reflection the minimum θ_2 should be equal to n_0/n_1 . Therefore, the fiber NA is given by Equation 1 and determines the acceptance cone of the optical fiber as it is often called.

$$NA = \sin \theta_{in,max} = n_1 \sin \theta_{1,max} = n_1 \cos \theta_{2,min} = n_1 \left(\sqrt{1 - \sin^2 \theta_{2,min}} \right) = n_1 \left(\sqrt{1 - \left(\frac{n_0}{n_1} \right)^2} \right)$$

$$\Rightarrow NA = \sqrt{n_1^2 - n_0^2}$$

Equation 1. Optical fiber numerical aperture

Maxwell's equations precisely determine the way that light propagates through an optical fiber. Because of the continuity of the electric and magnetic field at the boundaries between the core and the cladding the electromagnetic field equation needs to satisfy certain boundary conditions [35,36]. Consequently, there is only a finite number of incident angles θ_{in} within the acceptance cone of the fiber, which fulfil the boundary conditions. The different ways (angles) that the light can travel through the fiber are determined by the solutions of the Maxwell's equations in cylindrical coordinates and they are called fiber modes (more details can be found in the Appendix A1) and they propagate unchanged through the fiber length. Each fiber mode is characterized by a certain amplitude shape, phase profile and propagation constant (phase velocity). The Equation 2 describes the general electric field of a fiber mode E_{lm} , where $A_{lm}(r)$ is the amplitude distribution of the mode field, φ is the corresponding phase, β_{lm} is the propagation constant for this mode and ω is the angular frequency for monochromatic light. The numbers l, m are integers and they are related to the discrete solutions for the mode field obtained by the Maxwell's equations; m determines the order of the Bessel function that describes the amplitude distribution of the mode and it is called azimuthal index, while l is an integer that comes from the multiple solutions of the respective equations, representing the supported modes for a given fiber (Appendix A1) [36]. The Figure 2 shows the profile of three different propagation modes calculated for a 25 μm core radius fiber and an NA of 0.1; the first row depicts the intensity distribution of each mode field and the second the relative phase profile where white areas are of zero phase and gray areas of π phase.

$$E_{lm}(r, \varphi, z, t) = A_{lm}(r)e^{-im\varphi}e^{-i(\beta_{lm}z - \omega t)}$$

Equation 2. Electric field of the lm fiber mode

The total number of supported transverse propagation modes in the fiber is governed by the fiber parameter or as usually called the *V-number* of the fiber. The *V-number* is defined by the Equation 3 and depends on the physical characteristics of the optical fiber such as the NA , the core radius α as well as the wavelength of the incident light λ .

$$V = \alpha \frac{2\pi}{\lambda} NA$$

Equation 3. V-number of an optical fiber

From the mode analysis, it can be shown that for $V < 2.405$ the optical fiber supports only one spatial mode and this fiber type is called, single-mode fiber (SMF), while for larger V-numbers the optical fibers are called multimode fibers (MMFs). For step-index fibers that support a large number of propagation modes, the number of modes is approximated by Equation 4, which takes into consideration the degeneracy of the modes in the azimuthal term and the two polarizations.

$$N \approx \frac{4}{\pi^2} V^2$$

Equation 4. Number of propagation modes for a step-index optical fiber

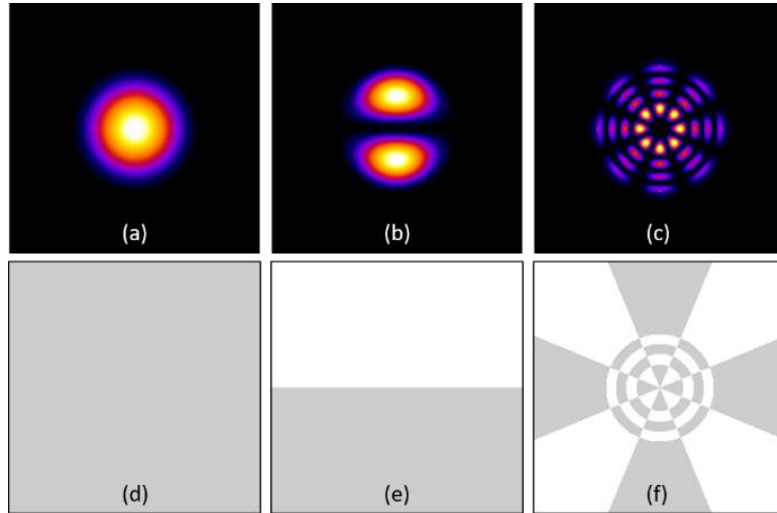


Figure 2. Intensity (first row) and phase (second row) profiles of fiber modes, LP01 mode (a,d), LP11 (b,e) and LP44 (c,f).

It is important to mention that the different propagation constants of the modes in step-index fibers can obscure the information delivery along the fiber. Travelling at different velocities along the fiber, the modes spread in time until they reach the fiber end, which would cause time broadening of a pulsed input light. Moreover, bending or slight defects of the refractive index profile along the fiber can induce mode coupling resulting in further mixing of the information encoded in each mode after propagation of some distance. Modal dispersion and coupling degrade the performance of step-index MMFs as imaging tools, as we will discuss further in the following sections.

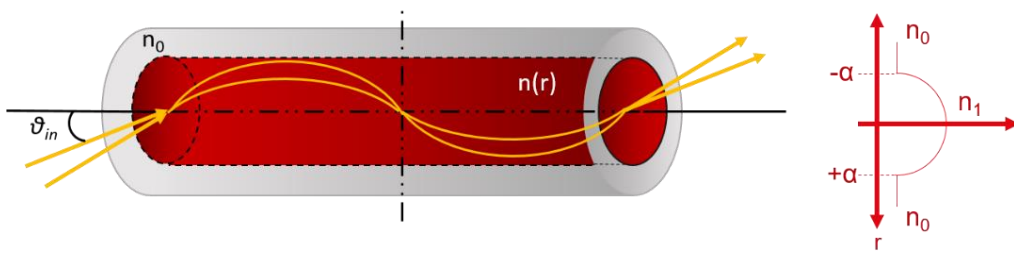


Figure 3. Light guiding in GRIN optical fibers (left) and the refractive index profile along the fiber core (right).

1.1.2 Optical fiber types

In the previous paragraph, we discussed about the way that light is guided in step-index optical fibers. Different fiber types became available with the advances of modern fiber drawing techniques. One example is graded-index (GRIN) fibers, for which the refractive index distribution of the core is designed in a parabolic shape to compensate the effect of modal dispersion (Figure 3). The refractive index profile is optimized in a way that the propagation constants β_{lm} of the

fiber modes are such that all the supported modes arrive at the distal end at the same time, implying zero modal dispersion for an ideal GRIN fiber [35,36]. Nevertheless, fabrication of a perfectly parabolic refractive index profile is not trivial and imperfections lead to a certain amount of modal dispersion for GRIN fibers too. Considering a quadratic refractive index profile GRIN fibers approximately support half of the number of modes N (Equation 4) compared to step-index MMFs for the same wavelength and refractive index difference between the cladding and the maximum refractive index value of the core [35,36]. The calculation of the spatial modes for the GRIN fibers is more complicated than for step-index fibers and more details can be found in the literature [35].

Another type of fibers commonly used in telecommunications and endoscopy is multicore fibers (MCFs) or fiber bundles as they are also called. MCFs consist of many small step-index cores densely packed in a common cladding or held together in a mesh, as they are schematically presented in the Figure 4. Each core of the bundle usually supports one mode at a certain wavelength range or few modes in others [23,26]. Depending on the core-to-core spacing, the MCFs can have minimal crosstalk and therefore transfer information independently in each channel/core [37–39]. Considerable effort has been dedicated to optimally design MCFs concerning arrangement and refractive index contrast of the individual step-index fiber cores to minimize the crosstalk [25,40]. The individual cores in a MCF can be either fixed in a common resin that works as the cladding material or they can each have a separate cladding and held together by a mesh. The first type is less flexible but of lower manufacturing cost, while the second type is highly flexible but expensive. The MCFs of the second type are usually called leached fiber bundles.

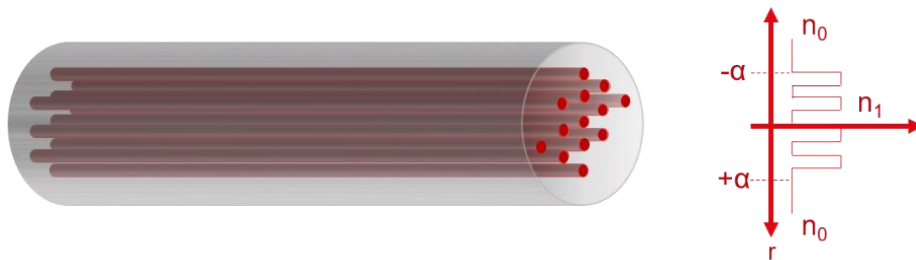


Figure 4. Fiber bundle scheme (left) and the refractive index profile along the bundle core (right).

Finally, it is important to mention another type of optical fibers, which guide the light in a different way than the solid core-cladding types of fibers discussed above. The light in these fibers propagates via the mechanisms of effective-index and photonic bandgap guidance and they are called photonic-crystal fibers (PCFs). The idea that light could be trapped in a central core by creating a Bragg like structure around it instead of a uniform cladding was already suggested in the early 70s [32,41–43]. That first proposed structure was generated by cylinders of lower and higher refractive index surrounding the central core creating a lower effective refractive index cladding.

Another example of effective-index guiding can be achieved by introducing wavelength-scale cylindrical holes in a pure silica-glass (Figure 5(a)), placed periodically around an unperturbed central area that acts as the fiber core. In this way, a lower effective index around the core is generated causing the light to trap in it. The cladding of air cylinders can also be regarded as a 2D periodic medium (photonic crystal) which is characterized by a dispersion diagram with photonic bandgaps which allow the guidance of selected optical frequencies. PCFs can be solid-core or hollow-core (Figure 5 (b-c)). In particular, hollow-core PCFs (HC-PCFs) show unsurpassed capability for delivering high peak power ultrashort pulses. Light guidance in an air core implies low losses, negligible pulse broadening and reduces the occurrence of nonlinear effects since air is characterized by orders of magnitude lower nonlinear response compared to silica [32–34,41–45]. HC-PCFs are thus our reference point for the studies of high peak intensity focusing through MCFs and MMFs as described in the following chapters. More details on the working principles of the PCFs is out of the scope of this thesis but can be found in the literature [32,41,43].

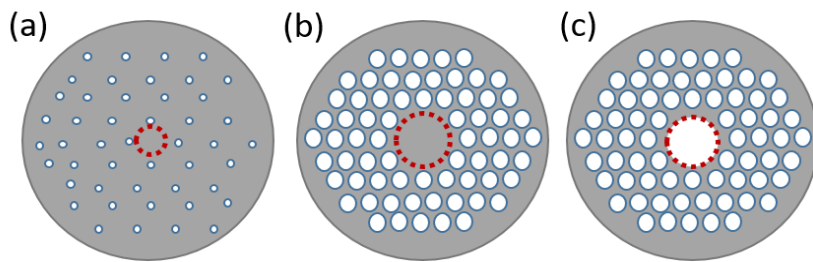


Figure 5. Other fiber types: a) Silica fiber with air-hole dopants, b) solid-core PCF and c) hollow-core PCF

1.1.3 Fiber imaging

Before describing the way that fiber imaging is accomplished through MMFs and MCFs, which is the main topic of the present work, it is important to underline some of the basic principles for the formation of an image through a linear optical system. An image is considered “good” when it preserves the information about the imaged object. Fourier analysis is widely used in signal processing for analyzing the information content of a digital signal. The basic idea of the Fourier analysis is that every signal can be expressed as a superposition of harmonic functions that are characterized by different frequencies. The same principle holds for an image, which is simply a 2D signal and thus the 2D Fourier transform can be used to decompose and study it [46]. The corresponding 2D harmonic functions to which an image can be decomposed using the Fourier analysis are described by the equations of plane waves propagating at different angles. The angles of propagation are called in this case spatial frequencies. Therefore, an ideal optical system does not introduce any loss of frequency components and forms an exact image of the object. However, because of the finite apertures of the optical systems, there is a limited bandwidth of frequencies that can be transmitted through and contribute to the final image formation. As a result, if the object is also described by frequencies outside this bandwidth, the final image misses this part of information and its quality is not optimal.

The complex field of a plane wave is given by Equation 5. The wavevector $\mathbf{k} = (k_x, k_y, k_z) = 2\pi/\lambda$ shows the direction of the propagation for a plane wave of light with wavelength λ . The k_x, k_y components of the wavevector are also known as transverse spatial frequencies or angular frequencies. For $k_x, k_y \ll k$, which implies small angles, we can consider $k_z \approx k$ and the final form of a plane wave is given as (Equation 5):

$$U(x, y, z) = Ae^{-j(k_x x + k_y y + k_z z)} = Ae^{-j(k_x x + k_y y + \sqrt{k^2 - k_x^2 - k_y^2} z)} = Ae^{-j(k_x x + k_y y)} e^{-jkz}$$

Equation 5. Plane wave function

Therefore, any complex field such an image $I(x, y)$ at a specific point in z can be described as a superposition of plane waves as described by the 2D Fourier transform (Equation 6). The frequencies f_x and f_y multiplied by 2π are equal to the corresponding spatial frequencies k_x, k_y . The amplitude components $F(f_x, f_y)$ of the Fourier transform describe the weight that each frequency pair (f_x, f_y) or (k_x, k_y) participates in the 2D signal/image formation. The frequency map $F(k_x, k_y)$ is known as the angular spectrum of the image/signal.

$$F\{I(x, y)\} = F(f_x, f_y) = \iint_{-\infty}^{+\infty} I(x, y) e^{-j2\pi(f_x x + f_y y)} dx dy = \iint_{-\infty}^{+\infty} I(x, y) e^{-j(k_x x + k_y y)} dx dy$$

Equation 6. 2D Fourier transform formula

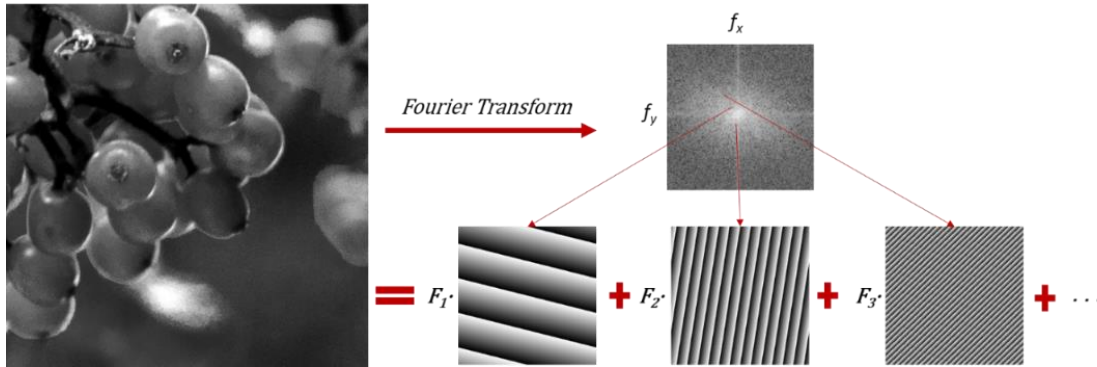


Figure 6. Image decomposition in plane waves via a 2D Fourier transform. The plane wave profiles shown correspond to the frequency components of the angular spectrum indicated by the red arrows.

A more intuitive explanation of the Fourier analysis of an image is depicted in Figure 6. We select an image and then we calculate its 2D Fourier transform, the amplitude of which we show in Figure 6. It can be seen that the image consists of different spatial frequencies, each one with a respective amplitude (F_1, F_2, F_3, \dots). By selecting different frequencies in the Fourier domain (f_x, f_y) , while setting the rest to zero each time and then calculate its inverse Fourier transform to observe the result in the spatial domain (x, y) , we get a field with the phase distribution of a blazed grating.

These fields correspond to plane waves propagating at certain angles. Therefore, it can be seen how an image is a superposition of 2D harmonic functions that correspond to plane waves.

Based on the brief Fourier analysis above, it is apparent that to get an image correctly propagated through an optical system we should be able to preserve all the spatial frequencies that it is composed of. However, as we mentioned before losses in the spatial frequency transmission of an image can be induced because of the non ideal nature of optical system, which results to “band-passing” of the frequency spectrum of the recorded object leading to degradation of the final image. The high spatial frequency components (large values in the Fourier domain) of an image correspond to the finest details (small features in the real space) of it and their absence will induce image blurring. As in the case of optical fibers discussed in paragraph 1.1.1, the number of angles that can be accepted and guided through is determined by the *NA* of the fiber. In a similar manner, all the imaging systems are characterized by a certain *NA* that allows only part of the spatial frequencies to pass through. Considering this point, in Figure 7, we observe how the image of the previous example (Figure 6) would appear after propagating through a system, which limits its spatial frequency content, by cutting off part of high-frequency components of the initial image. It is clear that the resolution of the image is lower and blurring of the image is induced. Consequently, it is now evident that the *NA* of an imaging system is a crucial parameter for the final resolution. Based on Rayleigh’s criterion the resolution of a system can be determined by the *NA* and the illumination wavelength as $\lambda/2NA$. In the following two paragraphs, we will translate the imaging analysis to the case of fiber endoscopes and we will specifically discuss the cases of two different fiber types used in the experiments presented in this thesis, the MMFs and MCFs.

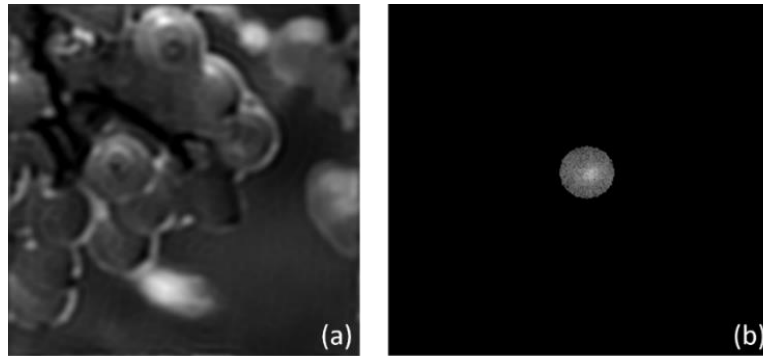


Figure 7. Image band-passing. a) Image result when limiting some of its high frequency components as shown in b)

1.1.3.1 MMF imaging

MMFs do not simply relay an image along the fiber length. An input image distributes among the supported spatial modes. The coupling of the image to the supported fiber modes is determined by the integral overlap between the image and the mode field. If we consider $z=0$ at the MMF entrance, then the coupling coefficients of an input image $I(r, \varphi)$ for the fiber modes E_{lm} (Equation 2) can be calculated as:

$$a_{lm}(z = 0) = \iint I(r, \varphi) E_{lm}^*(r, \varphi) dr d\varphi$$

Equation 7. Coupling coefficients of an image to the fiber modes

The image at a distance z along the fiber will be given as a linear combination of the supported fiber modes. Considering that each mode propagates with a different propagation constant the image $I(r, \varphi)$ at a point z along the fiber is:

$$I(r, \varphi, z) = \sum a_{lm}(z = 0) E_{lm}(r, \varphi) e^{-i\beta_{lm}z}$$

Equation 8. The field of the input image at a distance z along the fiber

Due to the different propagation constants of the fiber modes, the local information of the input decorrelates after few hundreds of micrometers after propagation in the MMF, forming a pattern at the fiber output that has a random intensity distribution. This interference pattern is called speckle and it is a result of the interference between the fiber modes emerging from the fiber as shown Figure 8. Therefore, a MMF completely scrambles the input image. Figure 8 was generated by calculating the fiber modes (Appendix A1) for a specific MMF and then decomposing the image to the available modes as described in Equation 8. The for different z values we calculate the speckle pattern by coherently adding the contribution by each excited mode. Based on the Equation 7 and Equation 8 the effect of image scrambling along an MMF could be undone by back propagating the output field. However possible for an ideal fiber, defects, bending or other perturbations affect the light propagation through the fiber inducing mode coupling and thus there is no deterministic way to process the output to recover the input. Nevertheless, MMFs are linear optical systems and the coupled information is not lost along propagation, but it can be recovered provided sufficient knowledge about the propagation through the system.

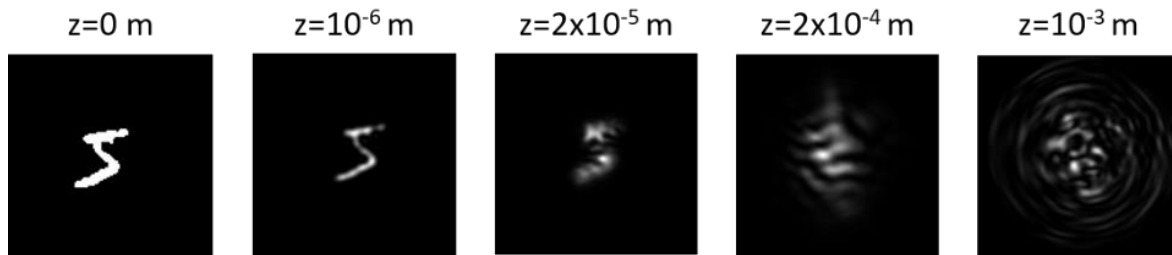


Figure 8. Image propagation through a MMF. Snapshots at different distances from the input facet.

Different optical techniques have been proposed to recover the information through MMFs. Some of the methods, such as phase conjugation and transmission matrix, are based on digital-holographic recording of the fiber output to recover the phase and the amplitude of the field aiming to calculate the desired wavefront to control the light through a MMF. It can be easily shown, that the complex conjugate field of the output is also solution to the propagation equation

through the system, meaning that if the output field propagates back through the system, forms again the input image. Spatial-light modulators can be used to shape the light and propagate the conjugate field back through an MMF so that the image will be translated at the other side. More details on the principles of digital holography as a wavefront shaping technique will be given in the next part of the thesis. Once the relationship between the input and output fields is known the MMF can be used as an imaging tool.

Other proposed methods to decipher the information through the scrambling caused by MMFs are based on deep learning and they are also discussed in the second part of the thesis. In these cases, an algorithm is trained with examples to learn the mapping between the input and the output of the MMF system without recording the full information (amplitude and phase) of the light field.

The resolution of the final image in a MMF endoscope is dependent on the *NA* of the fiber. This makes MMFs an attractive choice for ultrathin endoscopic tools since the *NA* depends only on the refractive index difference between core and cladding and it can be properly tuned for improved resolution while maintaining the size of the fiber core constant. Apart from the *NA*, the number of supported spatial modes is also an important parameter for the quality of the formed image. The fiber modes can be seen as the available degrees of freedom of the system that will contribute to the image formation and therefore they determine the signal-to-noise ratio (SNR) of the image [47–49]. Ultimately, a high *NA* MMF is capable to provide high resolution and SNR for an ultrathin endoscope. The details of fiber endoscopy through MMFs will be presented in the following chapters.

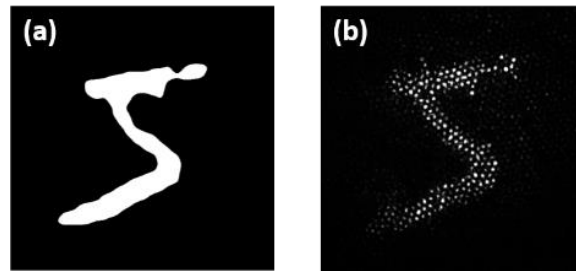


Figure 9. Image at the MCF a) input and b) output.

1.1.3.2 MCF imaging

As opposed to MMFs, MCFs are frequently used in endoscopic procedures because they can directly transfer images of a desired target illuminated with an incoherent light source without further processing. Each core of a MCF usually supports only one spatial mode (or few more depending on the wavelength) and acts as a pixel which captures the local information of the sample and delivers it to the far end of the fiber (Figure 9). As a result, the resolution in this case is defined by the core spacing, which is 5 μm on average for commercial endoscopes [17,23,24,26,38,50]. Reduction of the core spacing is limited because of the crosstalk which arises between neighboring cores. Light crosstalk in MCFs causes image blurring.

Consequently, the core spacing of the MCF imaging probes is chosen at a minimum value which is a compromise between resolution and core coupling. Higher resolution can be achieved by magnifying the object on the fiber facet by means of lenses attached at the fiber end, but it ultimately leads to a decrease of the field of view of the endoscope and an increase the final size. Another way to improve the resolution of the system is point scanning techniques using wavefront shaping to focus and scan the light through the MCF. In this way, the resolution will be given by the NA of each individual core. The analog of the modes for the MMF is the number of individual cores, which determines the total degrees of freedom available for shaping the light through an MCF. Consequently, the SBR of the focus spot will be proportional to the number of the available cores [37,51–58]. Focusing through an MCF using wavefront shaping is further discussed in the following chapters using the transmission matrix method.

Apart from wavefront shaping using digital-phase conjugation techniques, deep learning can also achieve high resolution imaging through an MCF imaging probe as in the case of MMFs. The pixelation artefact in the final image delivered by MCFs can be also overcome by properly training neural networks on desired datasets [59–66]. Both wavefront shaping and deep learning methods for fiber imaging through both fiber types are described in this thesis in the Parts A and B respectively.

1.2 State-of-the-art in fiber endoscopy

The concept of fiber endoscopy to create an image from confined places inside the body is considered a well-established medical approach of the 21st century, known to assist not only diagnosis but also treatment of various pathologies. It is interesting though to mention that the concept of endoscopy goes few centuries back, which underlines its importance in medicine.

In the early 19th century, the first “light guide”, which could transfer images from cavities inside the human body is reported by Bozzini. In this first demonstration the endoscope consisted of an open tube and an eye piece on the one side [20]. Bozzini is considered the first to separate the illumination from the sample path for the image formation using candle light as a source and a mirror in the visual tract. His invention was disapproved and dismissed by the medical community back then, but found some solid support by the military hospitals. However, more and more experts realized the importance of having endoscopic access for inspection of pathologies and the attempts to improve the endoscopes in terms of illumination, size and image quality continued for many years after Bozzini’s preliminary invention. Nitze’s pivoting work in 1877 led to the first cystoscope, which consisted of small lenses placed at certain intervals inside a rigid tube for relaying the image from the sample to the eyepiece. Hopkins improved the idea of the relay lenses based endoscope in the 20th century by substituting them with glass rods. The working principles of these early prototypes of subsequent image relays has been adopted by many endoscope designs up to nowadays. Karl Storz, who started one of the leading endoscopic companies, highly

valued to today, brought Nitze's cystoscope to industrial manufacturing. Many more scientists (too many to mention in this thesis) worked to develop endoscopy and bring it on the level existing today [11,19,21,22,24]. Image quality, flexibility and small size have been the primary parameters to optimize since these early works alongside with improvement of the final image recording devices.

Nowadays, the evolution of optical components, fibers and camera chips as well as the advances in microscopy methods have revolutionized the capabilities of commercial endoscopes. Coherent fiber bundles ushered flexible endoscopy in clinical practice. As briefly discussed before, coherent fiber bundles consist of many single-mode fibers (SMFs) that probe the light from the sample locally and transfer the scattered light from one side to the other of the endoscope forming an image. The imaging principle is the same of camera; each core corresponds to a pixel. One of the most discussed drawbacks of the fiber bundles is the limited resolution, which is determined by the core-to-core spacing, which is as aforementioned in the order of 5 μm . At the expense of smaller field of view the resolution can be improved using lenses at the distal side [26,50,67–69]. Wavefront shaping to focus the light through MCFs so that point scanning microscopy techniques can be applied has been reported to improve the imaging resolution [51,52,55,55,70]. The size of focused spot generated through the MCF depends on the individual core NA and not the core spacing and as a result, the resolution can be much higher. Alternatively, digital methods have also been proposed for removing the pixelated look from the images using computational imaging means such as compressive sensing and deep learning [71–73].

Other endoscope configurations combine a miniature camera chip placed at the distal tip of the flexible part and a SMF or a fiber bundle for illumination of the area of interest. The implementation of a housing part to hold the components at the distal end of the endoscope results in an average size of 10 mm for the commercial ones (by Pentax, Fujifilm, Olympus). A turning point in clinical medicine is currently expected by the development of scanning fiber endoscopes [17]. This kind of fiber endoscopes can be as small as 1-2 mm and consist of a rigid head that contains the end tip of a SMF and the optical lenses for high resolution as well as a piezo scanning mechanism to move the SMF core, so that the light spot scans the desired area of the sample. Scanning fiber endoscopes can be also combined with hollow-core photonics crystal fibers for high power delivery of ultrashort pulses opening new possibilities for endoscopic surgery [33,41,44]. The resolution of these endoscopes is limited by the low NA of the SMF and the focusing lens system. Tighter focusing of the beam implies diffraction of the fiber mode at longer distances, which imposes larger lens diameters and an increase of the final endoscope size.

While SMFs and MCFs are often used in the design of commercial endoscopes as mentioned in the previous paragraph, that is not the case for MMFs because of the image distortion after propagation along the MMF length. However, the information capacity of a MMF can be extremely high, while the size of the core remains few hundreds of micrometers. Therefore, image transmission through a MMF has been an area of research interest from many decades now. Assuming that the scrambling of the information along the fiber can be compensated, high

resolution endoscopes with only few hundred micrometers diameter can be realized. Various techniques based on either wavefront shaping [15,28–30,51,52,56,74–78] or computational recovery using various image processing algorithms, among which deep neural networks (DNNs) [59–64,66], have been employed showing promising results for MMF endoscopy.

Wavefront shaping is a well-known method originally used to control the propagation of light through scattering media so that the information is recovered [47–49,77,79–81]. Light control takes place in various areas in optics and acoustics such as in astrophysics for optimizing the image quality that is diminished when light travels through the atmosphere or in biomedical imaging of tissue samples, for which imaging in depth is hindered by the scattering of the light caused by tissues. The idea of controlling the light specifically through optical fibers was introduced in the 1960s. Optical phase conjugation was used for the first time to transmit an image through a MMF using holographic plates and nonlinear crystals [82–84]. The revolution on the field came with implementation of digital holography. In this case, the holograms are stored as an image in a computer and the light is modulated using spatial light modulators that consist of pixels that can locally assign a specific phase to the incoming wavefront related to the pixel value chosen by the user. Hence, focusing the light through a MMF for point scanning microscopy can be achieved by properly shaping the light input to the MMF. Iterative optimization algorithms or digital-holographic methods such as digital-phase conjugation and transmission matrix can be used for the calculation of the optimized wavefront to create a desired pattern or to focus the light at the fiber distal side [14,28,30,31,52,56,57,74,76,79,85]. Impressively MMFs can perform most of the leading microscopy techniques ranging from linear methods: fluorescence [76,86,87], photoacoustic [88,89], reflection [90], confocal [67,75], speckle scanning [91,92], to nonlinear optical methods which require higher power levels: two-photon fluorescence [13,15,29,52,69,87], coherent anti-Stokes Raman scattering [93] etc.

Image recovery through MMFs has been also investigated using purely computational means. Two early reports show that imaging is possible through MMFs without holographic methods but by training artificial neural network (ANN) algorithms to reconstruct the desired fiber input from its corresponding speckle intensity image recorded on a camera [94,95]. The ANN in this case, receives the intensity image of the speckle pattern as an input and it is trained to reconstruct the input image using backpropagation of the error between the reconstructed image and its ground truth. These early demonstrations used simple ANN architectures to decipher binary input patterns of handwritten digits from the corresponding speckle. The evolution of graphic units and computing memories opened the path for more complex problem solving using computational imaging. More complex ANNs can be realized to recover signals of high information content while handling large amounts of data for training [96–99]. Modern ANNs have shown impressive capabilities in image classification and reconstruction after kilometer long fibers [60,61,65,100] (see Part A: Chapter 6) as well as pattern projection through MMFs [61,101].

Each of the methods mentioned above, wavefront shaping and ANNs, offer certain advantages in imaging through MMFs and MCFs that we will be discussed further in this thesis and each can be selected based on the application and the imaging problem under study.

1.3 Motivation

Conventional microscopy can provide high-resolution images of biological samples but it is restricted in accessing information about the sample at depths in the centimeter range. On the other hand, optical fibers can recover information from confined places without fundamental depth limitations. Fast, accurate and non-invasive diagnosis has always been the main target of fiber endoscopy. The purpose of this thesis is to further investigate two main topics: the implementation of ultrathin endoscopes that not only accommodate imaging modalities but also deliver high enough power to perform local sample modifications in the same endoscopic tool and secondly, the improvement of the information quality of existing endoscopes and fibers using deep neural networks.

The term multimodal endoscope generally refers to a probe that can create multichannel images using different microscopy techniques (e.g. fluorescence, Raman, confocal etc.). Apart from diagnosis, fiber probes have been studied for delivering high laser power to perform local treatment of tissues [102,103]. Material ablation using laser pulses usually implies high peak power delivery and poses limitations for the conventional imaging endoscopes because of their low damage threshold. Therefore, in laparoscopic processes endoscopic imaging and laser treatment are separately performed by different tools. Following though the clinical trends for minimally invasive processes and faster recovery times for the patients, we propose a single endoscopic tool based on MMFs, which can provide a high-peak power focus spot, aiming for selective surgical modification of tissue, while obtaining images of the desired area with cellular resolution, using wavefront shaping techniques. Femtosecond laser ablation (FLA) is a well-known technique for high precision modifications with minimal collateral damage, which attracts significant interest various medical fields such as ophthalmology [104,105], but also in industry for micro-manufacturing purposes [106,107]. At the same time, resolution beyond the diffraction limit, high sectioning and improved penetration depth are characteristics that make two-photon fluorescence (TPF) microscopy a valuable tool for biological (and not only) sample inspection. Consequently, combination of the two methods in an endoscopic probe would offer important advantages for diagnosis and micro-surgery. The potential and limitations of a fiber endoscope that combines FLA for altering the sample locally and TPF imaging described in detail in the Chapter 4 of the Part A of the thesis.

Wavefront shaping for focusing the light through a MMF or MCF is calibration sensitive and the proposed endoscope needs to be designed in a rigid way. To overcome this obstacle the potential use of computational means is investigated for their potential robustness against perturbations of the system. Specifically, Deep Neural Networks (DNNs) are implemented and trained with examples of experimentally measured datasets to recover the information through MMFs. DNNs have presented many advances in conventional microscopy, which make them a promising solution for fiber endoscopy. Super-resolution, translation of information among microscopy modalities, robustness to distortions can be potentially integrated in fiber endoscopy via deep learning. Therefore, DNNs are studied aiming to demonstrate a multifunctional platform that digitally-only improves the imaging performance of fiber endoscopes. These advantages can lead

to digital and thus cost efficient upgrade of the existing probes instead of optically redesigning an endoscope from scratch.

PART A:

Wavefront-shaping in fiber endoscopy

Chapter 2 Digital holography

The word holography comes from the Greek word “Ολογραφία” which consists of two parts “όλος=holos” + “γραφή=graphy” that respectively mean total recording. The term was first proposed by Dennis Gabor in 1948 to describe a lensless imaging method, which records the total field information, amplitude and phase, by interfering the field scattered from an object with a reference coherent field [46]. Gabor called this interference pattern a *hologram* and received the Nobel Prize in Physics the 1971 for his revolutionary idea. Holography did not receive a lot of attention in Gabor’s time; it was only in 1960s that E.N. Leith and J. Upatnieks first applied and improved the idea of Gabor with their offset-reference hologram. At the same time, the invention of the laser provided the necessary coherence for the illumination source that revealed the full potential of holography to capture the full field information while using intensity-only recording media [46].

2.1 Principles of holography

Holography is an interferometric technique, which implies the summation of two fields (one unknown field and a reference field) at a defined plane in space. The interference pattern (hologram) is an intensity pattern that contains the phase and amplitude information of the unknown field. In a mathematical way, we can consider $S(x, y)$ the field coming from a sample of interest and $R(x, y)$ the known reference field, which are described as follows (Equation 9, Equation 10):

$$S(x, y) = |S(x, y)|\exp[j\varphi_S(x, y)]$$

Equation 9. Sample field

$$R(x, y) = |R(x, y)|\exp[j\varphi_R(x, y)]$$

Equation 10. Reference field

where (x, y) are the coordinates in space at the interference plane and $\varphi_S(x, y)$, $\varphi_R(x, y)$ are the phase distributions related to the sample and reference fields respectively.

The interference between the two fields results in an intensity pattern $I(x, y)$ that contains several terms related to the amplitude and phase information of the two fields:

$$\begin{aligned}
I(x, y) &= |S(x, y) + R(x, y)|^2 \\
&= [S(x, y) + R(x, y)][S(x, y) + R(x, y)]^* \\
&= |S(x, y)|^2 + |R(x, y)|^2 + S(x, y)R(x, y)^* + R(x, y)S(x, y)^*
\end{aligned}$$

Equation 11. Interference pattern intensity

In Equation 11, the first two terms are only related to the intensities of the two fields, while the last two terms also include information about the amplitude and phase of the sample and the reference fields. Once the intensity of the hologram is recorded (Figure 10a), we can remove the object and recreate its image at its initial position by illuminating the hologram. This step is often called as hologram readout (Figure 10b) since the illumination beam reveals the information content of the recorded hologram. The holograms created by a reference beam propagating in-line with the object wavefront are called Gabor holograms. In this thesis, we choose to use off-axis holography (also known as Leith-Upatnieks holography) in which the reference beam propagates at an angle relative to the direction of propagation of the object field as illustrated in Figure 10a. The Gabor hologram poses certain limitations due to the in-line reference, which creates twin images during reconstruction. In contrast, off-axis holography generates a clearly separated image of the object which does not suffer from the high background originating from the non-modulated light [46]. Off-axis holography will be further discussed in the following paragraphs.

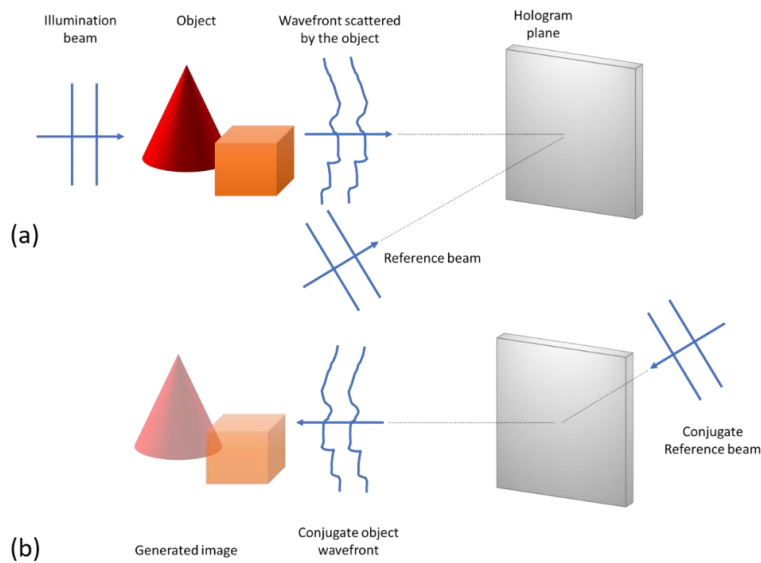


Figure 10. Holographic imaging. a) Hologram recording and b) hologram readout for image generation by illuminating the hologram with the conjugate reference beam

2.2 Digital holography

In this thesis, off-axis digital holography is used for recording the full field information (amplitude and phase) at the output of an optical fiber aiming to control the light propagating through. The

digital hologram is generated by the interference between the object (here the fiber output field) and the reference field on a camera detector. After the hologram is recorded, the phase and the amplitude can be digitally extracted. The recorded field can be then displayed using devices called spatial light modulators (SLMs). More information about the SLMs will be presented in the next paragraph. In this paragraph, we discuss the practical aspects that should be considered in digital holography related to the digital sampling due to the finite number of available pixels on the sensor, the pixel size, as well as the optimal reference angle for sufficiently recording the information in the digital off-axis hologram.

2.2.1 Minimum reference angle

We consider for simplicity the reference field to be a plane propagating at an angle $\theta=(\theta_x, \theta_y)$ with respect to the propagation direction of the object field. If $R(x, y)$ is the reference field and $k_0 = 2\pi/\lambda$ the amplitude of the wavevector in free space for a source emitting at wavelength λ , we can express the reference beam as a function of the transverse spatial frequencies, k'_x, k'_y , that determine the angle of propagation (Equation 12). For a plane wave reference R is a constant.

$$R(x, y) = R \exp[j(k_0 \sin \theta_x x + k_0 \sin \theta_y y)] = R \exp[j(k'_x x + k'_y y)]$$

Equation 12. Off-axis reference field

Substituting Equation 12 in the hologram equation (Equation 11), the intensity distribution on the camera sensor can be expressed as follows:

$$I(x, y) = \underbrace{|S(x, y)|^2 + |R|^2}_{DC \text{ term}} + \underbrace{S(x, y)R^* \exp[-j(k'_x x + k'_y y)]}_{Object \text{ term}} + \underbrace{R \exp[j(k'_x x + k'_y y)] S(x, y)^*}_{Conjugate \text{ object term}}$$

Equation 13. Off-axis hologram intensity

The first two terms of the Equation 13 propagate with zero angle with respect to the optical axis when the hologram is normally illuminated and contain no information concerning the phase component of the object field and they will be referred as DC component of the hologram. The last two terms are complex conjugates of one another and they contain information for both the amplitude and the phase of the field related to the object. In the case of using normal illumination for the hologram readout, the DC term propagates along the optical axis, while the two conjugate images of the recorded object deflect at angles θ and $-\theta$ with respect to the optical axis as a result of the exponential factor originating from the reference field. Here it is apparent now that the reference angle used needs to be optimized in order to have these three components well separated so that the object is nicely formed without interfering with part of the DC term.

As discussed in Chapter 1, each image is characterized by a specific frequency spectrum in the Fourier space, which should be taken into account in the calculation of the minimum reference

angle to avoid overlap between the hologram orders. The angular frequency spectrum of the recorded hologram is calculated by the Fourier transform of the hologram image (intensity profile). Assuming that the reference beam is a plane wave travelling at an angle, the respective Fourier transform is a δ function in Fourier space at the point $(k'_x = k_0 \sin \theta_x, k'_y = k_0 \sin \theta_y)$.

$$\begin{aligned}
 F\{I(x, y)\}(k_x, k_y) &= F\{|S(x, y)|^2\}(k_x, k_y) + F\{|R|^2\}(k_x, k_y) + \dots \\
 &\quad + F\{S(x, y)R^* \exp[-j(k'_x x + k'_y y)]\}(k_x, k_y) + \dots \\
 &\quad + F\{R \exp[j(k'_x x + k'_y y)]S(x, y)^*\}(k_x, k_y) \\
 &= F\{S(x, y)\}(k_x, k_y) \star F\{S(x, y)\}(k_x, k_y) + \delta(k_x, k_y) + \dots \\
 &\quad + R^* F\{S(x, y)\}(k_x + k'_x, k_y + k'_y) + RF\{S(x, y)^*\}(k_x - k'_x, k_y - k'_y)
 \end{aligned}$$

Equation 14. Fourier transform of the hologram intensity

Equation 14 gives the spatial frequency distribution of the hologram with respect to the reference and object field information. The first term comes from the autocorrelation theorem and is related to the object's frequency spectrum while the second term is a δ function related to the reference intensity as a plane wave. The last two terms are proportional to the object frequency spectrum shifted in frequency space because of the multiplication with the reference beam. The frequency shift of the object's spatial frequency spectrum is dependent on the reference angle. Figure 11 depicts the terms of Equation 14 to show more comprehensively how the reference angle needs to be determined.

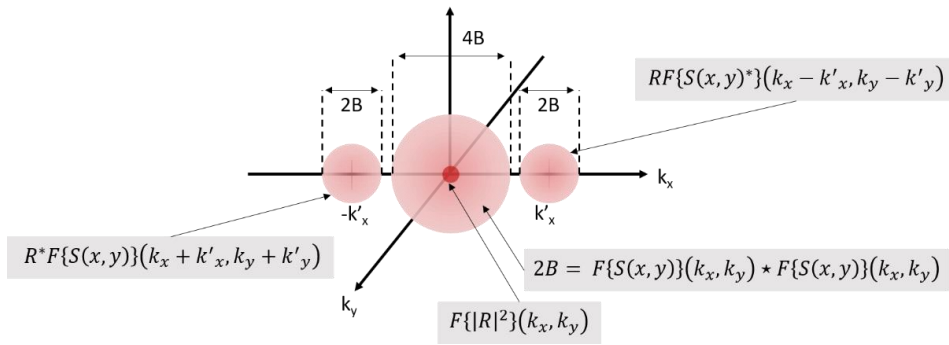


Figure 11. Fourier spectral components of the hologram

Without loss of generalization, k'_y is considered zero in Figure 11 and B is the half-bandwidth of the object's angular spectrum. The minimum $k'_{x,min}$ to avoid overlap between the DC component and the terms related to the object field should be $3B$ as it can be easily seen in Figure 11. Analyzing this further, we can determine the minimum reference angle as shown in Equation 15:

$$\begin{aligned}
 k'_{x,min} &= 3B \\
 k_0 \sin \theta_{min} &= 3B \\
 \sin \theta_{min} &= \frac{3B\lambda}{2\pi}
 \end{aligned}$$

Equation 15. Theoretical minimum reference angle

2.2.2 Digital sampling

When recording a hologram on a 2D detector of finite number of pixels (N_x, N_y) with pixel size d_{pix} , the frequencies of the object in the Fourier space must be correctly sampled based on the Nyquist theorem. Therefore, in order to calculate the experimentally implemented angle between the reference and the object, the effect of sampling effect must also be considered. Based on the pixel size, the frequencies that can be sampled in the 2D Fourier space according to the Nyquist theorem lie in the range $\left[\frac{-\pi}{d_{pix}}: \frac{\pi}{d_{pix}}, \frac{-\pi}{d_{pix}}: \frac{\pi}{d_{pix}}\right]$. To maximize use of the 2D space (from a geometrical perspective), we consider that the two conjugate orders lie in the diagonal of the 2D window. If we now consider the bandwidth of the object B , we can calculate the minimum angle θ_{min} which ensures both the separation of the DC term from the phase carrying terms (Equation 15) and the correct digital sampling using the Equation 16. The factor $\frac{3\sqrt{2}+2}{2}$ comes from the geometrical calculation of the diagonal with respect to the bandwidth B . According to the Equation 16, if we consider a detector of 8 μm pixel size and a wavelength of 1030 nm, which is the case for the experiments described in the following chapters, we get a minimum angle of 3.5°.

$$\begin{aligned} \frac{\pi}{d_{pix}} &= \left(\frac{3 + \sqrt{2}}{\sqrt{2}}\right) B \xrightarrow{\text{Equation 12}} \\ \frac{\pi}{d_{pix}} &= \left(\frac{3\sqrt{2} + 2}{2}\right) \left(\frac{2\pi}{3\lambda} \sin \theta_{min}\right) \Rightarrow \\ \sin \theta_{min} &= \frac{3\lambda}{(3\sqrt{2} + 2)d_{pix}} \end{aligned}$$

Equation 16. Minimum reference angle considering Nyquist theorem

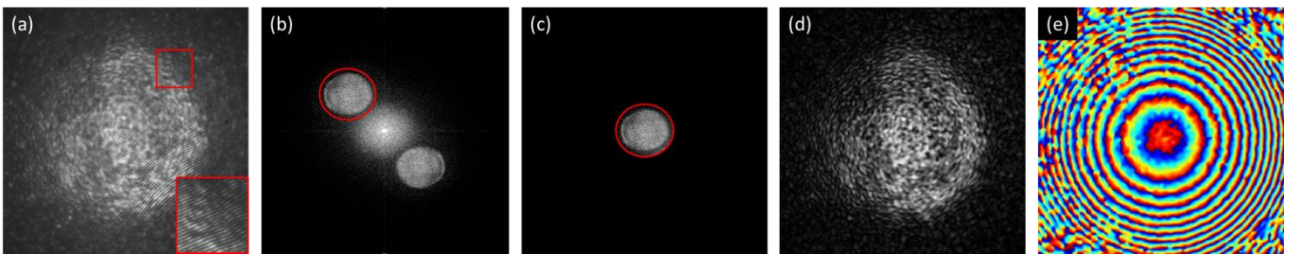


Figure 12. a) Digital hologram recorded on the camera (the inset shows a magnified area of the hologram to observe the fringes), b) Fourier transform of the hologram (a), c) the +1 order of the hologram cropped and shifted in the center to remove the reference angle contribution in the phase, d) the amplitude and e) the phase of the object field obtained by calculating the inverse Fourier transform of (c).

2.2.3 Reconstruction of the complex field

Digital holography has been extensively used in various areas of optics. In microscopy, digital holographic microscopes are able to generate images of transparent objects that are basically invisible in conventional bright-field microscopy. A further development of digital holography is

optical diffraction tomography, which uses digital holography to recover the scattered field from an object after illumination at different angles, generating the 3D refractive index distribution of the sample [108,109]. MMF imaging utilizes digital holography to characterize the propagation of light through the fiber by measuring its transmission matrix, which relates the input to the output field. Specifically, for point scanning based imaging, the light through the fiber needs to be focused at a certain distance from the distal fiber facet. Figure 12 provides an example extraction of the amplitude and phase from the hologram of a speckle pattern (Figure 12a) recorded in one of the experiments discussed in the following chapters based on the reconstruction steps below. In this case, the speckle hologram is recorded at the plane of the distal facet of a MMF. This example is obtained when focusing through a MMF, which is the main target of the work presented in the first part of the thesis (Figure 12). Therefore, a quadratic phase profile is generated at the distal facet of the MMF as shown in Figure 12e. The extraction of the phase and amplitude of the digital hologram is performed using the following steps:

1. Compute the Fast Fourier Transform (FFT) (Figure 12b) of the hologram image (Figure 12a).
2. Select the desired diffraction order of the FFT (red circle, Figure 12b).
3. Mask out all of the unwanted orders of the FFT (DC and conjugate order).
4. Shift the real order at the center of the Fourier space to remove the angle contribution in the phase coming from the reference beam (Figure 12c). This step is equivalent to illuminating the hologram with the conjugate reference, which would also result in normal propagation of the object term with respect to the optical axis (Equation 13) since the angle dependence would vanish.
5. Apply an inverse FFT to obtain the amplitude and phase of the field are obtained (Figure 12d, e respectively).

2.3 Spatial light modulators

In the discussion above, we analyzed the correct way to record the field of an unknown object $S(x, y)$ on a camera sensor with finite pixel number and pixel size using digital holography. How can we now use the recorded field in order to display the image of the object from the information extracted from the digital hologram? Historically photographic emulsions were used for holographic image projection. However the use of emulsions was severely limiting due to the long delays in the chemical processing and poor repeatability. The development of devices capable of modulating the phase and/or amplitude profile of an incident coherent wavefront, called spatial light modulators (SLMs), revolutionized the field of digital holography and displays. The different technologies are used for SLM devices including liquid crystal SLMs (LC-SLMs), magneto-optic SLMs, deformable mirror-based SLMs, digital micro-mirror devices (DMDs) and more [46,110]. Furthermore, new technologies for light modulators are emerging aiming for modulating devices at high speed and spatial resolution compared to the existing ones. One example are metasurface-

based light modulators for their potential of subwavelength pixel resolution and possibly unsurpassed speed [111].

For the work presented in this thesis, phase-only LC-SLMs are used. LC-SLMs can be designed to function in transmission or reflection depending on the needs of each experiment. The working principle of the LC-SLMs is based on the electrically induced orientation of the birefringent liquid crystal molecules contained in each pixel of the device. By applying a different electrical signal to each pixel of the liquid crystal screen, an incident wavefront experiences local phase retardation dependent on the orientation of the molecules in each pixel, which finally creates a desired modified output. By projecting, the recorded field information (amplitude and/or phase) of an object on the SLM array and illuminating it at the correct angle of incidence, the image of the object will appear at a distance far from the SLM that corresponds to the initial position of the object. An LC-SLM device always needs to be calibrated for the wavelength used so that the electric signal applied in each pixel results in the desired phase value.

LC-SLMs offer high resolution in terms of total pixel number (Full HD) and at least 8-bit dynamic range which favors the precise generation of complex phase profiles and linear gradients. In addition, they are characterized by high diffraction efficiency (modulated to unmodulated light ratio) which is crucial for high-power applications such as nonlinear imaging and laser ablation demonstrated in this thesis. Furthermore, amplitude information can be also displayed in these phase-only SLMs at the expense of diffraction efficiency [112]. One disadvantage of LC-SLMs is their low refresh rate because of the liquid crystal orientation time. The speed for displaying phase patterns ranges from 20Hz up to 700Hz. For low-power applications requiring high-speed pattern projection, DMDs can be used instead as they function in the kHz range.

2.4 Summary

In this chapter we explained the basic concepts of digital holography. The recording of a hologram corresponding to an unknown field on a camera sensor was discussed for the specific case of the off-axis holography. The determination of the optimal angle of the reference beam, which is critical for the optimal recovery of information related to the object field, is analyzed. In addition, it is shown that the correct digital sampling of the hologram on a finite pixel sensor is important for preserving the information. SLM devices are an essential tool in digital holography for light control applications such as wavefront shaping through optical fibers, which is investigated in this thesis. SLMs allow for the recorded information in a digital hologram to be afterwards displayed and generate the wavefront of an unknown object field at a certain location in space. Therefore, SLMs allow precise control of the light propagation through optical systems by shaping the light beam in a desired way. In the next chapters, the advantages of digital holography for wavefront shaping applications in microscopy and material manipulation are presented.

Chapter 3 Transmission Matrix

In this chapter, the basic theory behind the transmission matrix method is discussed. The transmission matrix of an optical system is not a new concept; it has been used to describe the propagation of light through simple or more complicated optical systems for many decades. The focus of the following section is the implementation of the method for modulating the light propagation through optical fibers. Part of the results presented can be also accessed in the following published articles in accordance with the publisher's agreement:

- D. B. Conkey, E. Kakkava, T. Lanvin, D. Loterie, N. Stasio, E. Morales-Delgado, C. Moser, and D. Psaltis, *"High power, ultrashort pulse control through a multi-core fiber for ablation"*, Optics Express 25, 11491 (2017)
- E. Kakkava, M. Romito, D. B. Conkey, D. Loterie, K. M. Stankovic, C. Moser, and D. Psaltis, *"Selective femtosecond laser ablation via two-photon fluorescence imaging through a multimode fiber"*, Biomed. Opt. Express, BOE 10, 423–433 (2019).

3.1 Introduction

3.1.1 Theoretical aspects

In acoustics and optics the “transmission” or “transfer matrix” is a subpart of the full scattering matrix, which is used to describe the propagation of acoustic and electromagnetic waves, respectively, through scattering media [47,113–115]. In this study, the transmission matrix theory will be used to study the light propagation through multimode fibers. Multimode fibers differ from scattering materials since no reflections and backscattering are taken into account and therefore, the transmission part of the full scattering matrix can sufficiently describe the light propagation through the fiber.

The transmission matrix T of the fiber relates an input complex field E_{in} to a corresponding output complex field E_{out} as described by the following equation.

$$E_{out} = T \cdot E_{in}$$

Equation 17: Transmission matrix

Theoretically, the conservation of energy dictates that the transmission matrix is a unitary matrix. However, experimental losses lead to an incomplete matrix, which is not strictly unitary. These losses are related to the optical system and could originate from fiber inhomogeneities, bending losses, etc. In addition, measurement of a single polarization (for experimental simplicity) introduces further loss in the measured transmission matrix. Another interesting property of the transmission matrix is reciprocity, which implies that the transmission matrix towards one direction $T_{1 \rightarrow 2}$ is the non conjugate transpose of the transmission matrix towards the opposite direction $T_{2 \rightarrow 1}^T$.

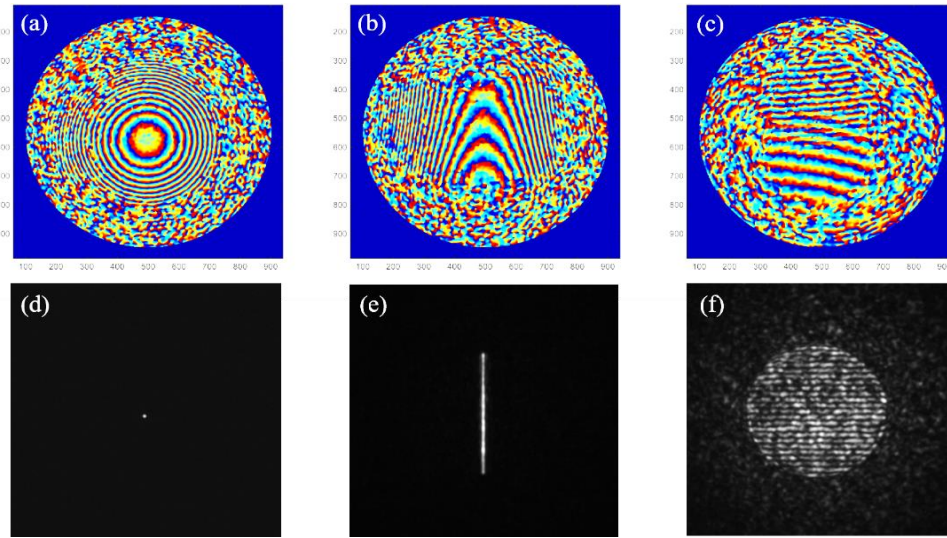


Figure 13. Visualization of the generation of patterns through the GRIN fiber using the measured TM of the optical system. a-c) The wavefront at the distal side of the fiber that results in the d-f) intensity distributions at 200μm distance from the fiber facet.

For many applications discussed in this thesis (e.g. nonlinear imaging, laser ablation), it is important to focus the light at the distal fiber end. Therefore, it is necessary to calculate the correct wavefront, which after propagation through the optical fiber, will render the desired quadratic phase at the distal fiber facet to create a focus spot at a chosen location in space. To do so, the inverse of the transmission matrix T^{-1} is required. However, the calculation of the algebraic inverse of the experimentally measured transmission matrix is often prone to measurement noise. Instead, the phase conjugate matrix is used. Phase conjugation refers to the backward propagation of light with respect to the optical axis z [14,52]. As discussed in the previous paragraph, because of the unitary property of the transmission matrix, the inverse matrix is the exact Hermitian transpose $T^{-1} = T^\dagger$ if experimental losses are not taken into account. Considering T^\dagger as the inverse matrix T^{inv} requires low computational power. Using matrix inversion by phase conjugation produces remarkable results in the experiments described in the following paragraphs. Generation of desired patterns at the distal fiber side is of high quality despite the existing losses and experimental limitations. In Figure 13, we present three examples

of projected patterns through a 400 μm core GRIN fiber using the measured transmission matrix to calculate the corresponding input wavefront.

3.1.2 Measurement basis

Measurement of the transmission matrix implies accurate recording of the input and respective output optical field. Digital holography is used to extract the amplitude and phase of the output field at the distal side of the fiber as it is described in the Chapter 2. The transmission matrix representation can be achieved using different basis as long as they completely describe the input and output fiber fields. Some examples found in literature describe the measurement of the transmission matrix using as input basis, plane waves in different angles, localized spots projected at the proximal facet, the fiber modes themselves or even random phase patterns [15,51,57,75,90,116–118]. There is no restriction to express and measure the transmission matrix using the same basis at the input and output since a linear transformation can be performed for a basis change. In the present work, we chose to measure the transmission matrix of the optical fibers studied in the basis of plane waves or “Fourier basis” as it is often called.

Measuring the transmission matrix in the Fourier basis is advantageous both optically and computationally. Expressing both the inputs and the outputs in the spatial frequency domain reduces the memory requirements. The fiber NA and the core size determine the number of the available Fourier components at each side of the fiber and therefore, the transmission matrix can be defined in a reduced size. On the contrary, if the values of the field are stored in the space domain (x, y) , the input and output vectors will be in the order of millions elements. Considering the resolution of the current spatial modulators and cameras, a single image is of 1000x1000 pixel size which suggests that the final transmission matrix size would end up being extremely large for a computer memory. In addition, the inversion of such a large matrix would require significantly higher computational power if not impossible to perform. On the other hand, saving the spatial frequency components of a wavefront is more efficient and the calculation of the field in space can be afterwards performed by a Fast Fourier Transform operation, which can be easily handled by a conventional computer.

Another advantage of using the Fourier basis is related to the modulation efficiency of the SLM devices. Plane waves at different angles are simply expressed by blazed grating phase images on the SLM screen. A blazed grating requires phase-only modulation (no amplitude encoding needed) and therefore a phase-only SLM (which is used in our experiments) can accurately project this kind of pattern. Phase-only basis are preferred when using a phase-only SLM. Low modulation efficiency is tolerated for low power applications but it can be a hindering factor when high peak power is needed at the distal fiber facet for material manipulation, as we will discuss in the next chapters.

3.1.3 Practical requirements

It was explained in the previous sections that the measurement of the transmission matrix of an optical system is constructed by recording the input and output fields. Interferometric measurements such as digital holography can accurately retrieve the information of an output field at the distal facet of the optical fiber but they are sensitive to instabilities of the measurement system. Phase drifts between the reference and object wave and intensity fluctuations of the light source over time can severely affect the measurement of the transmission matrix. Instabilities during the transmission matrix measurement can be caused by various factors such as mechanical and thermal perturbations in the environment of the experimental setup, wavelength drift of the light source etc. As discussed above, the inversion of the measured matrix of the system is performed using phase conjugation, which is based on the assumption of a time-invariant system. Therefore, it is important during the transmission matrix measurement to correct for the induced drifts as much as possible. To do so, phase drift monitoring during measurement of the matrix is performed by recording the phase of a standard test input every five input plane waves. At the end of the measurement, we correlate the phase of each test hologram recorded to the first one and the phase change is interpolated. In this way, the elements of the transmission matrix are corrected. Of course, correction by interpolation using the complex correlation coefficient measured from the test frames can be achieved only at a certain level. The simplest solution is the selection of a stabilized laser source and an isolated from vibrations optical setup.

Another parameter that needs further optimization for efficiently measuring the transmission matrix of the system is the synchronization between the camera that records the fiber output hologram for each input plane wave and the SLM. The phase-only LC-SLM (Pluto-NIR2, Holoeye) used in the experimental setup described in the following section is characterized by a certain refresh rate (60 Hz according to the specifications) and response time. In addition, as in most of the devices controlled by a computer interface, there is a time delay between the command execution on the computer and the action of the device. This delay time can last more than one refresh circle of the LC-SLM. Even when the command reaches the device the liquid crystals of the modulator need a certain time to orientate in the desired way in order to induce the assigned phase per pixel. Finally, in older SLM devices the electronic signal addressing the pixels causes flickering of the phase pattern projected on the device, which leads to intensity fluctuations of the shaped light after the SLM. A synchronization scheme is implemented to overcome the above timing restrictions between the detectors (cameras, photomultipliers etc.) and the SLM. To do so, every time a pattern is sent to the SLM a pulse is generated on the data acquisition card. From the time that the pulse is generated ($t=0s$), measurement of the SLM output is performed over time to determine the response curve of the SLM. Based on this curve, we define the specific delays in the system to ensure that the fiber output is recorded at the exact moment that the phase mask is projected by the SLM in an optimal way. More details about the synchronization method can be found in the work by D. Loterie [119].

3.2 Methods

3.2.1 Optical apparatus

For the transmission matrix measurement, we implemented the optical setup described Figure 14. A high pulse energy laser system (Satsuma, Amplitude Systèmes, $\lambda=1030$ nm, $E_{pulse,max}=40$ μ J) that provides tunable pulse duration ($\tau_{pulse,min}=300$ fs) and repetition rate was used. For the experiments presented in this thesis, the repetition rate is set at 20 kHz. The laser beam is expanded and collimated by combining two achromatic lenses (L1: $f=50$ mm, L2: $f=150$ mm) at the output of the laser before it is split into illumination and reference paths using a polarizing beam splitter (PBS). In the illumination path, the laser beam is reflected by a spatial light modulator (SLM, Pluto-NIR2, Holoeye) and the SLM plane is imaged using a 4f system at the proximal facet of the optical fiber. The 4f system (lens L3 and microscope objective O1) was optimized every time so that the SLM plane is demagnified approximately to fit the size of the fiber core under study. The size of the laser beam at the SLM plane was 8 mm to exploit most of the pixels available and to avoid damage of the device at the high pulse energy levels. Even at the maximum pulse energy used in the experiments ($E_p=15$ μ J) the peak intensity at the SLM plane is in the order of 60 MW/cm², which is much lower than the damage threshold specified for the device (~ 0.25 TW/cm²). A second 4f system is placed at the distal facet to image the output of the fiber on a CMOS (MV1-D1312IE-G2-12, PhotonFocus) detector array. For the second 4f system a 40x (0.65 NA, Newport) microscope objective (O2) was combined with an achromatic lens (L4) of 125 mm focal length. We chose a higher NA microscope objective at the distal side of the fiber to ensure the collection of most of the high frequency components of the output field. The reference beam is directed on the CMOS where it interferes with the fiber output resulting in a digital hologram. The use of a pulsed light source requires that the reference and illumination paths have the same optical length for interference to happen. A delay line placed in the reference path allows for adjusting the optical path length resulting in coherence gating of the fiber output.

During the calibration process, plane waves at different angles (blazed grating phase profiles) are sequentially projected by the SLM and the corresponding holograms of the speckle pattern at the distal end of the fiber is recorded on the CMOS detector for each input angle. The amplitude and phase of the distal field are then extracted using digital holography, as described in Chapter 2, and stored as a column vector of the transmission matrix. After the transmission matrix is calculated and inverted, any desired output can be generated by calculating the proximal phase on the SLM by simply multiplying the desired output field with the inverse matrix. The pulse energy used for the transmission matrix measurement ($E_p=0.2$ μ J) is chosen to avoid the presence of nonlinearities arising along the fiber length.

For the characterization of the focusing performance of the optical fibers used in the following experiments in the presence of high peak intensity pulses launched at the proximal fiber side, measurements of the output pulse duration and spectral profile are necessary. The pulse duration of the laser focus is determined by a custom-made two-photon fluorescence (TPF) interferometric autocorrelator that is integrated in the transmission matrix measurement setup. Two microscope

objectives of 40x magnification (0.65 NA, Newport) are used for focusing and collecting the light from the fluorescence sample (Rhodamine 6G in SU8) on CCD2 (Chameleon 3, Point Grey Research). The beam is redirected into the autocorrelator path by a flip mirror. The spectral characteristics of the distal output are measured using a spectrometer (HR4000CG-UV-NIR, Ocean Optics) placed in the conjugate plane of the focus spot.

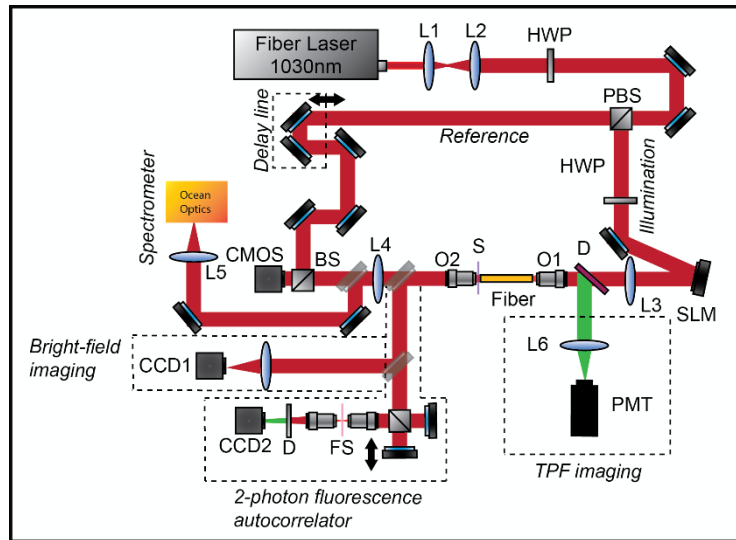


Figure 14. Transmission matrix experimental setup for focusing and scanning the light through the optical fibers, combined with a TPF autocorrelator and a spectrometer for characterizing the temporal and spectral profile of the pulse. The optical setup also includes a TPF imaging part for the creation of the sample images. (S = sample, D = dichroic, HWP = half-wave plate, FS = fluorescent sample for TPF autocorrelator, O = microscope objective, L = lens, PBS = polarizing beam splitter, BS = beam splitter) [15].

Additionally, a TPF detection system is implemented for the microscopy measurements. A dichroic mirror (D) and a photomultiplier (PMT, Hamamatsu, H11526-20-NN) are placed in the proximal side of the fiber so that the fluorescent signal emitted by the sample and collected back through the fiber is separated from the laser light and measured. For each focus scanned at the sample plane the fluorescence intensity measured by the PMT is saved as a pixel value for the image formation. The fluorescence signal is focused on the PMT by means of 100 mm lens. A bright field (BF) imaging block is also set in transmission from the sample to monitor the areas of interest consisting of a lens of 300 mm focal length to form an image on the CCD1 (Chameleon 3, Point Grey Research).

Different types of fiber are used for the transmission matrix experiments and they are listed below. The length of fibers in all cases is kept 10 cm for comparison.

- 200 μm core GRIN fiber (by Fiberware) with NA of 0.29, 7800 supported spatial modes
- 400 μm core GRIN fiber (by Fiberware) with NA of 0.29, and 31,200 spatial modes
- 460 μm core MCF (FIGH-10-500N, Fujikura) with NA of 0.34 for the individual cores

3.2.2 Biological sample preparation

Four ten-week old mice (NMRI strain) are used for the study presented in the paragraph 4.2. All animal procedures are approved by Ecole Polytechnique Fédérale de Lausanne (EPFL). The mouse is sacrificed, intracardially perfused and the cochlea is removed. Under the operating microscope, the middle turn of the cochlear bone is carefully opened with forceps while leaving the inner soft structures intact. The cochlea is then fixed in cold 4% paraformaldehyde (PFA) in 0.1 M phosphate-buffered saline (PhBS) for 30 min and rinsed with PBS three times for 5 minutes each. The remaining cochlear bone is gently removed and the pieces of Organ of Corti were microdissected. The Organ of Corti pieces are permeabilized in 0.3% Triton-X 100 for 15 minutes, rinsed with PhBS, and stained with propidium iodide (PI, Thermofisher Scientific, 1:100 dilution of 1.5 mM stock solution) which is a popular red-fluorescent ($\lambda_{em,max}=617$ nm) nuclear stain. After 30 minutes, the stained samples are rinsed with PhBS three times for 5 minutes each, mounted in ProLong Diamond media (from Life Sciences Technologies), sealed between two thin coverslips (145 μm) and imaged using the setup schematized in Figure 14. The prepared samples are extracted from the middle turn of the cochlea and their thickness is about 100 μm .

Chapter 4 High-intensity focusing through optical fibers

Ultrashort pulses have revolutionized conventional microscopy using nonlinear optical phenomena such as two-photon fluorescence (TPF) [120,121] and second-harmonic generation (SHG) [122,123] to improve the resolution and the quality of the obtained images. Ultrashort pulse delivery through fiber probes has made these imaging modalities endoscopically configurable [13,29,124,125]. Apart from nonlinear microscopy, femtosecond laser ablation (FLA) has been also extensively used in medicine for tissue manipulation due to its high precision and minimal collateral damage [104,106,126,127]. Integration of FLA in an imaging probe could be a pivoting advancement for high-precision fiber microsurgery. However, the peak intensities needed to perform FLA can exceed the tolerance of common endoscopes. With at least two-orders of magnitude more peak intensity than the one needed for TPF imaging ($\sim 10^9$ W/cm²), FLA may induce nonlinearities or even fiber damage if a single mode fiber is used [128].

The last decade has seen a rapid growth in the development of photonic crystal fibers [8], of which hollow core photonic crystal fibers (HC-PCFs) have shown an unsurpassed ability to deliver high energy, ultrashort pulses with minimal loss and dispersion [32–34,44,45]. Similarly to TPF imaging, FLA is also a point scanning technique which implies that, after propagation through a fiber, the emitted pulse needs to be focused and scanned onto the targeted material. Consequently, addition of distal end optics and scanning mechanisms is necessary for a HC-PCF-based probe, which results to a significant increase of the final endoscope size [33,129,130].

In the course of minimization of medical tools for less invasive clinical procedures, two different fiber types were studied (MCFs and GRIN MMFs) to investigate their potential capability to deliver a high peak intensity focus, aiming to enable both TPF and FLA through a single endoscope. Instead of opto-mechanical components, light focusing and scanning is achieved using wavefront shaping to control light propagation through the fibers, thus reducing the final size of the device. MCFs and GRIN MMFs present a promising choice for mitigation of fiber damage or nonlinearities because the input energy is spread among all the fiber cores or spatial fiber modes respectively. In the next paragraphs, the performance of different fiber probes for high-intensity focusing is characterized. The majority of the presented results and discussion in this chapter are published in the following journal articles and they are reproduced here according to the publisher's agreement:

- D. B. Conkey, E. Kakkava, T. Lanvin, D. Loterie, N. Stasio, E. Morales-Delgado, C. Moser, and D. Psaltis, *"High power, ultrashort pulse control through a multi-core fiber for ablation"*, Optics Express 25, 11491 (2017)
- E. Kakkava, M. Romito, D. B. Conkey, D. Loterie, K. M. Stankovic, C. Moser, and D. Psaltis, *"Selective femtosecond laser ablation via two-photon fluorescence imaging through a multimode fiber"*, Biomed. Opt. Express, BOE 10, 423–433 (2019).
- D. Psaltis, E. Kakkava, N. Stasio, D. B. Conkey, and C. Moser, *"Femtosecond pulse delivery through multi-core fibers for imaging and ablation"*, in Z. Liu, ed. (SPIE, 2017), p. 30.

4.1 Multicore fiber probes for endoscopy

4.1.1 Wavefront control through MCFs

MCFs are used for direct fiber imaging of an area of interest using each core as a pixel. This raises the question why wavefront shaping is necessary for this kind of fiber probes. As previously discussed, the imaging resolution of MCF-based endoscopes is dependent on the core spacing and it can be only improved by adding magnifying optics at the distal fiber side at the expense of field of view. Here, another approach is followed for improving the imaging resolution of the MCF probe, which uses wavefront shaping to generate a diffraction limited focus spot at the distal fiber side in order to perform point scanning methods. The size of the focus spot is only dependent on the NA of the individual cores, which implies that the final resolution can be significantly better than the core spacing distance ($\sim 5\text{ }\mu\text{m}$). In the presence of weak crosstalk between neighboring cores, a focusing wavefront launched at the proximal fiber facet is delivered at the distal end and generates at focus spot at a desired distance from the fiber facet. Small refractive index fluctuations and size non uniformities between the individual fiber cores have been reported to assist the reduction of crosstalk which ultimately allows the production of more dense MCFs [38,40]. However, the slight differences from core to core and the effect of core-to-core coupling result in a random phase distribution among the fiber cores at the distal facet, which implies that a desired phase profile would not propagate unchanged through the system and light focusing is not trivial. Due to the randomization of the phase distribution among the cores, a speckle pattern is generated at a distance far from the fiber facet. Previously reported work on MCF endoscopy has shown delivery of focused ultrashort pulses through a MCF for two-photon fluorescence imaging is possible using digital phase conjugation (DPC) [52]. Although DPC method can successfully control the light propagation through MCFs, low focusing efficiency is reported. We define focusing efficiency as the percentage of the total output power that is concentrated in the focus spot. It has been experimentally found, that the transmission matrix approach outperforms DPC in terms of light focusing efficiency. The main reason of the improved performance of the transmission matrix is related to the fact that it bypasses the complexity of alignment between the hologram capturing detector array and the SLM, which in turn is essential for DPC experiments. Like DPC, the transmission matrix accounts for the core-to-core coupling and allows the use of densely packed MCFs [37,52,55].

Nevertheless, the measurement of the transmission matrix for ultrashort pulses propagating through optical fibers presents further complexity, because of the coherence gating used in the case of pulsed laser sources. Based on the *NA* of the individual cores a certain number of spatial modes per core is supported. Therefore, modal dispersion and core-to-core coupling in addition to the group velocity dispersion induce temporal dispersion of the pulse, which lowers the amount of light that lies within the coherence length of the reference beam and consequently the overall focusing efficiency [51,52,124]. The temporal and spectral characteristics of the pulse propagation through the MCF system will be further discussed in the following paragraphs.

In the experiments described in this section, the transmission matrix is measured by capturing the output fields of orthogonal input modes after propagation through a 10 cm long MCF (FIGH-10-500N, Fujikura) as reported in the Methods paragraph 3.2 [51,75]. In order to explain in a comprehensive way the light focusing through the fiber probe, we define here the output field modes spatially, like pixels (where mode m represents a point in space of the output field), at a certain distance from the distal facet, while the input field is described at the fiber facet plane. The modes n are the fundamental modes of the individual cores. This definition is made for intuitive reasons since in the case of MCF we cannot talk about spatial modes as in the case of MMFs and each fiber core can be considered as a mode of the fiber. Once the transmission matrix is measured it can be used to calculate the necessary input wavefront to shape arbitrary intensity patterns at the output field [27]. In our case, we are interested in maximizing the focus intensity by creating a single focus spot at output mode m , whose intensity I_m is equal to:

$$I_m = C \left| \sum_n^N t_{mn} A_n e^{i\phi_n} \right|^2$$

Equation 18. Intensity of the focused spot after the MCF at the output point m

The parameter N is the number of cores, A_n is the amplitude at core n , ϕ_n is the phase on core n at the input, t_{mn} is the transmission matrix element which relates the input at core n to output mode m , and C is a factor which accounts for reduction of the focusing efficiency due to diffractive effects from the quasi-periodically spaced cores. The intensity I_m is maximized when ϕ_n compensates for the phase delay in core n measured in t_{mn} .

Figure 15 shows experimentally obtained data to demonstrate the measurement process. Figure 15a shows the facet of a MCF. As previously mentioned, the MCF is designed with high variability in the core size to reduce core-to-core coupling which allows a high core density [38,40]. Figure 15b shows the phase of the wavefront at the proximal facet of the MCF for generating a focus at the distal side. The input wavefront appears to be completely random on the facet. Figure 15c shows the quadratic wavefront created at the distal end of the MCF using Figure 15b as the input wavefront. The phase was extracted using digital holography. After free space propagation, the quadratic phase focuses the wavefront at some distance away from the fiber facet (Figure 15e, intensity in log scale). Because the quadratic phase is discretized by the MCF cores, diffracted light appears around the focus [37,51,52,55,131]. However, the relatively close spacing of the cores

and their quasi-periodic arrangement separates the diffractive effects far from the focus and spreads out the diffraction peak intensity [52]. In the following sections, we experimentally evaluate how the focus intensity depends on input pulse energy to understand the limitations imposed by nonlinearities.

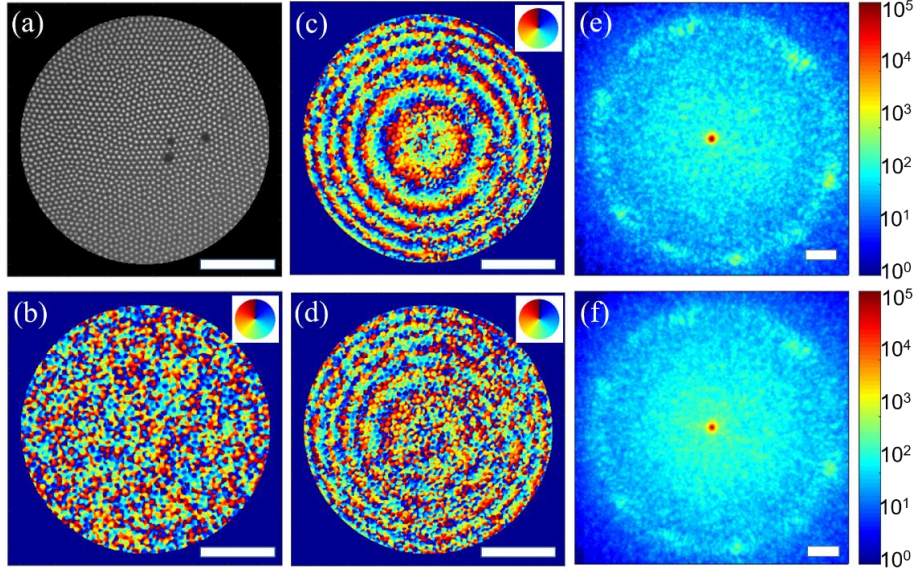


Figure 15. Visualization of wavefront shaping through an MCF. (a) The MCF facet with white light illumination. (b) The phase field input to the MCF facet to create a focus. The color wheel in the upper right indicates the relative phase value from zero (blue) to 2π (red). The wavefront at the distal end of the MCF with (c) low peak power and (d) high peak power input after the wavefront shown in (b) has propagated through the MCF. The focus formed 800 μm from the facet of the MCF with (e) low peak power and (f) high peak power input. The intensity is shown in a log scale. The scale bar in each figure is 50 μm [51].

4.1.2 Nonlinear degradation of focusing efficiency

The field of nonlinear optics refers to phenomena that take place when high intensity optical fields interact with matter strongly enough to modify the optical properties of the material system. For this reason, nonlinear phenomena were only observed after the invention of laser by Maiman in 1960. In the presence of strong optical fields, linear dependence of the polarization on the electric field is no longer accurate and nonlinear terms need to be taken into account as shown in Equation 19. The $\chi^{(j)}$ parameters correspond to the j th order susceptibilities which are tensors of $j+1$ order.

$$\mathbf{P} = \varepsilon_0(\chi^{(1)}\mathbf{E} + \chi^{(2)}\mathbf{E}^2 + \chi^{(3)}\mathbf{E}^3 + \dots)$$

Equation 19. Nonlinear polarization

In the Appendix A1 we obtain the wave equation in the linear case, but in the presence of nonlinear effects the wave equation is modified and the nonlinear part of the polarization appears as a source term, because it describes accelerating charges (Equation 20) [128,132]. The strength of the nonlinear phenomena depends on the peak intensity of the light source, the interaction

length within the material and naturally the properties of the material itself. As a result, nonlinear phenomena can be severe in the case of ultrashort pulses which are characterized by extremely high peak intensity even for relatively low pulse energy.

$$\nabla^2 \mathbf{E} - \frac{n^2}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{\partial^2 \mathbf{P}_{NL}}{\partial t^2}$$

Equation 20. Nonlinear wave equation

To understand the limitations imposed by nonlinearities in the MCF probe, when ultrashort pulses propagate through, we measure the focusing efficiency of the system across a range of input pulse energies and pulse widths. All the presented measurements are conducted using the optical setup presented in Figure 14. The Satsuma laser has the built in functionality to stretch the pulse width with a positive chirp, which also decreases the pulse peak power. The pulse energy coupled into the MCF is varied by an internal modulation device in the Satsuma as well as by adjusting the polarization of the beam incident on the SLM (which only modulates horizontally polarized light) with a half wave plate (Figure 14).

The experiment clearly illustrates the degradation of the focus efficiency due to nonlinearities in the MCF. Figure 16. shows the dependence of the focusing efficiency on the input pulse energy (measured at the input to the MCF) for four pulse durations. The exact value of the pulse duration is measured by the second order interferometric autocorrelation to be 500, 750, 1000, and 2000 fs before the MCF (autocorrelator shown in Figure 14). These measurements show that the focusing efficiency is highly dependent on the pulse width. As the input pulse energy increases, the focusing efficiency degrades more abruptly the shorter the pulses. The measurement of the transmission matrix and the creation of the focus is repeated 5 times for each pulse width and the error bars in Figure 16 indicate the standard deviation of the measurements. In addition, the measurement of the transmission matrix is performed in all cases using a low input pulse energy below 0.2 μJ to be on the linear optical response region.

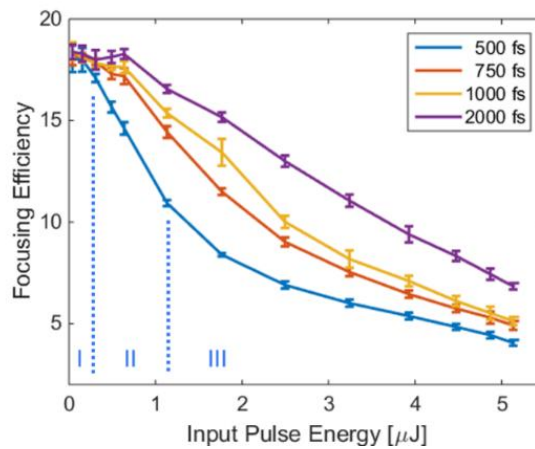


Figure 16. The focusing efficiency when focusing ultrashort pulses of varying pulse widths through a MCF compared to the input pulse energy. The three regimes of the 500 fs pulse focusing efficiency are marked by I, II, and III [51].

To describe the trend observed, we divide the data into three regimes. Using the 500 fs curve as an example (marked in blue on Figure 16), we see that at the region of input pulse energy (I), a linear regime exists, where the focusing efficiency is independent of pulse width and input pulse energy. In the input energy region (II), above 0.30 μJ , the focusing efficiency quickly degrades with increasing input energy due to the raise of nonlinear effects and more specifically because of self-phase modulation (SPM), as explained in the following section. (III) Above 1 μJ , the focusing efficiency decreases more slowly with increasing input pulse energy compared to the second regime. We suspect that in the third regime nonlinear effects have induced spectral broadening which, consequently, increases the effect of the group velocity dispersion (GVD) in each core leading to a broadened the pulse [128]. The broadened pulse is followed by a decreased intensity in the cores and induces less nonlinear phase shift with increasing input pulse energy.

The focal spot at the distal end of the MCF has maximum intensity I_m when the input wavefront compensates for the random phase delay differences between the individual cores to create the ideal quadratic wavefront at the distal facet (Figure 15c), which is the situation in the first regime (I) shown in Figure 16. To explain the behavior exhibited in the second regime, we will consider the consequence of the phase shift caused by SPM in the MCF. SPM is a third order nonlinear effect (related to $\chi^{(3)}$ susceptibility) that generates a nonlinear phase shift in the optical field because the refractive index of the material becomes dependent on the intensity of the optical field. The phase shift in each core is dependent on the nonlinear index n_2 , which has a material dependent characteristic value, and the intensity of the propagating mode. Assuming no pulse broadening and no core-to-core coupling for simplicity, the phase shift in core n is:

$$\Delta\phi_n^{SPM} = \frac{2n_2P_nL}{\lambda\omega_n^2}$$

Equation 21. Phase-shift per core because of SPM.

In Equation 21 λ is the wavelength, ω_n is the mode radius of core n , L is the length of the MCF and P_n is the peak power of the pulse in the core, determined by the location of the core in the MCF relative to the Gaussian shape of the beam incident on the MCF. By adding the SPM phase shift for core n from Equation 21 into Equation 18 the focal spot intensity becomes:

$$I_m = C \left| \sum_n^N t_{mn} A_n e^{i(\phi_n + \Delta\phi_n^{SPM})} \right|^2$$

Equation 22. Intensity of the focused spot after the MCF at the output point m in the presence of SPM

In Equation 22 ϕ_n is the phase of the wavefront at the input of core n that in the absence of nonlinearities would generate a focus after propagating through the MCF which is calculated by the measured transmission matrix. As stated in the previous paragraphs, for the MCFs used in this experiment, SPM induces to each core a unique $\Delta\phi_n^{SPM}$ due to the variability of the core characteristics (e.g. size, shape, material properties). In turn, the phase shift in the cores degrades the focus intensity. For example, Figure 15d shows how by increasing the input pulse energy, the phase shift in the cores degrades the quadratic phase profile by introducing random phase noise.

The SPM phase shift is more prominent near the MCF center where the intensity in the cores is highest due to the Gaussian profile of the input beam. The SPM induced in the cores results in the decrease in focus efficiency seen in Figure 16, meaning that part of the output light does not end up in the focus spot. Figure 15f shows how this causes a slight increase in the background intensity. Although obscured by the log scale, the maximum intensity of the focus in Figure 15f is half that of Figure 15e. In the above analysis, for simplicity we neglect the effect of core-to-core coupling in the MCFs. While the MCFs were designed to limit the core-to-core coupling in the fiber [40], in practice the coupling modulates power and phase along the length of the MCF. When the input power induces a SPM phase shift on top of the core-to-core coupling, the phase and intensity of the pulses propagating in neighboring cores changes, further degrading the quadratic phase at the distal facet.

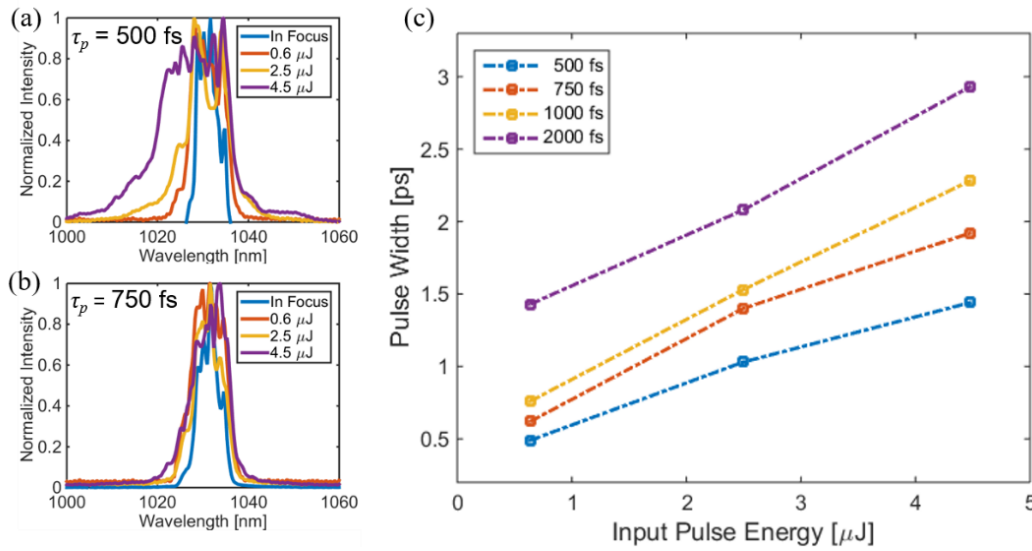


Figure 17. Spectral and temporal measurements in the focus (blue curve) and background (orange, yellow and purple curves) for different input pulse energies with (a) 500 fs and (b) 750 fs input pulse width, also showing the in focus spectrum which did not change with input energy. (c) The in focus pulse width dependence on input pulse energy [51].

As mentioned previously, the peak intensity in the cores becomes high enough to induce spectral and temporal broadening, which leads to the third regime shown in Figure 16. Measuring the pulse duration and spectrum of the output field both in and away from the focus provides some insight into the strength of the spectral and temporal broadening in the MCF. The pulse width and spectrum were measured with three different input pulse energies (0.6 μJ, 2.5 μJ and 4.5 μJ) at the four previous input pulse widths. For the spectroscopic measurement, a fiber input to a spectrometer (HR4000CG-UV-NIR, Ocean Optics) placed in a plane conjugate to the focus plane probes either the spectrum of the background or the focus (Figure 14). The pulse width of the focus is measured with an autocorrelator built after the MCF (shown in Figure 14). For temporally broadened pulses, the pulse durations are estimated by Fourier filtering of the second order interferometric autocorrelation trace to isolate the intensity autocorrelation of the pulse, as

explained in [133]. The laser emits a squared hyperbolic secant shaped pulse, so the pulse width was calculated by multiplying the FWHM of the intensity autocorrelation by a deconvolution factor of 0.65. However, it is important to note the temporal pulse shape of the wavefront shaping generated focus after the MCF is unknown, but we use the 0.65 deconvolution factor for consistency in all the reported values. Further pulse shape characterization could probably render more accurate values of the pulse broadening through the MCF because of nonlinearities and could be a topic to investigate in the future.

The spectral characterization verifies that at high intensities the nonlinear effects are strong enough to broaden the pulse spectrum. The spectral broadening could be induced by SPM, cross phase modulation (XPM), four wave mixing (FWM), or other nonlinear effects [128]. The spectrum of the background captured for the 500 fs pulse (Figure 17a) shows significant blue-shifted spectral broadening as the input pulse energy increases. For input pulse energies of 0.6 μJ , 2.5 μJ , and 4.5 μJ the respective linewidth (1/e) in the background are 10 nm, 14 nm, and 20 nm. In contrast, the linewidth of the in focus radiation remains unchanged (8 nm) regardless of input pulse energy. It is expected that the spectrum of the focus spot remains the same because the light that contributes to the focus formation is the one that is optimally modulated by the transmission matrix. Since the transmission matrix is wavelength dependent only the wavelengths within its validity range will be efficiently modulated and reach the focus spot [134]. Figure 17b shows the in and out of focus spectrum with 750 fs input pulse width, which shows that the SPM induced less spectral broadening than the 500 fs case. The spectral broadening for the 1000 fs and 2000 fs input pulse widths further decreases.

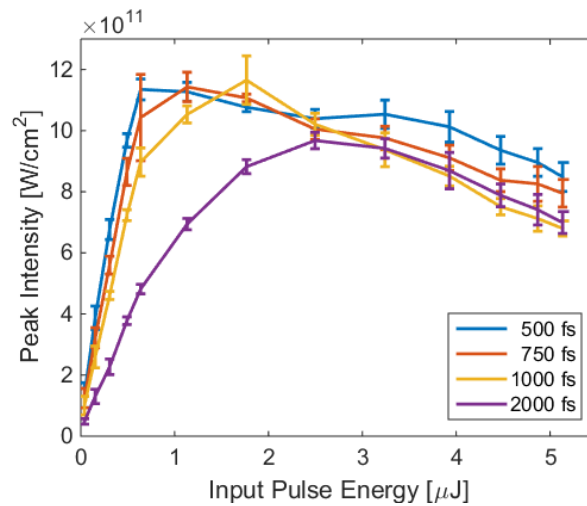


Figure 18. The estimated peak intensity in focus plotted for increasing input pulse energy and pulse width [51].

The pulse width of the focus displayed significant temporal broadening dependent on the input pulse energy. Figure 17c shows how the 500, 750, 1000, and 2000 fs pulses broaden with increasing input pulse energy. These results imply that the nonlinear effects modified the

broadening rate of the pulse in the cores, in other words the pulse broadened beyond the effect of GVD alone [128]. Interestingly, the pulse width of the focus at low input pulse energy was consistently shorter than the width of the pulse measured directly from the laser. It is unclear whether the pulse has compressed due to material dispersion in the MCF or whether it was an unexpected effect of the coherence-gated hologram. Further work needs to be conducted to identify the source of the compression.

Ultimately, the performance of the device for ablation depends on the peak intensity of the focused pulse. We estimate the peak intensity with the measured pulse durations and the energy in focus. Figure 18 shows the peak intensity of the created focus. Regardless of the input pulse width, the maximum intensity is around 1.1×10^{12} W/cm². Notably, the maximum obtained peak intensity occurs for all pulse widths at a focusing efficiency that is ~75% of the maximum obtained in regime (I) (comparing Figure 16 and Figure 18). It can be assumed that at this point the best compromise among focusing efficiency, input power and nonlinearities is obtained.

4.1.3 Comparison of two different MCFs

In the above presented description of the limitations of using a MCF for high power ultrashort pulse delivery only 4,400 cores of the 10,000 cores of the MCF were used. By choosing to illuminate part of the MCF cores, we simulate the case of an endoscope thinner than the actual MCF size. In this section, we present an optimized system that utilizes optimally the entire MCF and maximizes the peak intensity of the focus spot. We modified the optical system for optimal usage of all cores in the MCF. Referring to the optical setup diagram in Figure 14, the L3 lens was changed from a focal length of 200 mm to 125 mm, which caused the beam to slightly overfill the MCF facet. The second 4f system is also changed: L4 has a 150 mm focal length lens and O2 is a 40X objective (NA 0.65) instead of 200 mm and 20x (NA 0.4) used initially.

Table 1. A comparison of the optical parameters for the two MCFs [51].

System	Aperture Diameter	Working Distance	NA_{foc}	Focus size (FWHM)
4,400 cores	310 μ m	700 μ m	0.22	3.20 μ m
10,000 cores	460 μ m	800 μ m	0.28	2.65 μ m

To maximize the peak intensity, a balance between the focusing efficiency and the size of the focus spot must be found by selection of the distance between the MCF facet and the focus (working distance) [52,55]. In order to use of all of the delivered pulse energy from the MCF, the minimum working distance should be where all of the cores contribute to the focus, a distance

limited by the NA of the cores ($\sim NA=0.3$) and the diameter of the illuminated portion of the MCF [39]. Beyond this point, the NA of the created focus (NA_{foc}) is given by the ratio of the radius of the illuminated aperture of the MCF and the working distance. While a higher NA_{foc} results in a smaller focus spot size, a lower NA_{foc} increases focusing efficiency. In fact, the discretization of the wavefront by the cores results in undersampling near the edges of the quadratic phase profile that lowers the final focusing efficiency when the quadratic phase field is stronger [52]. Table 1 presents the working distance and NA_{foc} utilized with 4,400 cores and the optimized 10,000 core system, as well as the FWHM of the final focus spot size.

In contrast to the 4,400 core system, the 10,000 core system shows significantly reduced nonlinear degradation of the focus (Figure 19a). Of course, this is an expected result of dividing the pulse energy among more cores, which reduces the intensity of the light per individual core. Spreading the pulse energy into more cores decreased the spectral broadening (Figure 19b) and temporal broadening (Figure 19c). The focusing efficiency when using 10,000 cores was less than with 4,400 cores at low pulse energies, because of the higher focus NA (Table 1). Despite the lower focusing efficiency, the smaller focus size and decrease in nonlinear effects resulted in an increased maximum peak intensity (Figure 19d), in this case above $4 \times 10^{12} \text{ W/cm}^2$ with a 500 fs pulse, or nearly 4 times higher than the 4,400 core system [51,53]. Optimization between focusing efficiency and focus spot size can further increase the delivered peak intensity through the 10,000 core system.

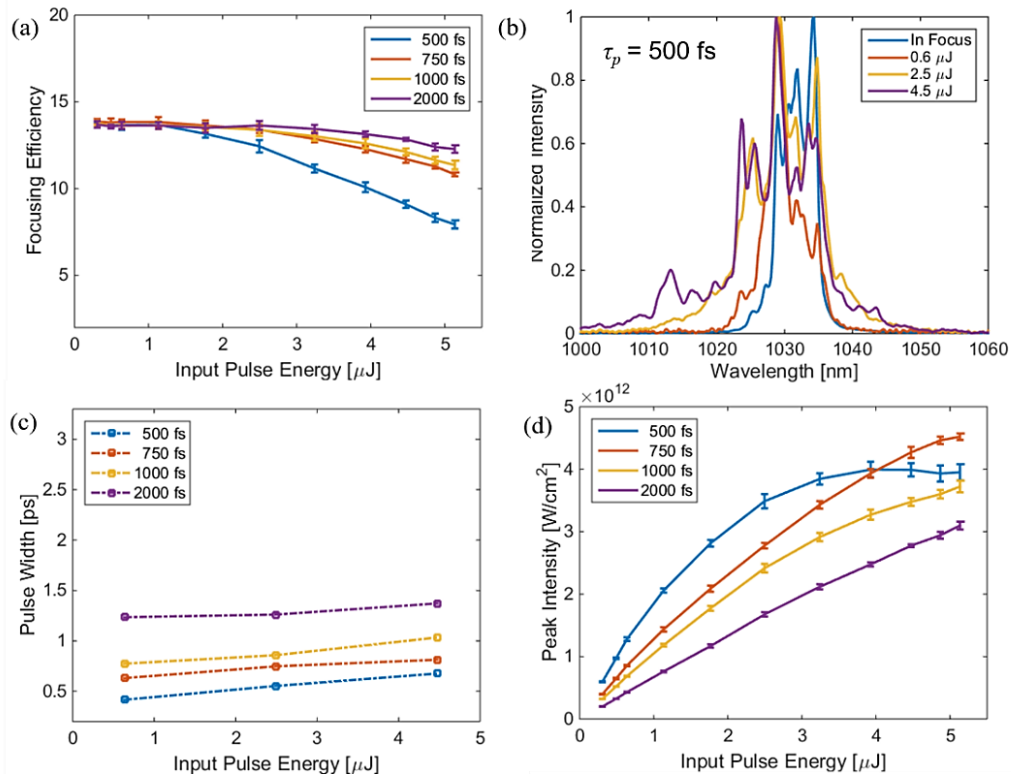


Figure 19. Characterization of the 10,000 core fiber system. (a) The focusing efficiency, (b) the measured spectrum of the in focus and background light for different input pulse energies with 500 fs pulse width, (c) the in focus pulse width dependence on input pulse energy, and (d) the estimated peak intensity [51].

4.1.4 Ultrashort pulse ablation through an MCF

Laser ablation refers to the removal of a portion of material from a target sample when a high peak intensity laser spot interacts with its surface (or volume). The high intensity laser pulse induces material breakdown, which means that the light-material interaction causes ionization resulting to a plasma state and as a result the atoms escape the lattice within the interaction area. Laser ablation is also a nonlinear optical process and it consequently depends on the intensity of the light source. The laser-matter interaction mechanisms differ for different laser source properties such as the pulse duration and wavelengths [107,135,136]. In the case of ultrashort pulses the interaction time between the pulse and the material is so short (pulse width range), which prevents the energy dissipation in the lattice of the material and therefore the interaction volume is rather directly evaporated than melted, which produces highly precise machining result. Every material is characterized by an ablation threshold, defined as the intensity at which the probability of the laser pulse to cause material breakdown is at least 50% [107]. The laser ablation threshold for a material is calculated at a certain pulse width and wavelength and differs when these parameters change.

As a demonstration of focusing and controlling high peak power ultrafast pulses through MCFs, we ablate a thin gold film (60 nm) deposited on silica. The ablation threshold of gold is 5.5×10^{11} W/cm² for a 750 fs pulse [135]. We utilize the 4,400 core system for this demonstration to show that gold ablation is possible even with the more constrained system, which implies that a thinner probe can also be employed. The output field of view for the measured transmission matrix enables proximal control of the light within an area with dimensions of 200 μ m x 200 μ m. Transmission matrix calculation and wavefront shaping, allow control of the distributed light intensity within the output field. In other words, a focus spot can be created anywhere within the measured output field, which allows for sequential focusing at designated locations to create an ablated pattern in the gold. Because the size of the ablated spot on the gold is intensity dependent [137], it is desirable that the focus intensity be invariant across the scanning range. However, when focusing away from the center of the output field, the focus intensity diminishes [52]. To compensate for the intensity variability, we modified the input wavefronts to make the focus intensities at the targeted locations uniform. To do this, we decreased the intensity of the higher intensity spots by adding noise to the input wavefront proportional to the targeted decrease in intensity. This is similar to adding random phase values to the input wavefront. Importantly, adding the noise to the phase mask does not affect the focus spot size.

Figure 20 shows two examples of gold ablation through MCF with proximal control of the beam. An EPFL logo ablated with input pulse energy of 2 μ J, 750 fs pulse width, and a repetition rate of 1 kHz is shown in Figure 20a. The size of the pattern is 106 μ m x 82 μ m with a 2 μ m separation between ablation focal spots. Figure 20b shows an image of the Matterhorn and a Swiss shield ablated with input energy of 3 μ J, 750 fs pulse width, and 1 kHz repetition rate. The image is 120 μ m x 120 μ m and has a 2 μ m pitch between focal spots. Both ablated images show good uniformity of ablation spot size across the image and show the capability of proximal control to scan a focus spot for ablation [51].

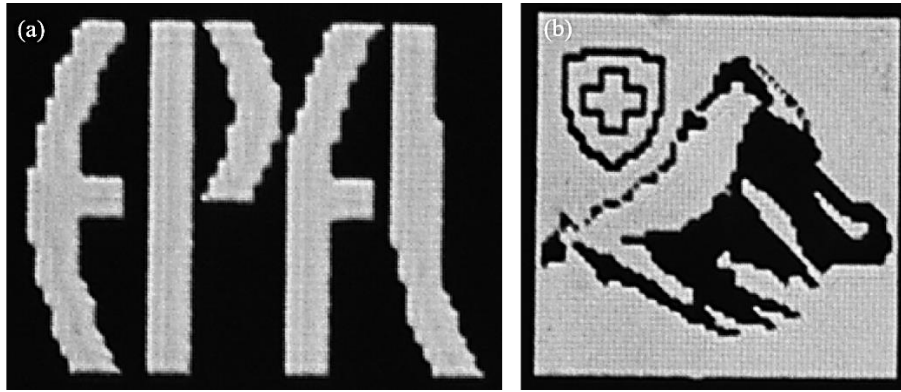


Figure 20. Widefield transmission optical images of ablated samples on thin gold films deposited on glass. (a) EPFL letters, $106\ \mu\text{m} \times 82\ \mu\text{m}$. (b) The Matterhorn and Swiss shield, $120\ \mu\text{m} \times 120\ \mu\text{m}$ [51].

Additionally, we show that the ablation spot size and therefore the resolution of the printed structures can be tuned. Specifically, the diameter of the focus spot created using the measured transmission matrix is $2.65\ \mu\text{m}$, but the laser ablation spot can be less than $1.9\ \mu\text{m}$ as shown in Figure 21, providing the possibility of high resolution micro-machining. Because of nonlinear nature of the laser ablation process discussed in the beginning of the paragraph, tuning the intensity of the focus spot close to the ablation threshold of the material allows for tuning the final resolution for micro-machining applications through an MCF (Figure 21).

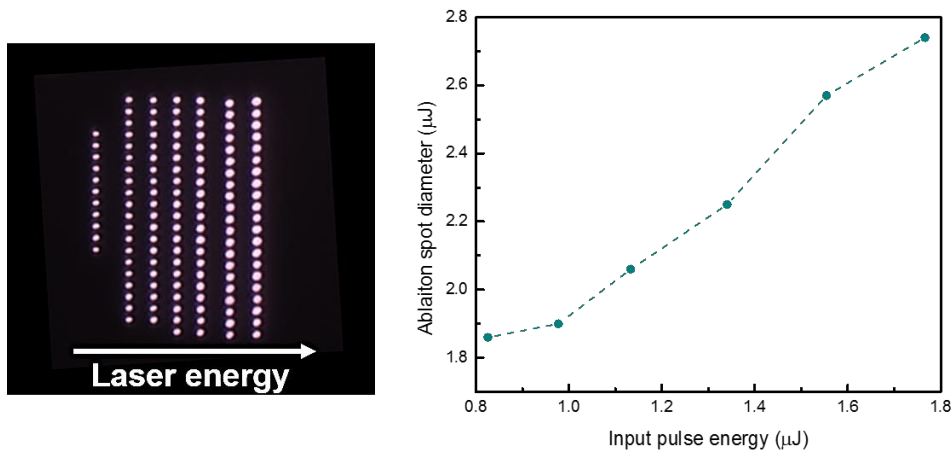


Figure 21. Laser ablation spots created using different pulse energies through the MCF. Each line of spots has been ablated using the same energy and corresponds to a point in the graph on the right which illustrates the spot size for each pulse energy [53].

4.1.5 Conclusions

In this section, high peak power, ultrashort pulse focusing and control through MCFs has been demonstrated and an investigation of the ways that the nonlinearities induced within the cores

limit the ultimate peak intensity generated in a focus is presented. Proximal light control through MCFs eliminates the need for distal end optical and mechanical components for focusing and scanning. However, the nonlinearities induced when high peak input power is launched through the system degrade the percentage of output power in the focus and contribute to the temporal broadening of the pulse, thus limiting the peak intensity of the focused pulse. Despite of the observed limitations a high peak intensity focus spot can be delivered even through an ultrathin MCF probe of only 4,400 cores with sufficient intensity to ablate gold.

To increase the intensity in a wavefront shaping generated focus a number of solutions are proposed. For example, a shorter MCF would decrease the nonlinear interaction length and increase focusing efficiency [53]. Additionally, as shown in the previous analysis, by spreading the pulse energy among more cores, the nonlinearity can be minimized and a higher intensity focus created. Increasing the MCF diameter or the density of the cores would allow the use of more cores. Although, with currently available MCFs it is not possible to use a smaller core spacing without significantly increasing the core-to-core coupling [38,40]. Alternatively, small distal components such as a GRIN lens could be included on the distal tip to focus with a higher NA and increase the intensity of the focus [124]. Other potential solutions involve further shaping of the wavefront and/or pulse to compensate for SPM [138]. As with other wavefront shaping through fiber techniques, the transmission matrix measurement is fiber conformation dependent. While the MCF shows some resilience to bending [55], new adaptive methods could overcome this obstacle [139].

4.2 Multimode fiber probes for endoscopy

In the previous section 4.1, the performance of MCFs in delivering high peak power focus spots was explored. Peak intensities in the order of 10^{12}W/cm^2 were observed. This intensity level is sufficient for both TPF imaging and for laser ablation of low threshold materials like metals as previously demonstrated, but not high enough for samples such as biological tissue or glass (intensity thresholds reported in the range of 10^{13}W/cm^2) [104,140]. One of the basic hindering factors for further increasing the light on the generated focus in the case of MCFs, even in the absence of nonlinear effects, is their quasi-periodic structure, which causes the light to diffract on a circle around the focus resulting in a decreased focusing efficiency. To improve peak power delivery of a focused ultrashort pulse through a fiber endoscope, use of different fibers types should be also considered. Step index MMFs are excluded because of the pulse broadening caused by the significant modal dispersion, which enables probing of only few modes at the output using coherent gating and consequently decreases the overall focusing efficiency. On the contrary, GRIN fibers are engineered with low modal dispersion and present a good alternative for pulse inputs. In this section, we use the transmission matrix technique to focus and scan the light through two GRIN fibers of 200 μm and 400 μm diameter core size and test their performance in delivering a

high peak intensity femtosecond focus spot for laser ablation applications. The results are also compared to those obtained by the MCFs in the previous section.

Multimodal endoscopes have been proposed [141,142] for acquiring images with several microscopy methods integrated in the same probe. In the following paragraphs, we report for the first time FLA through MMFs combined with TPF imaging at the same multimode fiber. We test the performance of our dual-modality endoscopic tool in the organ that hearing is taking place (Organ of Corti) from an extracted mouse cochlea. The Organ of Corti includes mechanosensory cells, named hair cells, which are responsible for the detection and transduction of sounds [143,144]. TPF images of the cochlear hair cells were acquired through our endoscope and used to define the area of interest, which is then ablated through the same tool. By developing a technique with imaging modalities integrated with selective femtosecond laser ablation capabilities, we provide a minimally invasive system for both cellular level investigation of distinct target areas and for cell manipulation.

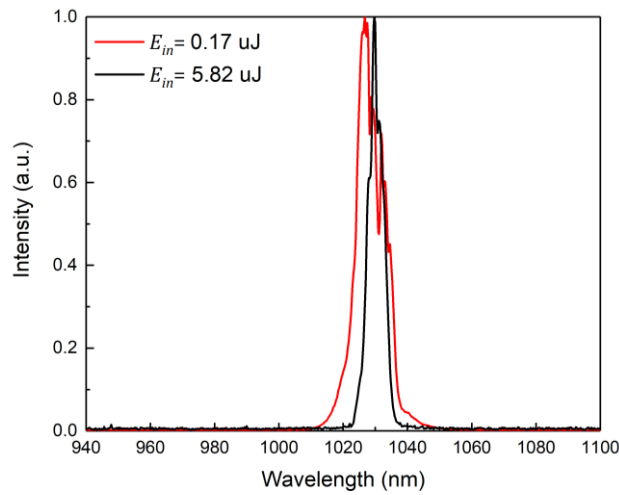


Figure 22. Spectral measurements for the lowest and highest values of input pulse energy in the case of the 200 μm GRIN fiber.

4.2.1 High peak intensity delivery through GRIN fibers

For each of the two fibers, we measure the dependence of the focusing efficiency on the input power for three different input pulse durations in a similar manner to the MCFs studies discussed before. A high intensity focus spot is generated at the distal fiber end after measuring the transmission matrix of the system using the experimental setup presented in Figure 14. The maximum value of the focusing efficiency was found to be $\sim 28\%$ for both probes, which is significantly higher than that achieved by the MCFs. In the case of an ideal Airy spot 84% of the total power should be concentrated in the central lobe. However, in our experiments, we probe only one polarization. By rotating a linear polarizer placed in front of the CMOS detector we observe that the output is elliptically polarized which implies that more than half of the modes are

probed in terms of polarization during the transmission matrix measurement. Apart from the loss of a part of the fiber modes because of the polarization selectivity, the coherence gating of the fiber output also reduces the number of contributing modes on the formation of a focus spot at the fiber distal end. The focusing efficiency achieved is limited to about 28%, which can be considered optimal if taking into account the losses explained above. The light is focused at 200 μm from the distal facet for both GRIN fibers to a spot with 2.3 μm FWHM. The spot size is slightly larger than the theoretical value (Rayleigh criterion) which is calculated to be 2.17 μm for the NA of the GRIN fibers ($NA=0.29$) and the laser wavelength (1030 nm) used in the experiments. Although modal dispersion is not considered significant for GRIN fibers because of the optimally designed refractive index profile of the core, in reality the existence of defects on the parabolic profile induces some modal dispersion. Hence, the deviation of the focus spot size from the theoretical value could be attributed to the fact that higher order modes are not probed by the system because of the coherent gating resulting in a lower effective NA [35,52,74,145].

For focusing the light through the GRIN fiber, the phase input on the SLM is calculated using the measured transmission matrix so that the ideal quadratic phase is created at the distal facet. When this phase profile is launched at the proximal side of the fiber, the light becomes distributed among the supported fiber modes and each one propagates through the system with a certain initial phase assigned by the wavefront-shaped input. As discussed in the previous paragraphs, wavefront shaping using the transmission matrix is a method applied to linear systems and its validity and performance deviate when nonlinear phenomena occur [47,51,80]. When the laser pulse energy exceeds a certain threshold then nonlinear phenomena such as self-phase modulation (SPM) and cross-phase modulation (XPM) can arise. The spectral measurements for the 200 μm core GRIN fiber shows spectral broadening of 9 nm confirming the presence of nonlinearities (Figure 22). However, no side bands were present in the spectrum, which would imply FWM or Raman scattering nonlinear effects [77,128].

Similarly to the MCFs, the maximum phase shift induced by these phenomena affects the relative phase among the modes when propagating along the fiber and as a result, the final phase profile at the distal facet is different from the optimized by the measured transmission matrix, reducing the final focusing efficiency [128]. Considering the dependence of the SPM and XPM phase shift on the fiber length [128], an increase of the fiber length would result in an increase of the nonlinear effect on the system's efficiency for focusing the light at the distal side. In our past work, we studied the dependence of the focusing efficiency of a MCF system for three different fiber lengths (10, 20 and 30 cm) and we observed a degradation for the longer segments [53]. The result is in agreement with the nonlinear optics theory that dictates that the strength of a nonlinear effect is proportional to the nonlinear interaction length [128,132]. Nevertheless, for practical applications, the fiber probe can be inserted in a cannula to avoid bending would be a part of a hand-held device, 10-20 cm length is enough to reach most of the areas of interest. Therefore, we selected the 10 cm long GRIN fiber for our experiments.

Based on the discussion above, at high energy input pulses, nonlinearities arise and affect the propagation of the input field along the fiber, resulting in scrambling of the optimized wavefront

at the distal facet and compromising the focusing efficiency (Figure 23) [51,128]. Because nonlinear phenomena are intensity dependent, the smaller the fiber core and the shorter the pulse duration for a given input pulse energy, the more significant the nonlinear effects that appear. We studied the performance of the two fiber probes as a function of pulse duration for three cases, 500 fs, 700fs and 1000fs pulses. The focusing efficiency of a 200 μm core diameter GRIN fiber diminishes rapidly after a certain threshold ($\sim 1 \mu\text{J}$) (Figure 23a), while the 400 μm core fiber (Figure 23c) shows a linear behavior for a wider range of input pulse energies [51]. As a result, increase in the input energy does not lead to a further increase in the final peak intensity in the focus for the thinner fiber (Figure 23b). The results agree with the findings of the MCF systems discussed in the previous paragraph.

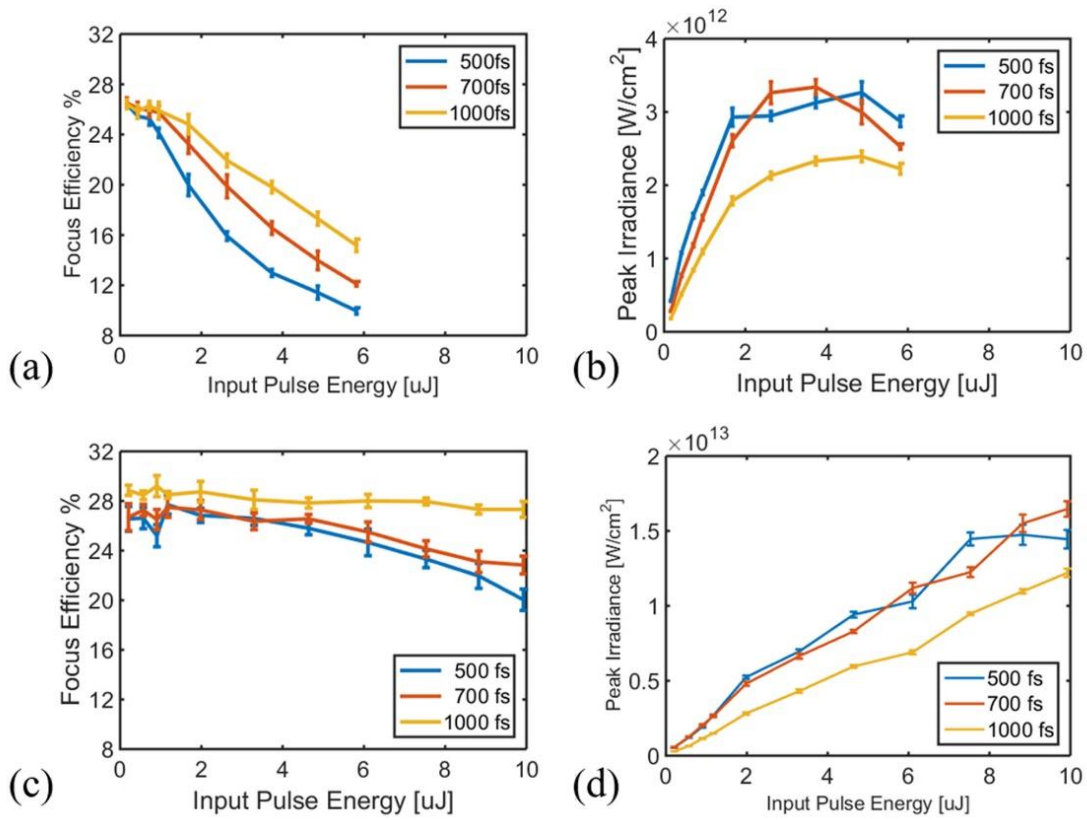


Figure 23. Focus delivery characterization of the 200 μm (a and b) and 400 μm (c and d) core diameter fibers. (a and c) Focusing efficiency for the 200 μm and the 400 μm core GRIN fiber respectively, and (b and d) the corresponding peak intensity delivered at the distal spot as a function of input pulse energy for each of the two systems respectively [15].

For calculating the peak intensity of the focus spot for the different input pulse energies, the pulse duration of the fiber output was measured using TPF-based autocorrelator in the same way as described for the MCFs. The spot size was measured on the CMOS for different input pulse energies (0.2, 1, 4.6 and 8.8 μJ) and no change was observed. For a 200 μm core fiber, nonlinearities cause temporal and spectral broadening when the power is increased above a certain level. For example, for 500 fs input pulse, the measured output pulse increases to about

700 fs when the input energy reaches 6 μJ , while the spectral width increases from 8 nm to 17 nm. Increase of the input power above 6 μJ was not possible for the thinner GRIN probe because of fiber damage (Figure 24). The quadratic profile of the refractive index in the core of a GRIN fiber generates focusing points with a certain periodicity along the fiber length when an plane wave is coupled in the core. However, the seemingly random input phase on the SLM imaged at the proximal facet, which results in focusing the light at the distal side, randomizes the input wavefront and self-focusing because of the refractive index profile is prevented. Nevertheless, when the input pulse energy is higher than a certain level there is enough amount of light still focusing at the entrance of the fiber that leads to local burns in the core. The picture shown in Figure 24 is captured at the experimental setup when the energy of the input pulse to the 200 μm core GRIN fiber is increased above 6 μJ . Self-focusing of the high power at the proximal facet is obvious from the luminescent spot appearing close to the proximal fiber facet possibly resulting from the optical breakdown of the fiber core (Figure 24), which in turn leads to fiber damage. The degradation of the focus spot efficiency because of the nonlinearities and the induced damage in the core for the 200 μm core GRIN fiber leads to a saturation of the peak intensity that can be delivered through the system to a maximum value of $\sim 2.5\text{--}3 \times 10^{12} \text{ W/cm}^2$. The maximum delivered peak intensity by this GRIN probe is comparable to the one achieved using the MCF of double the core size (460 μm) [51], which demonstrates that the ultrashort pulse delivery is more efficient for GRIN fibers.

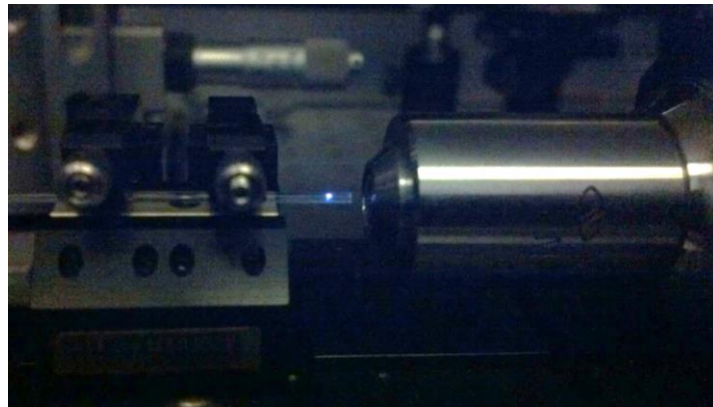


Figure 24. Fiber damage of the 200 μm core GRIN when the input pulse energy increases above a certain threshold.

For the 400 μm core GRIN fiber, no significant decrease in the focusing efficiency is observed for the range of input energies studied, including when the shortest pulse duration (500 fs) is tested Figure 23 c. In addition, temporal and spectral measurements of the focus spot showed no change even at the highest input pulse energy. Consequently, the peak intensity delivered by this probe shows a linear increase throughout the range studied for the 700 fs and 1000 fs input pulses. In the case of 500 fs input pulse duration a saturation of the delivered peak intensity is observed above 6 μJ . For the highest input power provided by the laser system used for the measurements, the maximum peak intensity measured is $1.5 \times 10^{13} \text{ W/cm}^2$ (Figure 23 d). Further increase of the

maximum peak intensity delivered by both systems could be achieved by pre-chirping the pulse at the input. In this way, pulse broadening because of group velocity dispersion (GVD) could be compensated and simultaneously, nonlinearities would be less significant for the stretched pulse (because of its lower peak power) [29].

Another attempt for increasing the performance of the imaging endoscope is made by measuring the transmission matrix of the fiber at the nonlinear regime aiming to compensate at a certain level the randomization of the quadratic phase profile at the output due to the nonlinear effects. However, no significant change in the focusing efficiency was observed though. This can be explained by considering the linearity of the system of a MMF transmission matrix. It is true that by increasing the input pulse energy to be in the nonlinear peak intensity region, the input plane waves at different angle will propagate through the fiber experiencing nonlinear optical modulation. Nevertheless, this would not compensate the nonlinearities that occur for an input wavefront which is the combination of these plane waves (formed to generate a focus spot at the distal side) because the addition in the input does not correspond to an addition of the propagation properties of the light until the fiber output for a nonlinear system. The random input phase for focusing the light through the GRIN fiber will give rise to nonlinearities along the propagation length in non predictable way by the linear transmission matrix.

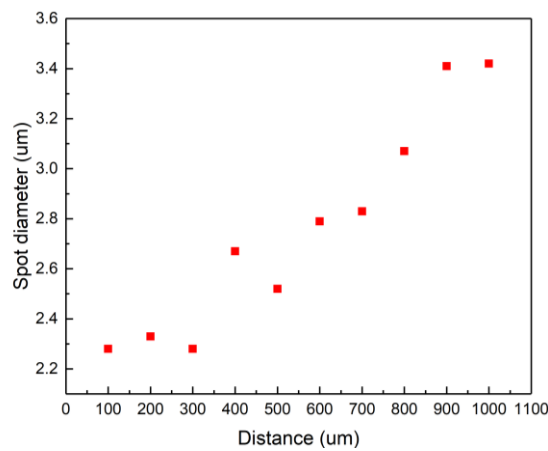


Figure 25. Focus spot diameter as a function of the focusing distance from the distal fiber facet.

Both GRIN probes studied in this work show high performance for delivery of focused femtosecond pulses for laser ablation or imaging applications that is comparable to the one achieved using HC-PCFs. Specifically, Kagome HC-PCFs have shown unsurpassed ability to deliver ultrashort pulses, reaching peak intensities of $4 \times 10^{13} \text{ W/cm}^2$ [146]. For endoscopic application, these fibers have been used to build a microsurgery scalpel by Subramanian et al. [33] and the maximum peak intensity delivered is reported to be about $5 \times 10^{12} \text{ W/cm}^2$ in a focus spot of $2.5 \mu\text{m}$ size. The need of distal lenses for focusing the light emerging from the fiber comprises the maximum intensity that the endoscope can handle due to the nonlinearities raised when high-energy pulses propagate through the optical elements. Recently, the use of different lens material

showed that the power throughput can be further increased for a HC-PCF endoscope [29,74,145]. However, the combination of distal lenses and a piezo-based scanning mechanism result in an increased size of the endoscope to 5mm if compared to the original fiber diameter, which is few hundred microns. In our case, both GRIN fiber probes provide more than 2×10^{12} W/cm² peak intensity pulses for microsurgery applications focused in a 2.3 μ m spot thus offering improved spatial resolution and smaller overall size. The lateral resolution of the system was also measured at difference distances away from the fiber facet and remains unchanged until about 300 μ m as shown in Figure 25. More details about the change of the 3D point spread function and the field of view (FOV) of a MMF as a function of distance from the fiber facet in previous works on imaging through fiber probes [29,74,145,147–149]. With a range of 300 μ m within which the resolution remains unchanged, we believe that imaging of thick samples up to 200-300 μ m would be possible. The FOV of the system is mainly dependent on the fiber core diameter and the focusing distance from the fiber facet, which implies a FOV of more than 200 μ m for the larger core GRIN fiber. Furthermore, the size of the probe can remain 200-400 μ m since lens-less focusing and mechanic-less scanning is demonstrated by means of wavefront shaping.

4.2.2 TPF imaging feedback-based selective FLA

FLA is a technique with many biological and material processing applications. In most cases, it is beneficial for the ablated area needs to be defined by a feedback mechanism. In this work, we demonstrate for the first time an MMF probe that can perform FLA with high selectivity based on a TPF image of the sample, which is obtained through the same probe. TPF imaging provides high penetration depth for inspection of scattering samples such as biological tissues due to the use of infrared photons. In addition, the nonlinear character of the process results in higher resolution and sectioning than single-photon fluorescence imaging. Therefore, in our endoscope we use the TPF images as a guide for the FLA of the sample.

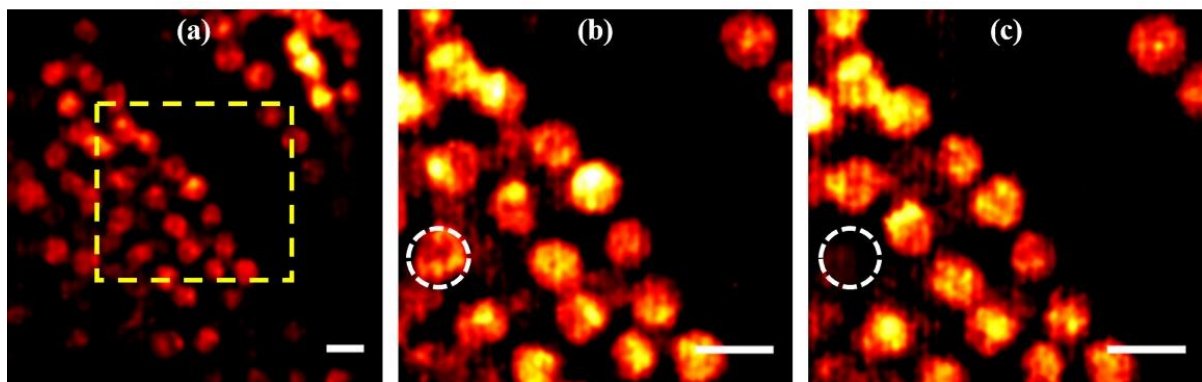


Figure 26. TPF imaging of the Cochlear hair cell sample stained with PI in a) a wide image of the sample, b) zoomed region of interest (yellow square) and c) the same area as (b) after selectively ablating a cell marked in the white circle. The scale bar in all figures is 10 μ m [15].

Biological tissues have significantly different properties than those of metals (for which we demonstrated ablation through MCF's in the previous paragraph [51,53]). The basic content of biological samples is water and the FLA intensity threshold is close to that of water, on the order of $\sim 10^{12}$ W/cm² [104]. Because of the scattering properties of tissue, higher peak intensities than those for metals are needed to achieve laser ablation. Here we demonstrate that the 400 μ m core size GRIN probe with a peak intensity focus of 10^{13} W/cm² can provide enough power for tissue processing.

In the TPF imaging and FLA experiments presented in this section, the maximum scanning window is 100 μ m x 100 μ m. In this observation area, there is some drop of the intensity at the corners of the window, which can be compensated by readjusting the phase masks on the SLM so that the intensity of the focus spots is equally distributed in the scanning area as previously described for the laser ablation of gold patterns through MCFs. Our probe can provide TPF images (using 7.4×10^{10} W/cm² intensity) of a wide field of view (FOV) of a fluorescently-stained Organ of Corti (prepared as detailed in the paragraph 3.2.2), which contains precisely arranged rows of hair cells (Figure 26 a). We show that cellular level modifications can be obtained by digitally zooming in the area of interest (Figure 26 b) based on the TPF image in Figure 26 a. By increasing the power above the ablation threshold for biological samples to 6.9×10^{12} W/cm² [104], the beam destroys the selected cell. Afterwards, a second TPF image (Figure 26 c) confirms ablation of the targeted cell. Importantly, the laser ablation selectively induced damage only in the desired area, without affecting the TPF signal from the surrounding cells. This suggests that the ablation process is limited to the cell area region approximately 7 μ m x 7 μ m. In our experiments, the TPF image acquisition time and the scanning time for laser ablation is mainly limited by the SLM refresh rate, which is at 20 Hz. As a result, to scan the area of a cell (7 μ m x 7 μ m) it takes about 1 second.

To confirm that FLA indeed results in specific cell removal as opposed to simple photo-bleaching of the specimen in the scanning area, we set a bright field imaging system in transmission at the distal side of the fiber (Figure 14). We observe that at the FLA intensity levels, a bubble is formed in the sample as shown in Figure 27 [104]. The bubble appears exactly at the location of the image scanned with the high intensity focus spot, but no bubble is present in the focus intensities used for TPF imaging. Since the sample is mounted in a supporting water-based medium, the bubble cannot easily escape and stays trapped until it collapses. As a consequence, in some cases, TPF imaging is not possible immediately after the ablation since the bubble prevents the light from the distal end of the fiber to be focused on the sample surface. We observed the time evolution, over 5 minutes, of the formation and disappearance of the bubble created after FLA using the BF imaging arm.

There have been several studies related to the cavitation bubble dynamics after laser ablation of a material [104,150]. The mechanical properties and the laser parameters can affect the time evolution of the bubble [104]. We observe that the time needed until the bubble collapses varies from few seconds to few minutes even for areas of the same sample. Further studies are needed to carefully evaluate the phenomena taking place in the sample when interacting with the laser beam.

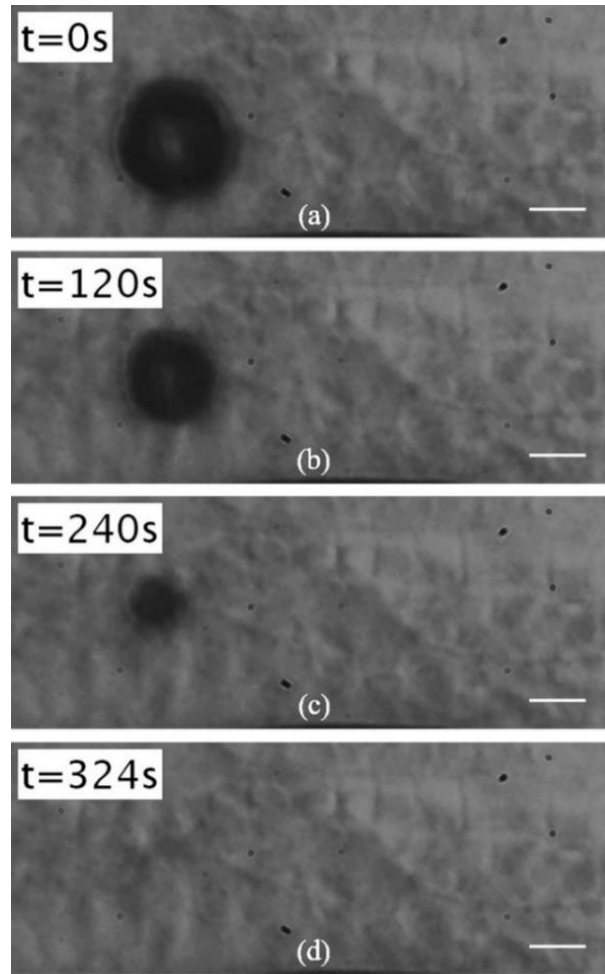


Figure 27. BF images that show the time evolution of the bubble generated after FLA of a hair cell within the organ of Corti. $t=0$ s is the time just after the FLA is finished. The scale bar in all the images is $10\text{ }\mu\text{m}$ [15].

4.2.3 FLA of a various materials through GRIN fibers

The high peak power focus $1.5 \times 10^{13}\text{ W/cm}^2$ delivered through the GRIN fibers discussed above, opens the path for processing materials with a wide range of ablation thresholds, from metals and polymers ($\sim 10^{10}\text{ W/cm}^2$), tissue ($\sim 10^{12}\text{ W/cm}^2$) and glass or ceramics ($\sim 10^{13}\text{ W/cm}^2$). For demonstration purposes, FLA is performed through the $400\text{ }\mu\text{m}$ to generate patterns on glass slides (Figure 28). Glass is among the material characterized by the highest FLA threshold. Glass laser ablation through fibers has only been reported using HC-PCFs. Debord et al. demonstrated glass processing using the fundamental mode of a Kagome HC-PCF with no distal components [146]. However, the feature size in their case directly depends on the fiber mode diameter, which is about $50\text{ }\mu\text{m}$ for a 19-cell Kagome HC-PCF design. On the other hand, the GRIN fiber probe cannot only ablate glass but the feature sizes can be also very small in the order of $2\text{ }\mu\text{m}$ or less as shown in the Figure 28 b [151]. The square pattern in Figure 28b is made by 21×21 spots with $10\text{ }\mu\text{m}$ distance. This findings bring the MMFs in the forefront as potential manufacturing tools of high precision.

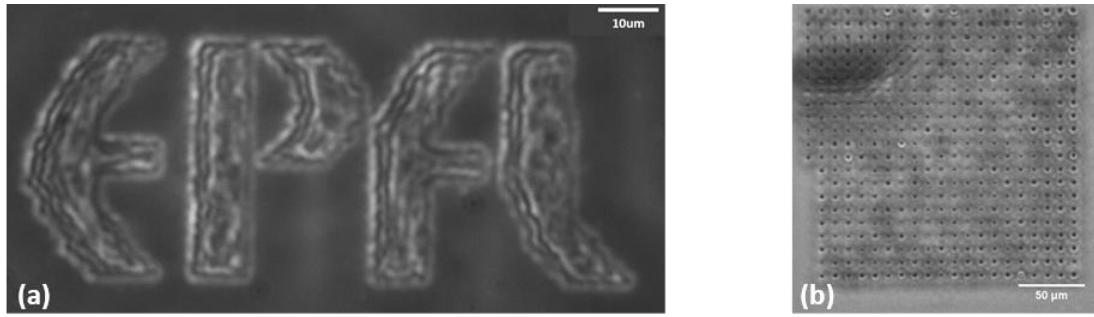


Figure 28. a) EPFL logo and b) a square of spots ablated on a glass slide

4.3 Conclusions

Focusing of femtosecond high-power pulses was demonstrated in this chapter using two different types of optical fibers, MCFs and MMFs. We explored the performance and limits of each system in terms of total peak intensity of the generated focus spot. GRIN fibers show higher focusing efficiencies than MCFs which allows the use of thinner cores (200 μm) to achieve the same peak intensity levels. This is an important finding for the realization of ultrathin non-invasive probes. In particular, the 400 μm core GRIN fiber shows impressive capability in delivering focused femtosecond pulses in space and time reaching peak intensities at the focus spot comparable to the ones reported only for HC-PCFs. By controlling the intensity field at the distal end of the proposed fibers by optimally modulating the wavefront at the proximal side using the transmission matrix method, the MCFs and MMFs essentially act as long, ultrathin objectives with a high aspect ratio of approximately 200:1. The main disadvantage of using a GRIN fiber endoscope is the post-calibration bending or misalignment of the system that would invalidate the measured transmission matrix of the fiber. To avoid that, the use of a fiber cannula could be a potential solution to prevent the fiber from bending and manipulate the endoscope as a thin needle. Even in this case, the final endoscope size would be in the order of about 1mm while offering FLA and TPF imaging capabilities with high spatial resolution. The ability of all the studied systems in providing enough power for micro structuring of materials together with the imaging modality brings MCFs and MMFs in the foreground of the endoscopic technology.

PART B:

Deep learning in fiber endoscopy

Chapter 5 Deep Neural Networks

“Inventors have long dreamed of creating machines that think. This desire dates back to at least the time of ancient Greece. The mythical figures Pygmalion, Daedalus, and Hephaestus may all be interpreted as legendary inventors, and Galatea, Talos, and Pandora may all be regarded as artificial life” (“Deep learning” by Ian Goodfellow, Yoshua Bengio, Aaron Courville [152]).

Artificial intelligence (AI) today refers to the execution of tasks that require human intelligence by a computer in order to achieve machine-based decision-making using data acquired from the environment through different sensor types [96,152]. Remarkably, AI algorithms were observed to easily perform tasks that a human would have difficulties to achieve when these tasks could be mathematically formulated, while more intuitive tasks, which were automatically perceived by a person were extremely difficult to teach to an AI system. Machine learning (ML) and AI are terms often used interchangeably but it is important to note that ML is a subset of AI. In contrast to hard-coding involved in AI, which is needed to fully describe the problem in a mathematical way so that it is solved by the computer, ML approaches the process in a different way, allowing the computer itself to learn and extract the important features of the given raw data [152]. As the definition suggests, *learning* makes the ML system dynamic and capable to optimize itself based on data to which it is exposed and enables problem-solving and decision-making for more subjective tasks.

Deep learning is a subset of the ML methods, which uses the idea of hierarchy in feature extraction and recognition. While the conventional ML techniques were initially based on a hand-crafted feature extraction that was afterwards used in the trainable algorithm for achieving various tasks, deep learning methods can learn to recognize the representations existing in the raw data and decompose the information using simple nonlinear functions. Therefore, deep learning algorithms remove the considerable work needed for engineering the data before they fed in the code and make possible the problem solving for abstract tasks that AI could not overcome for years [96].

The origin of the representation-learning idea on which deep learning methods are based, is the function of the visual cortex of the brain. When a human looks at a natural scene or object, the eye functions as a “sensor” that records “raw data”. During years of training and learning by being exposed to various signals, the brain synapses are formed and connected in a very complex network, which analyzes a “captured” raw image in a sequence of steps using different representations. Although all the details about brain functionality are not yet fully understood, the accepted idea is that the recognition action starts from decomposing a signal into low-level

features such as “edges” and gradually moves to higher-level features representing more abstract concepts such as “a human face”. The term “deep” in deep learning refers to the total number of the computational layers implemented for feature extraction in an algorithm. Deep learning algorithms are constructed by nested layers of mathematical operations that are connected to each other in a nonlinear way; these are widely known as deep neural networks (DNNs). Deep learning assists high-dimensional data interpretation in various fields of not only science but also, business, government and entertainment [96,152]. Someone could say that nowadays, deep learning is part of most of the social media and digital devices closely associated with the modern daily life. In this chapter, we first discuss the concepts of DNNs applied to computational imaging problems. In the subsequent chapters, we apply DNNs to fiber endoscopy as an alternative to the wavefront shaping methods that were investigated in the Part A of the thesis.

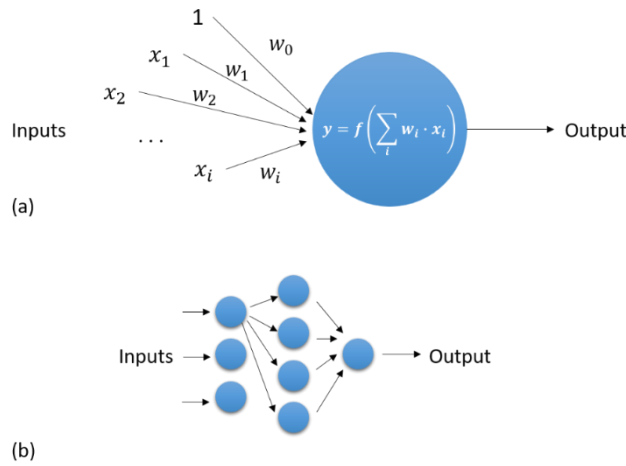


Figure 29. Schematic description of a) a single perceptron, b) a multilayer perceptron neural network.

5.1 Basic concepts in Artificial Neural Networks

5.1.1 The perceptron

In this paragraph, we introduce some of the key components of artificial neural networks (ANNs) and some concepts that are extensively used in Part B of this work. As mentioned in the previous paragraph, the idea of ANNs is inspired by biological neural networks and their functionality. Simply explained, brain neuron cells consist of a main cell body where the nucleus is found, dendrites which receive information from other neurons and axons, which transfer the signal to the axon terminal based on electrochemical processes. When the input signal to a neuron exceeds a certain threshold, the neuron fires and the signal is transmitted. Artificial neurons, often called perceptrons, are mathematical functions that try to model this process. A perceptron receives inputs, which have some assigned weight values, sums them up and then applies a nonlinear function that produces an output (or not) depending on the sum of the inputs, similarly to a

biological neuron firing (or not) based on its input signal. Figure 29a visually describes the perceptron concept: the perceptron maps an input vector $\mathbf{x} = (x_1, x_2, x_3, \dots, x_i)$ multiplied with a weight coefficient vector $\mathbf{w} = (w_1, w_2, w_3, \dots, w_i)$ to an output y through a function f . In addition to the weighted inputs, the function f also receives an extra value known as the bias (w_0), which is independent of the input values and creates a shift in the axis of the output.

A single perceptron is generally not sufficient to handle a complex nonlinear mapping problem. Usually, many perceptrons are organized in layers as shown in Figure 29b. Each perceptron or node of a layer can be connected to some or all of the nodes of the second layer, the nodes of the second layer can be fully or partially connected to the third layer and so on. The outputs of each layer are defined similarly to that of the single perceptron. Equation 23 gives the output of the k unit of the m layer of a multilayer structure. The summation is made over the total number of nodes of the previous layer to which a corresponding node is connected. Only the input layer is not weighted [99]. The ANN architectures for which each node of a layer is connected to all the nodes of the previous layer are called fully connected networks (FCNs). FCNs are computationally heavy when it comes to datasets of high-dimensionality. For example, an RGB image input with 512x512 pixels to a fully connected layer produces 786,432 weights all connected to the second layer. For this reason, convolutional neural networks (CNNs), a variant of DNNs described in details in a later sections are generally used. Unlike FCNs, in CNNs each node of a layer is connected to a small area of neurons in the previous layer sharing the same set of weights leading to a more computationally efficient algorithm.

$$y_k^m = f \left(\sum_i w_{km}^{m-1} y_i^{m-1} + w_{k0}^{m-1} \right)$$

Equation 23. Output of one node of a multilayer perceptron.

5.1.2 Activation functions

The nonlinear function f within each node is called the activation function and it can be defined using a range of mathematical equations. In these equations x is the summation of the weighed inputs and bias. The most commonly used activation functions, which are also those used in the DNNs implemented in this thesis, are the following:

Step function:

$$f(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

Softmax function:

$$f(x) = \frac{e^x}{\sum_{j=1}^N e^x}$$

Sigmoid function:

$$f(x) = \frac{1}{1 + e^{-x}}$$

Hyperbolic tangent:

$$f(x) = \tanh x$$

Rectified linear function:

$$f(x) = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

Leaky rectified linear function:

$$f(x) = \begin{cases} x & \text{if } x \geq 0 \\ 0.01x & \text{if } x < 0 \end{cases}$$

The activation function determines if a neuron fires or not and the value of the neuron output that will reach the corresponding nodes in the next layer of the stack. Intuitively, the activation functions limit the output values of each neuron within a range so that the computation is more efficient. For example, the sigmoid function imposes a limit for input values towards infinity to create an output of ± 1 . Therefore, the choice of the activation function is not trivial and has to be optimized based on the task the ANN needs to perform.

5.1.3 Error functions

A perceptron is considered “trained” when it produces realistic outputs. In order for it to do so, the weights and the bias need to be optimized. The process of optimization of the weights of a DNN is called training. Initially the weights of the inputs reaching the perceptron are random values. In supervised learning schemes, such as those studied in this thesis, when a first forward pass is attempted for an input through a perceptron, an output is y_{pred} is predicted. This output is compared to the true value y_{true} via an error function $E(y_{pred}, y_{true})$ to calculate how far the predicted value lies from the real values. The term supervised refers exactly to the case in which the real value, also known as ground truth is available for monitoring the performance of the algorithm. The weights of the perceptron are considered optimal when the error is minimized. Some of the common error functions or loss functions (in other words) are the following:

Mean squared error (MSE):

$$MSE = \frac{1}{N^2} \sum_i^{N^2} (y_{pred} - y_{true})^2$$

Mean absolute error (MAE):

$$MAE = \frac{1}{N^2} \sum_i^{N^2} |y_{pred} - y_{true}|$$

Cross entropy:

$$CE = \sum_{j=1}^N \sigma(p) \log(p) + (1 - p) \log(1 - p)$$

In the case of the cross entropy (CE) loss function, p is the probability that the values y_{true} and y_{pred} agree for each pixel in the case of an image processing problem. $\sigma(p) = e^{p_k} / \sum_j e^{p_j}$ for all the available cases k . The minimization of the error function is achieved using gradient-descent based error backpropagation. The backpropagation algorithm calculates the partial derivatives of the error function for all the weights of the specific neuron/perceptron or layer of neurons to find out which of the weights most affects the output predicted value. Therefore, in each backpropagation pass, the weights are updated in a way that minimizes the error function. Both the weights and the bias are trainable parameters that are tuned to optimize the performance of

the perceptron for a specific task. Each iteration for updating the weights of a learning unit is called an epoch. The change in the weights per epoch for one perceptron is described by Equation 24, where E is the error function and η the learning rate. The implementation of backpropagation is not a trivial task; however, it is already integrated in basic deep learning platforms such as Tensorflow or Pytorch. In this work, the implementation of the DNN algorithms is performed in Keras, a high-level application programming interface (API) that works on top of Tensorflow in a more intuitive and user-friendly way.

$$\mathbf{w}(\text{epoch} + 1) = \mathbf{w}(\text{epoch}) + \eta \left[\frac{\partial E}{\partial \mathbf{w}} \right]$$

Equation 24. Perceptron weights evolution per iteration step

5.2 Convolutional neural networks

CNNs, as implied by their name, use a convolution operation to extract information from an input and perform a desired mapping to an output. Instead of the weight vectors used by the perceptrons and fully-connected networks, CNNs instead use 2D weight matrices (also called convolution kernels). The elements of these small matrices are updated during the training process based on the same idea described for the fully-connected layer networks to minimize the value of the loss which is calculated by the employed error function. Once the CNN is trained, each kernel can be interpreted as extracting certain features from the input signal. The size of the kernels and the kernel stride used in the convolution should be tuned for each mapping problem. As kernel size we define the size of the 2D matrix of the kernel and the kernel stride refers to the number of pixels that the kernel shifts when convolved with the input image. For example, in the experiments described in the following chapters we use 3x3 kernels with stride 1 in rows and columns.

The most common layers of a CNN consist of a convolutional layer, an activation layer and a maxpooling layer. In the first convolutional layer the input is convolved with all the kernels in the first filter bank, generating another matrix known as a feature map [96,99,152]. In a second step, the activation layer applies a nonlinear function to each of the feature maps in a similar manner to the perceptron described above. The most commonly used activation function is a rectified linear unit (ReLU) that removes all the negative values in the matrix elements of each feature map. Finally, maxpooling is applied at each feature map. Maxpooling is a downsampling operation in which the maximum value of each local patch of elements of the feature map matrix is retained. This process flow is intuitively depicted in Figure 30 for two example kernels that produce two very different feature maps after convolution: edge enhancement (first row) and Gaussian blur (second row).

Following the first block of layers, a second block is performed using the same operations. New feature map sets are generated while the dimensionality is further decreased. As nicely explained

by *LeCun et al.*, each convolutional layer detects local conjunctions of features from the previous block of layers, while the pooling layer merges semantically similar features [96]. These blocks can be repeated several times depending on the needs of the problem, thus determining the depth of the desired CNN. Finally, the training step for a CNN is no different from the training used for the perceptron or FCNs: error backpropagation is used to update the weight elements of the kernels aiming to minimize the error between the prediction of the CNN and the ground truth.

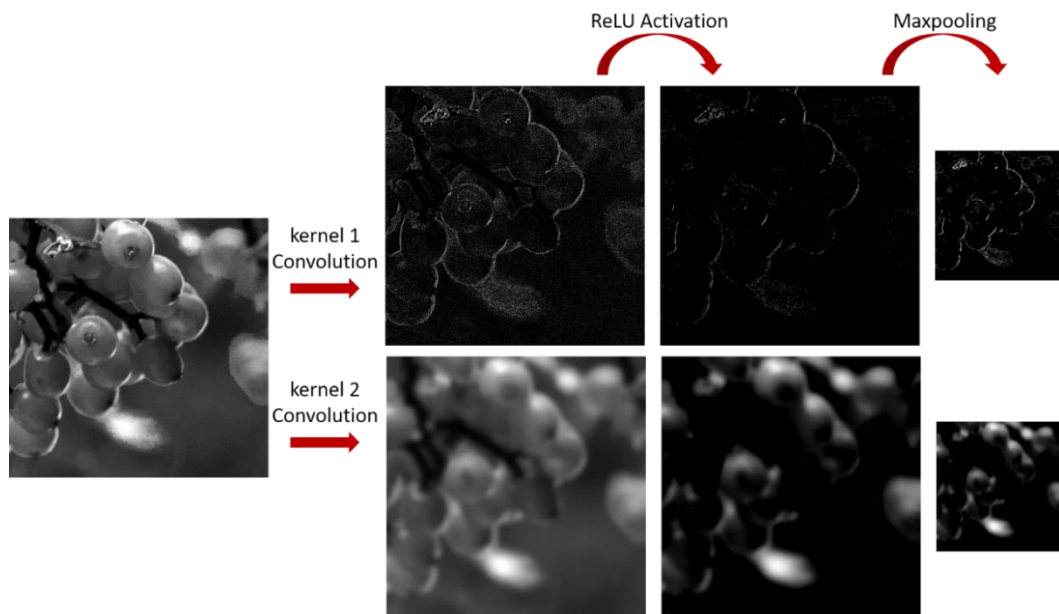


Figure 30. Example of a convolutional block. One input image is convolved with two different kernels, processed by a ReLU activation function and then downsampled to half the size using maxpooling.

5.3 Computational imaging

Computational imaging refers to the process of forming an image of an object using algorithms that rely on raw measurements [108,153,154]. The estimate of the object of interest is generated by solving an optimization problem. The idea of computational imaging algorithms is that inherent information exists in the raw image, which cannot be directly observed or quantified, can be recovered. Computational imaging offers multiple approaches for obtaining a high-quality representation of an object and has many applications in microscopy, sensing, photography and more. For the following discussion, we split the field of computational imaging into two categories: conventional computational imaging and ML-based computational imaging.

We use the term conventional computational imaging to describe an algorithm that receives the raw data obtained from a measurement and attempts to infer information about the object of the measurement using prior knowledge about the physics of the system to build a correct model. One example is the enhancement of an image recorded by an optical apparatus of poor quality

(e.g. containing aberrations and/or scattering) using an algorithm that models the propagation of light through the system in order to recover a high quality image. Another application of conventional computational imaging algorithms is optical diffraction tomography, where a physical forward model is used to determine the 3D refractive index distribution of an object from the measurements of the scattered field measured on a detector. The Rytov and Born approximations are two commonly used physical models in tomography [108,109]. Model-based computational imaging is characterized by computation time ranging from few minutes to hours depending on the complexity of the problem. However, this computation time corresponds to the optimization of a single imaging operation, meaning one set of raw data or one image, and a new optimization round is required for each new input [99].

ML-based computational imaging does not require knowledge about the physics priors of the problem; rather it is based on interpolating information trained on various known input-output example pairs. ML algorithms show significant robustness for interpreting images obtained in noisy conditions when compared to conventional computational imaging methods. Advances on a hardware level, in particular the development of GPUs (graphical processing units) have brought DNN-based machine learning in the forefront of current research for image processing. In terms of temporal performance, DNN training time can range from hours to days depending on the complexity of the problem and hence the DNN architecture; however, in contrast to conventional methods, once the training is completed the time to process a single new input is in the order on milliseconds, approaching real-time image processing for many applications [99]. It is important to mention that ML-based computational imaging often requires a large quantity of data for the training of the ML algorithm which is not the case for the conventional imaging models. Nevertheless, synthetic data augmentation can often provide a solution for digitally enlarging small datasets.

DNNs have recently been implemented for more and more computational imaging problems. In microscopy, one of the most common challenges in computational imaging is the improvement of resolution. Recent studies have obtained super-resolution from wide-field microscopy by training a DNN with corresponding pairs of images obtained from a low resolution system (low NA microscope objective) and from a high-resolution one (high NA microscope objective) [155]. In this ML approach the DNN produces high quality images by upsampling and deblurring the low resolution image. The same method can be used for other imaging methods such as fluorescence microscopy [99,156,157]. Another challenging problem in optical microscopy is the imaging of a phase object and the retrieval of its complex field information. Several iterative optimization solutions have been proposed in the past to recover the full field information from raw intensity measurements using conventional computational imaging algorithms such as Gerchberg-Saxton and Fienup [153]. Compared with these approaches for phase retrieval from intensity measurements, DNNs show high performance and, if combined with constraints imposed by the physics of the studied system, demonstrate impressive results [158,159]. In order to further promote the possibilities of current biological imaging inspection, deep learning based translation among microscopy methods has been proposed [156,158–160].

The potential of DNNs to extract desired information from raw images captured on a camera has been further extended in endoscopy. Imaging through multimode fibers using intensity measurements of their speckle patterns has been proposed by our research group and others and is discussed in second part of this thesis [59–61,63,65,94,95,100,161,162]. In addition, pixelation-free imaging through fiber bundles using DNNs is demonstrated [64,73]. We show in the following sections that the advances reported for DNNs in microscopy and fiber imaging can be combined to serve a powerful role in the evolution of endo-microscopy applications.

Chapter 6 Seeing through multimode fibers using deep learning

The great potential of the MMFs and MCFs for endoscopic imaging is widely recognized by the optics community, as it is clearly stated by the numerous works reported in literature. In the Part A of this thesis, we demonstrated that wavefront shaping methods can advance the role of MMFs and MCFs further than simple imaging tools by incorporating micro-surgical possibilities. However, perturbations of the system in such calibration-based techniques pose practical difficulties in the presence of temporal changes in the speckle patterns for a constant input, since it results in the decorrelation of the measured transmission matrix of the system rapidly with time.

Aiming to a more robust solution for recovering the information through the different types of endoscopes, the capabilities of DNNs are explored. There have been two previous reports on the use of artificial neural networks (ANNs) for recovering the images transmitted through MMFs in the previous decades [94,95]. In these early demonstrations, two-layer networks were trained and were able to recognize a few (~ 10) images after a 10 m long step-index fiber from the intensity images of their corresponding speckle patterns. In this work, we use state-of-the-art DNN architectures for not only classification purposes but also for reconstruction of the input image from the speckle output recorded by a camera. In addition, a number of parameters is discussed concerning their effect in the DNN performance such as the fiber length, the number of supported spatial modes, the content of the input image (amplitude and phase patterns) etc. This novel application of the DNNs in the field of optics for information transmission through MMFs is studied showing results with large potential impact in various fields. The studies described in this chapter are published in the following journal and conference articles:

- N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, "*Learning to see through multimode fibers*", *Optica* **5**, 960–966 (2018).
- E. Kakkava, B. Rahmani, N. Borhani, U. Tegin, D. Loterie, G. Konstantinou, C. Moser, and D. Psaltis, "*Imaging through multimode fibers using deep learning: The effects of intensity versus holographic recording of the speckle pattern*", *Optical Fiber Technology* **52**, 101985 (2019).
- E. Kakkava, N. Borhani, C. Moser, and D. Psaltis, "*Deep neural networks for seeing through multimode fibers*", in *High-Speed Biomedical Imaging and Spectroscopy IV*, K. Goda and K. K. Tsia, eds. (SPIE, 2019), p. 46.
- E. Kakkava, N. Borhani, B. Rahmani, U. Tegin, C. Moser, and D. Psaltis, "*Efficient Image Classification through a Multimode Fiber using Deep Neural Networks in presence of*

Wavelength Drifting," in *Imaging and Applied Optics 2019 (COSI, IS, MATH, PcAOP)* (OSA, 2019), p. CW1A.4.

- E. Kakkava, N. Borhani, B. Rahmani, U. Teğın, U. Teğın, C. Moser, and D. Psaltis, "Wavelength Independent Image Classification Through A Multimode Fiber Using Deep Neural Networks," in *2019 Conference on Lasers and Electro-Optics Europe and European Quantum Electronics Conference (2019), Paper Ci_2_1* (Optical Society of America, 2019), p. ci_2_1

6.1 Information recovery through multimode fibers using Deep Neural Networks

For the first time DNNs are introduced to reconstruct and classify input patterns through a MMF by receiving the intensity-only images of the corresponding speckle patterns. The dataset used to assess the performance of the DNNs was generated by projecting on the proximal fiber side images of handwritten digits [59,60,66]. These images are available online in the MNIST database, which is widely used for testing the capabilities of different neural network architectures [163]. DNN architectures with up to 14 hidden layers were trained on a database of 20,000 handwritten digits-speckle image pairs. Recognition or reproduction of an image launched at the proximal end of the fiber was achieved by detecting only the light intensity at the distal end facet. The performance was evaluated for different fiber lengths up to a maximum of 1 km.

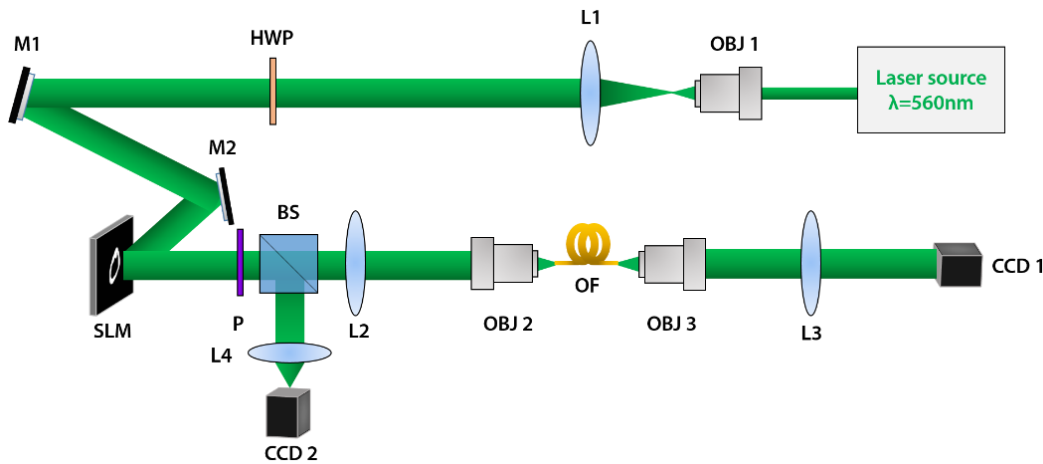


Figure 31. The experimental setup for pattern transmission through the MMF. Amplitude or phase images are projected on the fiber facet by imaging the SLM by means of a 4f system (L2-OBJ2). A CCD camera is recording the corresponding output magnified by the 4f system OBJ3-L3. Another camera (CCD2) records the output field from the SLM. (L = lens, OBJ = microscope objective, HWP = half-wave plate, M = mirror, PBS = polarizing beam splitter, BS = beam splitter)

6.1.1 Methods

6.1.1.1 Optical setup

The optical system used to collect the data is shown in Figure 31. The laser beam of a 560nm wavelength diode laser is used to illuminate a graded-index (GRIN) MMF with 62.5 μ m core diameter and numerical aperture (NA) of 0.275 (GIF625, Thorlabs). The fiber supports approximately 1800 spatial modes at the specific wavelength. A beam expansion is achieved by a set of lenses (L1: $f=200$ mm, OBJ1: 10x, Newport) placed at the laser output to ensure that the beam size is large enough to illuminate spatial light modulator (SLM). A phase-only SLM (SLM, 1920x1080 pixels, Pluto-Vis, Holoeye) is used to display the desired patterns and the SLM plane is imaged onto the proximal facet of the MMF by means of a 4f imaging system (lens L2: $f=400$ mm, microscope objective OBJ2: 60x, Newport). Another 4f system (same as the proximal one, OBJ3-L3) is placed at the distal end of the fiber to image the speckle pattern emerging from the distal facet on a CCD camera (CCD, Chameleon 3, 1024x1280 pixels, Mono, Point Grey). An additional lens (L4: $f=200$ mm) in 2f configuration allows a second camera at the proximal side to monitor the images reflected by the SLM. A halfwave plate (HWP) and a linear polarizer (P) are placed before and after the SLM (see Figure 31) respectively in order to test both phase and amplitude patterns as inputs to the GRIN fiber.

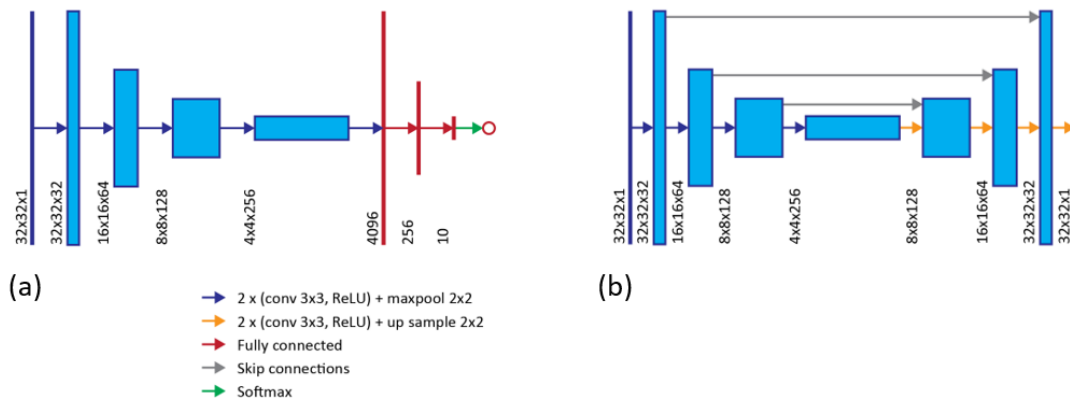


Figure 32. Details of the implemented (a) VGG type image classifier and (b) U-net type image reconstruction convolutional neural networks [60].

6.1.1.2 Deep neural network architectures

A ‘VGG’ type CNN, as developed by Simonyan & Zisserman [98] is used to classify the distal speckle images or the reconstructed SLM input images (Figure 32 a). This network consist of a convolutional front end with downsampling for encoding, and a fully connected back end for classification; see Figure 32a for details. The use of such deep CNN with very small filter kernels has been shown to provide high image classification accuracies. A ‘U-net’ type CNN with 14 hidden layers, as developed by Ronneberger et al. [164], was used to reconstruct the SLM input image from the recorded distal speckle intensity pattern (Figure 32 b). This nearly symmetric network

architecture comprises a convolutional encoding frontend with downsampling to capture content, and a deconvolutional decoding backend with upsampling for localization; see Figure 32 b for details. Skip connections copy feature layers produced in the contracting path with features layers in the expanding path of the same size, thus improving localization. For training both networks, the obtained 20,000 distal speckle pattern images were randomly split into 16,000 training, 2,000 validation, and 2,000 testing sets. The training sets were processed in 50 and 500 image batches for the reconstruction and classification networks, respectively, with batch shuffling to minimize over fitting. An Adam optimizer with a learning rate of 1×10^{-4} was used to minimize a mean square error loss function. The networks were trained for a maximum of 50 epochs. For each case, training was carried out 10 times to provide statistics for the training accuracies. The DNNs were implemented using the TensorFlow 1.5 Python library on a single NVIDIA GeForce GTX 1080Ti graphics processing unit.

6.1.2 Image reconstruction

In our experiments, the patterns generated by the SLM were handwritten digits from the MNIST database. Before processed by the DNN, each image recorded by CCD1 or CCD2 is cropped to a 1024×1024 pixels window centered on the digit and the speckle respectively. The cropped images were then downsampled to 32×32 pixels using bilinear interpolation and used as input for the DNNs. An example of the projected digits at the proximal fiber facet is shown in Figure 33, where the digits zero and four are shown for both amplitude (Figure 33 c-d) and phase modulation (Figure 33 e-f) along with the corresponding speckle patterns captured at the distal fiber end for the GRIN fiber with 2cm length. The speckle patterns (Figure 33 g-h) look similar to one another because their appearance is dominated by the DC component of the light from the SLM. However, when we subtract the intensity patterns (Figure 33 d and h) corresponding to the two digits, we reveal that there is a significant difference (Figure 33 i). It is the purpose of this study to understand if this small difference is sufficient for the DNN to distinguish the two inputs. The results presented in the remainder of the paragraph are obtained by adjusting the SLM so that the patterns entering the fiber are either phase only or amplitude modulated images of the digits.

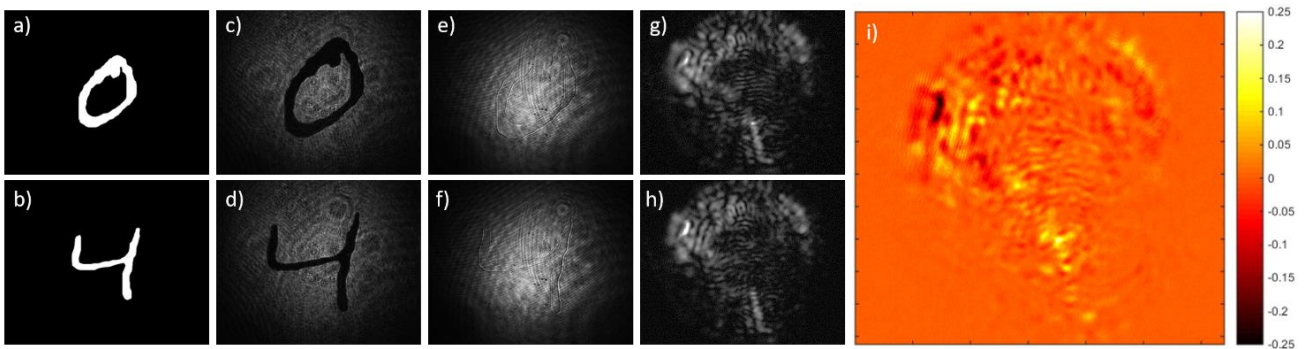


Figure 33. Images of the digits 0 and 4: a-b) input pattern on the SLM, c-d) amplitude modulated output from the SLM, e-f) phase modulated output from the SLM, g-h) speckle patterns of each digit respectively for amplitude inputs and i) the difference between the speckle patterns g and h. [60]

In a first step, the ability of our DNN to reconstruct the input digits from the distal speckle intensity patterns was tested. In Figure 34, we present the results of the reconstruction for the four different fiber lengths used in the experiments (0.02 m, 0.1 m, 10 m and 1 km) for amplitude and phase modulated inputs into the proximal facet of the GRIN fiber. Although appearing random, the speckle patterns contain information about the propagation of the input field through the fiber. In fact, the results confirm the above statement, showing that the recovery of the input is possible with an intensity-only image of the distal speckle pattern using the U-net CNN. Based on the reconstructed images obtained from our experiments (Figure 34), the fidelity of the reconstruction decreases from 97.4% for a 2 cm fiber to 90.0% for an 1 km fiber.

SLM input	0.02 m Fiber		0.1 m Fiber		10 m Fiber		1000 m Fiber	
	Amplitude	Phase	Amplitude	Phase	Amplitude	Phase	Amplitude	Phase
0								
1								
2								
3								
4								
5								
6								
7								
8								
9								
Accuracy	97.4%	97.6%	96.9%	97.2%	95.5%	96.2%	90.0%	88.0%

Figure 34. Examples and accuracies of the reconstructed SLM input images from the recorded distal speckle intensity patterns for amplitude modulated proximal inputs [60].

In the case of the 1 km long GRIN fiber, the speckle pattern at the distal end was unstable. Local temperature nonuniformities in the fiber induce changes in the optical path, due to both thermal expansion of the material and change of its refractive index. Thermal convection around the fiber can lead to drifting of the distal speckle pattern in time creating an extra “noise” on the acquired

speckle patterns acquired (Figure 35). Therefore, further care could be taken to thermally isolate the fiber and to maintain an isothermal environment, which might give an increase of performance. The high fidelity of the reconstructed SLM input images also show that this technique effectively denoises the system by removing artefacts associated with the optical setup. The fidelity was measured as the percent mean square error of the reconstruction compared to the input. For example, the network recovers the SLM input image shown in Figure 33a from the distal speckle intensity pattern shown in Figure 33g, while eliminating artefacts projected onto the proximal facet of the fiber, as shown in Figure 33 c and Figure 33 e, as well as subsequent artefacts due to flaws and dirt or even misalignments on the proximal facet of the fiber.

In the Figure 35, one can observe changes in the distal speckle pattern over a 2s period. A movie was recorded at a frame rate of 83fps for the 1km long GRIN fiber with a constant blank image as the SLM input. Although the proximal input does not change, the speckle intensity at the distal end of the fiber changes rapidly with time. This can be attributed to fluctuations of the ambient temperature or airflow over the optical setup that induce slight perturbations on the GRIN fiber that become significant over its 1km length. Therefore, changes on the distal output caused by the projection of different digits while the training dataset is acquired can be buried in the “noise” caused by the drifting of the speckle pattern. In order to test the effect of the drifting distal speckle patterns on the accuracy of the classifications, the Unet network was trained on the first 10,000 samples of images of an acquired dataset and tested on images from the second half (recorded several hours later); and vice versa. The results showed no significant change in the reconstruction fidelity. This suggests that the fluctuations seen in the video are not entirely random and the neural network has learned them.

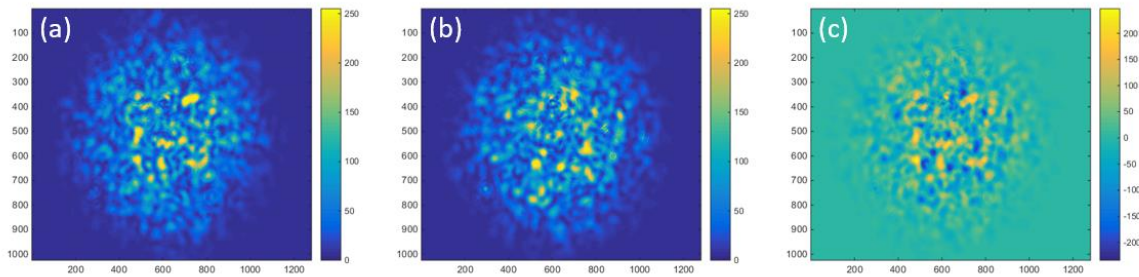


Figure 35. Speckle pattern decorrelation. a) Frame at 0s, b) frame at 2s and c) the difference between the two frames [60].

6.1.3 Image classification

Results for the classification of the distal speckle intensity patterns are presented in Table 2 and Figure 36. These show that the classification accuracy, defined as the percentage of correctly recognized digits, decreases with increasing fiber length for both amplitude and phase modulated proximal facet input modes. Generally, the accuracy decreases from 90% for a 2 cm fiber to 30%

for a 1 km fiber. For comparison, the classification accuracy of the VGG CNN used for the experimental data was calculated for the original SLM input digit images and found 98.4%. This decrease can be attributed to increased scattering losses, mode coupling, and drifting of the distal speckle pattern with increasing fiber lengths. The results also show that phase modulated input provides slightly better classification accuracies probably due to the more uniform distribution of the injected light across the fiber modes. In order to improve the classification accuracies, the neural network was also trained with the reconstructed SLM input images by the U-net. As shown in Table 2 and by the normalized confusion matrices in Figure 36a and b for the 1 km fiber, this provided a significant increase in classification accuracy. For the 1 km case, there is a general confusion between the 4 and 9s digits, and between the 3, 5, 6, and 8's. The similarities between these classes are also evident in the reconstructed SLM input images for the 1 km fiber shown in Figure 34. The confusion matrices for each fiber length and input type can be found in the Appendix A2.

Table 2. Classification accuracy for the four different fiber lengths using amplitude or phase input patterns. Classification was carried out on either the intensity image of the distal speckle patterns or on the reconstructed SLM inputs [60].

Fiber length [m]	Proximal input	Classification accuracy [%]	
		From distal speckle intensity	From reconstructed input
0.02	Amplitude	92.7 ± 0.5	98.1 ± 0.4
	Phase	95.1 ± 0.6	98.1 ± 0.3
0.1	Amplitude	90.7 ± 0.8	97.5 ± 0.5
	Phase	92.2 ± 0.7	97.5 ± 0.3
10	Amplitude	81.9 ± 1.6	96.5 ± 0.4
	Phase	87.2 ± 0.9	96.8 ± 0.5
1000	Amplitude	29.3 ± 5.5	69.9 ± 0.9
	Phase	22.4 ± 2.2	57.0 ± 1.0

Because of the instability of the speckle pattern intensity observed for the longer fiber system the mapping from input to output becomes a random mapping [165] and objects that are similar to one-another at the input are dispersed in the intensity measurement at the distal end. Therefore, the ability of the classifier-DNN to generalize (recognize objects it has not seen before) diminishes for longer fibers. When we first recover the input images with a U-net DNN, the random mapping is partially inverted and the classification network can recognize objects of the same class it has not seen before. This behavior is evidenced in where the classification performance of the VGG network when trained with the intensity of the raw speckle patterns is plotted as a function of iteration number during the learning process for the 10 m (Figure 37 a) and the 1 km (Figure 37 c) fibers. For the 10 m fiber the classification accuracy is the same for the training and validation sets. On the contrary for the 1 km fiber, in steady state, the training set is memorized well but the

validation set is classified accurately only 29.3 % of the time. The discrepancy in recognition rate between the training set and validation sets is an indication that the network is not able to generalize well. Also shown in Figure 37 are the learning curves when training the VGG network with the reconstructed images from the U-nets. In this case, the recognition rate is the same for the training and validation sets. In general, we can improve the recognition rate on the validation and test sets while decreasing the performance on the training set by reducing the number of weights in the network and/or increasing the size of the training set.

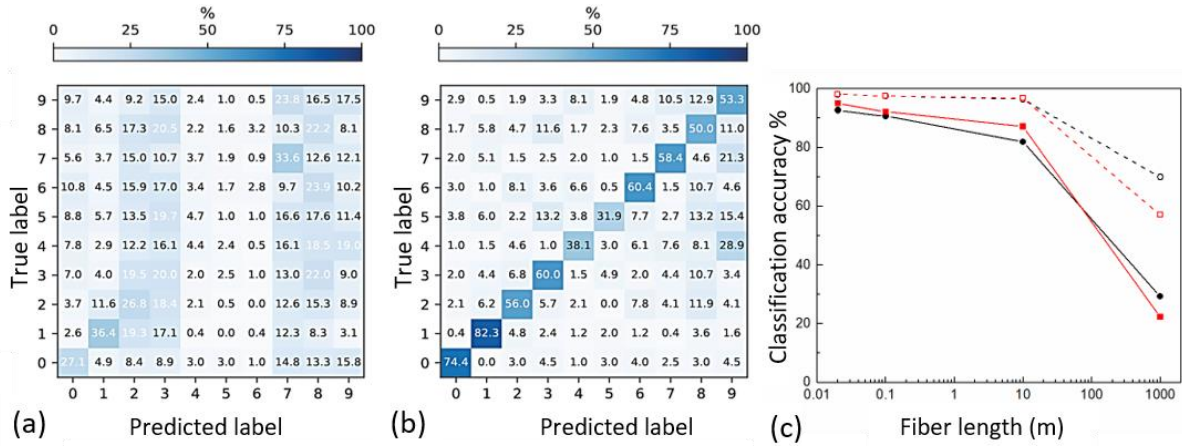


Figure 36. Normalized confusion matrices for the classification of the reconstructed SLM input images for the 1 km GRIN fiber for a phase modulated proximal input a) from the speckle pattern images and b) from the reconstructed by U-net images. c) Classification accuracy for different fiber lengths (Solid line: The inputs to the VGG-CNN are the recorded speckles at the distal fiber end, Dotted line: The inputs to the VGG-CNN are the reconstructed images obtained by the U-net-CNN, Circles: Amplitude input images Squares: Phase input images) [60].

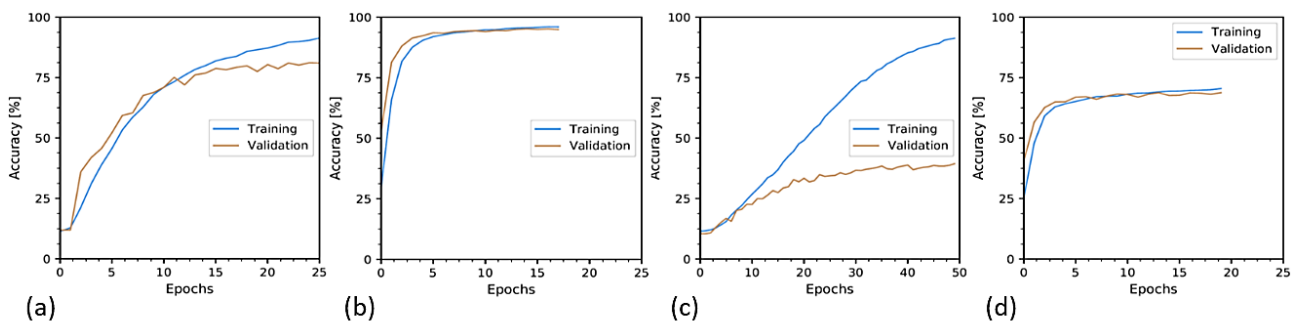


Figure 37. Training and validation classification accuracies as a function of epoch for the (a) 10m fiber distal speckle intensity pattern, (b) 10m fiber SLM reconstructed input, (c) 1km fiber distal speckle intensity pattern, and (d) 1km fiber SLM reconstructed input [60].

The recognition or reconstruction of the field at the proximal end of a MMF from complex field measurements at the distal end can be considered as an alternative to the DNN based inversion methods described in this chapter. The simplicity of intensity only detection though, is a clear

advantage in practice. At the same time, linear inversion methods (such as the transmission matrix) learn the fiber not the inputs. In other words, any input can be recognized or reproduced. DNNs on the other hand are trained on a class of objects and rely on statistical averaging within that class. In principle, the performance of the transmission matrix method should be independent of fiber length. However, as the fiber length increased additional background noise accumulates at the output because of scattering at the core for the fiber and the core-cladding interface. In addition, the temperature and mechanical instabilities that contaminate the measured data are to some extent learned by the DNN, whereas they directly degrade the reconstructions of coherent methods. Finally, the neural network can be directly trained to reproduce or recognize the versions of the input images as they are stored in the computer. Any nonlinearities, aberrations, speckle, pixelation, phase wrapping, or other distortions that are introduced before the light enters the input facet of the MMF (i.e. Figure 33 a versus Figure 33 c) are conveniently accounted for.

6.2 Intensity-only versus holographic data recording

In the previous discussion, we demonstrated that DNNs model nonlinear systems and therefore they can be used with intensity measurements rather than holographic recordings since they can accommodate the square law nonlinearity between input and output. In this section, more detail study comparing holographic to intensity-only recording is presented to verify if the lack of the full field information affects the performance of the DNNs to classify the input information after propagation through the MMF.

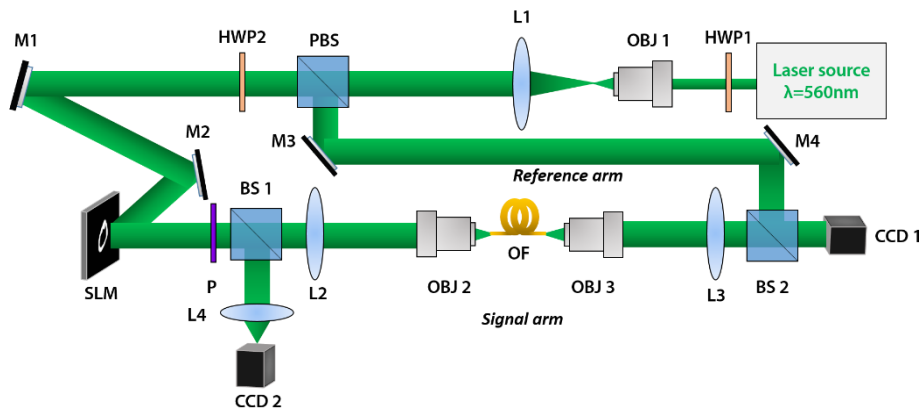


Figure 38. Optical set for the collection of the datasets consisting of fiber input-output image pairs. Amplitude or phase images are projected on the fiber facet by imaging the SLM by means of a 4f system (L2-OBJ2). A CCD camera is recording the corresponding output magnified by the 4f system OBJ3-L3. A reference arm is also directed to CCD1 where it interferes with the fiber output forming a digital hologram. Another camera (CCD2) records the output field from the SLM. (L = lens, OBJ = microscope objective, HWP = half-wave plate, M = mirror, PBS = polarizing beam splitter, BS = beam splitter) [59].

6.2.1 Methods

The optical setup described in the previous paragraph (Figure 31) was modified to allow holographic recording of the MMF output. A polarizing beam splitter (PBS) is added in the laser path after the beam collimation, splitting the laser beam into two paths; the illumination and the reference path as shown in Figure 38. A half waveplate (HWP1) placed before the PBS adjusts the energy ratio of each path for achieving the desired contrast for the speckle hologram. An extra beam splitter (BS2) is added to combine the reference and the fiber output field on the detector to generate the hologram image. The illumination path configuration is the same as in Figure 31. The laser beam in the signal arm is modulated by a phase-only spatial light modulator (SLM, Pluto-NIR2, Holoeye) and imaged onto the proximal fiber facet by means of a 4f-system. A second 4f system is placed at the distal side to image the fiber output on a CCD detector (CCD1, Chameleon 3, 1024x1280 pixels, Mono, Point Grey). For monitoring the input on the fiber, we also include a 4f-system to capture the SLM output on a second camera (CCD2).

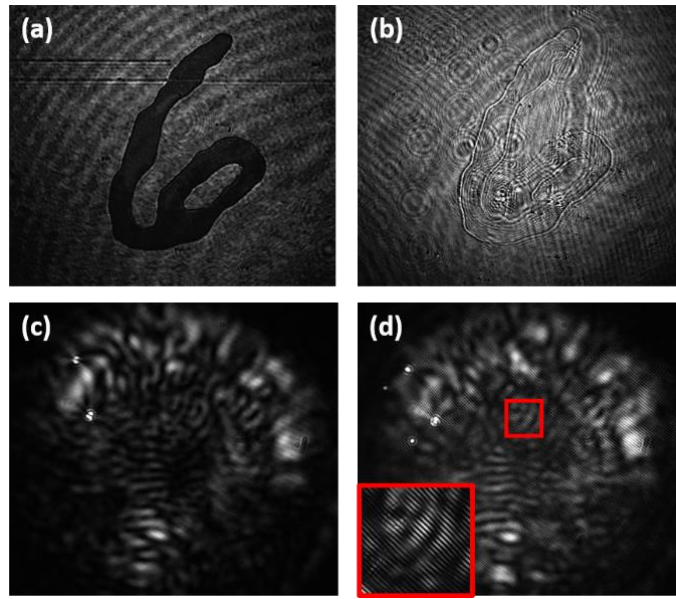


Figure 39. Image of the SLM output on CCD2 for the digit 6 for a) amplitude and b) phase modulation. Image of the speckle pattern generated at the fiber output for input (b) when the reference beam is c) blocked and d) interfering with the output resulting in a hologram. The inset in image (d) shows a magnified part of the hologram where the interference fringes are clear [59].

A 2 cm graded-index (GRIN) fiber of 62.5 μm core diameter and NA of 0.275 (by Thorlabs) is used in the experiments presented for the classification performance of the DNNs. The 2 cm fiber showed high performance for the intensity-only experiments introduced in the previous paragraph and kept the same for this part of the work to have a direct comparison. The inputs images are selected from the online available MNIST database of handwritten digits in this case too. In each experiment, we record 20,000 images of speckle patterns corresponding to each digit

projected on the SLM from which 16,000 are used for training the DNNs, 2,000 for validation and 2,000 for testing the DNN performance. The inputs of the fiber can switch from amplitude (Figure 39 a) to phase (Figure 39 b) and vice-versa by tuning the half waveplate (HWP2) and the polarizer before and after the SLM respectively. In addition, by blocking or not the reference arm we can either detect intensity-only images of the speckle patterns at the fiber output (Figure 39 c) or record the corresponding digital hologram (Figure 39 d) which is formed when the reference beam interferes with the speckle pattern on the CCD1. This configuration allows both amplitude and phase information of the speckle pattern to be obtained at the fiber distal side in order to compare DNNs performance with the intensity-only imaging case.

For the classification of the inputs a VGG-type DNN was used. The structure of the architecture and the layer parameters are the same as in the aforementioned intensity-only case as described in the Figure 32 a. The network input was either 1 or 2 channel 32x32 images. The 2 channel input case refers to the use of the complex field information, amplitude and phase, for the DNN training as it will be discussed below. The classification accuracies are reported as the mean and standard deviation of the results of 5 different training instances.

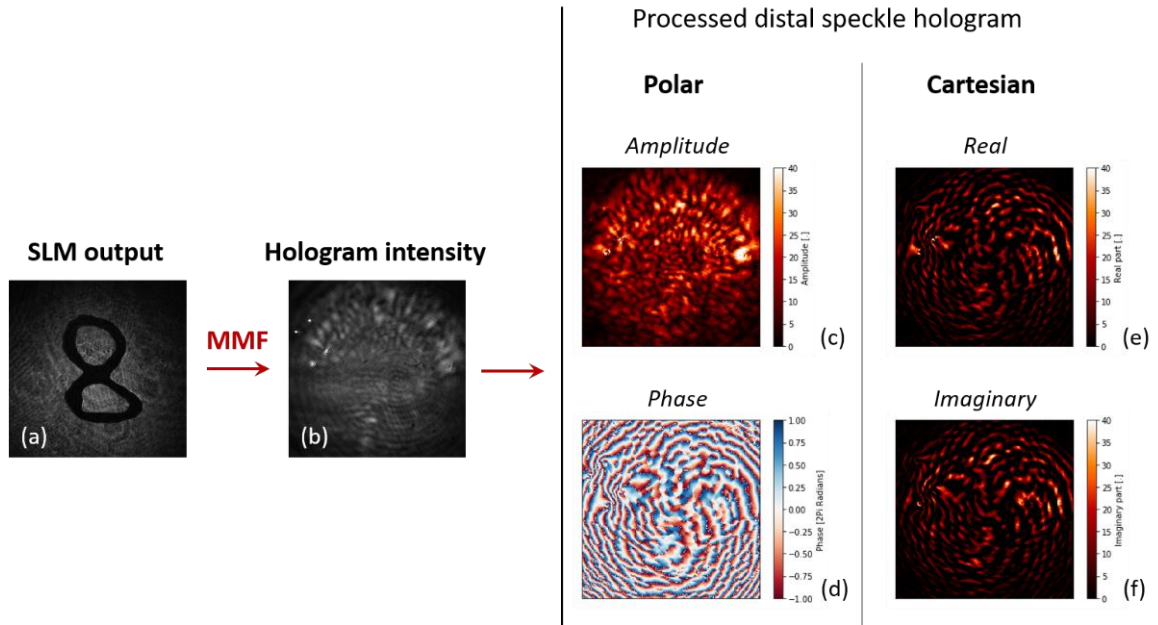


Figure 40. a) Amplitude modulated SLM output for the digit 8, b) hologram intensity of the fiber output for the input (a), c) amplitude and d) phase, e) real and f) imaginary part of the output field obtained after processing the hologram intensity (b) [66].

6.2.2 Speckle image classification

Using digital holography (as explained in Chapter 2), we can recover for each output speckle pattern, resulting from the different images of handwritten digits (Figure 40 a), the phase and amplitude (Figure 40c and d) or the real and imaginary parts of the output field (Figure 40 e and f)

and use them in order to train the DNN. The intensity image of the hologram (Figure 40 b) was also used to test the classification accuracy.

Table 3. Classification accuracies obtained after the DNN is trained with different input images [59].

VGG network input	Classification accuracy (%)		
	# Input channels	Amplitude modulated input	Phase modulated input
Hologram Intensity <i>Fig.33(b)</i>	1	91.8 ± 1.0	93.46 ± 0.6
Amplitude-only <i>Fig.33(c)</i>	1	94.3 ± 0.5	94.2 ± 0.7
Phase-only <i>Fig.33(d)</i>	1	75.2 ± 21.1	75.8 ± 21.8
Real part-only <i>Fig.33(e)</i>	1	91.0 ± 0.8	91.7 ± 0.3
Imaginary part-only <i>Fig.33(f)</i>	1	91.4 ± 0.8	91.7 ± 0.5
Complex <i>Fig.33(c&d)</i>	2	94.0 ± 0.5	94.4 ± 0.3
Complex <i>Fig.33(e&f)</i>	2	92.2 ± 1.1	93.4 ± 0.5
Speckle Intensity	1	92.7 ± 0.5	95.1 ± 0.6

The relative classification accuracies provided by intensity or holographic imaging of the speckle patterns, and the combination of different components of the recovered complex field reported in Table 3. The results indicate that intensity only speckle imaging provides similar classification accuracies as for holographic imaging. On the other hand, we find that using the phase-only information (Figure 40 d) as input to train the DNN leads to the least efficient image classification. We suspect that the main reason for that is that part of the information is affected by the wrapping of the phase in the image. In all the rest of the cases, the classification accuracy is almost equally good with a marginal improvement for amplitude inputs when using the complex field for training the network (Figure 40 c and d). Moreover, the classification results are better when phase modulated inputs are coupled into the GRIN fiber, which was also observed in the previous study for different fiber lengths [60]. Comparing the classification accuracies achieved between the speckle intensity and the hologram intensity images, we observe that the latter is always lower. It can be assumed that the additional information encoded in the fringes of the hologram intensity image is not preserved after the downsampling step.

6.2.3 Dataset processing effect on Deep Neural Networks performance

Since the information in the images projected onto the proximal fiber facet is distributed across the propagation modes supported by the fiber, the localized image information is spread across the fiber cross-section as it propagates along its length. This spreading of the localized information is saved in the speckle pattern generated when the light exits the fiber. As a result, the intensity of each pixel of the recorded distal speckle pattern is not related to single point only at the input image. This contribution can be higher for phase modulated inputs where a greater number of the

fiber modes participates in the speckle formation. To further evaluate this property of the MMFs, we follow different ways of processing the speckle images that form the DNN inputs to assess if there is sufficient information in a cropped area of a speckle image to classify the corresponding input images. In order to study the effects of cropping on the classification accuracy, the full resolution speckle patterns were center cropped to 32, 64, 128, 256, 512, and 1024 pixel sided square areas. These were then down-sampled to 32x32 pixel images; examples of these images are shown in Figure 41.

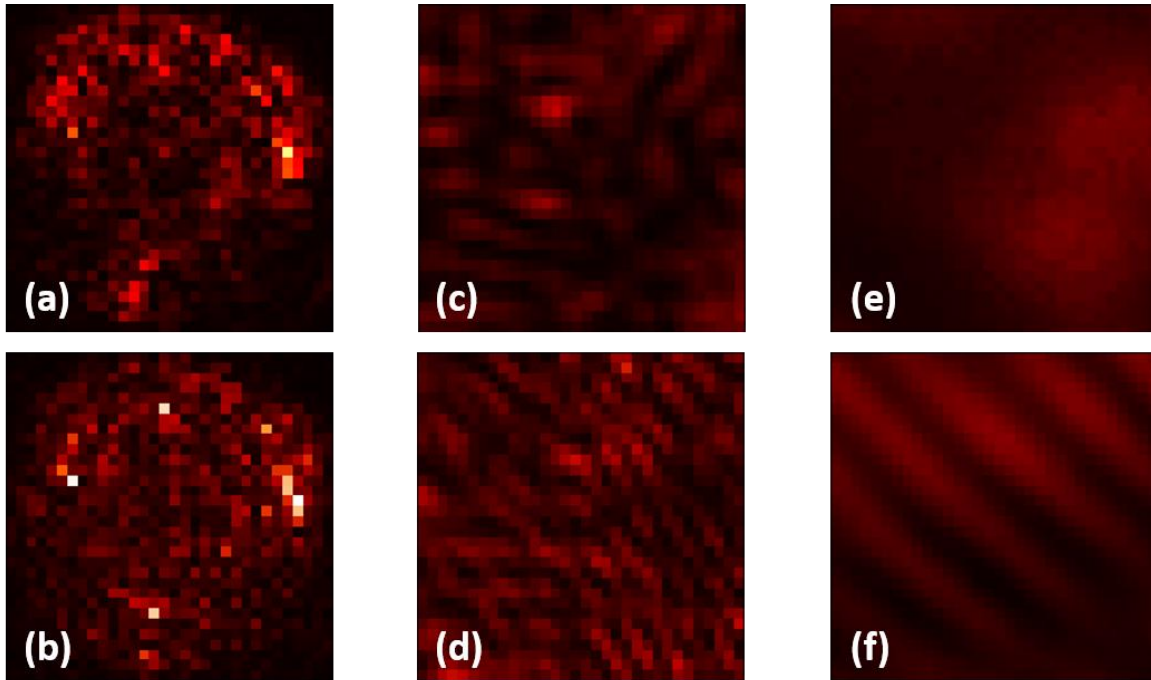


Figure 41. Cropped images of speckle patterns recorded on the camera CCD1 after downsampling to 32x32 pixels. The upper row of images corresponds to intensity-only recording while the lower row to holographic recording of the speckle output. The size of the square cropped area on the initial image changes from 1024x1024 pixels a) and b), 256x256 pixels c) and d) and 32x32 pixels for e) and f) [59].

Downsampling of the original speckle pattern beyond a certain level can introduce loss of information existing in the image and as a result, the classification success of different inputs can drop. For endoscopic applications this could be also the result of using a partially coherent illumination source. It has been already demonstrated that image reconstruction using an LED source that results in a blurred and almost uniform output of the MMF is possible using DNNs [100]. The results shown in Table 4 indicate that the classification accuracy decreases with decreasing crop sizes for all input modulation and speckle imaging combinations. It is apparent though, that the classification degradation is not linear with respect to the crop size, on the contrary DNNs seem to be able to successfully classify the inputs with an 83% accuracy using only a quarter area of the initial speckle pattern. In particular, for a crop size of 128x128 pixels, equivalent to a 1.5% area of the full speckle image, the accuracy decreases by around 12%; whilst

for a crop size of 32x32 pixels, equivalent to a 0.1% area of the full speckle image, the accuracy decreases by around 30% relative to an uncropped image. Therefore, we can conclude that it is better using the largest possible part of the image so that the speckle grain distribution is somehow preserved after the downsampling step rather than cropping a smaller part of the speckle in the first place. In addition, it is important to notice that even though cropping and then downsampling can preserve better the fringes structure in the case of hologram, intensity-only images of the speckle pattern still relate to higher classification performance in all cases.

Table 4. Classification accuracies obtained after the DNN is trained with differently sized cropped portions of the initial speckle image recorded either using intensity-only or holographic measurement [59].

Crop size [pixels]	Classification accuracy [%]							
	Amplitude input- Intensity recording		Phase input- Intensity recording		Amplitude input- Holographic recording		Phase input- Holographic recording	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1024x1024	93.2	0.7	94.9	0.4	91.8	1	93.5	0.6
512x512	91.2	0.8	93	3	89.6	2.2	92.2	1.1
256x256	83.2	1.8	91.6	0.4	85.5	1.4	88.3	0.6
128x128	76.7	1.4	85.2	0.8	79.8	1.1	83.7	1
64x64	66.3	1.9	77.5	1.5	69.2	1.8	73.1	2.3
32x32	59.2	2	70.5	2.1	52.9	1.8	63	2

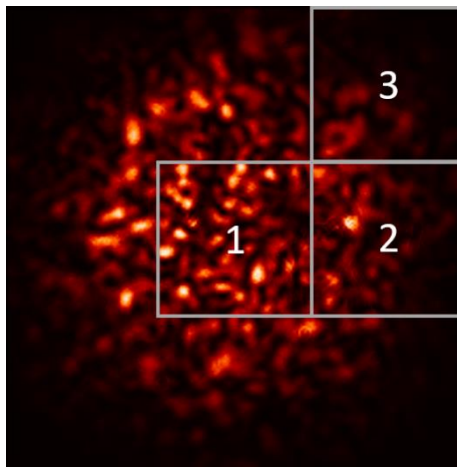


Figure 42. Image of a speckle pattern recorded at the fiber output. The white square indicate the locations of 340x340 pixels crop which are then down-sampled to 32x32 pixels for classification by the DNN [59].

Finally the effects of crop location on the classification accuracy are studied. The full resolution speckle images were cropped to 340x340 pixel patches at the three locations shown in Figure 42. These were then down-sampled to 32x32 pixel images for classification by the VGG network. The results, shown in Table 5, indicate that the classification accuracy is essentially independent of crop location on the speckle image, thus highlighting that the input image information is spread over the speckle pattern. Moreover, the values of the classification for the 340x340 pixel patches lie in between the values for the cropping sizes of 256x256 and 512x512 pixel images further verifying the results shown in Table 5.

Table 5. Classification accuracies obtained after the DNN is trained by cropping different locations on the initial speckle image recorded either using intensity-only or holographic measurement [59].

Crop location	Classification accuracy [%]							
	Amplitude input - Intensity recording		Phase input - Intensity recording		Amplitude input - Holographic recording		Phase input - Holographic recording	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	91.1	0.5	89.6	1.8	86.6	2.3	87.2	1.3
2	90.1	1.8	92.5	1.2	87.1	1	86.7	2.3
3	90.6	1.2	88.7	1.1	85.9	1	87.7	1.1

6.3 Effect of wavelength drifting on the deep learning performance

The robustness of the DNNs to recover information through a MMF even in the presence of instabilities was demonstrated in the paragraph 6.1 for the 1 km long fiber. In that case, random thermal and mechanical vibrations in the optical setup induced perturbations during the measurements, which were particularly severe for the 1 km fiber length. In order to further investigate the learning ability of DNNs, a perturbation is induced in the system during the measurements in a controllable way by introducing wavelength drifting. Since the speckle pattern intensity is highly sensitive to the light wavelength, using MMF as a spectrometer has been proposed as a way to take advantage of this property in previous works [134,166].

However, wavelength drifting is an impeding factor for multiplexing applications in telecommunications [167,168] and it is interesting to investigate to which extend DNNs can overcome the speckle decorrelation induced by it [161,162]. Efforts to overcome the noise in the data generated by wavelength drifting have turned towards the stabilization of the tunable laser sources themselves or introducing certain synchronization schemes to separate the respective fiber output signals [167,168]. The latter solution is usually not preferable because of it reduces the overall throughput. Other works have proposed the integration of a neural network circuit in the multiplexing process for dynamic drift correction [169]. In the work, the performance of state-

of-the-art DNNs to classify the input images in a MMF from their corresponding speckle patterns is investigated in the presence of wavelength drift of the light source. Remarkable results are obtained for wavelength drift bandwidth that reaches 100 nm and they are further discussed in the following paragraphs.

6.3.1 Methods

For the measurements presented in this paragraph, we integrated in the optical setup presented in the Figure 31 a wavelength tunable laser source (M-squared, SolisTiS 2000 PSX XF with Terascan software), emitting in the range 700-1000 nm. The laser source can shift the wavelength with 10 fm accuracy. To simulate a situation of wavelength drifting, a Matlab code was created to control the laser wavelength, the speckle acquisition in the camera CCD1 (Figure 31) and the input images on the SLM. We firstly choose the wavelength range of the drifting in the code and then a random array with 100 wavelength values within this range is generated. The wavelength values are sorted in an ascending order for two main reasons: firstly, it leads to a more stable operation of the tunable laser instead of tuning the wavelength to far spaced wavelength values and secondly, the laser stabilization at a specific wavelength is achieved much faster. The inputs are therefore projected on the proximal fiber facet in batches of 100, each one at a random wavelength in the array. For the next 100 inputs, the random wavelength array is recalculated so that randomly different values of wavelength are generated for the images of the full dataset. Phase modulated inputs from the MNIST database are used as in the previous experiments to generate a dataset of speckle patterns through the MMF. The fiber length was kept 10 cm to avoid the effect of further perturbations, which was the case of 1 km fiber and focus on the DNN performance changes induced only by the wavelength drifting.

We collect a total number of 10,000 intensity patterns from which 80% is used for training, 10% for validation and 10% for testing. For the classification of the data presented in the next paragraphs, we used the same DNN architectures described in the Figure 32 a.

6.3.2 Results

As discussed in the first part of this thesis, the number of fiber modes in MMFs determines the available degrees of freedom (channels) for transmitting the information. An input image decomposes to the fiber modes basis. Therefore, a MMF with few modes would not be able to transmit a complex signal. Therefore, for testing the performance of DNNs in the presence of wavelength drift, it is important that the classification accuracy is evaluated in for the different wavelengths within the drift bandwidth in no drift conditions. In this way, it can be ensured that any change in the classification performance is attributed to the induced “noise” in the dataset and not to the inherent incapability of the system to support the input images because of the low number of supported modes at a certain wavelength. Datasets of 10,000 images were created for wavelengths from 700 nm to 1,000 nm with an increment of 50 nm, by projecting phase-only

inputs of handwritten digits from the MNIST database on the SLM. In Figure 43, we observe that the DNN results do not differ for the different wavelengths in the 700-1,000 nm range. According to this observation, it can be assumed that if there is a difference in the classification accuracy of the DNN in the presence of drifting is only related to the perturbation of the system and not to an anyway reduced performance at a certain wavelength in the investigated bandwidth.

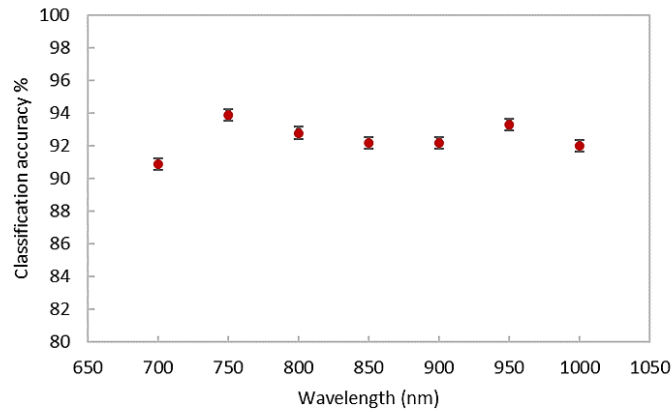


Figure 43. Classification accuracy calculated for datasets collected at different laser wavelengths for phase modulated inputs.

As a next step, we explore the effect of wavelength drifting on the DDN performance for different drifting bandwidths of 6 nm, 12 nm, 24 nm, 48 nm, 72 nm and 96 nm around the central wavelength that was chosen to be 800 nm. We investigate to different cases of drifting: a) the DNNs are trained with the data acquired with the laser wavelength fixed at 800nm and then tested on speckle patterns recorded at different wavelengths, b) the wavelengths within the drifting range are included in the training set by randomly acquiring speckle patterns at different wavelengths during the measurement as described in the Methods section before. For comparison, the speckle intensity correlation was calculated for a range of 100 nm around the central wavelength was calculated to study the DNN results as a function of the speckle decorrelation range. As seen from the Figure 44 a the speckle patterns decorrelate at 50% within a range of approximately 30 nm (FWHM).

In the case (a), DNN is trained with data recorded at 800 nm and we test their ability to recover the information from data measured at 803 nm, 806 nm, 812nm and 850nm (Figure 44 b, red circles). The classification accuracy drops at 69% for the shortest drift of 803nm, while the speckle correlation has only dropped to 90% at this wavelength distance. The results become completely random at 812nm showing 10% classification. Hence, it is shown that generalization of the VGG network for wavelengths further than the one used in the training is only efficient for a very short narrow range. On the other hand, the classification results show a decrease with increasing the wavelength drifting bandwidth but in a much lower rate in the case (b). For 6 nm drift, meaning ± 3 nm around the central wavelength, the VGG DNN shows no significant change of classification performance 88% if compared to the one in absence of drift ($\Delta\lambda = 0$) 90%, which is remarkably

better to the 69% observed in the experiments of the first case discussed before. Even for almost 100 nm drift the VGG DNN classifies correctly 70% of the inputs (Figure 44 b, blue circles). These results indicate high robustness of the DNNs in the presence of severe perturbations of the system as long as the training dataset is collected in the presence of the disturbance under study. In Figure 45, we show for comparison the confusion matrices for the two cases for a drift of 3nm and ± 3 nm respectively. It is evident at that many labels are mixed up of the case (a) studied and for example the digit 9 that is similar with at least two other digits such as 7s and 5s is highly misclassified. In spite of the high efficiency in recovering information for inputs within the drifting range, this is not the case for speckle patterns recorded at wavelengths outside this range. We trained the VGG classifier with data measured in a 12nm drifting situation and we tested afterwards the classification accuracy of the trained model for data outside the 12nm range, from 807nm to 850nm. The results obtained follow the trend observed in the case (a) discussed before as show in Figure 44 c. Therefore, the DNN cannot generalize outside the selected bandwidth for training even when different wavelengths are included in the learning process.

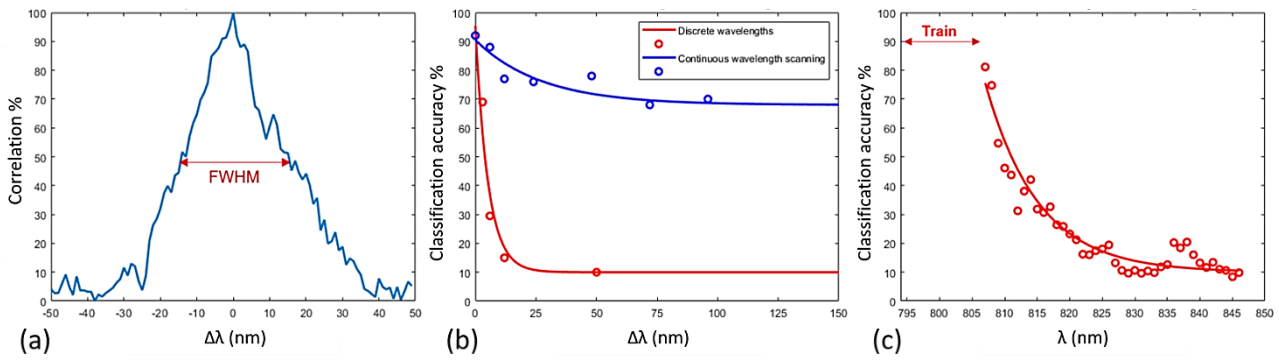


Figure 44. a) Speckle intensity correlation as a function of wavelength drifting bandwidth, b) classification accuracy for datasets collected in for the two cases discussed in the text: DNN trained at 800nm wavelength and tested at data collected at different wavelengths (red circles) and DNN trained with data that include the perturbation (blue circles), c) classification accuracy for a DNN trained in the presence of 12nm bandwidth drift and tested at data collected outside this range.

Once crucial parameter usually suggested for improving the performance of the DNNs is the number of available samples in the dataset. At the same time, this parameter is considered one of the drawbacks of deep learning, because there are cases for which the number of available data is limited. In order to evaluate if the dataset used to train the VGG classifier in the case of drifting is sufficiently large, we investigate the classification accuracy for different number of training datasets. The classification results for different dataset sizes are presented in the Figure 46. The classification accuracy varies with the size of the training dataset as expected. It is apparent from the Figure 46 that the more grave the wavelength drifting is the more samples are needed for achieving higher classification accuracies. Nevertheless, for all the tested wavelength bandwidths, a saturation of the classification improvement is observed while increasing the size of the dataset;

for the largest bandwidth (96 nm) this happens at about 6,000 to 8,000 samples as suggested by Figure 46.

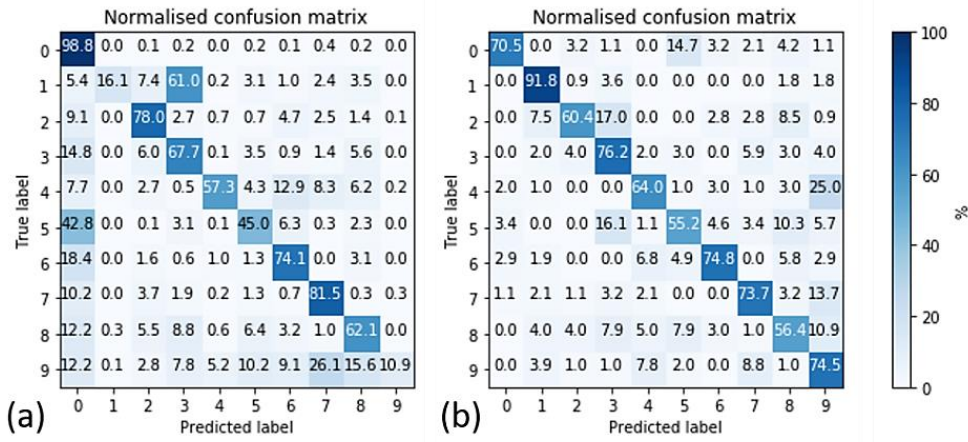


Figure 45. Normalized confusion matrices showing the classification results for the case that the drifting wavelength is a) not part of the training set and b) is included in the drifting wavelengths for which the VGG classifier is trained with [161,162].

Interestingly, the accuracies measured for each bandwidth saturate at values dependent on the bandwidth which implies that even larger dataset would not improve the results. One possible reason for this is the randomization of the input-output mapping. Because of the large wavelength drift, it can happen that the speckle pattern of a specific digit captured at a certain wavelength is very similar to the speckle pattern of another digit captures at another wavelength. Consequently, if the similarity between speckle patterns of different digits is higher than same digits at different wavelengths, wrong classification will occur.

Another interesting observation based on the measurements reported in the presence of wavelength drift, is that the classification accuracy does not notably deteriorate after increasing the drift bandwidth further than approximately 60 nm. It has been observed that MMFs support a type of modes that their shape is independent of the wavelength within a certain bandwidth, which are called principal modes [118,170]. Although further investigation is needed to conclusively prove it, it could be possible that part of the input images in the MMF is decomposed in the principal mode basis and therefore it remains unchanged within some drift bandwidth thus preserving the information at the output partially leading to the preservation of the classification accuracy.

One possible solution for increasing the performance of the system is to decrease the downsampling factor to assess if the reduction of lost information due to the image processing could help the DNN more accurately relate the speckle pattern features to the input information. Another possibility is to increase the weights of the current VGG network architecture so that

more learnable parameters are available and potentially accommodate the wavelength noise effect.

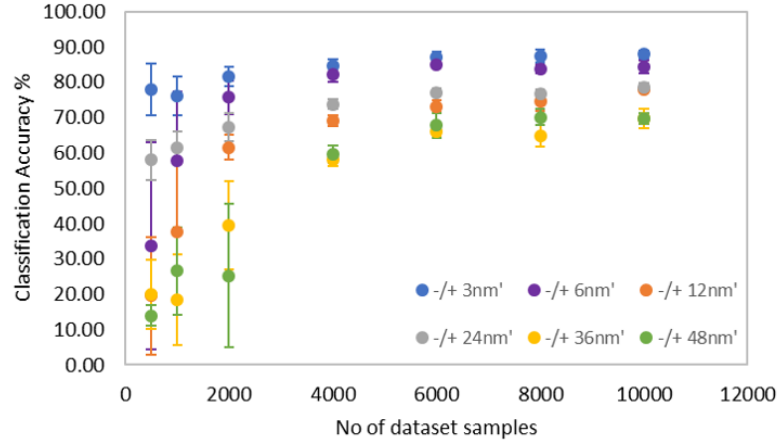


Figure 46. Classification accuracy of the VGG-type DNN classifier used in the experiment for increasing the number of available samples in the training dataset.

Other works have tried to evaluate the advantages of deep learning for information recovery under different types of disturbances such as bending [62]. *Fan et al.* showed that including data of different bending configurations of a MMF in the training set of the DNN algorithm makes it possible to learn the randomness imposed by bending the fiber. Therefore, the reported results concerning the wavelength perturbation are in agreement with these findings and establish that including in the training dataset samples of the perturbed state of the MMF makes the information recovery resilient to a certain extend.

Until this point, we have discussed the aspect of training a DNN algorithm to recover information through a MMF system despite the wavelength changes by including the noise in the dataset. Nevertheless, there are cases where sensitivity in the wavelength change is actually crucial; for example, MMFs have been proposed as suitable systems for high-resolution spectrometers [134,166]. In those previous works, the speckle decorrelation with wavelength is used for distinguishing different wavelengths and the resolution is dependent on the speckle decorrelation bandwidth of the specific fiber, which becomes narrower with increases in the fiber length. In the results presented in this work, we observed that the classification efficiency of a trained DNN with data recorded at a single wavelength is sensitive to wavelength changes and deteriorates within a significantly narrower bandwidth (Figure 44 b, red circles) than the decorrelation bandwidth of the system (Figure 44 a). Therefore, what is considered unwanted for information recovery through the system under a wavelength drift situation, it could potentially be desired for achieving a resolution which is higher than the one given by the decorrelation bandwidth and has been previously proposed in the literature for spectroscopic applications.

6.4 Conclusions

We have shown that DNNs can recognize distorted images at the output of a MMF from their corresponding intensity speckle patterns. In particular, the deep learning approach presents a remarkable robustness against wavelength drift of the light source within an extended bandwidth for which total decorrelation of the speckle pattern intensity is observed. Despite the randomization of the input–output mapping introduced by the various speckle drifting, image classification of high accuracy of about 70% is reported for almost 100 nm wavelength drift, while more than 80% classification accuracy is shown for a 50 nm bandwidth. The results presented in this article suggest that deep learning can overcome distortion in the signal after propagation through MMFs in the presence of severe wavelength drift. The potential of using a single trained DNN model for retrieving image information in multiple wavelengths is suggested by using the intensity-only images of the respective speckle patterns, which could allow multicolor illumination imaging without extensive system recalibration.

Chapter 7 Deep learning-enhanced imaging through fiber bundles

This chapter tackles the problem of reduced resolution when using fiber bundles because of the structure of the individual cores. In Part A of the thesis, wavefront shaping using the transmission matrix method was proposed to overcome the pixelation artefact in the images by creating a high intensity focus through a fiber bundle, which was used to perform point scanning-based imaging at a superior resolution. However, as previously discussed, the sensitivity of calibration-based methods to external perturbations can be an issue for the stability of the imaging system preventing the flexibility of the final endoscope. In this chapter, the robustness of deep neural networks (DNNs) in the presence of noise and their super-resolution ability are exploited to generate high resolution and pixelation-free images through fiber bundles, while allowing bending insensitive measurements. In addition, the capability of DNNs to translate information between different imaging modalities is discussed.

7.1 Introduction

Fiber bundles (also called multicore fibers (MCFs)) are widely used in clinical endoscopy and they can be categorized in the following general groups based on their drawing method. In the first kind of fiber bundles, the individual cores are fixed in a common polymer like cladding while in the second kind, each individual core has its own cladding and they are all fixed in a honey-comb structure by a mesh grid forming the final bundle; this kind of bundles are often called leached-type fiber bundles. The first kind of bundles tends to be more rigid because of the common cladding material. In contrast, fibers bundles of the second kind are highly flexible but, because of this property, they are slightly more prone to damage of individual cores leading to “burnt pixel” defects in the final image. In addition, leached fiber bundles cost significantly higher than the other type (1000\$/m) but they are often preferred for their flexibility and image quality [50].

It has been mentioned throughout the thesis that the image resolution of the fiber bundles is defined by the core spacing and further improvement is limited by the core-to-core coupling that would cause blurring in the final image. There has been a lot of work focused on the creation of fiber bundles with closely spaced cores but minimal coupling. It has been shown by *Chen et al.* [38,40] that inducing aperiodicity in the individual core positions as well as slight changes in the

core size and refractive index reduce the coupling between the neighboring cores. Addition of lens magnifying systems at the fiber distal facet can also be an option for increasing the resolution of the endoscope but it is followed by a reduction of the field of view and raise of optical aberrations [67]. As aforementioned, resolution improvement of the fiber bundle endoscopes has been also achieved by wavefront shaping techniques as described in Part A of the thesis by exploiting the *NA* of the fiber cores to create a diffraction limited spot for point scanning imaging. The calibration sensitivity of the holographic techniques to external perturbations limits the applicability of the fiber bundles to rigid endoscopic configurations. The image pixelation artefact generated by the fiber bundles can be also corrected using computational means which digitally increase the resolution of the image via image processing algorithms as proposed by *Shao et al.* [71,171]. In particular, deep learning has shown promising results for super-resolution in image processing and microscopy applications and it is explored in the following paragraphs as an alternative for removing the image discretization caused by fiber bundles [73,152,155–157,172].

Fiber bundles used in clinical practice mostly use conventional bright field microscopy for imaging the tissues and offer important information for diagnosis [24,26,68]. Many studies of fiber endoscopy have also shown that fiber bundles can support multiple imaging modalities, apart from bright field imaging, such as confocal endoscopy [67], wide-field and two-photon fluorescence imaging [26,68,69]. Most of the aforementioned techniques need a certain imaging contrast mechanism, which is not always feasible to use for in-vivo imaging of biological samples because of toxicity or interference with the biological processes under inspection. For this reason, phase imaging is often suggested as the most well-suited technique for label-free imaging of transparent, low contrast samples such as tissues and cells [108,173–176]. Therefore, combining phase imaging modalities in an endoscopic device presents many advantages for medical diagnosis [177]. In the following paragraphs, the use of deep learning is proposed for improving the resolution of an imaging fiber bundle and integration of phase contrast modality by properly trained DNNs.

7.2 Image reconstruction of different datasets through a fiber bundle

The evaluation of the DNN performance in reconstructing images transmitted through a fiber bundle is initially made using simple images from databases that are available online such as MNIST, MNIST-Fashion (MNIST-F) and CIFAR-10 [163,178,179]. Starting from the almost binary inputs obtained from the MNIST dataset to the gray-scale images of the CIFAR-10 and MNIST-F it is demonstrated that DNNs can reconstruct input images from their corresponding outputs after they propagated through a fiber bundle with a significantly superior quality to conventional interpolation algorithms.

7.2.1 Methods

7.2.1.1 Optical setup

The images were projected at the fiber facet by means of an SLM using the optical system described in paragraph 6.1.1 (Figure 31). Some lens modifications concerning the 4f systems were implemented so that the demagnification of the SLM screen matches the core size of the fiber bundles used for the experiments. Specifically, the microscope objectives OBJ2 and OBJ3 (Figure 31) were changed from 60x to 20x magnification (Newport, NA = 0.25). A 6,000 core Fujikura imaging bundle of 30 cm length (FIGH-06-400N, 400 μm core size) is used to deliver the images of the different datasets projected on the SLM. The fiber bundle is characterized by an average core size of 2.5 μm and an average core spacing 4.5 μm . The laser source coupled in the system is a tunable wavelength laser (M-squared, SolisTIS 2000 PSX XF with Terascan software, 700-1000 nm) for the experiments presented in this section (as in the case of the section 6.3).

7.2.1.2 Deep Neural Network architectures

A U-net type DNN is used for the image recovery through the fiber bundle. The input and output images were downsampled using bilinear interpolation to 32x32 and 64x64 pixel sizes for comparison. The U-net DNN used in the case of 32x32 pixel size images is the same as the one previously described in Figure 32 b. For the U-net trained with the 64x64 pixel size inputs, two extra convolutional and max pooling blocks for input and output were symmetrically added in the U-net architecture presented in the previous chapter (paragraph 6.1.1). The added layers contain 16 kernels each (Figure 47). An Adam optimizer with a learning rate of 1×10^{-4} is used to minimize a mean square error loss function, similarly to previous cases. A batch size of 10 and a maximum of 50 epochs is selected. The splitting ratios of the datasets studied in this chapter for creating the training, validation and test sets were 90%, 5% and 5% respectively.

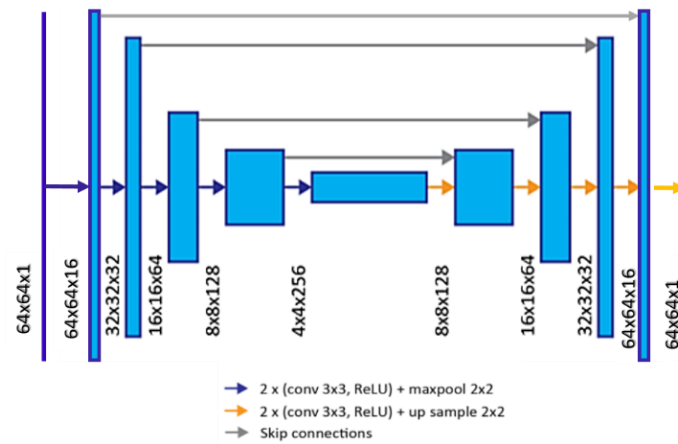


Figure 47. U-net type DNN architecture used to reconstruct the input images in the fiber bundle from the intensity images at the fiber output.

In order to compare the quality of the reconstruction results among the various datasets presented in this chapter, some quality metrics are defined to measure the similarity between the reconstructed images via DNNs and their corresponding ground truth. In the discussions to follow, this comparison is made using the mean squared error (MSE) and the mean structural similarity index (SSIM), which are calculated based on the following equations:

$$MSE = \frac{1}{N^2} \sum_i^{N^2} (y - x)^2$$

Equation 25. Mean squared error between two images

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

Equation 26. Formula for the calculation of the structural similarity index

In the above equations we consider x as the ground truth image vector and y the predicted images by the DNN. The two images have $N \times N$ pixel size which determines the averaging of the values. The rest of the variables used in the SSIM calculation are given below:

- μ_x : average of x
- μ_y : average of y
- σ_x^2 : the variance of x
- σ_y^2 : the variance of y
- σ_{xy} : the covariance of x and y
- $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$: two variables used to stabilize the division with a weak denominator
- L : the dynamic range of the images
- $k_1=0.01$, $k_2=0.03$ by default

7.2.2 Amplitude modulated inputs

The first dataset used to generate input-output pairs for the training of the U-net DNN and assessing its performance is the MNIST dataset that consists of images of handwritten digits, which we have been processed to contain either black or white pixels (binary). Amplitude modulated images are displayed on the SLM as shown in Figure 48 a and projected on the 30 cm long piece of the Fujikura fiber bundle by means of a demagnification optical system. The images

generated after propagation in the fiber bundle are collected by the camera at the distal fiber side (Figure 48 b). The distal image presented in the Figure 48 b is generated by illuminating the SLM at 800nm. The individual cores support two modes at this wavelength. Despite the aperiodic structure and core variation of the Fujikura imaging bundle, which is optimized to reduce the crosstalk between the neighboring cores, there is still part of light escaping from the illuminated cores to neighboring ones and as a result areas that are supposed to be dark, light up (Figure 48 b). Nevertheless, the image of the digit 6 is nicely represented through the fiber without further processing. However, it is important to note that in case of a more detailed image, light crosstalk can lead to image scrambling and the recognition of the output could be more challenging. This issue is investigated in the following paragraphs by comparing conventional interpolation and DNN image reconstruction.

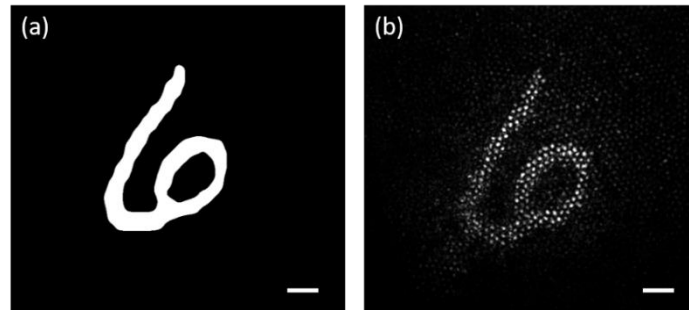


Figure 48. MNIST database. a) Input image on the SLM and b) the image (a) relayed through the Fujikura fiber bundle. Scale bar: 50 μ m

The DNN performance is evaluated by reconstructing the SLM input from the fiber bundle output for a dataset of 5000 samples of MNIST digits. The reconstruction results are presented in Figure 49. Remarkably, the reconstruction quality is quite good for the both image sizes. In the 32x32 case, the cores of the fiber bundle cannot be clearly resolved due to the downsampling step, while in the 64x64 case there are enough pixels to retain this detail of the image. This implies that the trained U-net model is able to find distinct features, which represent the images of the MNIST dataset after propagation through the fiber bundle even in the cases that the image information is highly compressed by the downsampling step. The quality metrics of the image reconstruction through the MCF bundle are presented in Table 6 where it can be observed that reduced downsampling of the input image can lead to better reconstruction results. In all cases, we can see that the DNNs can successfully result in good quality images even with after severe downsampling. The standard deviation values reported in the Table 6 are calculated among the images of the test set.

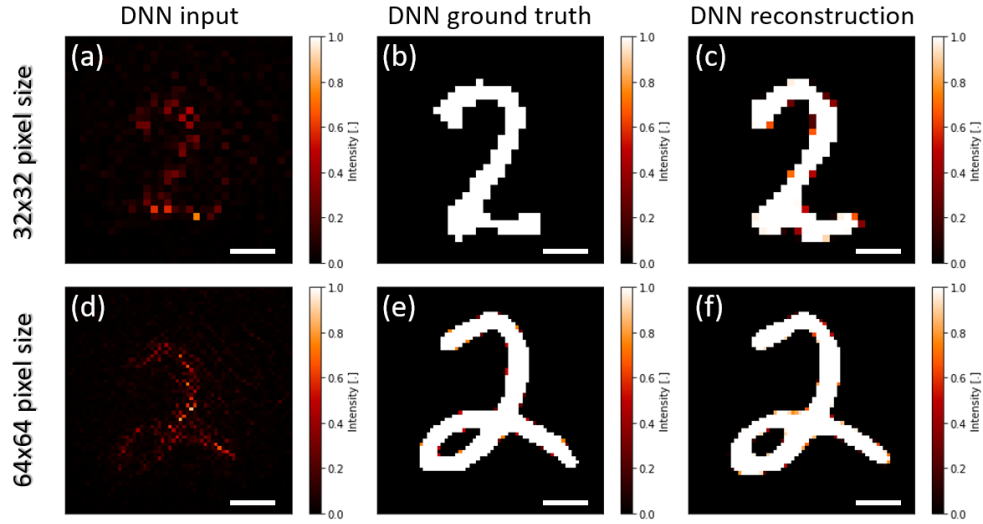


Figure 49. DNN reconstruction of MNIST dataset propagated through a fiber bundle for a-c) 32x32 and d-f) 64x64 image size. The first column shows the downsampled version of the image at the fiber output (a and d), which is the DNN input, the second column presents the corresponding input image on the SLM, which is the ground truth for the DNN training (d and e) and the last column shows the DNN reconstructed image (c and f) for the corresponding input on the SLM. Scale bar: 100 μm

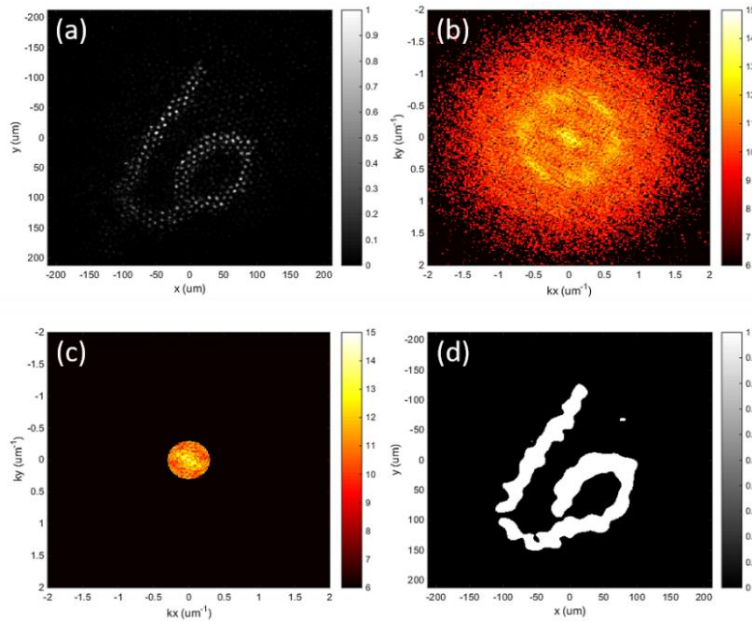


Figure 50. Fiber bundle imaging of the distal image of the digit 6. a) Image at the distal facet, b) the amplitude of the Fourier transform of (a) in log scale, c) filtering of the diffraction circle and d) inverse Fourier transform of (c) showing the final interpolated image.

For comparison, Fourier filtering is also used to remove the effect of the core sampling from the fiber bundle output using (Figure 50). To do so, the Fourier transform (Figure 50 b) of the fiber bundle generated image (Figure 50 a) is calculated. The Fourier transform shows a circle surrounding the central part of the frequency spectrum, which is a diffraction component

generated by the quasi-periodic structure of the fiber bundle core. By filtering out the diffraction circle (Figure 50 c) and calculating its inverse Fourier transform, the input image free from the high frequency sampling by the individual cores is obtained as shown in Figure 50 d. The final image (Figure 50 d) is further processed applying a thresholding step based on the prior knowledge that the input image is binary. The size of the circular mask applied in the Fourier space for filtering is defined based on the periodicity of the fiber bundle, which is equal to the core spacing Λ . Based on the Shannon sampling theory, the radius of the circular filter applied in the angular spectrum of the output corresponds a maximum diameter of $2\pi/\Lambda$. Considering the results of Fourier processing of the bundle output a comparison with the images obtained by the DNNs is presented. In both cases the recovered images is correlated with the SLM input (Figure 49 b). The results presented in the Table 6 show that simple interpolation and thresholding could provide nice images in the case of the MNIST database. More complicated sampling and image processing algorithms can be also used to improve further the image quality [171]. However, as discussed in the Chapter 5, conventional computational imaging algorithms optimize the output of a single image and they must be repeated for new images, which makes them time inefficient.

Table 6. MNIST and CIFAR-10 reconstruction metrics for amplitude inputs.

Reconstruction method	Metrics	
	MSE	SSIM
MNIST		
U-net (32x32)	0.025±0.009	0.85±0.06
U-net (64x64)	0.014±0.015	0.92±0.05
Interpolation + Thresholding	0.027	0.94
CIFAR-10		
U-net (64x64)	0.005±0.015	0.82±0.03
Interpolation + Smoothing	0.152	0.60

The results presented above do not show a clear advantage of using DNNs for recovering a good quality image from the corresponding pixelated output of a fiber bundle. Despite the presence of coupling, which is noticeable in Figure 50 a, the features of the input are still clearly distinguishable since the MNIST database of handwritten digits used consists of binary images. Consequently, the use of DNNs can be considered excessive for this reconstruction problem. What happens when imaging more complicated inputs though? To examine this aspect, we use the CIFAR-10 database that consists of gray-scale images of ten different classes of objects such as cats, cars, birds and others. Like MNIST, CIFAR-10 is a widely used database for the evaluation of classification DNN algorithms. Here, we use it to investigate the efficiency of DNNs in image reconstruction through a fiber bundle.

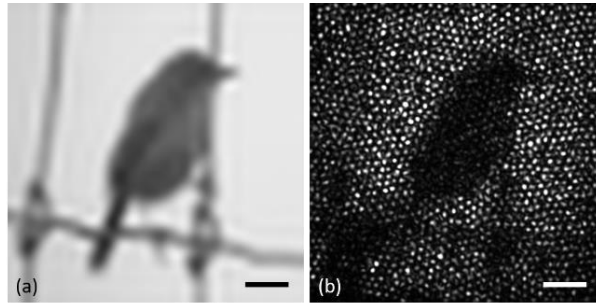


Figure 51. Example of an image from the CIFAR-10 database. a) Input image on the SLM b) the image (a) relayed through the Fujikura fiber bundle. Scale bar: 50 μm .

Figure 51 displays an example of an input image from the CIFAR-10 dataset and its corresponding output at the distal facet of the fiber bundle. It is already apparent that some of the details related to the gray scale variation as well as some fine features like the vertical thin lines are difficult to recognize after the sampling by the fiber bundle. The core-to-core coupling affects the light distribution across the image, distorting the gray level appearance of the input. In the previous case of the MNIST database for which the input images on the fiber bundle were processed to be binary, this effect can be compensated by thresholding the image to render two gray level values, 0 and 1. Conventional interpolation techniques such as Fourier filtering and thresholding fail to reproduce high quality reconstruction results images when examples of the CIFAR-10 dataset are projected (Figure 52). The gray level values are obscured. For this reason, CIFAR-10 dataset is more suitable for comparing the conventional image processing methods to DNNs.

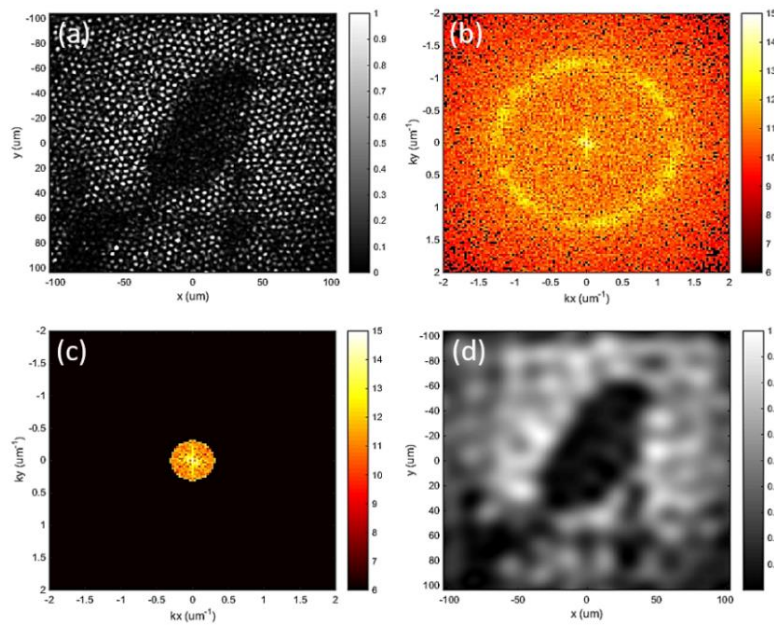


Figure 52. Fiber bundle deconvolution of the distal image from the CIFAR-10 database. a) Image at the distal facet, b) the amplitude of the Fourier transform of (a) in log scale, c) filtering of the diffraction circle and d) inverse Fourier transform of (c) showing the reconstructed image.

The same U-net DNN with 64x64 image size inputs is also trained to remove the discretization artefacts from the images of the CIFAR-10 dataset. Figure 53 a shows an example from the dataset images at the distal facet after downsampling, which is then used as an input to the DNN. The corresponding reconstructed image (Figure 53 c) seems to preserve most of the ground truth features (Figure 53 b) and is characterized by quality, evidently superior to the interpolation-based reconstructions such as the example presented in Figure 52 d. Table 6 summarizes the quality metrics calculated for the CIFAR-10 dataset concerning the image reconstruction results obtained by the two approaches; DNNs and Fourier filtering-based interpolation. The reported results for the U-net are calculated for the test set, meaning image examples that were not part of the training set of the DNN model. Impressively, the trained U-net model is capable of reconstructing details of the input image that are not noticeable at the intensity image recorded at the fiber bundle output. Specifically, features on the horse's face depicted in the example image in Figure 53 b are present in the DNN reconstruction (Figure 53 c) but not in the DNN input image of the fiber bundle distal facet (Figure 53 a).

Despite of the enhanced resolution achieved by removing the core sampling, certain limitations posed by the physics of the system are present. The U-net performance is also tested in conditions of better sampling of the input image by magnifying it on the proximal facet so that the physical resolution of the system is improved. To do so the microscope objective (Figure 31: OBJ2) of the experimental set is changed from 20x to 5x. The results after training the U-net on the new dataset of image pairs show MSE of 0.003 and 0.89 SSIM as opposed to the previous case reported in the Table 6 in which MSE and mean SSIM are found 0.005 and 0.82 respectively.

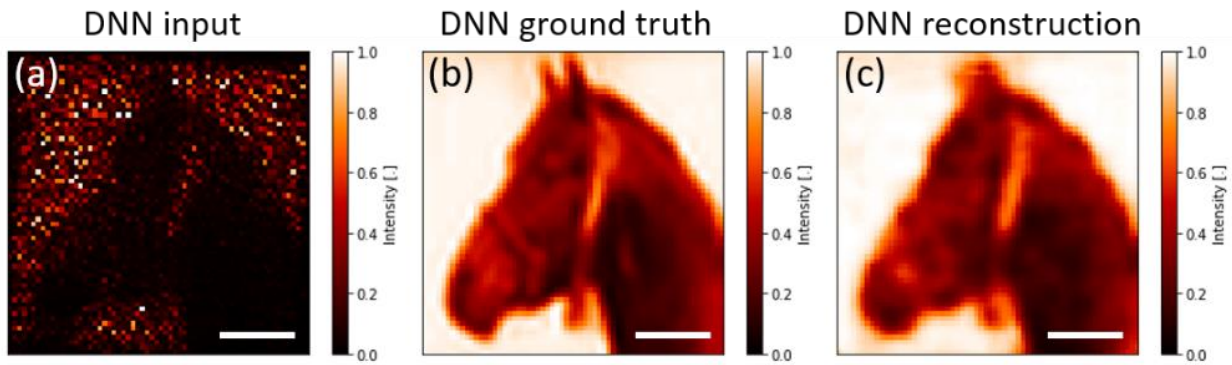


Figure 53. DNN reconstruction example from the CIFAR-10 dataset propagated through a fiber bundle. a) Downsampled version of the fiber output image at 64x64 pixels size, b) input image at the SLM projected at the fiber proximal side generating the output (a) and the DNN reconstruction of the input image (b) from the fiber output (a). Scale bar: 50 μm .

Based on the above discussion, it is concluded that DNNs can provide an efficient platform for quality enhancement of fiber bundle images. The U-net model can render accurate reconstruction results in the presence of core-to-core coupling and severe down sampling even for complex input shapes. For this outcome no particular hyperparameter optimization was made for the U-net

algorithm. More complicated DNN architectures, custom loss functions etc. can be also implemented to further enhance the final image quality as proposed by other works [64,73] and will be also demonstrated in the next paragraphs. However, for the test on the different online databases used to assess the imaging capabilities of the fiber bundles, a standard U-net configuration is used for direct comparison among the results.

7.2.3 Phase modulated inputs

In Chapter 6, we demonstrated that both amplitude and phase inputs can be efficiently reconstructed after propagating through a MMF. Even if the phase inputs have almost no distinct features in their corresponding intensity image (Figure 33 e & f), the MMF-induced mapping into speckle patterns enhances the contrast among different inputs. Therefore, the information content in the speckle patterns is enough to assist the DNN training to successfully reconstruct the phase inputs. It is interesting to study how phase image translation works in the case of a fiber bundle for which there is no particular mapping mechanism for contrast generation.

Figure 54 a and b present the fiber bundle output before and after downsampling respectively when the phase image of the digit 3 selected from the MNIST database (Figure 54 c) is displayed on the SLM. In Figure 54 c, the highest pixel value of the image corresponds to π phase and the lowest to zero phase. From the Figure 54 a and b, it is obvious that recognizing the digit from the intensity image of the fiber bundle output is not trivial, due to the almost uniform intensity distribution which characterizes a phase-only input.

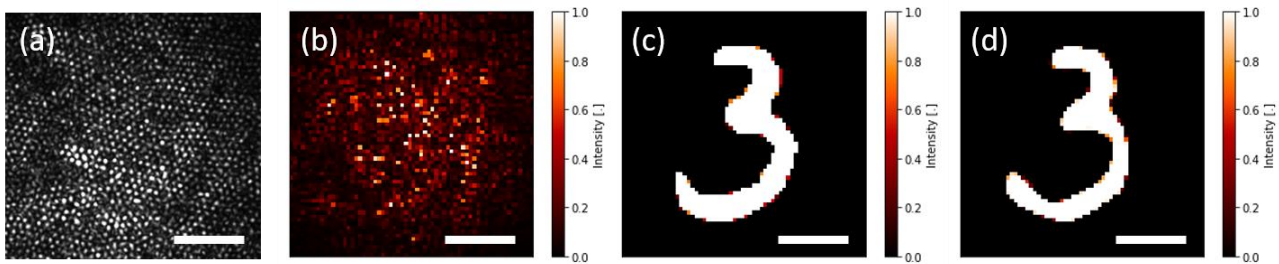


Figure 54. DNN reconstruction example from the MNIST dataset propagated through a fiber bundle. a) Fiber bundle output for the phase image of the digit 3, b) downsampled version of the fiber output image (a) at 64x64 pixels size, c) input image at the SLM projected at the fiber proximal side generating the output (a) and the DNN reconstruction of the input image (c) from the fiber output (b). Scale bar: 100 μm .

Similarly, to the amplitude input case discussed in the previous section, we generate a dataset of 5000 pairs of fiber outputs and corresponding phase-only SLM inputs. Without modifying the parameters of the U-net used for the amplitude images of the MNIST database in the previous section in the case of 32x32 pixel size inputs, a U-net DNN is trained to reconstruct the SLM phase input from the intensity image of the fiber bundle output. The results of the reconstruction show remarkable quality despite of the downsampling of the original data as shown in Figure 54 d, while

the MSE and SSIM metrics are calculated 0.04 and 0.82 respectively. Comparing the metric values to the corresponding ones reported in the Table 6 for the amplitude modulated images of the MNIST database (MSE: 0.025, SSIM: 0.85), a degradation in the DNN performance to produce the same quality of reconstructions is observed. Phase image reconstruction is a more demanding task due to the uniformity of the fiber output images. Better results can be possibly achieved by increasing the number of examples and/or reducing the downsampling factor. However, the scope of this preliminary experiment is to demonstrate that imaging of not only amplitude but also phase inputs can be achieved using the intensity-only output of a MCF. Further work on this, is presented in details in the following paragraphs.

Finally, it is important to note that in an ideal optical system for which there is no coupling among neighboring cores of the bundle and no high frequency loss due to the lens physical apertures or the *NA* of the cores, the intensity distribution of different phase-only images would show no variation. However, abrupt phase changes such as 0 to π correspond to high frequency components that are inevitably filtered out by the physical limitations of the experimental setup, thus creating small intensity variations in the intensity profile of a phase image. In addition, core-to-core coupling in the bundle and the coherence of the source can introduce fluctuations in the final intensity profile of the phase image through the fiber bundle. As a result, sufficient feature differences among the whole dataset exist, thus assisting the training of the DNN model to train and reconstruct the phase profile of the input from the intensity-only recording of the fiber bundle output. In the following sections, it is reported that core coupling is partially essential for phase imaging through fiber bundles, because it creates a contrast mechanism among the various outputs, but on the other hand implies that the final system is sensitive to bending. Bending changes the way that the light couples between the individual cores changing the intensity distribution at the output. Therefore, it can affect the efficiency of the image reconstruction of a trained DNN model. The effect of bending introduced in the imaging system is investigated further in the next section.

7.3 Bending sensitivity of the fiber bundle probe

Bending affects the light propagation through optical fibers and it is one of the fundamental obstacles in MMF endoscopy. Bending of a MMF after measuring its transmission matrix invalidates the calibration and destroys the imaging capability of the system. For this reason, MMFs are mainly limited to rigid endoscopic configurations. Although less bending sensitive, fiber bundles also suffer from bending because of the core coupling. Therefore, wavefront shaping techniques for improving resolution are also limited by bending in the case of fiber bundles. In this section, deep learning approach is investigated for image recovery through fiber bundles in cases that the bending configuration changes. DNNs have proved remarkably robust to noise in the data as discussed in the case of MMFs in Chapter 6 of the thesis, which makes them a promising approach for super-resolution and phase imaging through fiber bundles.

7.3.1 Methods

The same optical setup described in the section 7.2.1 is used to acquire the datasets of image pairs needed to train a U-net DNN. A 7,800 core leached fiber bundle by SCHOTT with a length of 1 m (IBULTRAVIEW, 450 μm core size) replaces the Fujikura imaging bundle used in the discussion above. Leached fiber bundles are more flexible and therefore suitable for bending experiments. In addition, for optimizing the system concerning the imaging quality and bending sensitivity, different light sources are integrated in the optical setup which was previously presented in Figure 31 and they are mentioned below:

- Tunable wavelength laser (M-squared, Solstis 2000 PSX XF with Terascan software, 700-1000 nm)
- Laser diode (Thorlabs, HL63163DG, 633 nm)
- LED (Thorlabs, M625L4, 625 nm)

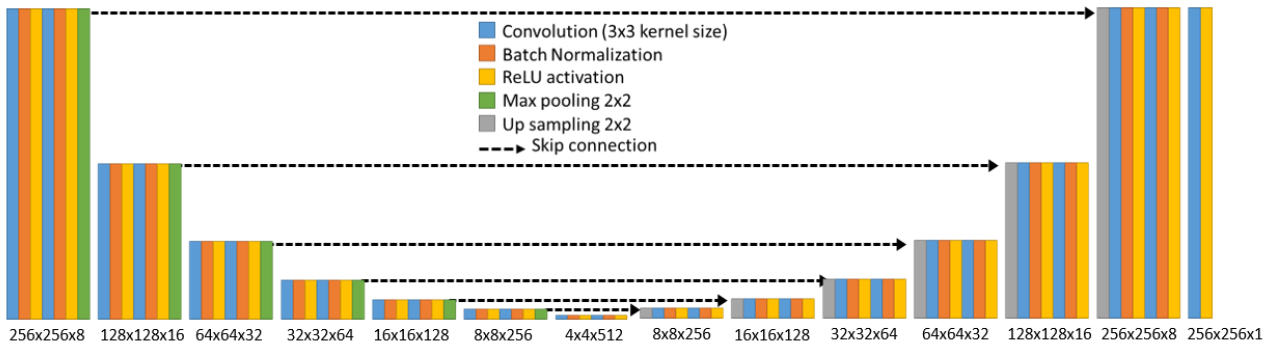


Figure 55. U-net DNN architecture for input images of 256x256 pixels size. Each operation is described with a different color and the number of kernels and size of the image matrix in each block is written below each block.

Datasets of image pairs of SLM inputs and fiber bundle output images were recorded for all the different cases of illumination sources described above. The projected images from the SLM are phase modulated, since they are the ones presenting a more challenging reconstruction problem for the DNNs. For each of the light sources, 10,000 image pairs are collected in one fiber configuration. Afterwards the fiber bundle configuration is modified and 500 extra image pairs are recorded to test the DNN performance in the new bending shape. In order to minimize the effect of downsampling in the DNN performance, the largest image size that the available computer memory can process considering the size of the dataset is chosen, which is 256x256 pixels. The U-net network architecture is thus expanded for the specific image size as described in the Figure 55. The hyperparameters of the DNN are kept the same for all the datasets for comparison; naturally, further tuning might improve the results. The batch size used for the training is 10 and the maximum number of epochs 200. The Adam optimizer is selected with a learning rate of 10^{-5} to minimize a MAE loss function. After the DNN is trained, its performance is evaluated firstly, on the

test data which belong in the same dataset as the training data and on the data collected in the new bending configuration. In this way, the dependence of reconstruction quality on bending is studied.

7.3.2 Results

Achieving high fidelity of image reconstruction using DNNs was successfully demonstrated for both amplitude and phase inputs in the previous section; yet, bending sensitivity is another important parameter to be evaluated for introducing a deep learning based endoscope. In the previously discussed experiments the wavelength was set at 800 nm, which corresponds to the propagation of two-modes per fiber bundle core (V-number = 3.4). Crosstalk between neighboring waveguides is wavelength dependent and specifically increases with the wavelength [35]. In Figure 56 a-c, the digit 0 is projected on the SLM and the respective output images at the distal side of the 1 m long SCHOTT fiber bundle are presented for three different wavelengths, 700 nm, 800 nm, 900 nm. It can be noticed that the light is better confined in the case of 700 nm input wavelength, resulting in a clear and sharp image of the input after its propagation along the fiber bundle. In contrast, for the longest wavelength (900 nm) the light seems to escape from neighboring cores resulting in blurred representation of the input. The main reason is that the mode field diameter increases for longer wavelengths and as a result, the coupling between closed spaced waveguides is more severe [35,38,40].

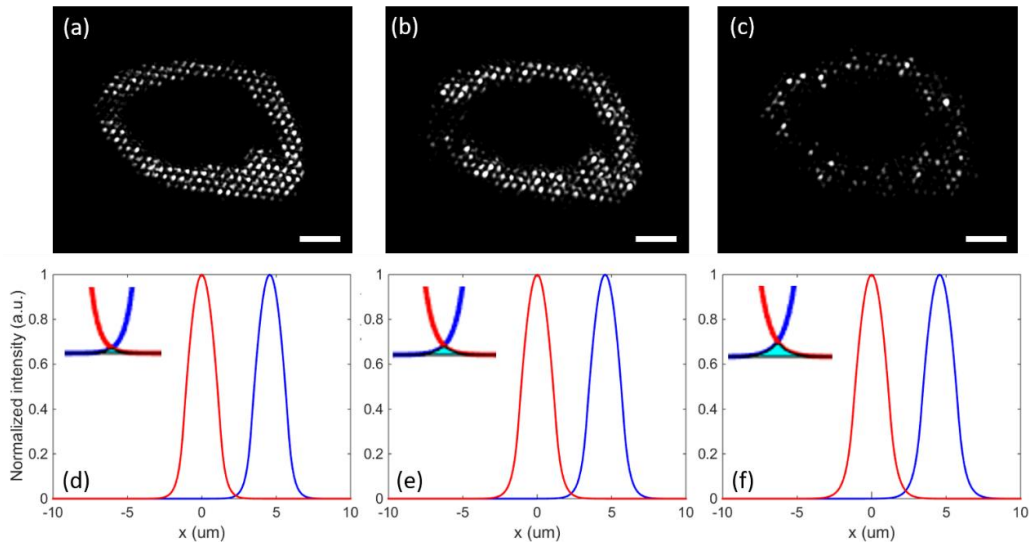


Figure 56. Fiber bundle output images digit 0 for input wavelengths of a) 700 nm, b) 800 nm and c) 900 nm. (d-f) Corresponding 1D waveguide simulation results of the mode coupling. The inset shows magnified the area of overlap. Scale bar: 25 μm .

Moreover, qualitative simulations for the case of 1D waveguides that have the same physical characteristics with the individual cores of the fiber bundle in terms of diameter (2.5 μm), spacing (4.6 μm) and NA (0.39) are shown in Figure 56 d-f. Based on the solution of the wave equation in

the case of 1D planar waveguides the intensity profile of the fundamental mode for the respective waveguides was calculated. In Figure 56 d-f, we plot the modes of two neighboring waveguides. The insets in the Figure 56 (d-f) show the area of overlap magnified for visualization purposes, which demonstrates the increased crosstalk at larger wavelengths. The presented results only serve the purpose of an intuitive correlation between the wavelength changes and the crosstalk observed between neighboring cores. For a more quantitative assessment simulation of the actual 2D model should be performed since the solutions of the wave equation for a cylindrical symmetry are based on Bessel functions and not cosine and sine functions as the planar waveguide case. Quantitative characterization of the crosstalk in fiber bundles has been extensively studied and more details can be found in literature [38]. Overall, selecting a shorter wavelength of illumination is suggested (if possible) since it leads to a better mode confinement and thus lower the bending sensitivity of the final endoscope.

Apart from the wavelength, the effect of the source coherence on the crosstalk level is investigated when delivering an image through the fiber bundle. Because of the better confinement of shorter wavelengths as discussed above (e.g. 700nm), bending of the fiber causes less intensity fluctuations in the individual cores, which form the image. Nevertheless, a decreased reconstruction efficiency is observed when the test inputs in the trained model are images recorded at the distal facet after the fiber bundle is bent in a different configuration. Coupling between the neighboring cores in which light accumulates different phase components, causes interference effects that change the light intensity in a single core. Because of the coherence of the light source, this interference can lead to drastic changes in the intensity of each core at the fiber output leading to black pixels/cores as observed in the case of the 900 nm illumination (Figure 56 c). This effect is even more evident in the case of phase modulated inputs for which the intensity is almost uniform over the fiber bundle proximal surface. Because of the uniform illumination, bending of the fiber causes changes of the intensity to a large number of individual cores/pixels of the output image inflicting “confusion” in the learnable features that the DNN is getting trained to recognize. On the contrary, for amplitude modulated inputs, the light is more localized to certain pixels and the degradation is slightly less severe.

To verify the coherence effect on bending sensitivity and thus the final reconstruction, the tunable laser source is replaced by laser diode emitting at 632 nm and then by an LED source emitting at 625 nm. The coherence length of the various sources is reported in Table 7 for comparison. By changing the light source, it was observed that movement of the fiber in random bending configurations was causing less change in the intensity image recorded at the bundle output in the case of laser diode and no noticeable change in the case of the LED illumination. For all the different light sources, the rest of the experimental setup was kept the same as described in the Figure 31.

In Table 7, the summary of the image reconstruction results is presented for the three datasets obtained for the difference sources evaluated on test images recorded before and after changing the bending configuration of the fiber bundle. First, it is important to notice that reduction of the downsampling applied on the fiber bundle output results in improved reconstruction of the MNIST

phase inputs; the SSIM is calculated 0.88 for the 256x256 image size as opposed to SSIM of 0.82 obtained for the 64x64 image size reported in the previous section (Figure 54). Secondly, a comparison among the reconstruction results reported in Table 7 concerning the MNIST database, shows increase of the U-net reconstruction performance, when the coherence of the light source and the wavelength are decreased. In the last column of Table 7, a qualitative measure of the degradation of the DNN capability to recover the proximal phase image is reported. The degradation metric is calculated as a division between the MSE and SSIM after bending over the MSE and SSIM before bending; values close to one imply high resistance of the reconstruction against bending induced noise in the data.

Table 7. Image reconstruction results for inputs obtained from MNIST and MNIST-Fashion databases in the case of three light sources. The results are presented for test data obtained without moving the fiber and after the fiber is set in a new configuration to study the effects of bending in the performance of the U-net DNN.

Laser source	Database	Metrics		
		Before bending	After bending	Degradation
700nm (CW tunable laser) $L_{\text{coherence}} \approx 1 \times 10^3 \text{ m}$	MNIST	MSE: 0.047 \pm 0.020 SSIM: 0.88 \pm 0.04	MSE: 0.103 \pm 0.031 SSIM: 0.78 \pm 0.04	MSE: 2.2 SSIM: 0.88
	MNIST-F	-	-	-
632nm (Laser diode) $L_{\text{coherence}} \approx 1 \times 10^2 \text{ m}$	MNIST	MSE: 0.043 \pm 0.019 SSIM: 0.89 \pm 0.03	MSE: 0.060 \pm 0.023 SSIM: 0.85 \pm 0.04	MSE: 1.4 SSIM: 0.95
	MNIST-F	MSE: 0.014 \pm 0.010 SSIM: 0.83 \pm 0.09	MSE: 0.049 \pm 0.024 SSIM: 0.64 \pm 0.13	MSE: 3.5 SSIM: 0.75
625nm (LED) $L_{\text{coherence}} \approx 1 \times 10^{-5} \text{ m}$	MNIST	MSE: 0.041 \pm 0.022 SSIM: 0.87 \pm 0.04	MSE: 0.046 \pm 0.018 SSIM: 0.88 \pm 0.03	MSE: 1.1 SSIM: 1.01
	MNIST-F	MSE: 0.008 \pm 0.006 SSIM: 0.87 \pm 0.07	MSE: 0.006 \pm 0.004 SSIM: 0.88 \pm 0.06	MSE: 0.75 SSIM: 1.01

In Figure 57, the reconstruction results generated by the U-net for one example of the digit 4 is presented for each light source. The three columns of images refer to the different laser sources; the first column corresponds to CW laser illumination at 700 nm (Figure 57 a, d, g), the second to the laser diode (Figure 57 b, e, h) and the third to the LED (Figure 57 c, f, i). In the case CW laser, for which the performance degradation is severe as shown in Table 7 (MSE almost twice worse), the reconstructed image of the digit by the U-net is not recognizable in Figure 57 g. Remarkable reconstruction quality is though achieved in the case of LED illumination. The fidelity of the reconstructed phase images is verified by both the images presented in Figure 57 c, f, i and by the metrics, MSE and SSIM reported in Table 7; no significant change is observed when the trained model is tested on data collected after bending the fiber bundle in a new configuration (Figure 57 i).

As discussed in the subsection 7.2.3 of the thesis more complex images (CIFAR-10) in terms of features but also intensity dynamic range (gray levels) present a more challenging task for

reconstruction by DNN algorithms in comparison to MNIST binary like digits. Hence, degradation of the image quality after bending is expected to be more severe the more complicated the input images in the U-net are. Phase inputs from the MNIST-F dataset are used to train and test the U-net performance in the case of gray scale inputs. The image quality of the reconstructions is severely lower after the fiber is bent when the laser diode source is used while it remains unchanged for the image data acquired with the LED illumination. The quality metrics of the reconstruction of MNIST-F phase images propagating through the fiber bundle are presented in Table 7 together with the results obtained for MNIST for comparison. Examples of the image reconstructions for the two light sources in the case of MNSIT-F can be found in more detail in the Appendix A3. The results summarized in Table 7, support the assumptions concerning the effect of both wavelength and coherence of the light source on the bending sensitivity of the fiber bundle endoscope.

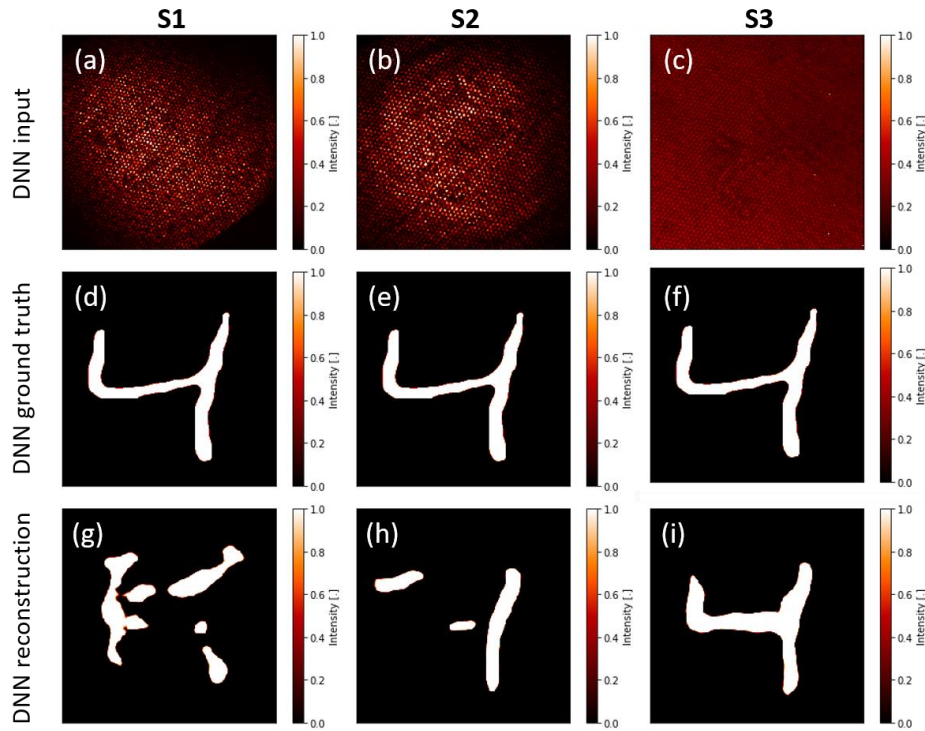


Figure 57. U-net DNN reconstruction results for the three light sources for data acquired after bending the fiber bundle. a-c) Fiber bundle output downsampled at 256x256 pixels, d-f) the corresponding ground truth image and g-i) the reconstructed images by the DNN for the respective inputs (a-c). The light sources are indicated as: S1 for the tunable laser at 700nm, S2 for the laser diode at 632nm and S3 for the LED at 625nm.

7.4 Deep learning assisted phase contrast endoscope

In the case of amplitude modulated inputs described in previous section information could be accurately retrieved by conventional image processing but significantly enhanced by DNNs. Nevertheless, in endoscopy tissue samples usually offer low contrast for bright field microscopy

because they are mostly transparent, phase objects. In clinical biopsy, contrast between different parts of the cell such as the nucleus or mitochondria is generated either by fluorescence dyes that attach selectively in the desired parts of the cell or tissue or by absorption-based staining chemicals. Staining protocols allow fluorescence or bright field microscopy to inspect the morphology of the biological sample. Despite of the well-established staining protocols in histopathology, certain shortcomings are present related to the tissue viability and result reproducibility. It has been observed that the final image of a stained biological sample can differ from one histopathologist to another who performs a specific staining protocol, or even from sample to sample treated by the same protocol and expert. As a result, artefacts generated during the sample preparation process can introduce uncertainty in the result interpretation by the final user [160,180]. In addition, most of the chemical reagents used for fluorescence or absorption-contrast microscopy are only suitable for *ex vivo* imaging of fixed biological specimens. Bio-compatible staining protocols partly solve possible toxicity problems but they might interfere with the natural processes in the tissue, causing an unpredictable biological response.

As a consequence, label-free imaging techniques have attracted great interest over the years for biomedical imaging [158,160,180–183]. Label-free microscopy can be achieved in different ways, namely taking advantage of the endogenous autofluorescence of the tissue, recording second-harmonic generation signal originating from the structural properties of the specific sample, Raman and phase microscopy [93,121,122,154,158,160,173–176,180,181,184,185]. In particular, phase contrast imaging and digital holographic microscopy [173,186–188] are widely used for imaging of transparent (phase) objects. Inexpensive illumination sources and low-phototoxicity make phase imaging a suitable microscopy method for both quantitative and qualitative phase measurements, which can be realized both *in vivo* and *ex vivo*. Moreover, the emergence of machine learning has managed to bridge the gap between different microscopy methods using data-driven DNNs to translate the information among different modalities such as: intensity to phase imaging [59–61,158], phase, autofluorescence and multiphoton microscopy to H&E (haematoxylin and eosin) stained bright field imaging [160,180,183], confocal to stimulated emission depletion microscopy [156] etc.

In this section, information translation via deep learning is introduced in the field of fiber endoscopy. This novel approach aims to enhance the image quality of a commercial endoscope by introducing further imaging modalities. Specifically, it is shown that bright field images obtained at the distal facet of the fiber bundle can be converted to phase images revealing details that are not easily (if not at all) detected in the original bright field image. It has been already demonstrated that phase imaging from intensity-only image recordings is possible by using the intensity variations caused by the Fresnel propagation [158]. In this study, the information content in the intensity image comes mainly from the inherent scattering of the tissue, the high-frequency filtering by the finite *NA* of the optical components and partially induced by the core-to-core coupling. Therefore, DNNs can be used to not only to improve the intensity image resolution beyond the core sampling, but also to reconstruct the phase information of the biological sample under examination. The presented technique could be generalized for different microscopy

methods such as fluorescence or confocal by properly training the DNN model to include further information channels allowing for digitally integrated multimodal endoscopic imaging.

7.4.1 Methods

7.4.1.1 Optical apparatus

Figure 58 shows the experimental setup that is built to generate the necessary datasets for training a DNN to translate intensity images of biological samples recorded by fiber bundle endoscope to phase images. The optical setup consist of two parts: a digital holographic microscope part implemented using a diode pumped solid state (DPSS) laser emitting at 532 nm and the fiber bundle imaging system using a red LED (Thorlabs, M625L4, 625 nm) for sample illumination. The proposed configuration is built in a way to exploit the LED incoherent illumination, which shows good quality reconstructions results as confirmed in the previous paragraph even in the presence of bending, and the coherent laser source to generate high contrast holograms of the sample, from which the phase image of the sample is extracted. The phase images of the biological samples obtained by the holographic recording serve as ground truth for the DNN training, while the intensity images of the sample collected through the fiber bundle will be the inputs in the DNN algorithm.

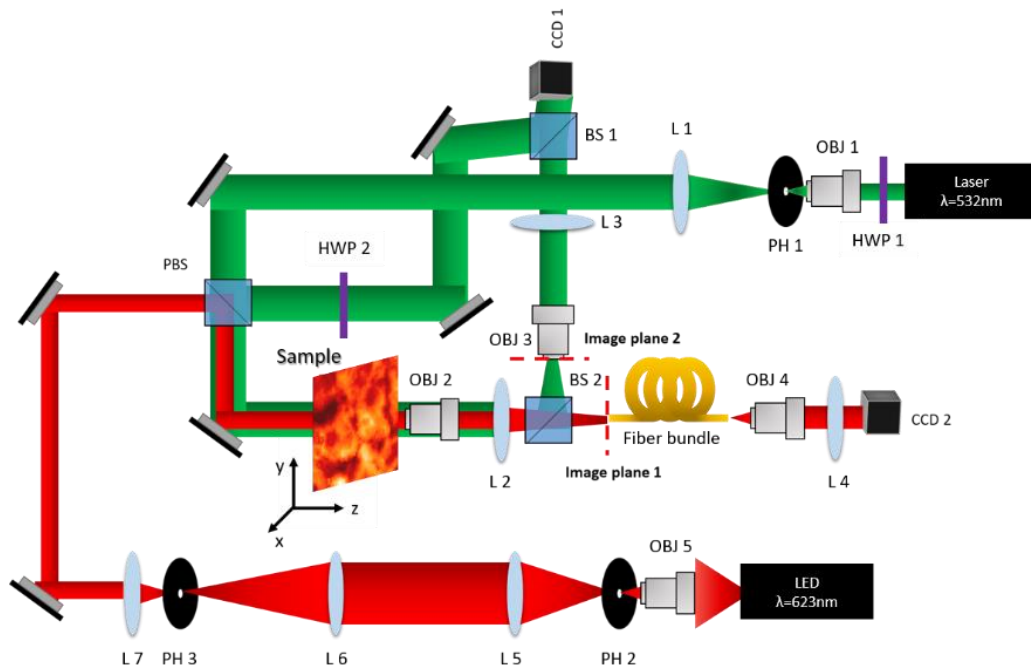


Figure 58. Experimental setup for collecting bright field images of tissue samples through the SCHOTT fiber bundle (red light path) and the digital holographic microscope used to obtain the phase information from the sample (green light path). The digital holography is performed using the coherent beam of a DPSS laser emitting at 532 nm. The fiber imaging is performed by illuminating the sample with a low power red LED source emitting at 625 nm. The two light sources have been properly aligned to illuminate the same area of the sample and spectral filters are used to separate the signal recorded by each source on the two CCD cameras. (L = lens, OBJ = microscope objective, HWP = half-wave plate, PBS = polarizing beam splitter, BS = beam splitter, PH: pinhole)

For the digital holographic microscope the beam of the green DPSS laser is spatially cleaned using a pinhole (PH 1: 100 μm) in front of a microscope objective of 5x magnification (Newport, NA=0.10) and then collimated by a 200 mm lens (L1). Similarly to the previously described holographic setups, the expanded beam is split into two arms, the reference and illumination arm, by means of a polarizing beam splitter (PBS). A halfwave plate (HWP 1) placed after the laser and combined with the PBS controls the laser power ratio between the two arms. A second halfwave plate (HWP 2) after the PBS ensures that the reference and illumination path have the same polarization when they recombine through the beam splitter (BS 1: Thorlabs, BS013, 50:50) to form digital off-axis hologram on the camera (CCD1: Chameleon 3, 1024x1280 pixels, Mono, Point Grey). In the illumination arm, the sample is placed in front of the laser beam and it can be moved in space by a 3D motorized stage (Thorlabs, T-cube). A 4f system (OBJ2: Newport, 10x, NA=0.25, L2: f=80 mm) magnifies the sample on the proximal fiber facet. The image plane of this first 4f system is split into two by another beam splitter (BS 2: Thorlabs, BS028, 90:10) as illustrated in Figure 58. The image plane 1 corresponds to the proximal fiber facet plane and the image plane 2 is located at the focal distance of a second 4f system (OBJ3: Newport, 20x, NA=0.45, L3: f=150 mm) placed in the illumination arm to further magnify the sample on the CCD1 where the digital hologram is formed. This part of the experimental setup is recording the digital hologram of the area of the sample so that its phase image can be afterwards extracted.

The fiber imaging of the tissue sample is performed using the red LED source (Thorlabs, M625L4, 625 nm) that is integrated in the same path with the green laser illumination arm. In order to be well collimated the LED is spatially filtered twice as shown in Figure 58. The LED light is tightly focused using a 40x microscope objective (OBJ5: Newport, NA=0.65) onto a pinhole of 200 μm (PH 2) and then recollimated by a lens (L5: f=100 mm). Then the beam size is demagnified by a telescope of two lenses (L6: f=200 mm and L7: f=50 mm) and further cleaned by adding another pinhole (PH 3) with 500 μm diameter at the focal plane of L6. The red beam is aligned in the same path with the green beam to ensure that they both illuminate the same area of the sample. The 4f system (OBJ2-L2) which is located just after the sample forms the image on the fiber bundle facet as described above. The fiber bundle used for this experiment is an 1 m long SCHOTT leached fiber bundle with 7,800 individual cores (IBULTRAVIEW, 450 μm core size). The diameter of the cores is approximately 2.5 μm , NA of 0.35 and the core separation 4.5 μm . The fiber bundle output is imaged on a second camera (CCD2: Chameleon 3, 1024x1280 pixels, Mono, Point Grey) by means of a magnification system (OBJ4: Newport, 10x, NA=0.25, L4: f=200 mm). A long pass filter (at 550 nm) is placed in front of the proximal fiber facet to block the green laser beam from entering the fiber bundle so that the imaging of the sample is performed using the LED light only. Likewise, a bandpass filter with a central wavelength at 532nm (20nm bandwidth) is added in front of the CCD1 to prevent the LED incoherent light from generating extra background noise in the hologram recording.

An important parameter for the data collection in the above system is to verify that the image of the sample delivered by the fiber bundle and recorded by CCD2 is aligned with the corresponding digital hologram generated on CCD1. This can be achieved according to the following steps:

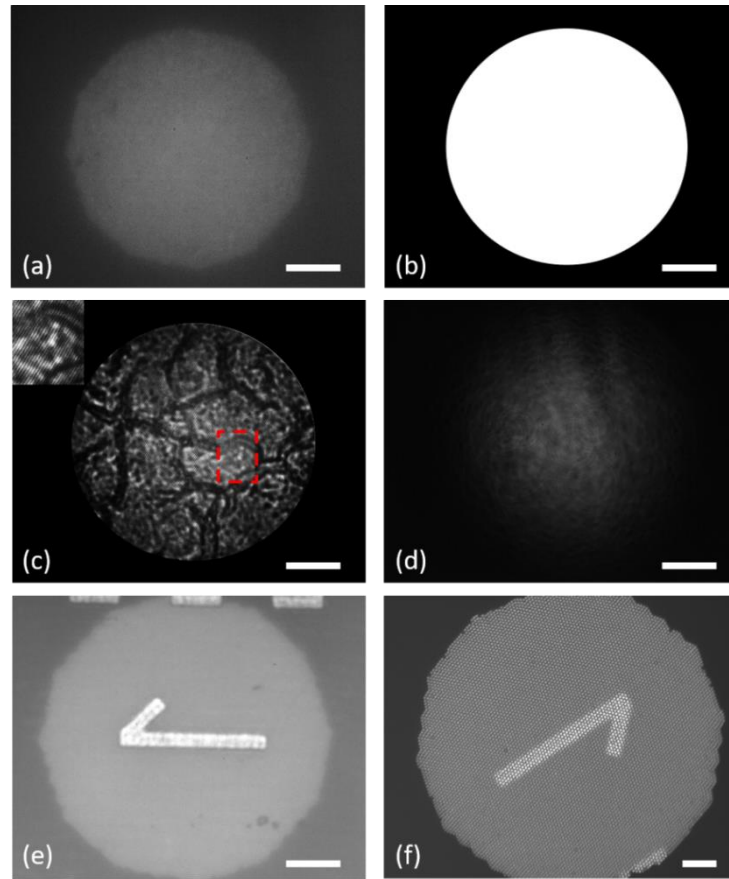


Figure 59. Description of the physical and digital alignment between the proximal and distal images of the dataset. a) Image of the proximal fiber facet when white light illumination is used at the distal side at CCD1 in Figure 58, b) the amplitude mask that corresponds to the facet area at CCD1, c) the reference beam aligned at the center of the proximal facet region indicated in image (a), an example of a digital hologram recorded at CCD1 multiplied by (b), g) pattern image at the proximal camera and f) the corresponding image of (g) through the fiber bundle captured by CCD2. Scale bar: 100 μm .

1. The sample is placed in the right position so that it is in focus at both cameras, CCD1 and CCD2.
2. The laser and LED are blocked and a white light source is illuminating the fiber from the distal side backpropagating in the system.
3. The sample is moved so that an empty from sample region, where only the glass coverslip reflects the white light, which is then imaged on the CCD1 (Figure 59a). In this way, an image of the core at the proximal side is formed at the same camera that records the hologram.
4. A digital circular mask is then generated (Figure 59 b), based on the proximal facet image (Figure 59 a) and it is used to select the area of the hologram (Figure 59 c) which is the one that corresponds to the one imaged by the fiber bundle.
5. Then the reference arm is unblocked and the reference is aligned to overlap exactly with the proximal facet image (Figure 59 d).
6. After the proximal side is correctly aligned, there is one final step to take into account. The individual cores in the SCHOTT bundle are twisted and this results in a rotation of the

image by the distal fiber side. In addition, mirrors, beam splitters and lenses, flip up-down and left-right the images so it is important to make sure that the fiber bundle output is processed properly to match the proximal input. To do so we project a known pattern at the proximal side (Figure 59 e) from the USAF-1951 resolution target and record the fiber bundle output at CCD2 (Figure 59 f). Then the output image is transformed digitally so that it matches the input image captured by the CCD1.

Following the procedure as described in these steps, alignment of both fiber facets physically can be achieved. Additionally, all the information needed to process and correct for the rotation or slight shift induced misalignments between the image pairs can be performed digitally. The data acquisition is implemented in Matlab. The Matlab code controls and synchronizes the 3D stage with the sample, the proximal hologram recording on CCD1 and the distal fiber output image recording on CCD2. A digital phase extraction code was also demonstrated to provide a continuous visual feedback of the sample phase by processing the CCD1 frames of the corresponding hologram images. Using this Matlab code and moving in an area outside the sample, it is also possible to minimize the aberrations and misalignments existing in the digital holographic part of the optical setup by properly moving the positions of the optical components while monitoring the real-time phase feed from the Matlab window, thus achieving a flat phase profile. The digital phase extraction is made following the same process described in the Chapter 2 of the thesis.

7.4.1.2 Sample preparation

In the experiments presented in the following paragraph the samples are thin histological tissue sections of mouse liver placed on thin glass coverslips (175 μm thickness). The 10 μm thick tissue sections are obtained by the EPFL histology core facility (HCF) using a cryostat microtome device and stored in the -20°C . Before the experiments, the samples were fixed using cold 4% paraformaldehyde (PFA) in 0.1 M phosphate-buffered saline (PhBS) for 30 min and rinsed with PBS three times for 5 minutes each. Afterwards the sample is sealed by another thin square coverslip using double sided tape as spacer to avoid dehydration during the measurement.

7.4.1.3 Deep Neural Network architectures

For the phase imaging endoscope demonstrated in this section two different DNN architectures are used: a U-net type and a Generative Adversarial Network (GAN) type of DNN as described in Figure 60. The U-net architecture used in the experiments presented in the next paragraph has the same structure as the one presented previously (Figure 55). GANs are advanced DNN algorithms that are capable of learning deep representations from data that are not annotated in detail and they find applications in various tasks such as image super-resolution, image classification, image synthesis and so on [189]. GANs consist of two DNNs that are known as generator and discriminator. The generator is a DNN that generates output data from noise or certain inputs and the discriminator is a DNN classifier that receives the generator output and characterizes them as fake or as true if they correspond to ground truth data. The generator and discriminator networks

are trained simultaneously in an adversarial process. During the training process the two networks compete in a way that provides high quality results, since the generator gets optimized to produce more realistic data, while the discriminator on the other hand tries to improve its classification accuracy. Another intuitive way to understand the GAN training principle is to consider the discriminator as a highly complex error function that enhances the quality of the images produced by the generator. After the training process is completed the generator model is saved and it is the one used to produce the desired images depending on the task.

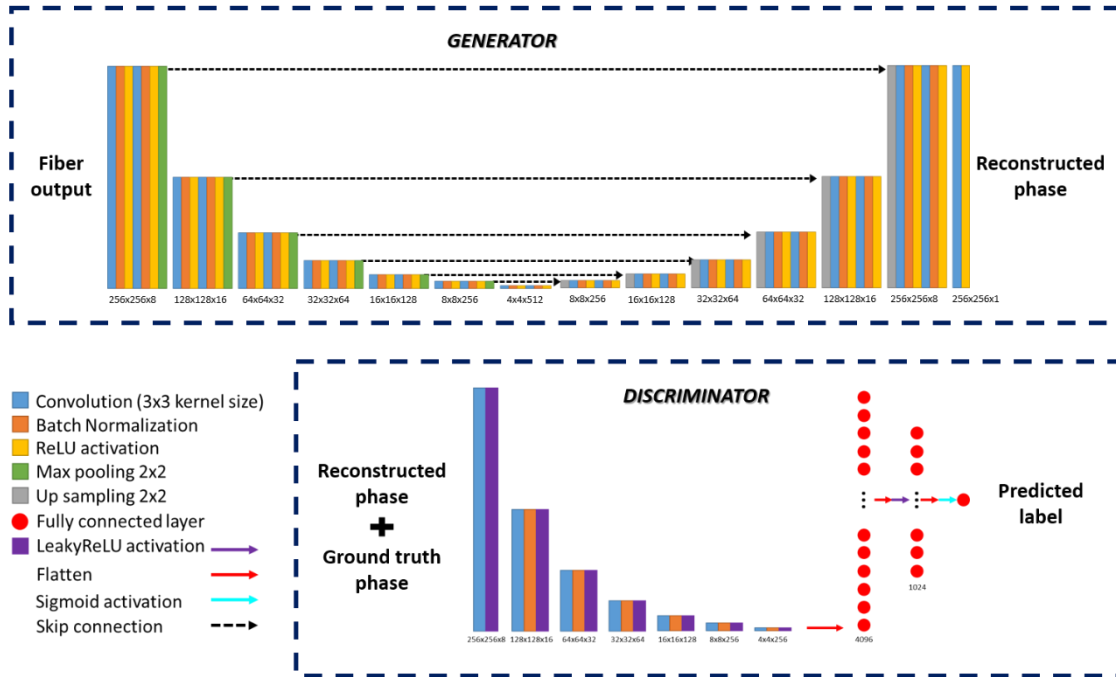


Figure 60. GAN-type DNN used for the reconstruction of phase images from the intensity output of the fiber bundle.

In this work, a special case of GAN that is called conditional GAN because during training instead of receiving random noise as input, it receives the intensity images of the fiber bundle output. The U-net architecture described in Figure 55 is integrated unchanged as the generator part of the GAN DNN. The discriminator part of the GAN is a classifier type DNN similar to the VGG DNN described in the Chapter 6. The main difference is that each convolutional block of the discriminator DNN consist of only one set of 2D convolution and batch normalization operations instead of two in the previous case (Figure 32 a). Moreover, the activation function in this case is changed from ReLU to Leaky ReLU as suggested in most of the works using GAN-type DNNs for super-resolution to achieve more stable training [190,191]. Different types of loss functions are used for the U-net and the GAN are explained in more detail in the following paragraphs. An Adam-type optimizer is used with a learning rate of 10^{-5} for the U-net and 10^{-4} for the GAN. For the implementation of the DNNs a computer equipped with the NVIDIA GeForce GTX 1080Ti graphics processing unit is used as in the previous cases presented in the previous chapter. More details on the GAN used in for the results presented in this section can be found in the Python code of the Appendix A4.

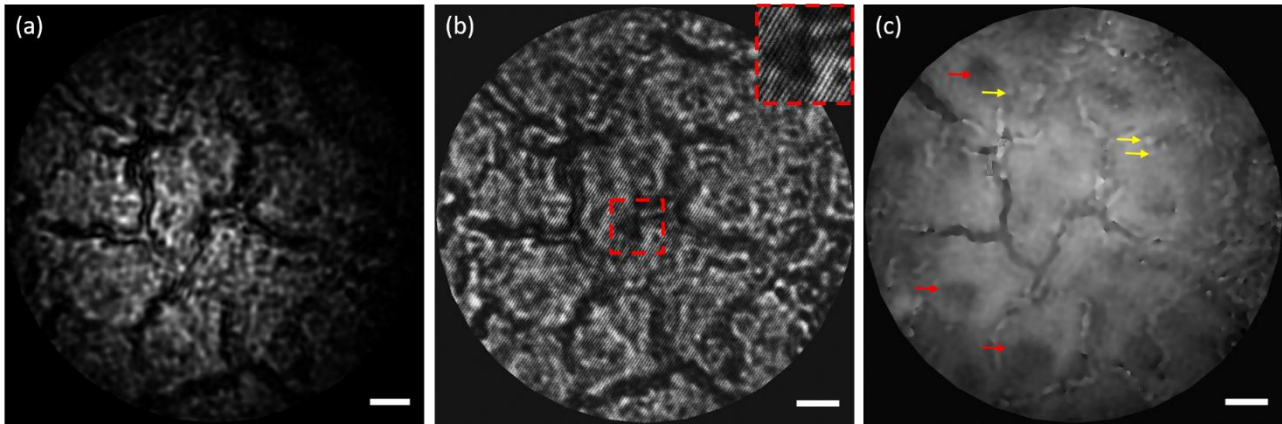


Figure 61. a) Intensity-only image of a liver tissue sample, b) intensity image of the hologram of the same liver tissue (the red square shows a magnified a region of the hologram to visualize the interference fringes) and c) the extracted phase from (b). The red arrows show low phase areas that correspond to the cell nucleus and the yellow arrows show small features of high phase values attributed to lipids. The scale bar is 10 μm

7.4.2 Phase endoscopy of liver tissue samples using deep learning

For many pathologies, detailed understanding is based on information available at a cellular level of a tissue instead of its macroscopic structural image [175,182,192,193]. Bright field endoscopy can only provide limited information about the cell image of a certain tissue sample based on the low contrast generated by the scattered and diffracted light. However, as stated before phase images offer sufficient contrast for the inspection of cellular structure details despite the low refractive index contrast among the cell components. Phase-contrast endoscopy has been attempted through fiber bundles using dark-field based techniques by illuminating the sample with sharp angles in order to detect high frequency components of the image that bright-field microscopy cannot because the small refractive index variations do not modify the intensity of the collected light through the sample [177]. In this last section of the thesis, a novel method to obtain not only phase contrast images but also quantitative phase measurements through a fiber bundle endoscope is proposed using deep learning. In the next paragraphs, without loss of generalization, only the normalized phase profiles between 0 and 1 are presented for the studied tissue samples for simplicity. The proposed method can be also translated to multiple imaging modalities and opens the way of low-cost digital modification of current fiber bundle based endoscopes to expand their imaging content.

The biological samples investigated in the experiments described in this section consist of histological slices of mouse liver tissue fixed on a glass slide and placed in the experimental setup depicted in the Figure 58. In the digital holographic microscope part of the optical setup, both the intensity-only image (Figure 61 a) and the digital hologram (Figure 61 b) of the sample can be recorded by blocking or not the reference arm, respectively. Using the principles of digital holographic phase extraction, the corresponding phase map of the sample can be visualized (Figure 61c). The images shown in Figure 61 are masked by a circle that corresponds to the fiber bundle core diameter (as described in the Methods paragraph above). Because of the thickness of

the sample, phase wrapping can be present in the phase map extracted from the hologram. Therefore, an extra processing step is used for phase unwrapping using the PUMA algorithm as proposed by *Bioucas-Dias and G. Valadao* [194]. Noise or aliasing in the raw extracted phase of the hologram can affect the performance of the phase unwrapping and create artefacts of very high or very low phase values in the unwrapped image. Although further optimization of the phase extraction is needed to achieve accurate quantitative measurements, this method is sufficient for training the DNNs to perform phase imaging through the fiber bundle endoscope, as it is further explained in the following paragraphs.

Figure 61 c shows that phase imaging reveals details of the sample that are otherwise undetectable in the intensity-only images. Specifically, the cell nuclei are well defined in the phase image and the red arrows indicate some of them for clarity. In addition, some bright round parts shown in the cell indicated by the yellow arrows are most probably related to lipid droplets often found in a cell [109,195].

The training dataset is created by cropping a rectangular region of the unwrapped phase in a way that matches the corresponding fiber bundle output intensity image and form a training pair as shown in Figure 62. Different image pairs are collected by raster scanning in x-y plane of the sample in front of the imaging objective (OBJ2) using a 3D motorized stage. Every step of 2 μm a hologram and a distal intensity image is captured and 4000 images of different sample regions are collected. Moreover, data augmentation is performed to ensure that enough pairs are available for the network training. The data augmentation includes 4 rotations of 90° and a vertical flip for each rotation. Data augmentation is necessary in most cases to avoid long measurements, which could cause degradation of the sample.

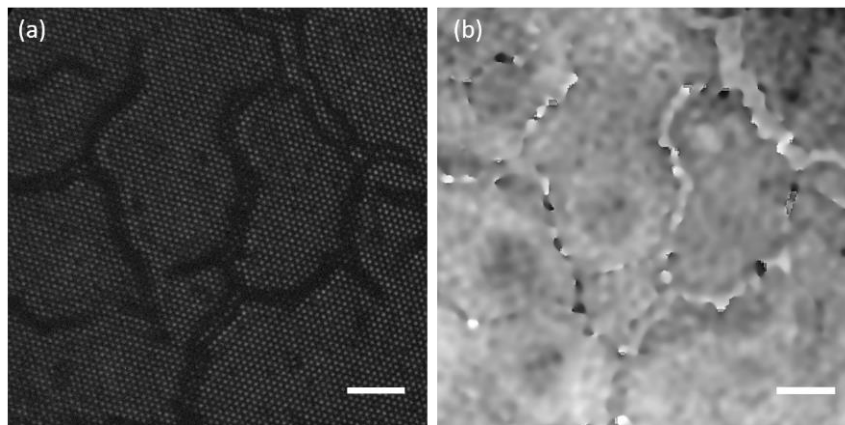


Figure 62. DNN training image pair, a) fiber bundle output and b) its proximal phase extracted by the holographic setup at the proximal side. Scale bar: 10 μm .

7.4.2.1 Phase imaging using U-net DNNs

After the images have been processed and cropped to have a co-registered image pair (as much as possible), they are downsampled to 256x256 pixels to train the DNN. The U-net type DNN

described in the previous paragraph (Figure 55) is used to reconstruct the phase image of the tissue sample having as inputs the corresponding intensity images as they are recorded at the distal facet of the fiber bundle. The hyperparameters of the U-net are kept similar to the previous paragraph (loss function=MSE, Adam optimizer: learning rate= 10^{-5} , batch size=10, epochs=400). 25,000 image pairs comprise the dataset used for the DNN training, from which 90% were part of the training set, 5% used for validation and 5% for testing the DNN performance. The average time of the training lies in the range of 5-6 h on the available computer.

The results of the reconstruction using the above parameters render a smoothed output image of the phase (Figure 63 a-iii). Despite of the smoothing effect, the information content of the reconstructed image shown Figure 63 a-iii is much higher concerning the cellular morphology of the tissue sample if compared to the pure intensity image sampled by the fiber bundle (Figure 63 a-i). Darker regions of low phase correspond to the cell nucleus. The smoothing effect of the reconstruction is often observed when MSE is selected as a loss function. MSE is not a pixel-wise loss and calculates a global error in the image reconstruction and therefore large errors in some pixels are compensated by small error in other areas of the image. MSE tends to keep the overall content of the image but it weakly assists the training of the network to improve structural details of the reconstructed image.

It has been proposed by *Li and Barbastathis* [196] that pre-processing of an image to enhance the high frequency components can help to achieve a better reconstruction, because it compensates for the high frequency loss. In this work, an alternative way to deal with the reconstruction smoothing is proposed by introducing a custom loss function. The suggested custom loss consists of two error-determining components, the MSE and the mean Sobel error (MSoE) that are weighted based on the quality of the reconstruction results. To calculate the MSoE, the Sobel filter is applied to both the predicted image from the U-net and the ground truth and then the mean squared error of the Sobel convolved images is calculated.

The Sobel filter is well-known in image processing because it is often used in edge detection applications. In fact, it is a 3x3 kernel matrix operator that calculated the gradient of intensity in an image point by point. Convolution of an image with the Sobel operator results in an edge enhanced version of the image that contains the areas of high intensity gradient. The Sobel operator consists of two 3x3 kernel matrices g_x and g_y (Equation 27), one for the x and one for the y derivative of the image. If I is the image under processing, the final intensity gradient magnitude of the image $I_{gradient}$ is given by the square root of the sum of the convolutions of I with each kernel, as described by the following equations:

$$g_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} \quad g_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad I_{gradient} = \sqrt{g_x * I + g_y * I}$$

Equation 27. Definition of the Sobel filter operator

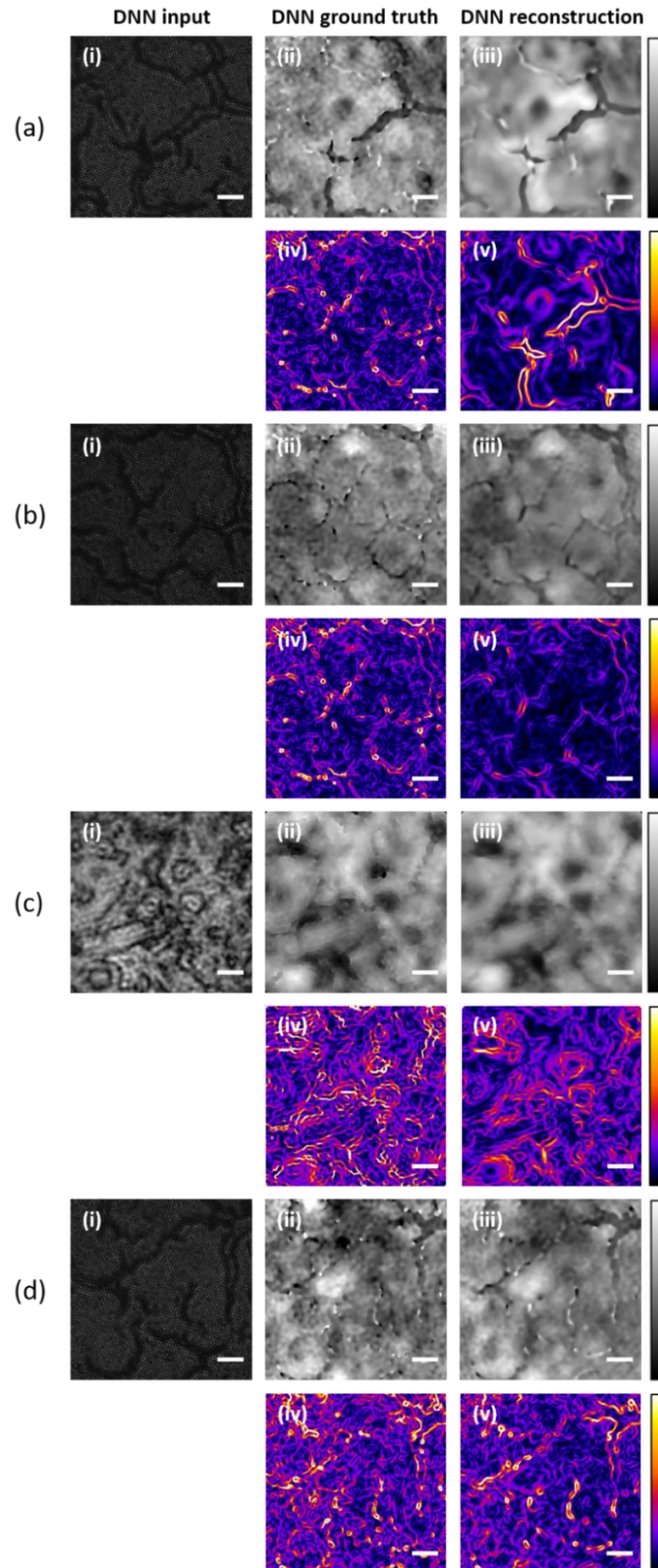


Figure 63. The reconstruction results for different learning parameters are presented. In all cases i) the DNN input, ii) the ground truth phase image, iii) the reconstructed phase from (i), iv) and v) the maximum intensity gradient map after applying the Sobel operator at the phase images (ii) and (iii) respectively. The phase image is reconstructed a) from distal fiber output using U-net with MSE loss and b) U-net with MSE and Sobel error custom loss, c) from the intensity at the proximal fiber side using U-net with MSE and Sobel error custom loss and d) from distal fiber output using a GAN-type DNN with MSE and Sobel error custom loss for the generator. Scale bar: 10 μm .

In Figure 63, the high frequency gradient maps are included for both the ground truth and the reconstructed images calculated using the convolution with the Sobel operator. The rows iv and v of Figure 63 present the high frequency gradient maps for the different DNN results, in order to directly compare the effect of various loss functions on the high frequency content of the reconstructions performed by the U-net DNN. The quality metrics of the phase reconstruction through the fiber bundle are also summarized in Table 8.

In the case of phase imaging using a MSE loss functions the smoothing effect is evident not only comparing the ground truth (Figure 63 a-ii) and the reconstructed image (Figure 63 a-iii) but also by their corresponding high frequency maps shown in Figure 63 a-iv and Figure 63 a-v, respectively. However, implementing a custom loss related to the high frequencies of the image, such as the MSoE, improves significantly the quality of the U-net generated images reducing the MSE to 0.004 and reaching a SSIM of 0.90 instead of the 0.82 achieved by simply using an MSE loss function. This improvement is even more evident in the respective Sobel edge maps of the ground truth and reconstruction presented in the Figure 63 b-iv and Figure 63 b-v. After testing different combinations of weights between the two error functions, the best results are reached when MSE and MSoE are weighted in the loss function with a ratio of 1:20 respectively.

For further evaluation of the results of the deep learning-based phase imaging through the fiber bundle, it is interesting to compare the performance of U-net to reconstruct the sample phase from the intensity images of the sample recorded on the CCD1, before the fiber bundle (Figure 58). The training parameters of the U-net are kept the same to correlate the results. A dataset of image pairs that consist of intensity-only images of the sample as they would have recorded by a bright-field microscope and their respective phase map is generated by analyzing the holograms. Specifically, the intensity images are obtained from the amplitude part of the processed hologram recorded at the proximal end (Figure 63 c-i) and serve as inputs to the U-net DNN instead of the fiber bundle outputs, while the phase part of the hologram is used again as the ground truth (Figure 63 c-ii).

The results presented in the third row of Table 8 show that the reconstruction of the phase from the corresponding intensity image is of higher quality (Figure 63 c-iii) than the one obtained when the intensity of the sample is recorded at the distal facet of the fiber bundle (second row of the table). The pixelated imaging caused by the fiber bundle sampling is found to hinder the reconstruction performance. It is important to acknowledge though, that the quality of the DNN generated phase images in the two cases is not significantly different considering the complexity of the inversion problem. Furthermore, in the case of reconstruction based on the intensity images of the sample before the fiber bundle, the co-registration of the information between intensity and phase image pairs is ideal, since they are part of the same image. On the contrary, in the case of the fiber bundle intensity images small shifts can exist increasing the difficulty of the reconstruction further to core sampling effect. Although these errors are present in the phase recovery from the fiber distal image, the Sobel error factor in the custom loss seems to improve the phase reconstruction of the endoscope in a comparative fidelity to the one reported in the case of proximal intensity images as DNN inputs.

Table 8. Phase reconstruction results through the fiber bundle endoscope for different inputs and DNN learning parameters.

DNN architecture	Loss function	DNN input	Metrics
U-net (256x256)	MSE	Fiber distal Intensity image	MSE: 0.006±0.003 SSIM: 0.82±0.02
U-net (256x256)	MSE + SOBEL	Fiber distal intensity image	MSE: 0.004±0.004 SSIM: 0.89±0.02
U-net (256x256)	MSE + SOBEL	Proximal sample Intensity image	MSE: 0.003±0.003 SSIM: 0.91±0.04
GAN (256x256)	MSE + SOBEL (G)	Fiber distal Intensity image	MSE: 0.003±0.002 SSIM: 0.92±0.03

It is important to note that part of the high frequency content in the phase images, which serve as ground truth for the training, is attributed to the speckle noise induced by the coherence of the laser source. Therefore, it can be assumed that the U-net reconstruction is closer to the real sample phase than indicated by the comparison with the “non-perfect” ground truth. In the future, this problem can be addressed by using less coherent illumination in the digital holographic setup like a diode laser or an LED.

7.4.2.2 Phase imaging using GANs

Finally, the advantage of using a GAN-type DNN for the realization of a fiber bundle phase endoscope is investigated. GANs have shown remarkable results in super-resolution problems due to the pixel wise error monitoring during training instead of the global image one achieved by conventional architectures such as the U-net presented before. The GAN architecture which is used in the experiments discussed is described in the previous section 7.4.1.3 and it consists of a U-net type generator and a VGG-type classifier discriminator. Based on the promising reconstruction quality achieved using the custom loss of MSE and MSoE implemented in the U-net DNN, the same parameters are kept for the generator. For the discriminator, binary cross-entropy is used as a loss function. The final loss consists of the content and adversarial loss with a weighting ratio 1000:1. The training of the GAN takes about 3 days using a batch size of 10, 400 epochs.

The results of the GAN phase reconstruction are included in Figure 63 d for comparison with those achieved by the U-net for the same custom loss. The average MSE and SSIM among the 500 test images are 0.003 and 0.92 obtained by the generator after GAN training is finished. A significant improvement in the image quality of the reconstruction is observed with respect to the U-net with custom loss which is apparent from differences between the Figure 63 d-iii and the Figure 63 b-iii. Moreover, the quality of the recovered phase image from the intensity-only fiber bundle output using the GAN is equally good to the one obtained in the absence of fiber, namely using the intensity image of the sample at the proximal side (Figure 63 c). Conceptually, the superior quality of phase imaging using the GAN architecture can be better perceived looking at the related high

frequency maps presented in Figure 63 d-iv and Figure 63 d-v, which confirm that the majority of the high frequency components of the ground truth phase image are preserved and they are present in the reconstruction. GAN networks are usually difficult to handle in terms of optimization of the training parameters and memory requirements. During the training the generator and discriminator networks compete with each other to generate high quality data and therefore, their loss functions need to stay stable between the epochs. This is an important detail to monitor in order to determine if the GAN training was successful. In Figure 64, the losses of the GAN trained model, which generates the results presented in Figure 63 d, show that the model is properly trained and there is no unbalanced behavior between the generator and the discriminator over the epochs.

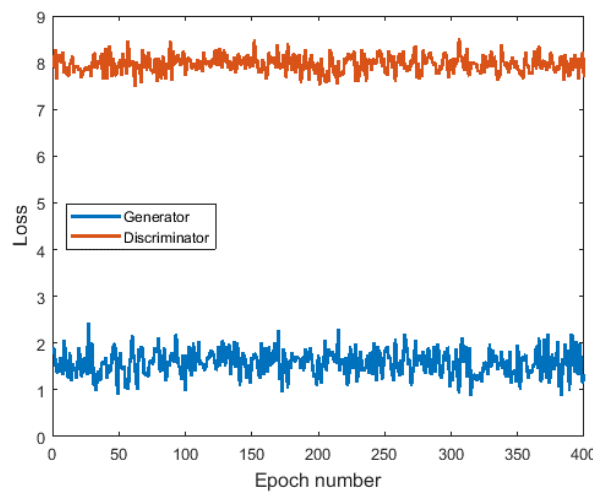


Figure 64. Generator and discriminator network losses over the epochs.

7.4.2.3 Bending studies for the phase imaging endoscope

So far, it has been demonstrated that DNNs can be trained to transform an intensity image collected through a commercial fiber bundle to the respective phase image map of the studied sample, which implies that phase imaging can be integrated in current fiber bundle endoscopes in a digital way. Finally yet importantly, evaluation of the deep learning-based phase imaging process is made with respect to the bending of the fiber bundle. Preliminary results suggesting bending insensitivity of our system were introduced in the previous paragraph 7.3 using the MNIST and MNIST-Fashion datasets. In the remainder of the section, the bending tolerance of the phase endoscope is also verified for the image datasets of the liver tissue sample.

A liver tissue section is prepared and a new dataset is generated for training the DNN. Taking into account the remarkable results observed using a GAN to reconstruct the phase from intensity images recorded through the fiber bundle, the same DNN architecture is kept to explore the bending effect on the reconstruction quality. After the generator is trained with image pairs collected with the fiber bundle fixed at a certain bending configuration, the fiber bundle is set to a different bending shape and 40 new images of a different area of the liver tissue are collected.

Furthermore, aiming to mimic the bending effect in real conditions of using an endoscope, which is more dynamic, namely the fiber configuration constantly changes, 10 more images were acquired while the fiber bundle simultaneously moves in a random way. These new sets of test images are then used as inputs to the trained model of the generator to test the reconstruction performance related to the bending.

In Figure 65, the three rows of images correspond to the phase reconstruction before bending the fiber, after changing the fiber to a new configuration and in case of real time random bending. The values of the MSE and SSIM, reported next to the images for each case, show some degradation of the phase reconstruction when bending of the fiber bundle changes. The reduced performance of the DNN can be attributed to the degradation of the sample itself, since the samples for different bends were acquired almost a day after the end of the training dataset collection. The experiment could be repeated by acquiring the bending test images in the beginning of the measurement and the training dataset which takes longer to complete can be collected afterwards. In this way, GAN would probably provide more valid results.

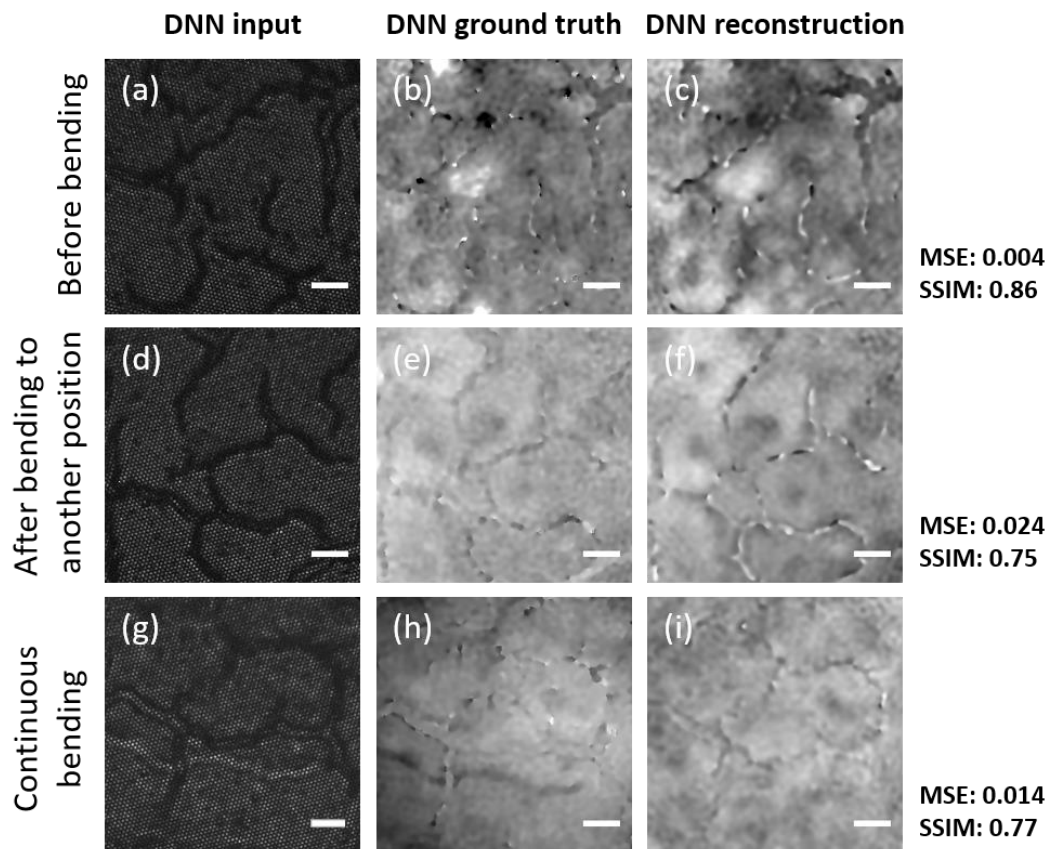


Figure 65. GAN reconstruction results in the presence of bending. a-c) Reconstruction from data obtained in one fiber bundle bending configuration, d-f) data obtained after bending the fiber at another position without retraining the GAN, g-i) data obtained while continuously bending the fiber in random ways without retraining the GAN. Scale bar: 10 μm .

Overall, the phase reconstruction through the system is good and it actually provides good phase images, which in some cases are better than the available ground truth images. This result actually verifies the hypothesis of a faulty ground truth possibly because of sample degradation. Another solution would be also to acquire the training data while the fiber is bent in different configurations. In this way, whatever perturbation of the system in case of moving the fiber will be included in the training dataset information and be accounted for as it was proposed for the MMFs [197].

7.5 Conclusions

In this chapter, it is demonstrated for the first time that DNN can be used not only to improve the imaging through a fiber bundle but also to provide a flexible platform to assign further imaging modalities in an existing endoscope. Specifically, phase imaging through a commercial fiber bundle endoscope is accomplished using the DNNs to translate the bright field intensity images relayed at the proximal camera to the corresponding phase map of the sample. Investigation of different light sources shows that incoherent illumination helps to have bending tolerant imaging through the bundle. In addition, the importance of the DNN architecture is underlined, especially for delicate reconstruction processes of biological samples such as the ones presented in this thesis. GAN-type DNNs result in high resolution images with minimal loss of information, but they are computationally more heavy and slow in terms of training time. At a small expense of reconstruction quality a simpler DNN architecture such as the U-net type, combined with a custom loss can achieve sufficiently good results in shorter training time. Finally, the findings of this chapter show a novel way of digital upgrade of fiber bundle endoscopes, which can be further expanded to fluorescence imaging or other microscopy techniques.

Summary

Fiber endoscopy through multimode and multicore fibers was investigated in the framework of the present thesis in two different approaches. In the case of multimode fibers, the distribution of an input image among the spatial modes of the fiber does not allow the information to simply relay through the fiber length but results in the formation of speckle patterns. Multicore fibers on the other hand, can image directly an object using the individual single mode cores as pixels. However, the crosstalk between the cores and the inherent image pixelation result in a low resolution image. Therefore, the fiber imaging is attempted, by a first approach, which applies linear imaging transformations by calculating the transmission matrix of the fiber system using digital holography and wavefront shaping, while the second method follows a data driven interpretation based on deep learning.

In the first part of the thesis (Part A), the principles of the transmission matrix technique for shaping the light through the two fiber types is explained. Using the transmission matrix method, we show that femtosecond pulses can be spatial and temporally focused through multimode and multicore fibers with high efficiency. The limitations in the focusing efficiency imposed by the nonlinear effects rising when high power femtosecond pulses propagate in fibers were explored for both fiber systems. The detailed characterization performed demonstrates that a high peak intensity focus spot can be generated despite the nonlinearities in the fiber reaching impressive peak intensity of $1.5 \times 10^{13} \text{ W/cm}^2$ (Chapter 4). These results manifest for the first time the potential of multimode fibers to integrate in novel applications such as material manipulation. Moreover, a dual modality endoscope based on multimode fiber, which provides not only imaging but also micro-surgery capabilities, was presented. Selective laser ablation of cochlear tissue is demonstrated through an ultrathin multimode fiber probe guided by the high resolution two-photon fluorescence images obtained through the same tool.

In the second part (Part B), we explore deep learning as an alternative approach for fiber endoscopy. The basic concepts of the deep neural networks are initially discussed to offer the required background for understanding the experimental findings. In the Chapter 6, intensity speckle patterns are used to train a deep neural network algorithm to reconstruct and classify both amplitude and phase images. Impressive results were reported for multimode fibers up to 1 km length, despite the instabilities observed in the system, which are causing the speckle intensity to rapidly change even for a constant input. The recovery of information from the intensity only images of the speckle patterns is also compared to the case in which the full field information is recorded using digital holography. For the datasets tested based on the MNIST database of handwritten digits, no difference in the performance of the input image reconstruction and

classification was observed, which verifies that deep learning is suitable for nonlinear mapping problems in computational imaging such as phase to intensity. Furthermore, the resilience of the deep neural networks in the presence of perturbations induced in the system was also evaluated for the specific case of the wavelength drifting. Two cases of wavelength drifting were studied. In the first case, the training set only consists of data obtained at a certain wavelength while the speckles formed in the drifting wavelength range were not included in the training. In the second case, the training set includes data recorded in all random wavelengths within the drifting bandwidth. While the results in the first attempt show poor information recovery from the speckle patterns at wavelengths far from the training one, the classification of the input images in the second case shows interestingly superior results even for almost 100 nm of wavelength drift. Notably, the correlation of a speckle pattern for a certain input drops to 50% within 30 nm.

Apart from the multimode fibers, fiber bundles or also called multicore fibers show improved imaging results when deep learning is used to remove the core sampling from the output image. We show that using an incoherent source helps to reduce the effects of core-to-core crosstalk on the image formation leading to a more bending insensitive system. Phase imaging through a fiber bundle is also reported in this thesis by training the network to translate the pixelated output image of a fiber bundle recorded in a bright field mode to its corresponding phase profile using digital holography to obtain the ground truth. High fidelity of the reconstructed phase is evaluated by metrics of MSE and SSIM, which reach values up to 0.003 and 0.92 respectively. GAN-type DNN is implemented to push the performance of the phase endoscope providing exceptional reconstruction results even when bending the fiber. The implementation of a phase contrast-based fiber bundle endoscope using deep neural networks is concluding the results of the current thesis and suggests a machine learning way to digitally improve the current clinical tools.

Overall, the results discussed in the present work propose two different ways to recover the information through different types of fibers that induce a kind of distortion when the input propagates from the object to the observer side. Each technique is characterized by advantages and disadvantages and in the end depends on the problem we need to solve which one is the correct pathway to follow. In brief, if working with a certain category of data and access in multiple examples is available, deep learning is a suitable technique to recover the information and it shows robustness against external perturbations such as the fiber bending, thermal, mechanical and wavelength drifts. On the other hand, generalization of the image reconstruction among samples that poses completely different features is not always achieved in a deep learning approach. On the other hand, wavefront shaping methods like the transmission matrix presented in this work need more demanding optical apparatus and they suffer from misalignments, mechanical drifts over time and generally perturbations, because they are calibration based techniques. However, for a stable system configuration, that implies no fiber bending among others, wavefront shaping can reproduce any desired field at the fiber end without loss of performance. If any of the system components change though, recalibration of the system is essential, hindering the real-time inspection in most of the cases. For endoscopic approach light control through a fiber system implies a needle-like rigid design which is feasible with the write optical engineering and packaging of the essential components.

Outlook

During the research work in the context of this thesis we explored possible ways towards the ideal fiber endoscope based on multimode or multicore fibers. Considering the fiber imaging implementations described throughout the previous chapters, it is important to close with some suggestions for the future work on these projects.

Particularly, we identified that nonlinearities harm the focusing efficiency achieved when using the measured transmission matrix of the system at high input pulse energies. The main reason is that the optimized wavefront to generate a high intensity focus is getting scrambled by the additional nonlinear phase factors generated in the system. Consequently, reducing the effect of nonlinearities or compensating for them should be targeted in the future in order to increase the energy throughput in the multimode fiber endoscope. One simple step is to stretch the pulse to account for the group velocity dispersion through the fiber length, which reduces the peak intensity on the focus. Doing so we gain in two ways: firstly, we achieve a shorter pulse at the focus spot at the fiber distal side and secondly the stretched pulse will inflict less strong nonlinear response along the fiber thus increasing the final focusing efficiency. Another possibility to improve the focusing efficiency through the fiber endoscope in the nonlinear regime is to modify the transmission matrix of the system. The transmission matrix is a linear transformation and even if it is measured in the presence of the nonlinearities cannot compensate the losses because once the focusing wavefront is projected on the proximal fiber facet the phase of each mode will be changed in a nonlinear way different from the measured on in case of the calibration database. Nevertheless, if we can compute this perturbation component of the matrix we can invert this degradation effect. For example measurements of input and output wavefronts with and without nonlinearities could help us calculate the perturbation component of the linear transmission matrix and it could be interesting to investigate.

Concerning the bending sensitivity, inserting the fiber in a cannula has been proposed as an idea to prevent it from changing its configuration and therefore, sustain the imaging efficiency of the calibrated system. It would be indeed interesting to realize a rigid optical design that includes the basic components of the calibration in place, namely the SLM and the 4f system projecting the SLM wavefront on the proximal facet, so that afterwards the whole device can be moved and inserted in tissue for inspection. In addition, multimode fibers with specially designed refractive index profile between the core and the cladding could also be a solution to reduced bending sensitivity.

Finally, deep learning potential in fiber endoscopy gave answers to many problems in imaging through different types of fibers but at the same time raised many ideas for further development. Starting from the final result, which demonstrates a fiber bundle-based phase endoscope that uses deep neural networks to translate the intensity only recording to a phase image, we can think of adding more modalities in the currently available endoscopes by suitably choosing the training ground truth. For example, autofluorescence has been already demonstrated to translate to H&E stained images that are more familiar to the medical doctors for interpretation. Another thought that we aim to implement is a compact lens system in front of the fiber bundle endoscope to access real-time phase imaging on a trained model.

Furthermore, classification through multimode fibers showed impressive results for the MNIST database used as an input. Application in real tissue samples would be a fascinating follow up to investigate if the deep learning algorithms can perform tissue classification based on the speckle pattern intensities recorded at the user's camera. Finally yet importantly, further understanding on how the physical system, meaning the fiber probe itself, and the number of trainable parameters are interrelated and how they need to be optimized according to the dataset in a more consistent way, is needed.

Appendix

A1. Fiber modes

Endoscopy through optical fibers was described in this thesis. A better understanding of the light propagation through the optical fibers is given in this Appendix section by solving the Maxwell's equations considering the weakly guiding approximation for simplicity. Knowing the propagation constants and the mode profiles of a fiber is important to simulate and predict at a certain level the distal field distribution for a desired input in the fiber.

Firstly, we start by writing the Maxwell's equations in their general form:

$$\nabla \cdot \mathbf{D} = \rho_{free}$$

A1. 1

$$\nabla \cdot \mathbf{B} = 0$$

A1. 2

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

A1. 3

$$\nabla \times \mathbf{H} = J_{free} + \frac{\partial \mathbf{D}}{\partial t}$$

A1. 4

In the equations above, \mathbf{E} is always the electric field of the electromagnetic wave and \mathbf{B} is the magnetic field, ρ_{free} is the density of free charges and J_{free} the surface free current density. We also define below \mathbf{D} , the electric field displacement and \mathbf{H} the magnetizing field. In the equations, we indicate the vectorial values with bold. \mathbf{P} expresses the polarization of the medium in which the electromagnetic wave is travelling and is linearly related to the electric field $\mathbf{P} = \epsilon_0 \chi \mathbf{E}$ (where χ is the electric susceptibility) in the case of a dielectric medium, \mathbf{M} is the magnetization, ϵ_0 is the free space permittivity and ϵ the permittivity of the medium and μ_0 the free space permeability and μ the permeability of the medium.

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} = \epsilon_0 \mathbf{E} + \epsilon_0 \chi \mathbf{E} = \epsilon \mathbf{E}$$

A1. 5

$$\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M})$$

A1. 6

We consider propagation in a dielectric medium, meaning that the medium is non magnetic, homogenous and isotropic, then \mathbf{M} , ρ_{free} , J_{free} are zero and the polarization is linearly related to the electric field the four Maxwell's equations are simplified to the following forms:

$$\nabla \cdot \mathbf{E} = 0$$

A1. 7

$$\nabla \cdot \mathbf{H} = 0$$

A1. 8

$$\nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t}$$

A1. 9

$$\nabla \times \mathbf{H} = \varepsilon \frac{\partial \mathbf{E}}{\partial t}$$

A1. 10

We now apply the curl operator on the equations A1. 9 and A1. 10 and taking into consideration the curl of the curl identity we obtain and the zero gradients of \mathbf{E} and \mathbf{H} we obtain the wave equations for the electric and magnetization fields:

$$\nabla^2 \mathbf{E} - \frac{1}{\varepsilon \mu} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0 \xrightarrow{n = \sqrt{\frac{\varepsilon \mu}{\varepsilon_0 \mu_0}}} \nabla^2 \mathbf{E} - \frac{n^2}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0$$

A1. 11

$$\nabla^2 \mathbf{H} - \frac{n^2}{c^2} \frac{\partial^2 \mathbf{H}}{\partial t^2} = 0$$

A1. 12

The equation above are satisfied for all the components of the electric and magnetic field. We choose to solve the equation for the electric field in the case that the medium is a step index optical fiber with a refractive index n_1 in the core and n_0 in the cladding. The system is characterized by cylindrical symmetry and therefore cylindrical coordinates are used to define the solutions. The equation A1. 11 can be rewritten for each field component as:

$$\frac{\partial^2 E}{\partial r^2} + \frac{1}{r} \frac{\partial E}{\partial r} + \frac{1}{r^2} \frac{\partial^2 E}{\partial \varphi^2} + \frac{\partial^2 E}{\partial z^2} - \frac{n^2}{c^2} \frac{\partial^2 E}{\partial t^2} = 0$$

A1. 13

In addition, considering the weakly guiding approximation ($n_1 \cong n_0$), the longitudinal components of the fields are negligible, meaning $\mathbf{E}_z, \mathbf{H}_z \rightarrow 0$. To solve the wave equation for each of the components of the electric field we use the method of separation of variables by expressing the

electric field components as $E = R(r)\Phi(\varphi)Z(z)T(t)$. Plugging this expression in the equation A1. 113 for the electric field we end up with the following equation:

$$R(r)\Phi(\varphi)Z(z)T(t) \left(\frac{1}{R} \frac{\partial^2 R}{\partial r^2} + \frac{1}{rR} \frac{\partial R}{\partial r} + \frac{1}{r^2\Phi} \frac{\partial^2 \Phi}{\partial \varphi^2} + \frac{1}{Z} \frac{\partial^2 Z}{\partial z^2} - \frac{n^2}{Tc^2} \frac{\partial^2 T}{\partial t^2} \right) = 0$$

A1. 14

Since the electric field cannot be zero the term " $R(r)\Phi(\varphi)Z(z)T(t)$ " of the equation above cannot be zero either and therefore the term in the parenthesis is zero. Since we assume variables to be independent and the only way for the derivatives based on time to be equal to the rest of the terms dependent on r, ϕ, z , is if both are equal to a constant number. Therefore, we have:

$$\frac{1}{R} \frac{\partial^2 R}{\partial r^2} + \frac{1}{rR} \frac{\partial R}{\partial r} + \frac{1}{r^2\Phi} \frac{\partial^2 \Phi}{\partial \varphi^2} + \frac{1}{Z} \frac{\partial^2 Z}{\partial z^2} = k^2$$

A1. 15

$$\frac{n^2}{Tc^2} \frac{\partial^2 T}{\partial t^2} = -k^2 \Rightarrow \frac{\partial^2 T}{\partial t^2} + \frac{k^2 c^2}{n^2} T = 0$$

A1. 16

Continuing in the same way separating the independent variables we end up with the three harmonic oscillator equations for the terms $\Phi(\varphi), Z(z), T(t)$ of the electric field and their solutions can be written as shown below:

$$T(t) = e^{i\omega t}$$

A1. 17

$$Z(z) = e^{-i\beta z}$$

A1. 18

$$\Phi(\varphi) = e^{il\varphi}$$

A1. 19

If we now plug the equations A1. 17, A1. 18 and A1. 19 in the equation A1. 154 and multiplying by R we end up with a more complicated equation for the radial distribution term of the electric field:

$$\frac{\partial^2 R}{\partial r^2} + \frac{1}{r} \frac{\partial R}{\partial r} - \left(\frac{1}{r^2} l^2 - \beta^2 + \frac{n^2 \omega^2}{c^2} \right) R = 0$$

A1. 20

Though not obvious at a first sight, the equation A1. 20 is a Bessel's equations and its solutions are the Bessel functions of first and second kind. To select the correct Bessel function kind we have to consider that for $r \rightarrow \infty$ the radial distribution should drop to zero and for $r \rightarrow 0$ does not

increase to infinity. Taking also into account the distribution of the refractive index for a step-index fiber, we separate the solution in two parts: inside and outside the core area (with a radius a) as follows:

$$R(r) = \begin{cases} AJ_l(k_T r) & r < a \\ BK_l(\gamma r) & r > a \end{cases}$$

A1. 21

The values k_T, γ describe the decay of the field amplitude in the two areas of the sample and they are the transverse parameters related to the propagation constant β as follows:

$$k_T^2 = n_1^2 k_0^2 - \beta^2$$

A1. 22

$$\gamma^2 = \beta^2 - n_0^2 k_0^2$$

A1. 23

Finally, using the continuity of the electric field and its first derivative at the boundary between the core and the cladding ($r = a$) we obtain the characteristic equation for the fiber. For simplicity we set $k_T a = X$ $\gamma a = Y$.

$$X \frac{J_{l\pm 1}(X)}{J_l(X)} = \pm Y \frac{K_{l\pm 1}(Y)}{K_l(Y)}$$

A1. 24

If we consider that:

$$X^2 + Y^2 = (k_T a)^2 + (\gamma a)^2 = k_0^2 (n_1^2 - n_0^2) a^2 = V^2$$

A1. 25

, where V is the fiber parameter related to the physical properties of the system (NA, core size, wavelength), it is obvious that the equation A1. 24 has only one unknown and can be solved graphically. For each value $l = 0, \pm 1, \pm 2, \dots$, we get multiple values for the X and thus the Y in the intersection between the equation A1. 24 and A1. 25 while we always need to consider that $0 \leq X, Y \leq V$. After determining the X or Y we can also calculate the propagation constant for each mode from one of the equations A1. 22, A1. 23 and the modes of the fiber are fully determined.

A2. Classification results

The results presented in the Figure A2.1 and Figure A1.2 are complementary to the classification performance of DNNs as a function of fiber length presented in the paragraph 6.1.3.

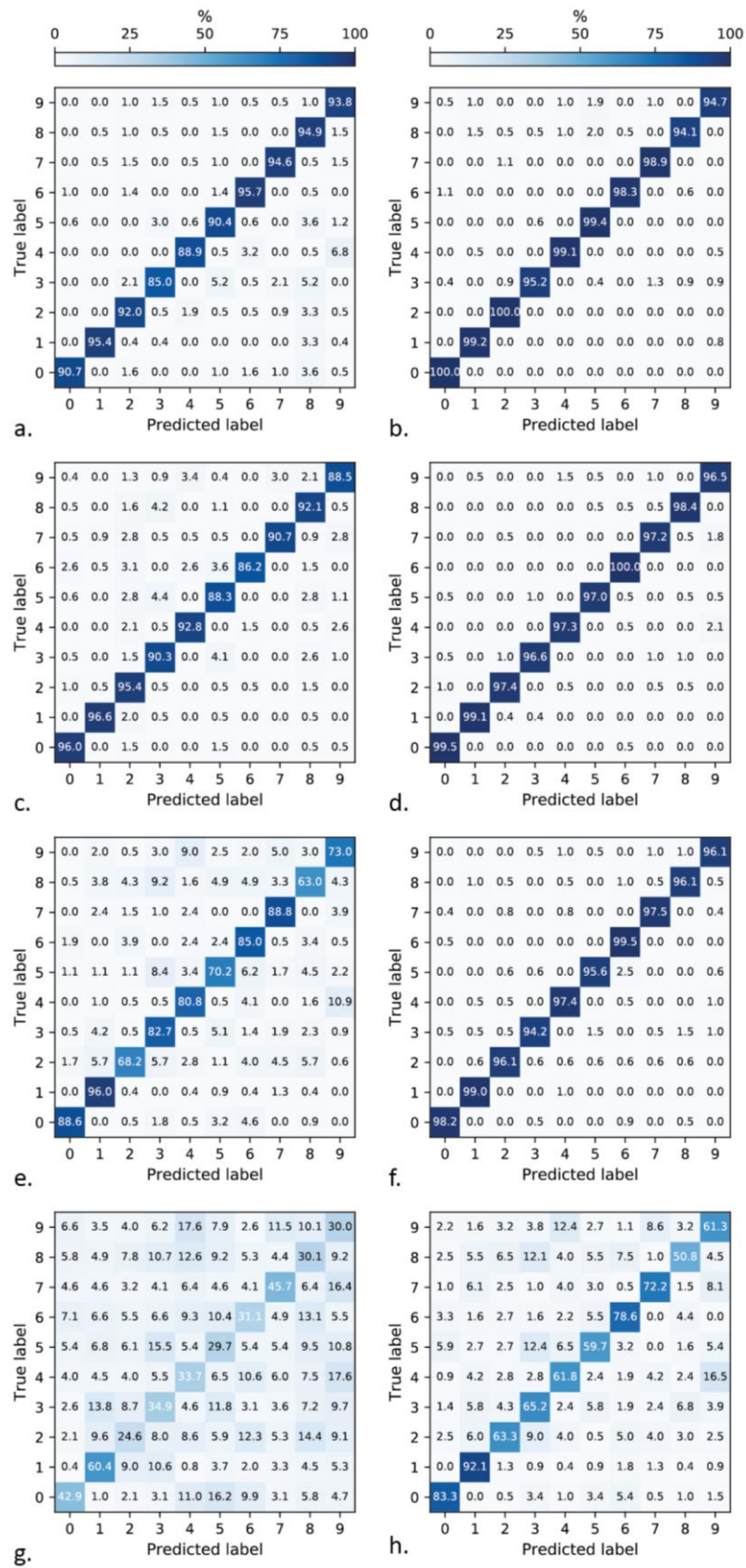


Figure A2.1. Normalized confusion matrices for classification with amplitude modulated proximal inputs for a) 2cm fiber speckle patterns, b) 2cm reconstructed SLM inputs, c) 10cm fiber speckle patterns, d) 10cm reconstructed SLM inputs, e) 10m fiber speckle patterns, f) 10m reconstructed SLM inputs, g) 1km fiber speckle patterns, and h) 1km reconstructed SLM inputs.

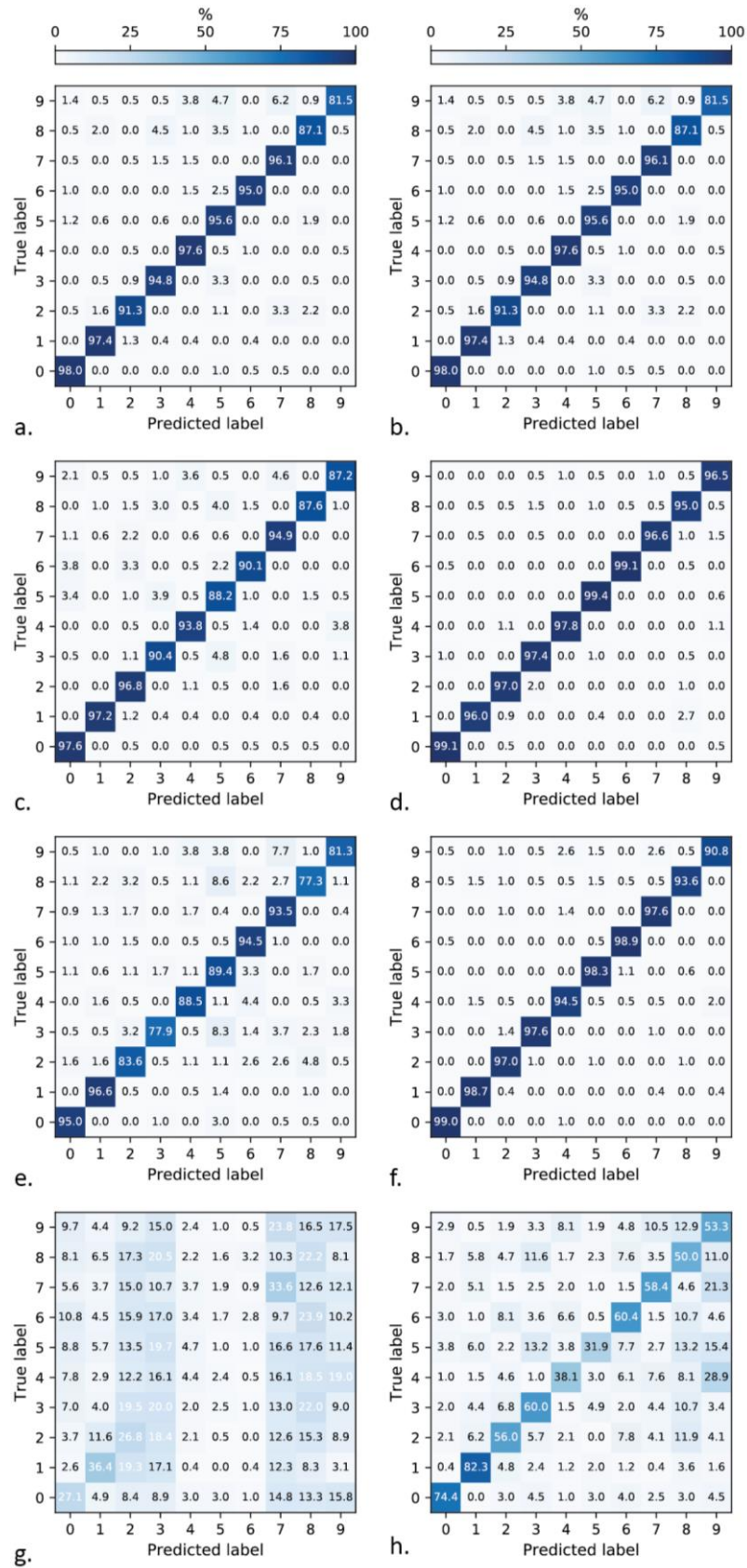


Figure A1.2. Normalized confusion matrices for classification with phase modulated proximal inputs for a) 2cm fiber speckle patterns, b) 2cm reconstructed SLM inputs, c) 10cm fiber speckle patterns, d) 10cm reconstructed SLM inputs, e) 10m fiber speckle patterns, f) 10m reconstructed SLM inputs, g) 1km fiber speckle patterns, and h) 1km reconstructed SLM inputs

A3. Image reconstruction results

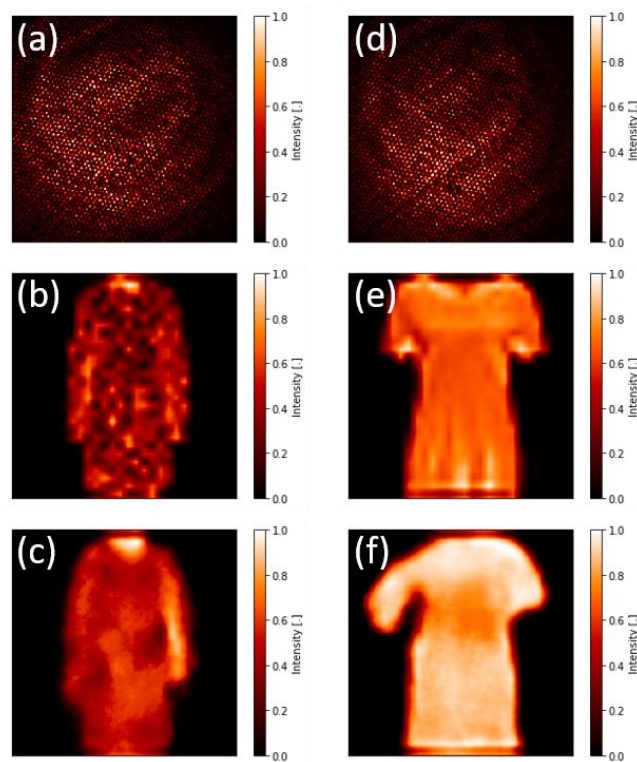


Figure A3.1. Phase imaging using deep learning using laser diode illumination at 632 nm, before bending the fiber bundle (a-c) and after bending (d-f) using the already trained U-net. a,d) Fiber bundle intensity output, b,e) ground truth phase image on the SLM and c,f) DNN reconstruction of a) and c) respectively.

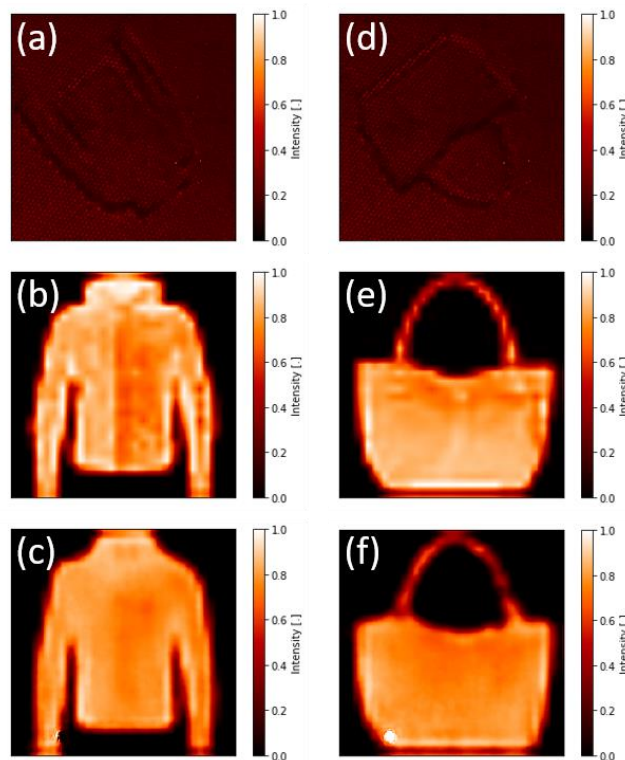


Figure A3.2. Phase imaging using deep learning using LED illumination with 625 nm central wavelength, before bending the fiber bundle (a-c) and after bending (d-f) using the already trained U-net. a,d) Fiber bundle intensity output, b,e) ground truth phase image on the SLM and c,f) DNN reconstruction of a) and c) respectively.

In this paragraph of the Appendix some more examples of the reconstruction results of phase images from the intensity output of the fiber bundle are provided. The image reconstruction is achieved using the U-net DNN as described in the Chapter 7 of the thesis. The figures below correspond to the phase imaging in the case of MNIST-F database using the laser diode illumination at 632nm (S2) and the red LED at 625nm (S3) as described in the relevant paragraph of the thesis.

A4. Python code for the GAN model used in the present thesis

```
"""
```

Created on Tue Mar 26 16:45:32 2019

author: Eirini Kakkava

```
"""
```

```
import numpy as np
```

```
import pandas as pd
```

```
import h5py # HDF5 data file management
```

```
import matplotlib.pyplot as plt
```

```
from tqdm import tqdm
```

```
import keras.backend as K
```

```
from numpy.random import randint
```

```
from tensorflow.python.keras.layers import Dense, Dropout, Input, LeakyReLU, Conv2D, BatchNormalization, Flatten
```

```
from tensorflow.python.keras.layers import concatenate, UpSampling2D, MaxPooling2D, Activation, LeakyReLU
```

```
from tensorflow.python.keras.models import Model, Sequential
```

```
from tensorflow.python.keras.optimizers import Adam
```

```
from sklearn.model_selection import train_test_split
```

```
from tensorflow.python.keras.models import load_model
```



```
# load the data

### Choose the folders and filenames for the experimental data

hf = h5py.File(input_data_filename_full, 'r')

temp = hf.get(input_data_name)

input_data = np.array(temp[0:25000,:,:,:])

hf.close()

hf = h5py.File(output_data_filename_full, 'r')

temp = hf.get(output_data_name)

output_data = np.array(temp[0:25000,:,:,:])

hf.close()

# Split the data

data_number_total = len(input_data)

data_number_train = 0.8*data_number_total #16000

data_number_test = 0.2*data_number_total #2000

# set [train-valid] : test split state for sklearn's train_test_split function :

data_split_state = None

X_train, X_test, Y_train, Y_test = \

    train_test_split(input_data, output_data, \

        test_size=data_set_test_ratio, random_state=data_split_state)

# Define auxiliary functions

def show_2dmatrix_cb(matrix, imin, imax, clabel, colour_map):

    plt.figure()
```

```
plt.imshow(matrix, cmap=colour_map)
```

```
plt.xticks([])
```

```
plt.yticks([])
```

```
plt.colorbar(label=clabel)
```

```
plt.clim(imin, imax)
```

```
plt.show()
```

```
plt.close()
```

```
# Define custom loss
```

```
#this contains both X and Y sobel filters in the format (3,3,1,2)
```

```
#size is 3 x 3, it considers 1 input channel and has two output channels: X and Y
```

```
sobelFilter = K.variable([[[[1., 1.], [0., 2.], [-1., 1.]],  
                           [[2., 0.], [0., 0.], [-2., 0.]],  
                           [[1., -1.], [0., -2.], [-1., -1.]]]])
```

```
def sobelLoss(yTrue,yPred):
```

```
    filt = sobelFilter
```

```
    #calculate the sobel filters for yTrue and yPred
```

```
    #this generates twice the number of input channels
```

```
    #a X and Y channel for each input channel
```

```
    sobelTrue = K.depthwise_conv2d(yTrue,filt)
```

```
    sobelPred = K.depthwise_conv2d(yPred,filt)
```

```
    # K.mean(K.square(sobelTrue - sobelPred)) +
```

```
    #now you just apply the mse:
```

```
    return K.mean(K.square(sobelTrue - sobelPred)) + 0.05*K.mean(K.square(yTrue - yPred))
```

```
verbose_fit = 2
```

```
verbose_early_stop = 0

verbose_evaluate = 0


#Create a Unet generator

xsize = 256

ysize = 256

nn_input_num_channels = 1

input_shape = (xsize, ysize, nn_input_num_channels)


import keras.backend as K


def get_optimizer():

    adam = Adam(lr=1E-4, beta_1=0.9, beta_2=0.999, epsilon=1e-08)

    return adam


def create_generator():

    inputs = Input(shape=input_shape)


# u-net convolution block [down0] input: 256x256x1, output: 128x128x8

# linear feature size is halved but the number of features is doubled


    down = Conv2D(8, (3, 3), padding='same')(inputs)

    down = BatchNormalization()(down)

    down = Activation('relu')(down)

    down = Conv2D(8, (3, 3), padding='same')(down)

    down = BatchNormalization()(down)
```

```
down = Activation('relu')(down)
```

```
down_pool = MaxPooling2D((2, 2), strides=(2, 2))(down)
```

```
# u-net convolution block [down0] input: 128x128x8, output: 64x64x16
```

```
# linear feature size is halved but the number of features is doubled
```

```
down0 = Conv2D(16, (3, 3), padding='same')(down_pool)
```

```
down0 = BatchNormalization()(down0)
```

```
down0 = Activation('relu')(down0)
```

```
down0 = Conv2D(16, (3, 3), padding='same')(down0)
```

```
down0 = BatchNormalization()(down0)
```

```
down0 = Activation('relu')(down0)
```

```
down0_pool = MaxPooling2D((2, 2), strides=(2, 2))(down0)
```

```
# u-net convolution block [down1] input: 64x64x16, output: 32x32x32
```

```
down1 = Conv2D(32, (3, 3), padding='same')(down0_pool)
```

```
down1 = BatchNormalization()(down1)
```

```
down1 = Activation('relu')(down1)
```

```
down1 = Conv2D(32, (3, 3), padding='same')(down1)
```

```
down1 = BatchNormalization()(down1)
```

```
down1 = Activation('relu')(down1)
```

```
down1_pool = MaxPooling2D((2, 2), strides=(2, 2))(down1)
```

```
# u-net convolution block [down2] input: 32x32x32, output: 16x16x64
```

```
down2 = Conv2D(64, (3, 3), padding='same')(down1_pool)
```

```
down2 = BatchNormalization()(down2)
```

```
down2 = Activation('relu')(down2)
```

```
down2 = Conv2D(64, (3, 3), padding='same')(down2)
down2 = BatchNormalization()(down2)
down2 = Activation('relu')(down2)
down2_pool = MaxPooling2D((2, 2), strides=(2, 2))(down2)
```

```
# u-net convolution block [down3] input: 16x16x64, output: 8x8x128
```

```
down3 = Conv2D(128, (3, 3), padding='same')(down2_pool)
down3 = BatchNormalization()(down3)
down3 = Activation('relu')(down3)
down3 = Conv2D(128, (3, 3), padding='same')(down3)
down3 = BatchNormalization()(down3)
down3 = Activation('relu')(down3)
down3_pool = MaxPooling2D((2, 2), strides=(2, 2))(down3)
```

```
# u-net convolution block [down4] input: 8x8x128, output: 4x4x256
```

```
down4 = Conv2D(256, (3, 3), padding='same')(down3_pool)
down4 = BatchNormalization()(down4)
down4 = Activation('relu')(down4)
down4 = Conv2D(256, (3, 3), padding='same')(down4)
down4 = BatchNormalization()(down4)
down4 = Activation('relu')(down4)
down4_pool = MaxPooling2D((2, 2), strides=(2, 2))(down4)
```

```
# u-net layer [center]
```

```
# input: 4x4x256 [down2_pool], convoluted to 4x4x512 [center]
```

```
# linear feature size remains the same but the number of features doubles
```

```
center = Conv2D(512, (3, 3), padding='same')(down4_pool)
```

```
center = BatchNormalization()(center)
```

```
center = Activation('relu')(center)
```

```
center = Conv2D(512, (3, 3), padding='same')(center)
```

```
center = BatchNormalization()(center)
```

```
center = Activation('relu')(center)
```

```
# u-net convolution block [up4] :
```

```
# inputs 4x4x512 from [center] which is upsampled to 8x8x256,
```

```
# this is then concatenated with the 8x8x128 [down4] feature layers
```

```
# via a skip connection to give 8x8x384 for convolution, which is outputted as 8x8x128 [up4]
```

```
up4 = UpSampling2D((2, 2))(center) # output 8x8x256
```

```
up4 = concatenate([down4, up4], axis=3) # skip connection, output 8x8x384
```

```
up4 = Conv2D(256, (3, 3), padding='same')(up4) # output 8x8x128
```

```
up4 = BatchNormalization()(up4)
```

```
up4 = Activation('relu')(up4)
```

```
up4 = Conv2D(256, (3, 3), padding='same')(up4)
```

```
up4 = BatchNormalization()(up4)
```

```
up4 = Activation('relu')(up4)
```

```
# u-net convolution block [up3], output: 16x16x172
```

```
up3 = UpSampling2D((2, 2))(up4)
```

```
up3 = concatenate([down3, up3], axis=3) # output 16x16x172
```

```
up3 = Conv2D(128, (3, 3), padding='same')(up3)
```

```
up3 = BatchNormalization()(up3)
```

```
up3 = Activation('relu')(up3)
```

```
up3 = Conv2D(128, (3, 3), padding='same')(up3)
```

```
up3 = BatchNormalization()(up3)
```

```
up3 = Activation('relu')(up3)
```

```
# u-net convolution block [up2], output: 32x32x96
```

```
up2 = UpSampling2D((2, 2))(up3)
```

```
up2 = concatenate([down2, up2], axis=3) # output 16x16x96
```

```
up2 = Conv2D(64, (3, 3), padding='same')(up2)
```

```
up2 = BatchNormalization()(up2)
```

```
up2 = Activation('relu')(up2)
```

```
up2 = Conv2D(64, (3, 3), padding='same')(up2)
```

```
up2 = BatchNormalization()(up2)
```

```
up2 = Activation('relu')(up2)
```

```
# u-net convolution block [up1], output: 64x64x48
```

```
up1 = UpSampling2D((2, 2))(up2)
```

```
up1 = concatenate([down1, up1], axis=3)
```

```
up1 = Conv2D(32, (3, 3), padding='same')(up1)
```

```
up1 = BatchNormalization()(up1)
```

```
up1 = Activation('relu')(up1)
```

```
up1 = Conv2D(32, (3, 3), padding='same')(up1)
```

```
up1 = BatchNormalization()(up1)
```

```
up1 = Activation('relu')(up1)
```

```
# u-net convolution block [up0], output: 128x128x24
```

```
up0 = UpSampling2D((2, 2))(up1)
```

```
up0 = concatenate([down0, up0], axis=3)
```

```
up0 = Conv2D(16, (3, 3), padding='same')(up0)
```

```
up0 = BatchNormalization()(up0)
```

```
up0 = Activation('relu')(up0)
```

```
up0 = Conv2D(16, (3, 3), padding='same')(up0)
```

```
up0 = BatchNormalization()(up0)
```

```
up0 = Activation('relu')(up0)
```

```
# u-net convolution block [up], output: 256x256x8
```

```
up = UpSampling2D((2, 2))(up0)
```

```
up = concatenate([down, up], axis=3)
```

```
up = Conv2D(8, (3, 3), padding='same')(up)
```

```
up = BatchNormalization()(up)
```

```
up = Activation('relu')(up)
```

```
up = Conv2D(8, (3, 3), padding='same')(up)
```

```
up = BatchNormalization()(up)
```

```
up = Activation('relu')(up)
```

```
outputs = Conv2D(nn_input_num_channels, (1, 1), activation='sigmoid')(up)
```

```
# Compile neural network
```

```
model = Model(inputs=inputs, outputs=outputs)
```

```
return model
```

```
def create_discriminator():
```

```
    inputs = Input(shape=input_shape)
```

```
#convolution block [down] input: 256x256x1, output: 128x128x8
```



```
down = Conv2D(8, (3, 3), strides=(1, 1), padding='same')(inputs)
```

```
down = LeakyReLU(alpha=0.2)(down)
```

```
# convolution block [down0] input: 128x128x8, output:64x64x16
```

```
down0 = Conv2D(16, (3, 3), strides=(2, 2), padding='same')(down)
```

```
down0 = BatchNormalization()(down0)
```

```
down0 = LeakyReLU(alpha=0.3)(down0)
```

```
# convolution block [down1] input: 64x64x16, output: 32x32x32
```

```
down1 = Conv2D(32, (3, 3), padding='same')(down0)
```

```
down1 = BatchNormalization()(down1)
```

```
down1 = LeakyReLU(alpha=0.3)(down1)
```

```
# convolution block [down2] input: 32x32x32, output: 16x16x64
```

```
down2 = Conv2D(64, (3, 3), strides=(2, 2), padding='same')(down1)
```

```
down2 = BatchNormalization()(down2)
```

```
down2 = LeakyReLU(alpha=0.3)(down2)
```

```
# convolution block [down3] input: 16x16x64, output: 8x8x128
```

```
down3 = Conv2D(128, (3, 3), padding='same')(down2)
```

```
down3 = BatchNormalization()(down3)
```

```
down3 = LeakyReLU(alpha=0.3)(down3)
```

```
# convolution block [down4] input: 8x8x128, output: 4x4x256
```

```
down4 = Conv2D(256, (3, 3), strides=(2, 2), padding='same')(down3)
```

```
down4 = BatchNormalization()(down4)
```

```
down4 = LeakyReLU(alpha=0.3)(down4)

# convolution block [down4] input: 4x4x256, output: 4096x1
down5 = Conv2D(256, (3, 3), strides=(2, 2), padding='same')(down4)
down5 = BatchNormalization()(down5)
down5 = LeakyReLU(alpha=0.3)(down5)
down5 = Flatten()(down5)

# fully connected layers
dense_01 = Dense(1024)(down5)
dense_02 = LeakyReLU(alpha=0.3)(dense_01)
dense_03 = Dense(1)(dense_02)
dense_03 = Activation('sigmoid')(dense_03)
outputs = dense_03
model2 = Model(inputs=inputs, outputs=outputs)

return model2

def get_gan_network(discriminator, shape, generator, optimizer, sobelLoss):
    discriminator.trainable = False
    gan_input = Input(shape=shape)
    x = generator(gan_input)
    gan_output = discriminator(x)
    gan = Model(inputs=gan_input, outputs=[x, gan_output])
    gan.compile(loss=[sobelLoss, "binary_crossentropy"], loss_weights=[1., 1e-3],
optimizer=optimizer)
    return gan
```

```

# Remember to change image shape if you are having different size of images

image_shape = input_shape

def train(epochs, batch_size, X_train, X_test, Y_train, Y_test, model_save_dir, number_of_images,
train_test_ratio):

    # Loads training and test data

    # # Create custom function, where hr_images and lr_images function can be used to get high
    resolution and low resolution images

    # x_train_lr, x_train_hr, x_test_lr, x_test_hr = load_training_data(input_dir, '.jpg',
    number_of_images, train_test_ratio)

    x_train_lr = X_train

    x_test_lr = X_test

    x_train_hr = Y_train

    x_test_hr = Y_test

    batch_count = int(x_train_hr.shape[0] / batch_size)

    generator = create_generator()

    discriminator = create_discriminator()

    optimizer = get_optimizer()

    generator.compile(loss=sobelLoss, optimizer=optimizer)

    discriminator.compile(loss="binary_crossentropy", optimizer=optimizer)

    gan = get_gan_network(discriminator, image_shape, generator, optimizer, sobelLoss)

    loss_file = open(model_save_dir + 'losses.txt' , 'w+')

    loss_file.close()

    for e in range(1, epochs+1):

        print ('-'*15, 'Epoch %d' % e, '-'*15)

        for _ in tqdm(range(batch_count)):

            rand_nums = np.random.randint(0, x_train_hr.shape[0], size=batch_size)

```

```
image_batch_hr = x_train_hr[rand_nums]

image_batch_lr = x_train_lr[rand_nums]

generated_images_sr = generator.predict(image_batch_lr)

real_data_Y = np.ones(batch_size) - np.random.random_sample(batch_size)*0.2

fake_data_Y = np.random.random_sample(batch_size)*0.2

discriminator.trainable = True

d_loss_real = discriminator.train_on_batch(image_batch_hr, real_data_Y)

d_loss_fake = discriminator.train_on_batch(generated_images_sr, fake_data_Y)

discriminator_loss = 0.5 * np.add(d_loss_fake, d_loss_real)

rand_nums = np.random.randint(0, x_train_hr.shape[0], size=batch_size)

image_batch_hr = x_train_hr[rand_nums]

image_batch_lr = x_train_lr[rand_nums]

gan_Y = np.ones(batch_size) - np.random.random_sample(batch_size)*0.2

discriminator.trainable = False

gan_loss = gan.train_on_batch(image_batch_lr, [image_batch_hr, gan_Y])

print("discriminator_loss : %f" % discriminator_loss)

print("gan_loss :", gan_loss)

gan_loss = str(gan_loss)


loss_file = open(model_save_dir + 'losses.txt' , 'a')

loss_file.write('epoch%d : gan_loss = %s ; discriminator_loss = %f\n' %(e, gan_loss,
discriminator_loss) )

loss_file.close()

if e % 400 == 0:

    generator.save(model_save_dir + 'gen_model%d.h5' % e)

    discriminator.save(model_save_dir + 'dis_model%d.h5' % e)

generator.save(model_save_dir + 'gen_model%d.h5' % e)
```

```
discriminator.save(model_save_dir + 'dis_model%d.h5' % e)  
train(400,10, X_train, X_test, Y_train, Y_test, save_path, 25000, 0.2)
```


References

1. M. Azadeh, "Fiber Optic Communications: A Review," in *Fiber Optics Engineering*, M. Azadeh, ed., Optical Networks (Springer US, 2009), pp. 1–27.
2. D. J. Richardson, J. M. Fini, and L. E. Nelson, "Space-division multiplexing in optical fibres," *Nature Photonics* **7**, 354–362 (2013).
3. S. (Shizhuo) Yin and P. Ruffin, "Fiber Optic Sensors," in *Wiley Encyclopedia of Biomedical Engineering* (American Cancer Society, 2006).
4. L. Thévenaz and M. A. Soto, "Novel Concepts and Recent Progress in Distributed Optical Fiber Sensing," in *Optical Fiber Communication Conference (OSA, 2016)*, p. M2D.5.
5. M. N. Zervas and C. A. Codemard, "High Power Fiber Lasers: A Review," *IEEE Journal of Selected Topics in Quantum Electronics* **20**, 219–241 (2014).
6. L. G. Wright, Z. Liu, D. A. Nolan, M.-J. Li, D. N. Christodoulides, and F. W. Wise, "Self-organized instability in graded-index multimode fibres," *Nature Photonics* **10**, 771–776 (2016).
7. F. Wise, *Ultrafast Fiber Lasers*, A Wiley-Science Wise Co-Publication (Wiley, 2016).
8. U. Teğın, E. Kakkava, B. Rahmani, D. Psaltis, and C. Moser, "Spatiotemporal self-similar fiber laser," *Optica* **6**, 1412 (2019).
9. A. Katzir, *Lasers and Optical Fibers in Medicine* (Elsevier, 2012).
10. G. Berci and K. A. Forde, "History of endoscopy: What lessons have we learned from the past?," *Surg Endosc* **14**, 5–15 (2000).
11. S. F. Elahi and T. D. Wang, "Future and advances in endoscopy," *J. Biophoton.* **4**, 471–481 (2011).
12. C. M. Wilcox, "Fiberoptic Endoscopy: The Singular Transformative Event of Our Time," *Dig Dis Sci* **59**, 2619–2622 (2014).
13. E. R. Andresen, G. Bouwmans, S. Monneret, and H. Rigneault, "Two-photon lensless endoscope," *Opt. Express* **21**, 20713–20721 (2013).
14. I. N. Papadopoulos, S. Farahi, C. Moser, and D. Psaltis, "Focusing and scanning light through a multimode optical fiber using digital phase conjugation," *Opt. Express* **20**, 10583–10590 (2012).
15. E. Kakkava, M. Romito, D. B. Conkey, D. Loterie, K. M. Stankovic, C. Moser, and D. Psaltis, "Selective femtosecond laser ablation via two-photon fluorescence imaging through a multimode fiber," *Biomedical Optics Express* **10**, 423 (2019).
16. Y. Choi, C. Yoon, M. Kim, T. D. Yang, C. Fang-Yen, R. R. Dasari, K. J. Lee, and W. Choi, "Scanner-Free and Wide-Field Endoscopic Imaging by Using a Single Multimode Optical Fiber," *Phys. Rev. Lett.* **109**, 203901 (2012).
17. C. M. Lee, C. J. Engelbrecht, T. D. Soper, F. Helmchen, and E. J. Seibel, "Scanning fiber endoscopy with highly flexible, 1 mm catheterscopes for wide-field, full-color imaging," *Journal of Biophotonics* **3**, 385–407 (2010).
18. J. Baillie, "The endoscope," *Gastrointestinal Endoscopy* **65**, 886–893 (2007).
19. R. Schindler, "Gastrosocopy with a flexible gastroscope," *American Journal of Digestive Diseases and Nutrition* **2**, 656–663 (1935).
20. G. Berci and K. A. Forde, "History of endoscopy: What lessons have we learned from the past?," *Surg Endosc* **14**, 5–15 (2000).
21. J. D. Kaunitz, "The Fruits of Fiber: The Invention of the Flexible Fiberoptic Gastroscope," *Dig Dis Sci* **59**, 2616–2618 (2014).

22. B. I. Hirschowitz, "Development and application of endoscopy," *Gastroenterology* **104**, 337–342 (1993).
23. M. Kudou, "(73) Assignee: Fujikura Ltd., Tokyo (JP)," 10 (n.d.).
24. M. Hughes, T. P. Chang, and G.-Z. Yang, "Fiber bundle endocytoscopy," *Biomed. Opt. Express*, **BOE** **4**, 2781–2794 (2013).
25. K. L. Reichenbach and C. Xu, "Numerical analysis of light propagation in image fibers or coherent fiber bundles," *Optics Express* **15**, 2151 (2007).
26. A. Orth, M. Ploschner, E. R. Wilson, I. S. Maksymov, and B. C. Gibson, "Optical fiber bundles: Ultra-slim light field imaging probes," *Sci. Adv.* **5**, eaav1555 (2019).
27. D. Loterie, S. Farahi, I. Papadopoulos, A. Goy, D. Psaltis, and C. Moser, "Digital confocal microscopy through a multimode fiber," *Opt. Express*, **OE** **23**, 23845–23858 (2015).
28. M. Plöschner, T. Tyc, and T. Čížmár, "Seeing through chaos in multimode fibres," *Nature Photonics* **9**, 529 (2015).
29. E. E. Morales-Delgado, D. Psaltis, and C. Moser, "Two-photon imaging through a multimode fiber," *Optics Express* **23**, 32158 (2015).
30. S. Bianchi and R. Di Leonardo, "A multi-mode fiber probe for holographic micromanipulation and microscopy," *Lab Chip* **12**, 635–639 (2012).
31. R. Y. Gu, R. N. Mahalati, and J. M. Kahn, "Design of flexible multi-mode fiber endoscope," *Optics express* **23**, 26905–26918 (2015).
32. F. Benabid, "Hollow-core Photonic Crystal Fibers: Guidances and Applications," in *Advanced Photonics*, OSA Technical Digest (Online) (Optical Society of America, 2014), p. SoM4B.5.
33. K. Subramanian, I. Gabay, O. Ferhanoglu, A. Shadfan, M. Pawlowski, Y. Wang, T. Tkaczyk, and A. Ben-Yakar, "Kagome fiber based ultrafast laser microsurgery probe delivering micro-Joule pulse energies," *Biomedical Optics Express* **7**, 4639 (2016).
34. B. Beaudou, F. Gerôme, Y. Y. Wang, M. Alharbi, T. D. Bradley, G. Humbert, J.-L. Auguste, J.-M. Blondy, and F. Benabid, "Millijoule laser pulse delivery for spark ignition through kagome hollow-core fiber," *Opt. Lett.* **37**, 1430–1432 (2012).
35. K. Okamoto, *Fundamentals of Optical Waveguides, Second Edition*, 2 edition (Academic Press, 2005).
36. B. E. Saleh, M. C. Teich, and B. E. Saleh, *Fundamentals of Photonics* (Wiley New York, 1991), Vol. 22.
37. E. R. Andresen, S. Sivankutty, V. Tsvirkun, G. Bouwmans, and H. Rigneault, "Ultrathin endoscopes based on multicore fibers and adaptive optics: a status review and perspectives," *JBO* **21**, 121506 (2016).
38. X. Chen and C. Xu, "Numerical Analysis of the Role of Core-Clad Index Contrast in a Multicore Fiber Bundle," in (OSA, 2007), p. FThG7.
39. Stasio, Nicolino, Conkey, Donald B., C. Moser, and D. Psaltis, "Light control in a multicore fiber using the memory effect," *Optics Express* **23**, (2015).
40. X. Chen, K. L. Reichenbach, and C. Xu, "Experimental and theoretical analysis of core-to-core coupling on fiber bundle imaging," *Optics Express* **16**, 21598 (2008).
41. J. C. Knight, "Photonic crystal fibres," *Nature* **424**, 847–851 (2003).
42. G. J. Pearce, G. S. Wiederhecker, C. G. Poulton, S. Burger, and P. St. J. Russell, "Models for guidance in kagome-structured hollow-core photonic crystal fibres," *Opt. Express* **15**, 12680–12685 (2007).
43. P. Russell, "Photonic Crystal Fibers," *Science* **299**, 358–362 (2003).
44. H. Choi and P. T. C. So, "Improving femtosecond laser pulse delivery through a hollow core photonic crystal fiber for temporally focused two-photon endomicroscopy," *Scientific Reports* **4**, (2015).

45. Y. Y. Wang, X. Peng, M. Alharbi, C. F. Dutin, T. D. Bradley, F. Gérôme, M. Mielke, T. Booth, and F. Benabid, "Design and fabrication of hollow-core photonic crystal fibers for high-power ultrashort pulse transportation and pulse compression," *Opt. Lett.* **37**, 3111–3113 (2012).
46. J. W. Goodman, *Introduction to Fourier Optics* (Roberts and Company Publishers, 2005).
47. S. Popoff, G. Lerosey, M. Fink, A. C. Boccarda, and S. Gigan, "Controlling Light Through Optical Disordered Media : Transmission Matrix Approach," 1107.5285 (2011).
48. J. Bertolotti, E. G. van Putten, C. Blum, A. Lagendijk, W. L. Vos, and A. P. Mosk, "Non-invasive imaging through opaque scattering layers," *Nature* **491**, 232–234 (2012).
49. V. M. LagendijkA., and M. P., "Exploiting disorder for perfect focusing," *Nat Photon* **4**, 320–322 (2010).
50. Z. A. Steelman, S. Kim, E. T. Jelly, M. Crose, K. K. Chu, and A. Wax, "Comparison of imaging fiber bundles for coherence-domain imaging," *Appl. Opt.* **57**, 1455 (2018).
51. D. B. Conkey, E. Kakkava, T. Lanvin, D. Loterie, N. Stasio, E. Morales-Delgado, C. Moser, and D. Psaltis, "High power, ultrashort pulse control through a multi-core fiber for ablation," *Optics Express* **25**, 11491 (2017).
52. D. B. Conkey, N. Stasio, E. E. Morales-Delgado, M. Romito, C. Moser, and D. Psaltis, "Lensless two-photon imaging through a multicore fiber with coherence-gated digital phase conjugation," *J. Biomed. Opt.* **21**, 045002–045002 (2016).
53. D. Psaltis, E. Kakkava, N. Stasio, D. B. Conkey, and C. Moser, "Femtosecond pulse delivery through multi-core fibers for imaging and ablation," in Z. Liu, ed. (SPIE, 2017), p. 30.
54. A. Orth, M. Ploschner, I. S. Maksymov, and B. C. Gibson, "Extended depth of field imaging through multicore optical fibers," *Optics Express* **26**, 6407 (2018).
55. N. Stasio, D. B. Conkey, C. Moser, and D. Psaltis, "Light control in a multicore fiber using the memory effect," *Optics Express* **23**, (2015).
56. E. R. Andresen, G. Bouwmans, S. Monneret, and H. Rigneault, "Toward endoscopes with no distal optics: video-rate scanning microscopy through a fiber bundle," *Opt. Lett.* **38**, 609–611 (2013).
57. D. Kim, J. Moon, M. Kim, T. D. Yang, J. Kim, E. Chung, and W. Choi, "Toward a miniature endomicroscope: pixelation-free and diffraction-limited imaging through a fiber bundle," *Opt. Lett.*, **OL 39**, 1921–1924 (2014).
58. R. Kuschmierz, E. Scharf, N. Koukourakis, and J. W. Czarke, "Self-calibration of lensless holographic endoscope using programmable guide stars," *Optics Letters* **43**, 2997 (2018).
59. E. Kakkava, B. Rahmani, N. Borhani, U. Teğin, D. Loterie, G. Konstantinou, C. Moser, and D. Psaltis, "Imaging through multimode fibers using deep learning: The effects of intensity versus holographic recording of the speckle pattern," *Optical Fiber Technology* **52**, 101985 (2019).
60. N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, "Learning to see through multimode fibers," *Optica*, **OPTICA 5**, 960–966 (2018).
61. B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, and C. Moser, "Multimode optical fiber transmission with a deep learning network," *Light: Science & Applications* **7**, 69 (2018).
62. P. Fan, T. Zhao, and L. Su, "Deep learning the high variability and randomness inside multimode fibres," 19 (n.d.).
63. U. Kürüm, P. R. Wiecha, R. French, and O. L. Muskens, "Deep learning enabled real time speckle recognition and hyperspectral imaging using a multimode fiber array," *Opt. Express* **27**, 20965 (2019).
64. J. Shao, J. Zhang, R. Liang, and K. Barnard, "Fiber bundle imaging resolution enhancement using deep learning," *Opt. Express*, **OE 27**, 15880–15890 (2019).
65. P. Wang and J. Di, "Deep learning-based object classification through multimode fiber via a CNN-architecture SpeckleNet," *Appl. Opt.* **57**, 8258 (2018).

66. E. Kakkava, N. Borhani, C. Moser, and D. Psaltis, "Deep neural networks for seeing through multimode fibers," in *High-Speed Biomedical Imaging and Spectroscopy IV*, K. Goda and K. K. Tsia, eds. (SPIE, 2019), p. 46.
67. A. F. Gmitro and D. Aziz, "Confocal microscopy through a fiber-optic imaging bundle," *Optics letters* **18**, 565–567 (1993).
68. A. Shinde, S. M. Perinchery, and V. M. Murukeshan, "A targeted illumination optical fiber probe for high resolution fluorescence imaging and optical switching," *Scientific Reports* **7**, (2017).
69. W. Gobel, J. N. D. Kerr, A. Nimmerjahn, and F. Helmchen, "Miniaturized two-photon microscope based on a flexible coherent fiber bundle and a gradient-index lens objective," *Optics Letters* **29**, 2521 (2004).
70. E. R. Andresen, S. Sivankutty, G. Bouwmans, L. Gallais, S. Monneret, and H. Rigneault, "Measurement and compensation of residual group delay in a multi-core fiber for lensless endoscopy," *Journal of the Optical Society of America B* **32**, 1221 (2015).
71. J. Shao, W.-C. Liao, R. Liang, and K. Barnard, "Resolution enhancement for fiber bundle imaging using maximum a posteriori estimation," *Opt. Lett.*, OL **43**, 1906–1909 (2018).
72. J. Shin, B. T. Bosworth, and M. A. Foster, "Compressive fluorescence imaging using a multi-core fiber and spatially dependent scattering," *Optics Letters* **42**, 109 (2017).
73. J. Shao, J. Zhang, X. Huang, R. Liang, and K. Barnard, "Fiber bundle image restoration using deep learning," *Opt. Lett.*, OL **44**, 1080–1083 (2019).
74. I. N. Papadopoulos, S. Farahi, C. Moser, and D. Psaltis, "High-resolution, lensless endoscope based on digital scanning through a multimodeoptical fiber," *Biomed. Opt. Express* **4**, 260–270 (2013).
75. D. Loterie, S. Farahi, I. Papadopoulos, A. Goy, D. Psaltis, and C. Moser, "Digital confocal microscopy through a multimode fiber," *Opt. Express*, OE **23**, 23845–23858 (2015).
76. S. Ohayon, A. Caravaca-Aguirre, R. Piestun, and J. J. DiCarlo, "Minimally invasive multimode optical fiber microendoscope for deep brain fluorescence imaging," *Biomed. Opt. Express* **9**, 1492 (2018).
77. O. Tzang, A. M. Caravaca-Aguirre, K. Wagner, and R. Piestun, "Adaptive wavefront shaping for controlling nonlinear multimode interactions in optical fibres," *Nature Photonics* **12**, 368 (2018).
78. R. N. Mahalati, D. Askarov, J. P. Wilde, and J. M. Kahn, "Adaptive control of input field to achieve desired output intensity profile in multimode fiber with random mode coupling," *Opt. Express*, OE **20**, 14321–14337 (2012).
79. I. M. Vellekoop and A. P. Mosk, "Focusing coherent light through opaque strongly scattering media," *Opt. Lett.* **32**, 2309–2311 (2007).
80. D. J. McCabe, A. Tajalli, D. R. Austin, P. Bondareff, I. A. Walmsley, S. Gigan, and B. Chatel, "Spatio-temporal focusing of an ultrafast pulse through a multiply scattering medium," *Nat Commun* **2**, 447 (2011).
81. Y. Choi, T. D. Yang, C. Fang-Yen, P. Kang, K. J. Lee, R. R. Dasari, M. S. Feld, and W. Choi, "Overcoming the Diffraction Limit Using Multiple Light Scattering in a Highly Disordered Medium," *Phys. Rev. Lett.* **107**, 023902 (2011).
82. E. Spitz and A. Werts, "Transmission des images à travers une fibre optique," *COMPTEs RENDUS HEBDOMADAIRES DES SEANCES DE L ACADEMIE DES SCIENCES SERIE B* **264**, 1015–+ (1967).
83. E. N. Leith and J. Upatnieks, "Holographic Imagery Through Diffusing Media," *J. Opt. Soc. Am.* **56**, 523 (1966).
84. B. Fischer and S. Sternklar, "Image transmission and interferometry with multimode fibers using self-pumped phase conjugation," in (1985).

85. A. M. Caravaca-Aguirre, E. Niv, D. B. Conkey, and R. Piestun, "Real-time resilient focusing through a bending multimode fiber," *Optics Express* **21**, 12881 (2013).
86. B. A. Flusberg, E. D. Cocker, W. Piyawattanametha, J. C. Jung, E. L. M. Cheung, and M. J. Schnitzer, "Fiber-optic fluorescence imaging," *Nat Meth* **2**, 941–950 (2005).
87. C. J. Engelbrecht, R. S. Johnston, E. J. Seibel, and F. Helmchen, "Ultra-compact fiber-optic two-photon microscope for functional fluorescence imaging in vivo," *Opt. Express*, OE **16**, 5556–5564 (2008).
88. J.-M. Yang, R. Chen, C. Favazza, J. Yao, C. Li, Z. Hu, Q. Zhou, K. K. Shung, and L. V. Wang, "A 2.5-mm diameter probe for photoacoustic and ultrasonic endoscopy," *Opt. Express* **20**, 23944–23953 (2012).
89. N. Stasio, A. Shibukawa, I. N. Papadopoulos, S. Farahi, O. Simandoux, J.-P. Huignard, E. Bossy, C. Moser, and D. Psaltis, "Towards new applications using capillary waveguides," *Biomedical Optics Express* **6**, 4619 (2015).
90. T. Čižmár and K. Dholakia, "Exploiting multimode waveguides for pure fibre-based imaging," *Nat Commun* **3**, 1027 (2012).
91. A. M. Caravaca-Aguirre, S. Singh, S. Labouesse, M. V. Baratta, R. Piestun, and E. Bossy, "Hybrid photoacoustic-fluorescence microendoscopy through a multimode fiber using speckle illumination," *APL Photonics* **4**, 096103 (2019).
92. N. Stasio, C. Moser, and D. Psaltis, "Calibration-free imaging through a multicore fiber using speckle scanning microscopy," *Opt. Lett.*, OL **41**, 3078–3081 (2016).
93. F. Légaré, C. L. Evans, F. Ganikhanov, and X. S. Xie, "Towards CARS Endoscopy," *Optics Express* **14**, 4427 (2006).
94. S. Aisawa, K. Noguchi, and T. Matsumoto, "Remote image classification through multimode optical fiber using a neural network," *Optics Letters* **16**, 645 (1991).
95. R. K. Marusz and M. R. Sayeh, "Neural network-based multimode fiber-optic information transmission," *Applied Optics* **40**, 219 (2001).
96. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
97. J. Ker, L. Wang, J. Rao, and T. Lim, "Deep Learning Applications in Medical Image Analysis," *IEEE Access* **6**, 9375–9389 (2018).
98. K. Simonyan and A. Zisserman, "VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION," **14** (2015).
99. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica*, OPTICA **6**, 921–943 (2019).
100. N. Shabairou, E. Cohen, O. Wagner, D. Malka, and Z. Zalevsky, "Color image identification and reconstruction using artificial neural networks on multimode fiber images: towards an all-optical design," *Optics Letters* **43**, 5603 (2018).
101. P. Caramazza, O. Moran, R. Murray-Smith, and D. Faccio, "Transmission of natural scene images through a multimode fibre," *Nat Commun* **10**, 2029 (2019).
102. N. M. Fried and K. E. Murray, "High-Power Thulium Fiber Laser Ablation of Urinary Tissues at 1.94 μm ," *Journal of Endourology* **19**, 25–31 (2005).
103. S. O. Konorov, V. P. Mitrokhin, A. B. Fedotov, D. A. Sidorov-Biryukov, V. I. Beloglazov, N. B. Skibina, A. V. Shcherbakov, E. Wintner, M. Scalora, and A. M. Zheltikov, "Laser ablation of dental tissues with picosecond pulses of 1.06- μm radiation transmitted through a hollow-core photonic-crystal fiber," *Appl. Opt.*, AO **43**, 2251–2256 (2004).
104. A. Vogel and V. Venugopalan, "Mechanisms of Pulsed Laser Ablation of Biological Tissues," *Chem. Rev.* **103**, 577–644 (2003).
105. H. K. Soong and J. B. Malta, "Femtosecond Lasers in Ophthalmology," *American Journal of Ophthalmology* **147**, 189–197.e2 (2009).

106. J. Qiu, J. Neev, T. Wang, and T. E. Milner, "Deep subsurface cavities in skin utilizing mechanical optical clearing and femtosecond laser ablation," *Lasers Surg. Med.* **45**, 383–390 (2013).
107. X. Liu, D. Du, and G. Mourou, "Laser ablation and micromachining with ultrashort laser pulses," *IEEE Journal of Quantum Electronics* **33**, 1706–1716 (1997).
108. U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, A. Goy, C. Vonesch, M. Unser, and D. Psaltis, "Learning approach to optical tomography," *Optica* **2**, 517 (2015).
109. J. Lim, A. B. Ayoub, E. E. Antoine, and D. Psaltis, "High-fidelity optical diffraction tomography of multiple scattering samples," *Light: Science & Applications* **8**, 1–12 (2019).
110. C. Maurer, A. Jesacher, S. Bernet, and M. Ritsch-Marte, "What spatial light modulators can do for optical microscopy," *Laser & Photonics Reviews* **5**, 81–101 (2011).
111. A. Leitis, A. Heßler, S. Wahl, M. Wuttig, T. Taubner, A. Tittl, and H. Altug, "All-Dielectric Programmable Huygens' Metasurfaces," *Advanced Functional Materials* **n/a**, 1910259 (n.d.).
112. J. A. Davis, D. M. Cottrell, J. Campos, M. J. Yzuel, and I. Moreno, "Encoding amplitude information onto phase-only filters," *Appl. Opt.*, AO **38**, 5004–5013 (1999).
113. M. Born, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th expanded edition (Cambridge University Press, 1999).
114. J. Aulbach, B. Gjonaj, P. M. Johnson, A. P. Mosk, and A. Lagendijk, "Control of Light Transmission through Opaque Scattering Media in Space and Time," *Phys. Rev. Lett.* **106**, 103901 (2011).
115. O. Katz, E. Small, Y. Bromberg, and Y. Silberberg, "Focusing and compression of ultrashort pulses through scattering media," *Nat Photon* **5**, 372–377 (2011).
116. S. Sivankutty, E. R. Andresen, R. Cossart, G. Bouwmans, S. Monneret, and H. Rigneault, "Ultra-thin rigid endoscope: two-photon imaging through a graded-index multi-mode fiber," *Optics Express* **24**, 825 (2016).
117. J. Carpenter, B. J. Eggleton, and J. Schröder, "110x110 optical mode transfer matrix inversion," *Opt. Express* **22**, 96 (2014).
118. W. Xiong, P. Ambichl, Y. Bromberg, B. Redding, S. Rotter, and H. Cao, "Principal modes in multimode fibers: exploring the crossover from weak to strong mode coupling," *Opt. Express*, OE **25**, 2709–2724 (2017).
119. D. C.-M. Loterie, "Microscopy and digital light shaping through optical fibers," 139 (n.d.).
120. F. Helmchen and W. Denk, "Deep tissue two-photon microscopy," *Nat Meth* **2**, 932–940 (2005).
121. S. W. Perry, R. M. Burke, and E. B. Brown, "Two-Photon and Second Harmonic Microscopy in Clinical and Translational Cancer Research," *Ann Biomed Eng* **40**, 277–291 (2012).
122. E. Capitaine, N. O. Moussa, C. Louot, S. M. Bardet, H. Kano, L. Duponchel, P. Lévêque, V. Couderc, and P. Leproux, "Fast epi-detected broadband multiplex CARS and SHG imaging of mouse skull cells," *Biomed. Opt. Express*, BOE **9**, 245–253 (2018).
123. P. J. Campagnola and L. M. Loew, "Second-harmonic imaging microscopy for visualizing biomolecular arrays in cells, tissues and organisms," *Nature Biotechnology* **21**, 1356–1360 (2003).
124. E. E. Morales-Delgado, S. Farahi, I. N. Papadopoulos, D. Psaltis, and C. Moser, "Delivery of focused short pulses through a multimode fiber," *Optics Express* **23**, 9109 (2015).
125. E. R. Andresen, S. Sivankutty, G. Bouwmans, O. Vainvinq, L. Gallais, S. Monneret, and H. Rigneault, "Towards two-photon lensless endoscopes: inter-core group delay compensation in a multi-core fiber," in E. Beaufort, P. T. C. So, F. Pavone, and E. M. Hillman, eds. (2015), p. 953605.
126. L. He, K. Sheehy, and W. Culbertson, "Femtosecond laser-assisted cataract surgery:," *Current Opinion in Ophthalmology* **1** (2010).

127. J. Serbin, T. Bauer, C. Fallnich, A. Kasenbacher, and W. H. Arnold, "Femtosecond lasers as novel tool in dental surgery," *Applied Surface Science* **197–198**, 737–740 (2002).
128. G. Agrawal, *Nonlinear Fiber Optics* (Academic Press, 2001).
129. C. L. Hoy, N. J. Durr, P. Chen, W. Piyawattanametha, H. Ra, O. Solgaard, and A. Ben-Yakar, "Miniaturized probe for femtosecond laser microsurgery and two-photon imaging," *Opt. Express* **16**, 9996–10005 (2008).
130. S. P. Mekhail, G. Arbuthnott, and S. N. Chormaic, "Advances in Fibre Microendoscopy for Neuronal Imaging," *Optical Data Processing and Storage* **2**, (2016).
131. E. R. Andresen, G. Bouwmans, S. Monneret, and H. Rigneault, "Two-photon lensless endoscope," *Opt. Express* **21**, 20713–20721 (2013).
132. R. W. Boyd and D. Prato, *Nonlinear Optics* (Elsevier Science, 2008).
133. M. Wollenhaupt, A. Assion, and T. Baumert, "Femtosecond Laser Pulses: Linear Properties, Manipulation, Generation and Measurement," in *Springer Handbook of Lasers and Optics*, F. T. Prof, ed. (Springer New York, 2007), pp. 937–983.
134. B. Redding, S. M. Popoff, and H. Cao, "All-fiber spectrometer based on speckle pattern reconstruction," *Opt. Express* **21**, 6584 (2013).
135. B. C. Stuart, M. D. Feit, S. Herman, A. M. Rubenchik, B. W. Shore, and M. D. Perry, "Optical ablation by high-power short-pulse lasers," *J. Opt. Soc. Am. B, JOSAB* **13**, 459–468 (1996).
136. J. Ilgner, M. Wehner, J. Lorenzen, M. Bovi, and M. Westhofen, "Morphological effects of nanosecond- and femtosecond-pulsed laser ablation on human middle ear ossicles," *J. Biomed. Opt* **11**, 014004–014004–7 (2006).
137. P. P. Pronko, S. K. Dutta, J. Squier, J. V. Rudd, D. Du, and G. Mourou, "Machining of sub-micron holes using a femtosecond laser at 800 nm," *Optics Communications* **114**, 106–110 (1995).
138. M. Tsang, D. Psaltis, and F. G. Omenetto, "Reverse propagation of femtosecond pulses in optical fibers," *Opt. Lett.*, OL **28**, 1873–1875 (2003).
139. Y. Kim, S. C. Warren, J. M. Stone, J. C. Knight, M. A. A. Neil, C. Paterson, C. W. Dunsby, and P. M. W. French, "Adaptive Multiphoton Endomicroscope incorporating a Polarization-Maintaining Multicore Optical Fibre," *IEEE Journal of Selected Topics in Quantum Electronics* **PP**, 1–1 (2015).
140. A. Ben-Yakar and R. L. Byer, "Femtosecond laser ablation properties of borosilicate glass," *Journal of Applied Physics* **96**, 5316–5323 (2004).
141. Y. Li, J. Jing, J. Yu, B. Zhang, T. Huo, Q. Yang, and Z. Chen, "Multimodality endoscopic optical coherence tomography and fluorescence imaging technology for visualization of layered architecture and subsurface microvasculature," *Optics Letters* **43**, 2074 (2018).
142. J. Xi, Y. Chen, Y. Zhang, K. Murari, M.-J. Li, and X. Li, "Integrated multimodal endomicroscopy platform for simultaneous en face optical coherence and two-photon fluorescence imaging," *Opt Lett* **37**, 362–364 (2012).
143. R. Traynor, *Incidence of Hearing Loss Worldwide* \textbar Bob Traynor, *Hearing International* (n.d.).
144. M. Romito, Y. Pu, K. M. Stankovic, and D. Psaltis, "Imaging hair cells through laser-ablated cochlear bone," *Biomed Opt Express* **10**, 5974–5988 (2019).
145. S. Sivankutty, E. R. Andresen, R. Cossart, G. Bouwmans, S. Monneret, and H. Rigneault, "Ultra-thin rigid endoscope: two-photon imaging through a graded-index multi-mode fiber," *Optics Express* **24**, 825 (2016).
146. B. Debord, M. Alharbi, L. Vincetti, A. Husakou, C. Fourcade-Dutin, C. Hoenninger, E. Mottay, F. Gérôme, and F. Benabid, "Multi-meter fiber-delivery and pulse self-compression of milli-Joule femtosecond laser and fiber-aided laser-micromachining," *Optics Express* **22**, 10735 (2014).

147. M. Plöschner, V. Kollárová, Z. Dostál, J. Nylk, T. Barton-Owen, D. E. K. Ferrier, R. Chmelík, K. Dholakia, and T. Čižmár, "Multimode fibre: Light-sheet microscopy at the tip of a needle," *Scientific Reports* **5**, (2016).
148. A. Descloux, L. V. Amitonova, and P. W. H. Pinkse, "Aberrations of the point spread function of a multimode fiber due to partial mode excitation," *Optics Express* **24**, 18501 (2016).
149. S. Rosen, D. Gilboa, O. Katz, and Y. Silberberg, "Focusing and Scanning through Flexible Multimode Fibers without Access to the Distal End," **8** (n.d.).
150. K. Nakamura, Y. Sora, H. Y. Yoshikawa, Y. Hosokawa, R. Murai, H. Adachi, Y. Mori, T. Sasaki, and H. Masuhara, "Femtosecond laser-induced crystallization of protein in gel medium," *Applied Surface Science* **253**, 6425–6429 (2007).
151. P. V. W. Sasikumar, G. Blugan, N. Casati, E. Kakkava, G. Panusa, D. Psaltis, and J. Kuebler, "Polymer derived silicon oxycarbide ceramic monoliths: Microstructure development and associated materials properties," *Ceramics International* **44**, 20961–20967 (2018).
152. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (Cambridge, Massachusetts, London, England, n.d.).
153. J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**, 2758–2769 (1982).
154. J. Lim, K. Lee, K. H. Jin, S. Shin, S. Lee, Y. Park, and J. C. Ye, "Comparative study of iterative reconstruction algorithms for missing cone problems in optical diffraction tomography," *Opt. Express* **23**, 16933 (2015).
155. Y. Rivenson, Z. Göröcs, H. Günaydin, Y. Zhang, H. Wang, and A. Ozcan, "Deep learning microscopy," *Optica*, *OPTICA* **4**, 1437–1443 (2017).
156. H. Wang, Y. Rivenson, Y. Jin, Z. Wei, R. Gao, H. Günaydin, L. A. Bentolila, C. Kural, and A. Ozcan, "Deep learning enables cross-modality super-resolution in fluorescence microscopy," *Nature Methods* **1** (2018).
157. E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Deep-STORM: super-resolution single-molecule microscopy by deep learning," *Optica* **5**, 458 (2018).
158. A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica* **4**, 1117 (2017).
159. Y. Rivenson, Y. Zhang, H. Günaydin, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light: Science & Applications* **7**, 17141 (2018).
160. N. Borhani, A. J. Bower, S. A. Boppart, and D. Psaltis, "Digital staining through the application of deep neural networks to multi-modal multi-photon microscopy," *Biomed. Opt. Express*, *BOE* **10**, 1339–1350 (2019).
161. E. Kakkava, N. Borhani, B. Rahmani, U. Teğın, U. Teğın, C. Moser, and D. Psaltis, "Wavelength Independent Image Classification Through A Multimode Fiber Using Deep Neural Networks," in *2019 Conference on Lasers and Electro-Optics Europe and European Quantum Electronics Conference (2019), Paper Ci_2_1* (Optical Society of America, 2019), p. ci_2_1.
162. E. Kakkava, N. Borhani, B. Rahmani, U. Tegin, C. Moser, and D. Psaltis, "Efficient Image Classification through a Multimode Fiber using Deep Neural Networks in presence of Wavelength Drifting," in *Imaging and Applied Optics 2019 (COSI, IS, MATH, PcAOP)* (OSA, 2019), p. CW1A.4.
163. L. Deng, "The MNIST Database of Handwritten Digit Images for Machine Learning Research [Best of the Web]," *IEEE Signal Processing Magazine* **29**, 141–142 (2012).
164. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, eds. (Springer International Publishing, 2015), Vol. 9351, pp. 234–241.

165. A. Saade, F. Caltagirone, I. Carron, L. Daudet, A. Drémeau, S. Gigan, and F. Krzakala, "Random projections through multiple optical scattering: Approximating Kernels at the speed of light," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2016), pp. 6215–6219.
166. B. Redding, M. Alam, M. Seifert, and H. Cao, "High-resolution and broadband all-fiber spectrometers," *Optica* **1**, 175–180 (2014).
167. E. Connolly, F. Smyth, A. K. Mishra, A. Kaszubowska-Anandarajah, and L. P. Barry, "Cross-Channel Interference Due to Wavelength Drift of Tunable Lasers in DWDM Networks," *IEEE Photon. Technol. Lett.* **19**, 616–618 (2007).
168. S. Yoo, J. K. Lee, and K. Kim, "Suppression of thermal wavelength drift in widely tunable DS-DBR laser for fast channel-to-channel switching," *Opt. Express* **25**, 30406 (2017).
169. S. Aisawa, K. Noguchi, and H. Miyao, "Learning method for neural networks using weight perturbation of orthogonal bit sequence and its application to adaptive WDM demultiplexer," *Journal of Lightwave Technology* **15**, 1997–2005 (1997).
170. S. Fan and J. M. Kahn, "Principal modes in multimode waveguides," *Optics Letters* **30**, 135 (2005).
171. T. N. Ford, D. Lim, and J. Mertz, "Fast optically sectioned fluorescence HiLo endomicroscopy," *J. Biomed. Opt.* **17**, 021105 (2012).
172. C. Dong, C. C. Loy, K. He, and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **38**, 295–307 (2016).
173. E. Cuche, F. Bevilacqua, and C. Depeursinge, "Digital holography for quantitative phase-contrast imaging," *Opt. Lett.*, **OL 24**, 291–293 (1999).
174. W. Choi, C. Fang-Yen, K. Badizadegan, S. Oh, N. Lue, R. R. Dasari, and M. S. Feld, "Tomographic phase microscopy," *Nat Meth* **4**, 717–719 (2007).
175. Y. Park, C. Depeursinge, and G. Popescu, "Quantitative phase imaging in biomedicine," *Nature Photonics* **12**, 578–589 (2018).
176. M. Rostykus, F. Soulez, M. Unser, and C. Moser, "Compact lensless phase imager," *Opt. Express*, **OE 25**, 4438–4445 (2017).
177. C. Ba, M. Palmiere, J. Ritt, and J. Mertz, "Dual-modality endomicroscopy with co-registered fluorescence and phase contrast," *Biomed Opt Express* **7**, 3403–3411 (2016).
178. H. Xiao, K. Rasul, and R. Vollgraf, "Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms," *arXiv:1708.07747 [cs, stat]* (2017).
179. A. Krizhevsky, "Learning Multiple Layers of Features from Tiny Images," in (2009).
180. Y. Rivenson, H. Wang, Z. Wei, K. de Haan, Y. Zhang, Y. Wu, H. Günaydin, J. E. Zuckerman, T. Chong, A. E. Sisk, L. M. Westbrook, W. D. Wallace, and A. Ozcan, "Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning," *Nature Biomedical Engineering* **3**, 466–477 (2019).
181. C. W. Freudiger, W. Min, B. G. Saar, S. Lu, G. R. Holtom, C. He, J. C. Tsai, J. X. Kang, and X. S. Xie, "Label-Free Biomedical Imaging with High Sensitivity by Stimulated Raman Scattering Microscopy," *Science* **322**, 1857–1861 (2008).
182. B. Rappaz, B. Breton, E. Shaffer, and G. Turcatti, "Digital Holographic Microscopy: A Quantitative Label-Free Microscopy Technique for Phenotypic Screening," *CCHTS* **17**, 80–88 (2014).
183. Y. Rivenson, T. Liu, Z. Wei, Y. Zhang, K. de Haan, and A. Ozcan, "PhaseStain: the digital staining of label-free quantitative phase microscopy images using deep learning," *Light Sci Appl* **8**, 23 (2019).
184. M. A. Kara, F. P. Peters, F. J. W. ten Kate, S. J. van Deventer, P. Fockens, and J. J. G. H. M. Bergman, "Endoscopic video autofluorescence imaging may improve the detection of early

- neoplasia in patients with Barrett's esophagus," *Gastrointestinal Endoscopy* **61**, 679–685 (2005).
185. P. J. Campagnola and L. M. Loew, "Second-harmonic imaging microscopy for visualizing biomolecular arrays in cells, tissues and organisms," *Nature Biotechnology* **21**, 1356–1360 (2003).
 186. E. CuChe, P. Marquet, and C. Depeursinge, "Simultaneous amplitude-contrast and quantitative phase-contrast microscopy by numerical reconstruction of Fresnel off-axis holograms," *Appl. Opt.*, AO **38**, 6994–7001 (1999).
 187. F. Zernike, "Phase contrast, a new method for the microscopic observation of transparent objects," *Physica* **9**, 686–698 (1942).
 188. C. J. Mann, L. Yu, C.-M. Lo, and M. K. Kim, "High-resolution quantitative phase-contrast microscopy by digital holography," *Opt. Express*, OE **13**, 8693–8698 (2005).
 189. H. Alqahtani, M. Kavakli-Thorne, and G. Kumar, "Applications of Generative Adversarial Networks (GANs): An Updated Review," *Arch Computat Methods Eng* (2019).
 190. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," arXiv:1609.04802 [cs, stat] (2017).
 191. H. N. Pathak, X. Li, S. Minaee, and B. Cowan, "Efficient Super Resolution For Large-Scale Images Using Attentional GAN," arXiv:1812.04821 [cs] (2019).
 192. K. Lee, K. Kim, J. Jung, J. Heo, S. Cho, S. Lee, G. Chang, Y. Jo, H. Park, and Y. Park, "Quantitative Phase Imaging Techniques for the Study of Cell Pathophysiology: From Principles to Applications," *Sensors (Basel)* **13**, 4170–4191 (2013).
 193. C. J. Mann, L. Yu, C.-M. Lo, and M. K. Kim, "High-resolution quantitative phase-contrast microscopy by digital holography," *Opt. Express*, OE **13**, 8693–8698 (2005).
 194. J. M. Bioucas-Dias and G. Valadao, "Phase Unwrapping via Graph Cuts," *IEEE Transactions on Image Processing* **16**, 698–709 (2007).
 195. K. Kim, S. Lee, J. Yoon, J. Heo, C. Choi, and Y. Park, "Three-dimensional label-free imaging and quantification of lipid droplets in live hepatocytes," *Scientific Reports* **6**, 1–8 (2016).
 196. S. Li and G. Barbastathis, "Spectral pre-modulation of training examples enhances the spatial resolution of the phase extraction neural network (PhENN)," *Opt. Express*, OE **26**, 29340–29352 (2018).
 197. P. Fan, T. Zhao, and L. Su, "Deep learning the high variability and randomness inside multimode fibers," *Opt. Express*, OE **27**, 20241–20258 (2019).

Curriculum Vitae

Eirini Kakkava

PERSONAL INFORMATION

Born in Patras, Greece, 8th of July 1991

Email: eirini.kakkava@epfl.ch

Phone: +41 21 69-33742

Address: EPFL STI IMT LO
BM 4106 (Bâtiment BM)
Station 17
CH-1015 Lausanne

EDUCATION

PhD – Photonics

Nov 2015 – Present

École Polytechnique Fédérale de Lausanne, Optics Laboratory, Lausanne, Switzerland

Doctoral Thesis: "Wavefront shaping and deep learning in fiber endoscopy"

Advisor: Prof. Demetri Psaltis

MSc – Photonics

Sep 2013 – Sep 2015

University of Patras, Department of Physics, Greece

Master Thesis: "Preparation and characterization of Polymer/Metallic Nanoparticle based systems by laser ablation and investigation of their nonlinear optical properties"

Advisor: Prof. Stelios Couris

GPA: 9.3/10.0

BSc – Physics**Sep 2009 – Jul 2013**

University of Patras, Department of Physics, Greece

Major: Lasers and Photonics

Bachelor Thesis: "Synthesis and characterization of Copper/Copper Oxide Nanoparticles by laser ablation in liquids"

Advisor: Prof. Stelios Couris

GPA: 7.7/10.0

High School**Sep 2009**

Major: Technology Sciences

GPA: 19.7/20.0, High School of Diakopto, Aigio, Greece

RESEARCH EXPERIENCE**PhD Candidate-Optical Scientist****Nov 2015 – Present**

École Polytechnique Fédérale de Lausanne, Optics Laboratory, Lausanne, Switzerland

Advisor : Prof. Demetri Psaltis

- ❖ Demonstration of multimodal endoscopic probes using multimode optical fibers supporting femtosecond laser ablation based microsurgery and two-photon fluorescence microscopy using wavefront shaping techniques.
- ❖ Investigation of deep learning in fiber endoscopy for improved imaging resolution and translation of information between imaging modalities, namely bright field to phase imaging.
- ❖ Optical setup design, data acquisition using Matlab interface, wavefront shaping via spatial light modulators, image processing using different modules (ImageJ, Matlab, Python), deep learning (Keras API, Python), light propagation simulations using the beam propagation method.
- ❖ Biological sample manipulation: cell necrosis and apoptosis monitoring using fixation and staining protocols, hydrogel preparation, cell line care.

Bachelor Intern

Apr 2012 – Jun 2012

FORTH-ICEHT, Patras, Greece

Supervisors : Prof. Stelios Couris, Dr. Vassilios Dracopoulos

- ❖ Preparation of metallic nanoparticle solutions using laser ablation of metals in solutions of various solvents.
- ❖ Material characterization techniques (SEM, XRD, DLS, Zeta Sizer) for measuring the structural and chemical properties of metallic nanoparticles in thin films and solutions.

TEACHING EXPERIENCE

École Polytechnique Fédérale de Lausanne, Switzerland

Sep 2017– Dec 2017

Master courses:

- ❖ Optical waves propagation

Semester projects:

- ❖ Fiber mode multiplexing by designing and 3D printing polymer structures,
MSc. Student: Dinç Niyazi Ulas
- ❖ Enhancement of imaging resolution through Multicore fibers using Deep Neural Networks,
MSc. Student: Barthe Lancelot

University of Patras, Department of Physics, Greece

Sep 2014– Jun 2015

Bachelor Courses:

- ❖ Laboratory of Geometrical and Wave optics
- ❖ Electromagnetism
- ❖ Laser Physics & Lasers' Laboratory

PROFESSIONAL QUALIFICATIONS

Technical skills

- ❖ **Optics:** Microscopy, digital holography, spatial light modulators, optical fibers, laser ablation, spectroscopy, nonlinear optics, metrology
- ❖ **Biology:** Cell cultures, tissue phantoms, necrosis and apoptosis studies, histology, sample preparation protocols (staining, fixation)
- ❖ **Materials:** Characterization techniques (UV-Vis spectroscopy, Raman, AFM, SEM), additive manufacturing

Computer Skills

- ❖ **Programming:** Matlab, Python
- ❖ **Data processing:** Origin Pro, Microsoft Excel
- ❖ **Other software:** Adobe Illustrator, Lightroom, Zemax, ImageJ

Languages

- ❖ **Greek:** native language
- ❖ **English:** very good (C2)
- ❖ **German:** good (B2)
- ❖ **French:** basic (A2/B1)

AWARDS

- ❖ Scholarship by the State Scholarships Foundation-IKY (Greece), Best score of the year for the entry exams to the Physics Department of the University of Patras, Greece.
- ❖ Master thesis scholarship co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF)-Research Funding Program: Thales. Investing in knowledge society through the European Social Fund, Research Project PHOTOPOLIS at the University of Patras.

PUBLICATIONS

Peer-reviewed articles:

1. E. Kakkava, N. Borhani, B. Rahmani, U. Teğın, C. Moser, D. Psaltis, "Deep Learning-Based Image Classification through a Multimode Fiber in the Presence of Wavelength Drift", *Applied Sciences* 10, 3816 (2020)
2. G. Konstantinou, E. Kakkava, L. Hagelüken, P. V. W. Sasikumar, J. Wang, M. Grazyna Makowska, G. Blugan, N. Nianias, F. Marone, H. Van Swygenhoven, J. Brugger, D. Psaltis, C. Moser, "Additive micro-manufacturing of crack-free PDCs by two-photon polymerization of a single, low-shrinkage preceramic resin", *Additive Manufacturing* 101343 (2020)
3. P. V. W. Sasikumar, E. Mueller, P. Clement, J. Jang, E. Kakkava, G. Panusa, D. Psaltis, K. Maniura-Weber, M. Rottmar, J. Brugger, G. Blugan, "In vitro Cytocompatibility Assessment

- of Ti-modified Silicon-oxycarbide Based Polymer-derived Ceramic Implantable Electrodes under Pacing Conditions”, ACS Applied Materials & Interfaces (2020)
4. U. Teğın, B. Rahmani, E. Kakkava, N. Borhani, C. Moser, D. Psaltis. “Controlling spatiotemporal nonlinearities in multimode fibers with deep neural networks”, APL Photonics 5 (3), 030804 (2020)
 5. U. Teğın, E.Kakkava, B. Rahmani, D. Psaltis, C. Moser , “Spatiotemporal self-similar fiber laser”, Optica 6, 1412-1415 (2019).
 6. E. Kakkava, B. Rahmani, N. Borhani, U. Teğın, D. Loterie, G. Konstantinou, C. Moser, and D. Psaltis, "Imaging through multimode fibers using deep learning: The effects of intensity versus holographic recording of the speckle pattern," Optical Fiber Technology 52, 101985 (2019).
 7. E. Kakkava, M. Romito, D. B. Conkey, D. Loterie, K. M. Stankovic, C. Moser, D. Psaltis, "Selective femtosecond laser ablation via two-photon fluorescence imaging through a multimode fiber", Biomedical optics express 10 (2), 423-433 (2019).
 8. P. Vallachira Warriam, G. Blugan, N. Casati, E. Kakkava, G. Panusa, D. Psaltis, J. Kuebler, " Polymer derived silicon oxycarbide ceramic monoliths: Microstructure development and associated materials properties ", Ceramics International 44 (17), 20961-20967 (2018).
 9. N. Borhani, E. Kakkava, C. Moser, D. Psaltis, "Learning to see through multimode fibers", Optica 5, 960–966 (2018).
 10. DB. Conkey, E. Kakkava, T. Lanvin, D. Loterie, N. Stasio, E. Morales-Delgado, C. Moser, D. Psaltis, "High power, ultrashort pulse control through a multi-core fiber for ablation", Optics Express 25 (10), 11491-11502 (2017)

Conference articles and presentations:

1. 2019-02 SPIE Photonics West, San Francisco, California, USA, “Two-photon imaging and selective laser ablation of cochlea hair cells through a multimode fiber probe”, Eirini Kakkava, Marilisa Romito, Damien Loterie, Konstantina Stankovich, Christophe Moser, Demetri Psaltis (Oral presentation)
2. 2019-02 SPIE Photonics West, San Francisco, California, USA, “Deep neural networks for seeing through multimode fibers”, Eirini Kakkava, Navid Borhani, Christophe Moser, Demetri Psaltis (Poster presentation)
3. 2019-06 OSA Computational Optical Sensing and Imaging, Munich, Germany, “Efficient Image Classification through a Multimode Fiber using Deep Neural Networks in presence of Wavelength Drifting”, Eirini Kakkava, Navid Borhani, Babak Rahmani, Ugur Tegin, Christophe Moser, Demetri Psaltis (Oral presentation)
4. 2019-06, The European Conference on Lasers and Electro-Optics (CLEO-Europe), Munich, Germany, “Wavelength Independent Image Classification Through A Multimode Fiber Using Deep Neural Networks”, Eirini Kakkava, Navid Borhani, Babak Rahmani, Uğur Teğın, Christophe Moser, Demetri Psaltis (Oral presentation)
5. 2019-07, 7th Advanced Electromagnetics Symposium, Lisbon, Portugal, “Optical imaging using Deep Neural Networks”, Eirini Kakkava, Navid Borhani, Babak Rahmani, Uğur Teğın, Christophe Moser, Demetri Psaltis (Oral presentation)
6. 2018-02 SPIE Photonics West, San Francisco, California, USA, “Wavefront shaping for ultrashort pulse delivery through optical fibers for imaging and ablation”, E. Kakkava, N. Stasio, D. B. Conkey, Damien Loterie, C. Moser, D. Psaltis (Oral presentation)

7. 2018-06 OSA Computational Optical Sensing and Imaging, Orlando, Florida, USA, "Seeing through Multimode Fibers with Deep Learning", Eirini Kakkava, Navid Borhani, Christophe Moser, Demetri Psaltis (Oral presentation)
8. 2018-09, OSA Frontiers in Optics, Washington DC, USA, "Deep Neural Networks for Information Recovery Through Multimode Fibers", Eirini Kakkava, Navid Borhani, Christophe Moser, Demetri Psaltis (Oral presentation)
9. 2018-09, OSA Frontiers in Optics, Washington DC, USA, "Two-photon fluorescence microscopy and selective laser ablation through multimode fibers", E Kakkava, BD Conkey, D Loterie, Christophe Moser, Demetri Psaltis (Poster)
10. 2017-08 SPIE Optics and Photonics, San Diego, California, USA, 6-10August 2017, "Femtosecond pulse delivery through multi-core fibers for imaging and ablation", E. Kakkava, N. Stasio, D. B. Conkey, C. Moser, D. Psaltis (Invited oral presentation)
11. 2017-06 CLEO/Europe-EQEC 2017 Munich, Germany, 25 - 29 June 2017, "Ultrashort pulse laser ablation through a multi-core fiber", E. Kakkava, D. B. Conkey, T. Lanvin, D. Loterie, N. Stasio, E. Morales-Delgado, C. Moser, D. Psaltis (Oral presentation)

PROFESSIONAL MEMBERSHIPS

- ❖ Former delegate PhD representative for the Photonics doctoral school, EPFL (2017-2018)
- ❖ Vice-president of the EPFL Photonics Chapter committee funded by OSA and SPIE (2018-2019)
- ❖ OSA and SPIE member (2017-2020)