Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Reference-less decomposition of highly multimode fibers using a physics-driven neural network

Qian Zhang

qian.zhang@tu-dresden.de

Laboratory of Measurement and Sensor System Technique, TU Dresden https://orcid.org/0000-0002-1951-2722

Yuan Sui

Laboratory of Measurement and Sensor System Technique, TU Dresden

Stefan Rothe

Department of Applied Physics, Yale University

Nektarios Koukourakis

TU Dresden

Juergen Czarske

TU Dresden https://orcid.org/0000-0001-7280-0523

Article

Keywords: multimode fiber, mode decomposition, deep learning, physics-informed, phase retrieval, fiber laser

Posted Date: October 14th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-5110336/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Reference-less decomposition of highly multimode fibers using a physics-driven neural network

Qian Zhang[1*†], Yuan Sui[1†], Stefan Rothe[2†], Nektarios Koukourakis[1], Juergen W. Czarske[1]

[1*]Laboratory of Measurement and Sensor System Technique, TU Dresden, Helmholtzstrasse 18, 01069, Dresden, Sachsen, Germany.
[2]Department of Applied Physics, Yale University, 06520, New Haven, CT, USA.

*Corresponding author(s). E-mail(s): qian.zhang@tu-dresden.de;
†These authors contributed equally to this work.

Abstract

Multimode fibers (MMFs) play an increasing role in the optical communication field, as well as in ultra-thin imaging systems and fiber lasers. The characterization of light propagation properties through active and passive MMFs attracts high interest as many linear and nonlinear optical phenomena fundamentally depend on the interplay of many different spatial modes. Access to the exact modal amplitude and phase weights, namely mode decomposition (MD), helps to describe the present state and brings a better understanding of physical effects, driving further technological advancements on MMF-based telecommunication, endoscopy, or amplifiers. Here we show that using an untrained neural network, which assumes a simple physical model as a prior, we achieve MD with more than **98%** accuracy with up to 5,796 spatial modes. For the first time, we demonstrate that a deep learning-based MD method can work on a 1-km-long MMF for communication. We present our network framework and results using synthetic data and experimental data measured on passive MMFs as well as on an MMF amplifier. This method holds great promise for applications in fiber lasers, endoscopic metrology, and reference-free calibration techniques in fiber-based classical communication and single-photon quantum key distribution.

Keywords: multimode fiber, mode decomposition, deep learning, physics-informed, phase retrieval, fiber laser

1 Introduction

The investigation of multi-dimensional optical fields is a fundamental interest. Multi-dimensional optical fields are used in their corresponding multi-dimensional optical systems with which huge scientific potential can be unleashed. Numerous optical degrees of freedom relate to the existing *modes* the control of which goes hand in hand with the achievable performance and scalability of the underlying system and its application. For instance, the capacity in multimode fiber (MMF) communication systems that rely on spatial and wavelength multiplexing scales linearly with the number of transverse fiber modes and frequency modes used [1]. The resolution attained in multimode and multicore fiber endoscopes scales linearly with the number of controlled transverse modes or cores [2, 3]. In the realm of fiber lasers and amplifiers, the achievable threshold enhancement of limiting nonlinear interactions inside the gain medium like Stimulated Brillouin Scattering [4] or Transverse Mode Instability scales linearly with the number of modes. Optical computers that support multiple modes are expected to even achieve a cubic scaling of the computational complexity they can process.

The growing interest in multi-dimensional optical systems has stimulated the need for measuring and controlling the propagation of light through them. Technological advancements on this frontier benefit from the control of the *entire* light field, i.e. amplitude and particularly the phase and polarization. For this objective, it is crucial to provide quantitative feedback from a determined plane on the distal side, i.e. where the control process becomes recognizable. However, especially feedback of the phase in multi-dimensional photonic systems poses major challenges. Usually, one very convenient way to obtain phase feedback is via an iterative measurement of the intensity. When power transmission with many spatial degrees of freedom shall be optimized, e.g. focusing light through an MMF, only the intensity inside the focus region on the distal side requires observation [5]. Within at least three defined phase increments shifted from the proximal side using a spatial light modulator (SLM), the intensity evolution is recorded, and the optimal SLM phase pattern is found by scanning the phase of each pixel [6]. For more complex tasks, e.g. the transmission of images through the MMF, phase feedback of the entire distal plane is often required. One prominent method to provide full-plane phase retrieval is given by heterodyne interferometry, e.g. digital phase-stepping holography [7]. Phase stepping holography exploits mutual light paths between object and reference beam, but requires four defined phase shifts and therefore four measurements for only one phase feedback. Digital off-axis holography, in turn, enables the measurement of the distal phase in single-shot, but at the expense of external reference beam provision which may not be suitable for some applications. With Fourier optics, distal phase information can be retrieved based on measuring the intensity and additonal optimization algorithms like Gerchberg-Saxton. This reference-less method, however, comes with the cost of at least two intensity measurements in both the near-field and far-field [8–10].

In the context of MMF, the field analysis providing separation into the MMF's eigenmode basis is called mode decomposition (MD). MD enables a quantitative description of the distal field from which both practical experiments and comparison with theory benefit [11]. Trivially, an MD can be performed as an additional

step after the distal field measurement, whereby practically no gain was achieved in terms of measurement efficiency. However, with the premise that the field that is an output of an MMF consists of the linear superposition of the MMF mode basis, the MD approach is suitable for estimating the full field by a single reference-less intensity measurement. Manuylovich et al. propose an analytical model that carries out a full-field MD using one single intensity image [12]. Despite its fast response rate, the scalability of this method is greatly hampered by measurement noise, which limits this MD approach to 8 modes. MPLC is an emerging technique for mode converters in telecommunication [13–15]. However, this technique is currently only applied to pure mode categorization and lacks the ability to decompose complex mode combinations. Thanks to its popularity and the following rapid growth of artificial intelligence, deep neural networks (DNNs) emerged as MD method with the two-fold advantage being (i) reference-less and (ii) single-shot [16–22]. Such intelligent MD methods based on DNNs outperform all other methods in terms of inference time. However, supervised training of DNNs typically requires high-performance GPUs, large data sets, and extremely long pre-training times [19]. Probably the most significant detriment, however, is that currently there is a strong limitation of the number of modes that can be processed with supervised learning-based MD. Although theoretically an infinite number of modes could be processed with infinite computational power, the counter of the current state of research is $\lessapprox 30$ modes. Recently, Zhang et al. trained a mode transformer neural network (MTNet) with a data set comprising $3,000,000$ samples for 34 hours [20]. The MTNet shown can process 23 modes on simulation data with 98.24% accuracy, and 10 modes in the experiment with 97.35% accuracy with a high-performance GPU in 1.1 milliseconds inference time.

The recent appearance of physical model-driven neural networks has brought new insights to MD [23–35]. In these works, the integration of physical laws enables the possibility of reducing the dependence on massive labeled data, which is critical for the decomposition of a large number of modes.

Here, we introduce a novel deep learning-based MD approach that harnesses the physical premise that a MMF field comprises the linear superposition of the eigenmodes. Our physics-driven mode decomposition neural network (PhyMoNet) thereby greatly reduces the complexity of the task that shall be learned. Huge data sets and long training sessions are no longer necessary, yet vast numbers of modes can be analyzed. Thus, our relatively shallow network can perform MD with over 5000 modes within a few minutes of inference time. We show that our approach can be applied to both passive and active MMF systems. This demonstrates the wide-ranging potential of intelligent MD for photonic applications.

2 Results

2.1 Concept of PhyMoNet

The concept of our PhyMoNet is illustrated in Fig. 1. The MMF can be regarded as a disordered optical medium and tends to scatter light during its propagation towards the output. Due to manufacturing tolerances, fiber bending or input coupling deviations, a light beam that has concentrated the optical power in one single mode

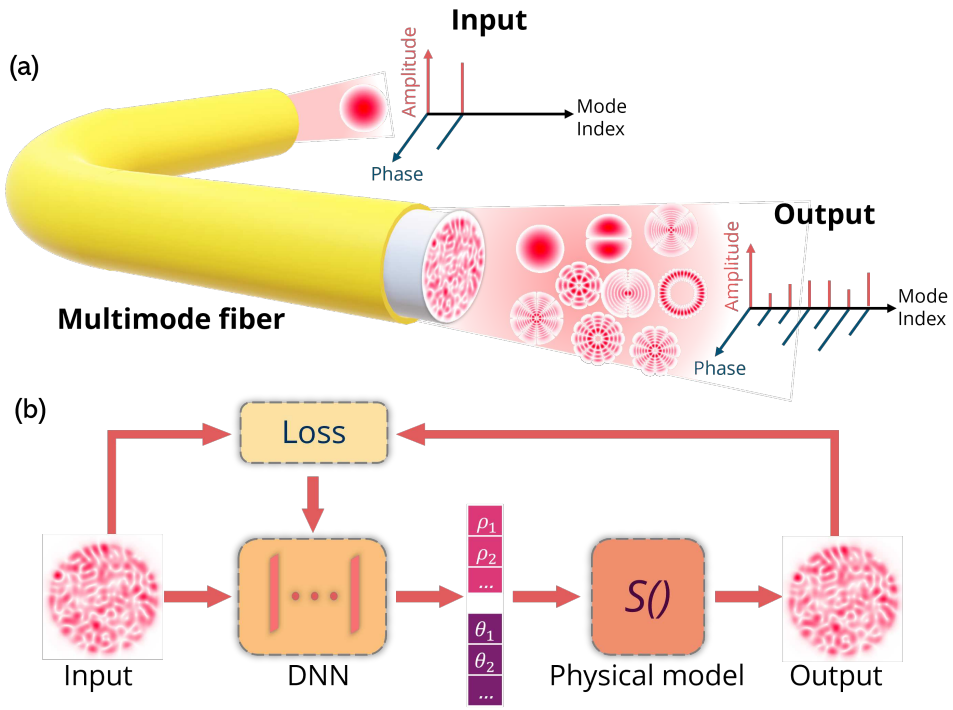


Fig. 1 (a) Schematic illustration of the multimode fiber. When a Gaussian mode is at the input side, a speckle pattern composed of different transverse modes appears at the output since mode conversations happens. (b) Schematic diagram of the PhyMoNet.

will seemingly arbitrarily re-distribute its power to other modes and produce a speckle pattern (see Fig. 1a). Such mode mixing randomizes both modal amplitude and modal phase coefficients at the fiber output (note, that we neglect polarization mixing effects). As such, the transverse optical field emerging from the MMF (e.g. $E_{\mathrm{MMF}}(x, y)$) results from the linear superposition of all individual modes supported by the MMF for one polarization with certain mode weights:

$$E_{\mathrm{MMF}}(x, y) = \sum_{i=1}^{N} \rho_i e^{i\theta_i} \Psi_i(x, y), \tag{1}$$

where ρ_n and θ_n are the modal amplitude and modal phase coefficients, respectively. N is the total number of supported modes in the fiber. For a standard step-index MMF, the total number of modes can be calculated as $N \approx \left(\frac{\sqrt{2}\pi a \mathrm{NA}}{\lambda_0}\right)^2$ assuming step-index profile, with a is the core radius, NA is the numerical aperture and λ_0 is the wavelength in vacuum. The mode field profile of the ith propagating mode $\Psi_i(x, y)$ can be calculated by solving the wave equation that is derived from Maxwell's electromagnetism theory [36]. For standard step-index fibers and a small refractive index

difference between core and cladding, the solution of the wave equation can be relaxed using the *weakly guiding fiber* approximation [37]. For many cases, it is then sufficient to use linearly polarized scalar fields as mode basis, which are calculated using Bessel and modified Hankel functions. A more accurate approach to theoretically describe the modes propagating in a MMF takes into account the space-dependent polarization that results in minor corrections of the set of solutions to the wave equation [38]. From (1), it becomes evident that a 2D optical field that outputs an N-mode MMF can be fully determined by a 1D set of $2N-1$ modal coefficients comprising N amplitude weights and $N-1$ phase weights, since the phase of a reference mode can be treated as global offset and consequently set to zero. Thus, the task of our PhyMoNet is to predict only the mode weights (ρ_i, θ_i) corresponding to a speckle pattern.

The architecture of the PhyMoNet is shown in Fig. 1b. PhyMoNet comprises a DNN unit and a physical model $S(\cdot)$. When an image, e.g. a speckle pattern, is fed into PhyMoNet, the learnable parameters are updated based on a loss function, like in most supervised training examples. Here, instead, the loss is not only calculated based on the features extracted from the input image, but also on the interaction between DNN and model $S(\cdot)$. The DNN block outputs the modal weights (ρ_i, θ_i) we are intentionally looking for and this layer usually serves as output layer for most previous DNN architectures working on MD [17–20]. In PhyMoNet, however, the modal weights are input to our model, that maps them back to a 2D-speckle pattern $S(\sqrt{\rho_n^2}, \theta_n) = \sqrt{\left|\sum_{n=1}^{N} \rho_n e^{i\theta_n} E_n\right|^2} = \sqrt{I(x,y)} = A(x,y)$, where $A(x,y)$ is the amplitude profile and $I(x,y)$ is the intensity profile. As such in the presented problem, we calculate a 2D amplitude distribution using the model weights provided from the DNN and compare it with the input image by calculating the loss between them. Note, that we include two additional constraints. (i) the amplitude weights must be positive, and (ii) energy conservation must be satisfied: $\sqrt{\sum \rho_i^2} = 1$. As a consequence, by reducing the loss between input image and model output, we are optimizing the learnable parameters and thereby the modal weights for each input image, individually. Other architectures [17–20], instead, consider a comprehensive training phase *a priori* the actual application and learn a specific MD for a certain MMF — offline to the later experiment. Here, since we carry out an individual training process for each speckle pattern, we therefore use an *untrained* DNN as ground state for our MD and optimize the learnable parameters of the DNN online.

It is worth emphasizing that we are not aware of any preferred DNN architectures that would significantly influence the later MD performance. Creative solutions can certainly be used here. In our work, we use a simple multilayer perceptron (MLP) containing 3 fully connected layers: input layer, hidden layer and output layer [39]. Such architecture has already been successfully implemented for MMF-based experiments [40]. Here, the input layer has the same size as the pixel number of the test image. In the hidden layer, *LeakyReLU* is used as the activation function and the number of output channels is 256. The output layer has a dimension of $(2N-1)\times 1$ including N amplitude coefficients and $N-1$ phase coefficients. The normalized N

amplitude coefficients can be directly forwarded to the model. Afterward, the normalized amplitude coefficients and the phase coefficients with zero-padding to the first mode are directly forwarded to the superposition model to generate the light field.

It should be clarified here that unlike many methods where the range of the phase is limited or the absolute value of the phase is used, we do not process anything with the phase. The intensity image of an optical field is obtained by multiplying the field E by its complex conjugate E^*, resulting in a real-valued quantity: $I \sim |E|^2 = E \cdot E^*$. As a consequence, the conjugate field has the same intensity distribution as the original field. However, it is not possible to distinguish between the original field and its conjugate by using only the intensity information. Unlike the previous solution used in [20], pre-and post-processing of phase signs are not required. In the proposed method, PhyMoNet automatically converges towards the phase or the conjugate phase.

2.2 Performance of PhyMoNet

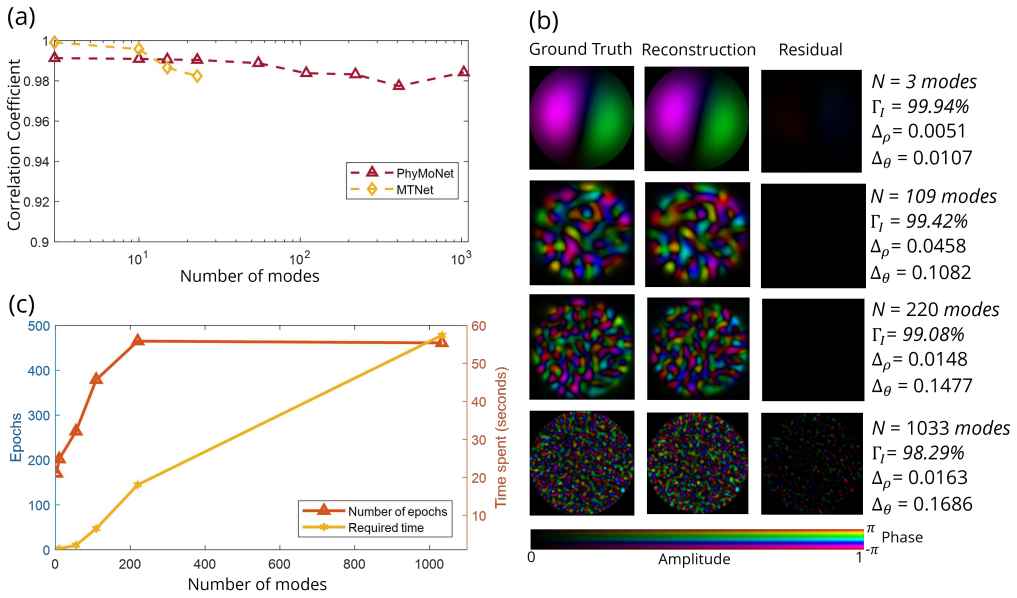


Fig. 2 (a) Comparison between the performance of PhyMoNet and MTNet[20] on synthetic data without noise. Each point is the mean value of 100 test data. Number of modes for PhyMoNet: 3-, 10-, 15-, 23-, 55-, 109-, 220-, 1033-mode; for MTNet: 3-, 10-, 15-, 23-modes. The image size is 64×64. (b) Examples of decomposition on different number of modes: 3 modes, 109 modes, 220 modes, and 1033 modes. Intensity correlation coefficients (Γ_I) and decomposition errors are shown on the right of each set of images. (c) Relationship between convergence rate (here, E_{MD}) and number of modes. The shown results are the average for 100 test data.

The PhyMoNet is first tested using synthetic noise-free data. For a fair and comprehensive comparison, 100 images with the resolution of 64×64 have been generated

for all test cases. Figure 2a shows the correlation coefficient γ plotted over varying mode numbers. *For a change in mode number, we varied the fiber core diameter in our simulation.* A detailed overview of the MD results is presented in **Supplementary Material Table S1**. It can be found that the average correlation coefficients $\bar{\Gamma}$ exceed 98% over three orders of magnitude increasing mode number ranging from 3 modes to 1033 modes. Compared to Zhang's MTNet [20], a displacement can be observed. For a few-mode (≤ 10) case, MTNet shows better performance. For more modes, it is obvious that PhyMoNet is more advantageous. As the number of modes gradually increases, $\bar{\Gamma}$ shows a slight performance drop from 99.9% to 98%. This may be enhanced by changing the DNN architecture. Note, that for highly multi-mode MMFs, the speckle patterns become extremely complex. When the image resolution is too low, it fails to extract detailed image information. When increasing the image resolution for the highly multi-mode MMFs cases, the results for deteriorate first when maintaining the number of training epochs. However, by also increasing the number of epochs, the accuracy of decomposition can be improved again. In **Supplementary Material Table S2**, the performance of the PhyMoNet on 1033 modes is given for different input image sizes and epoch numbers.

Temporal effort typically scales proportionally with the number of modes [10, 12]. As the PhyMoNet needs not to be trained, We also measure the decomposition epoch E_{MD}, which reflects the number of training epochs required for the reconstruction to reach the target correlation (here, 98%). To more accurately reflect the relationship between convergence speed and the number of modes, E_{MD} was recorded and is shown in Fig. 2 (c). As the number of modes increases, the required number of training epochs also increases, corresponding to the exponential growth of decomposition complexity. Compared to data-driven deep learning approaches, where the required training data grows exponentially, this nearly proportional relationship demonstrates a significant advantage [20]. In addition, because different numbers of modes differ in the computational requirement, the individual epoch times required vary. Fig. 2 (c) also gives the time required to reach the target correlation coefficient. These values here are for reference only, as they will vary with changes in the DNN architecture and the number of modes.

In experimental scenarios, the accuracy of MD is often affected by noise from the environment. Deep learning-based MD may experience a performance degradation of approximately 3% to 10% [17, 20, 41] due to the presence of noise. Another analytic approach has also low robustness [12]. Therefore, testing with data containing noise can reflect PhyMoNet's applicability for MD in real-world scenarios. For a comprehensive investigation, a Gaussian noise model has been introduced. The signal-to-noise ratio varies from 1 dB to 30 dB representing different noise levels (see **Supplementary Material Section 2**).

It can be found in Table S3 that when the SNR is lower than 5 dB, the correlation between the reconstructed light fields and the original noise-free light fields, an average correlation coefficient of around 95% can be achieved. This indicates that PhyMoNet is able to extract valuable information for MD under noise conditions even below 5 dB. When the SNR of the optical field is greater than 10 dB, the accuracy of MD can exceed 95%. As the SNR increases to 20 dB, the results of MD can already reach 98%, which is

very close to the results of perfect synthetic data. The comprehensive investigation of different cases from 3-mode to 220-mode shown in **Supplementary Material Table 3** has demonstrated that the PhyMoNet is strong against noisy data. The lower the noise level, the higher the decomposition accuracy. For all test cases, as the noise level decreases, the standard deviation is also reduced, which means PhyMoNet is also stable on noisy data with a large number of modes, which is the same as non-noisy data. Additionally, compared to amplitude error, phase error increases more rapidly as the SNR decreases.

A further study found that the impact of noise data on the convergence speed of the network model is almost negligible. **Supplementary Material Fig. S3** illustrates the training process with noisy data for 55-mode. For a more intuitive analysis, correlation coefficients were computed for the reconstructed light field and the noise light field. The results represent the average of 100 samples. It can be observed that all training processes converge rapidly within the first 200 epochs and gradually stabilize.

We used the optical system shown in Fig. 3(a) to apply the PhyMoNet to a real-world environment. We use CW lasers that are spatially filtered using a SMF. After collimation, the wavefront of the beam phase is modulated using a spatial light modulator (SLM, Holoeye Pluto 2.1). We apply random phase masks to generate randomized test data for our PhyMoNet [42]. The phase-modulated beam is coupled to the MMF using a 4-f system. Note, that we couple the SLM near-filed into the MMF. However, conjugate-plane coupling would produce sufficient data, as well. The output near-field of the MMF is imaged onto a camera, where we capture the intensity profile. Considering the different application scenarios of MMFs, various fibers are utilized in this work for a comprehensive test. As for fiber-based endoscopes, a large core diameter is needed, a fiber with a core diameter of 200 μm and a numerical aperture of NA = 0.39 is used. At the given wavelength (@1550 nm), 5796 modes are supported per polarization stat. For the usage scenarios in communications, another 2 fibers, one with a core diameter of 10 μm and one with a core diameter of 19 μm, are studied. At the given wavelength (@532 nm), they support 8 and 65 modes per polarization state, respectively. MMF_{65} has a length of 1km. Since polarization diversity is not considered in this work, two polarizers are placed at both ends of the MMF. Before being sent into the PhyMoNet, each intensity-only image is cropped.

For evaluation, using 100 different random phase masks all intensity images are directed into PhyMoNet for MD. The accuracy of MD was evaluated by calculating the correlation coefficient between the reconstructed light distribution and captured light field distribution. In a comparative study, we analyze the MD performance of different network architecture types on 8 modes in the experiment. Besides testing our PhyMoNet, we feed experimental data to both VGG [18] and MTNet-type DNNs [20] and evaluate their responses by calculating the 2D cross-correlation with the ground truth on as much as 100 different samples. All networks are tested with the same data and each sample corresponds to a different MMF output profile resulting from a random SLM phase mask at the input. Table S4 shows the result of the 2D correlation together with the standard deviation calculated over the 100 data points, which indicates the network robustness. It turns out that PhyMoNet outperforms the

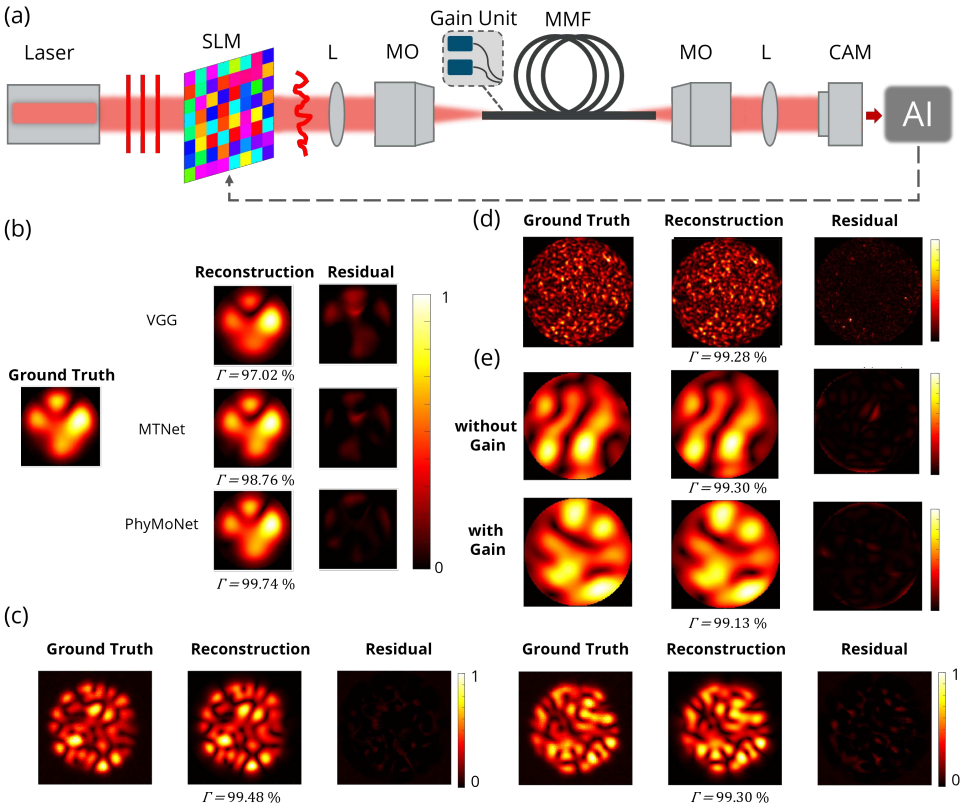


Fig. 3 (a) Scheme of the experimental setup. Abbreviations: CAM, camera; L, lens; MMF, multi-mode fiber; MO, microscope objective; P, polarizer; and SLM, spatial light modulator. In this work, normal MMF supporting 8 modes and 5796 modes @1550 nm and MMF laser supporting 40 modes @1064 nm were tested. (b) Comparison between PhyMoNet and other DNN models [18, 20] on the experimental examples (8 modes). (c) Examples of MD on 1-km-long MMF. (d) Examples of MD on experimental data (5796 modes). (e) Examples of MD on experimental data from the MMF amplifier.

other 2 network types, whilst MTNet is only slightly inferior on 8 modes. For the 8-mode case, both MTNet and PhyMoNet show great robustness, while MTNet shows an even lower standard deviation than PhyMoNet. However, only with PhyMoNet we are able to carry out MD at mode numbers greater than 23. For MTNet, in turn, only as much as 23-mode MD has been achieved on synthetic data so far. It can be noticed that in the experimental data, PhyMoNet has degraded performance compared to the synthetic data. This is due to the imperfection of the optical system and the rescaling of capture intensity images, which brings noise to the measured data. For the 8-mode case, images are rescaled to 64×64. We further investigate the performance of PhyMoNet on the 65 modes fiber with kilometer length. A 2D correlation coefficient with a value of 98.01% has been achieved with the resolution of 128×128, see Fig. 3 (c).

When a MMF has over a thousand modes, it can be used as an imager for imaging systems [43, 44]. With 5796 modes in the experiment, with PhyMoent we achieve a 2D cross-correlation of 98.77% with the resolution of 128×128. Examples of reconstructed mode fields and corresponding ground truths including their residuals are shown in Fig. 3 (d). As mentioned in the testing of the synthetic data, more modes require greater resolution to preserve more complex scattering structures. which also implies an increase in the amount of computation required.

Another study on PhyMoNet is the use of an MMF amplifier. Fig. S5 illuminates the schematic of the system. A CW single-frequency laser @1064 nm is expanded as the input. A SLM is used to manipulate the input wavefront. The laser is then coupled into an MMF amplifier. At the distal end of the MMF amplifier, a CMOS is placed to capture the near-field beam profile, and a power meter is used to measure the power. Using the SLM, 200 different speckle patterns at the amplifier output with random mode excitation at the input are collected. Although there are nonlinear variations, the neural network can still perform MD of scattered patterns. In this test, all images were recorded with the original 128×128 pixels. The holographic MD is performed as a reference. When the amplifier is turned off, a mean cross-correlation coefficient of 98.23% is achieved using PhyMoNet, whereas the holographic MD yields a mean correlation coefficient of 97.22%. When the amplifier pumps to a power of 10W, the PhyMoNet has a mean correlation coefficient of 98.16%. In this case, a mean correlation coefficient of 97.70% is achieved by using the holographic MD. The results from both methods are nearly the same, see Fig. 3 (e).

3 Discussion

In summary, we propose a novel method that enables complex mode decomposition (MD) of passive and active multimode fibers (MMFs) using only intensity images towards the characterization and calibration of fiber communication, MMF-based endoscopy, and fiber amplifiers. MMFs have attracted much attention in different fields. Due to the complex transmission behavior, pre-calibration is normally needed, where MD plays a role. Once the mode content of the output from MMF is obtained using MD, distortion behavior can then be provided as the feedback to the proximal end of a fiber for pre-coding, which enables the correction of channel for classical and quantum communication [45]. As investigated in previous work, high-fidelity mode transmission can be achieved by employing distortion-free channels with the help of SLM [20]. Furthermore, through the measuring of the whole transmission matrix, physical layer security can then be implemented [46].

Thanks to the straightforward experimental setup, neural networks have increasingly become a primary focus for numerical MD. However, achieving precise decomposition for a larger number of modes remains a significant challenge. To address MD in highly MMFs, we introduced a physics-informed mode decomposition neural network (PhyMoNet) that incorporates the overlap of modes in MMFs into the framework of a deep neural network model. In comparison to previous methods with data-driven neural networks, PhyMoNet does not require massive datasets for training, which is one of the major challenges when dealing with a large number of modes. As a result,

PhyMoNet is capable of decomposing over thousands of modes in the experimental environment, which is two orders of magnitude more than existing numerical methods. Additionally, the PhyMoNet current used contains only 3 fully connected layers, requiring significantly less computational power compared to more complex models like DenseNet[19], Transformer[47], or ResNet[48].

By using synthetic data for testing, it is easy to explore the impact of the number of modes. Under low-mode conditions, PhyMoNet exhibits performance comparable to both neural network and mathematical methods. As the number of modes increases, the advantage of PhyMoNet becomes more pronounced. Additionally, PhyMoNet's resistance to high noise makes it highly suitable for practical scenarios. We confirm that even under an SNR of only 1dB, PhyMoNet can still extract useful information from noisy data. In order to quantitatively analyze the accuracy of the pattern decomposition, the predicted mode weights error was calculated to comprehensively and accurately evaluate the results of PhyMoNet. Additionally, to provide a more intuitive assessment, the correlation coefficient of the reconstructed optical field from the mode weight predicted to the ground truth was also evaluated. From Table S2, it can be found that the amplitude error first increases and then decreases. The increase in amplitude error is consistent with a decrease in the correlation coefficient, which is mainly due to the increasing number of modes. As the total energy of the light field is fixed, more modes involved means that the average amplitude weight of each mode becomes smaller. This results in the amplitude error becoming relatively small for a larger number of modes. Therefore, the relative amplitude error is used. The results show that the relative error increases as the number of modes increases, which is also consistent with the change in phase error. In addition to the greater complexity of the problem itself, another reason that brings about an increase in error is that the likelihood of establishing intensity with similar shapes increases [10]. Simulation results indicate that the phase error is greater than the amplitude error. This discrepancy is may mainly due to the fact that the neural network's optimization is based on intensity or amplitude distribution. When the amplitude weight of a particular mode is very small, even significant phase differences have minimal impact on the overall amplitude distribution of the optical field. This suggests that when the amplitude distribution has a very high correlation with the target, the amplitude error is smaller compared to the phase error.

In the experimental tests, considering different application scenarios, MMFs supporting different numbers of modes were involved: a few modes or tens of modes for mMMF network in communication, thousands of modes for endoscopy, and MMF amplifier. An important feature that should not be forgotten is that compared to the data-driven-based method, physically driven networks have a great advantage in the experimental scenario, where labels are not available. The inclusion of physical models makes MD for highly MMF possible. In this work, we successfully decomposed a 1-km-long MMF supporting 65 modes. This is the first experimental demonstration of deep learning-based MD being able to run on kilometer-scale MMF. The nature of incoherent measurement makes our method not limited by the coherence length of the laser source, which is easy to realize. Although we are using the step-index MMF in this work, our method is suitable for GRIN fibers. Additionally, as the mode crosstalk

in a GRIN fiber mainly occurs within the mode group, MD has the potential to be accelerated using parallelized decomposition of different mode groups.

The more challenging work is the decomposition of fiber with a large core diameter. Through experimentation, we successfully demonstrate PhyMoNet's ability to decompose patterns in experimental data containing 5796 modes. MMF-based endoscopy usually needs to be pre-calibrated by measuring the transmission behavior [2, 43]. Calibration without reference laser beam greatly reduces the complexity of the optical system and the loss of illumination laser.

In addition to being used on passive MMF, our method can also be used in active MMF amplifiers which operate in the highly nonlinear gain saturation regime. The comparable performance of PhyMonet with the DH method suggests that our method is applicable to any coherent photonic system with beams that are within the spectral correlation width. It should be noted that, unlike few-mode fibers, the light field of highly MMF contains more complex information, where a higher image resolution is necessary. This increases the learning difficulty of PhyMoNet to achieve precise decomposition. In this case, more training loops are required. It is also expected that a more advanced network, such as DenseNet[19], Transformer[47], or ResNet[48], could improve the accuracy. Although the current work considers only a single polarization, polarization-diverse measurements can be easily created by adding polarization beam splitters and the same performance is expected.

It is worth noting that in tests ranging from 3 to 5796 modes, PhyMoNet exhibits minimal performance degradation. Therefore, the MD of the reference-free method theoretically has the potential to surpass the limitation on the number of modes. Further improvement in accuracy can be achieved by increasing the resolution of input images, though this would extend the time for MD. An expedient approach to accelerate this process is pre-training the network. Additionally, with the rapid advancement of chips, sophisticated computing units, or FPGA-accelerated neural networks, have the potential for fast MD. As deep learning continues to evolve, more advanced networks and training strategies are expected to bring breakthroughs in accuracy and the number of modes. Additionally, PhyMoNet can find suitable solutions even in the presence of phase ambiguity which provides a new perspective for reference-free phase retrieval. Collectively, we are convinced that our proposed novel reference-free optical mode decomposition opens new possibilities for applications based on MMFs. This method can be employed to realize reference-less digital optical phase conjugation via scatterer [49, 50, 50–53], enhance classical communication capacity, physical layer security through MMF [46], secure communication for quantum key distribution, or transferred for use in fiber-based endoscopes [3, 54–58].

4 Methods

Training of PhyMoNet

The neural network was implemented in Python with the PyTorch platform. The learning process is performed on a GPU NVIDIA RTX A6000 and a CPU AMD EPYC 7452. As the initial investigation, PhyMoNet has been tested under the same hyperparameters: the input size and the reconstructed field are 64×64. For the experimental

data, the resolution varies with the number of modes. Each sample was trained for 1000 epochs. The learning rate is initially set to 0.0001 and decreases by 10% every 200 epochs. The optimizer is Adam. The mean absolute error (MAE) loss function was applied to evaluate the difference between the input and reconstructed patterns.

Evaluation of PhyMoNet

To evaluate the performance and accuracy of our PhyMoNet, we calculate the correlation coefficient (Γ) between the reconstructed image and the input image:

$$\Gamma = \frac{\sum_R (I_t - \overline{I_t})(I_r - \overline{I_r})}{\sqrt{\left(\sum_R (I_t - \overline{I_t})^2\right)\left(\sum_R (I_r - \overline{I_r})^2\right)}}, \tag{2}$$

where I_t refers to the intensity of the input image and I_r is the reconstructed intensity. The weight error between the predicted mode weights and the actual mode weights is defined as: $\Delta\rho = \left|\sqrt{\rho_p{}^2} - \sqrt{\rho_t{}^2}\right|, \Delta\theta = \frac{|\theta_p - \theta_t|}{2\pi}$. Here, the subscript p indicates a predicted value, and subscript t denotes a target value. Note, that the error between predicted and actual mode weights can only be calculated for synthetic simulation data as there is no label for experimental data. $\Delta\rho$ refers to the amplitude error. In order to provide a more comprehensive response to the accuracy of the predicted weights for all modes, relative amplitude error was calculated as $\Delta\rho_r = avg(\left|\sqrt{\rho_p{}^2} - \sqrt{\rho_t{}^2}\right|)/avg(\sqrt{\rho_t{}^2})$, where the $avg(\sqrt{\rho_t{}^2})$ is the average value of the mode weights. Additionally, we calculate the standard deviation σ_Γ of the correlation coefficient to study the robustness of PhyMoNet.

Supplementary information. Please find the supplementary material in a separate file.

Declarations

Acknowledgements. This research was funded by the Federal Ministry of Education and Research of Germany with the project 6G-life (grant identification number: 16KISK001K) and QUIET (project identification number: 16KISQ092). Furthermore, this work was partially supported by the German Research Foundation for funding (grant number: CZ 55/42-2).

We thank Prof. Hui Cao (Department of Applied Physics, Yale University, New Haven, USA) for providing data from the MMF amplifier. We thank Dennis Pohle (Laboratory of Measurement and Sensor System Technique, TU Dresden, Dresden, Germany) for helpful discussions..

Competing interests. The authors declare no competing interests.

Author contributions statement. Q.Z., Y.S., and S.R. conceived the idea and designed the experiments, Q.Z. and Y.S. conducted the experiment(s), Q.Z., Y.S. and S.R. analysed the results. Q.Z., S.R. and Y.S. contributed to the manuscript writing. Q.Z., Y.S. and S.R. contributed equally to this work. J.W.C. and S.R. acquired the

research projects. J.W.C. and N.K. supervised the research work. All authors have read and agreed to the published version of the manuscript.

Availability of data and materials. The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

[1] Richardson, D.J., Fini, J.M., Nelson, L.E.: Space-division multiplexing in optical fibres. Nature photonics **7**(5), 354–362 (2013)

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

[3] Wen, Z., Dong, Z., Deng, Q., Pang, C., Kaminski, C.F., Xu, X., Yan, H., Wang, L., Liu, S., Tang, J., *et al.*: Single multimode fibre for in vivo light-field-encoded endoscopic imaging. Nature Photonics **17**(8), 679–687 (2023)

[4] Czarske, J., Müller, H.: Heterodyne detection technique using stimulated brillouin scattering and a multimode laser. Optics letters **19**(19), 1589–1591 (1994)

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

[6] Mastiani, B., Vellekoop, I.M.: Noise-tolerant wavefront shaping in a hadamard basis. Optics express **29**(11), 17534–17541 (2021)

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

[8] Shapira, O., Abouraddy, A.F., Joannopoulos, J.D., Fink, Y.: Complete modal decomposition for optical waveguides. Physical review letters **94**(14), 143902 (2005)

[9] Forbes, A., Dudley, A., McLaren, M.: Creation and detection of optical modes with spatial light modulators. Advances in optics and photonics **8**(2), 200–227 (2016)

[10] Brüning, R., Gelszinnis, P., Schulze, C., Flamm, D., Duparré, M.: Comparative analysis of numerical methods for the mode analysis of laser beams. Applied optics **52**(32), 7769–7777 (2013)

[11] Wright, L.G., Wu, F.O., Christodoulides, D.N., Wise, F.W.: Physics of highly multimode nonlinear optical systems. Nature Physics **18**(9), 1018–1030 (2022)

[12] Manuylovich, E.S., Dvoyrin, V.V., Turitsyn, S.K.: Fast mode decomposition in few-mode fibers. Nature communications **11**(1), 5507 (2020)

[13] Morizur, J.-F., Nicholls, L., Jian, P., Armstrong, S., Treps, N., Hage, B., Hsu, M., Bowen, W., Janousek, J., Bachor, H.-A.: Programmable unitary spatial mode manipulation. JOSA A **27**(11), 2524–2531 (2010)

[14] Fontaine, N.K., Ryf, R., Chen, H., Neilson, D.T., Kim, K., Carpenter, J.: Laguerre-gaussian mode sorter. Nature communications **10**(1), 1865 (2019)

[15] Rademacher, G., Puttnam, B.J., Luís, R.S., Eriksson, T.A., Fontaine, N.K., Mazur, M., Chen, H., Ryf, R., Neilson, D.T., Sillard, P., *et al.*: Peta-bit-per-second optical communications system using a standard cladding diameter 15-mode fiber. Nature Communications **12**(1), 4238 (2021)

[16] Gao, H., Hu, H., Zhao, Y., Li, J.: A real-time fiber mode demodulation method enhanced by convolution neural network. Optical Fiber Technology **50**, 139–144 (2019)

[17] An, Y., Huang, L., Li, J., Leng, J., Yang, L., Zhou, P.: Learning to decompose the modes in few-mode fibers with deep convolutional neural network. Optics express **27**(7), 10127–10137 (2019)

[18] Rothe, S., Zhang, Q., Koukourakis, N., Czarske, J.W.: Deep learning for computational mode decomposition in optical fibers. Applied Sciences **10**(4), 1367 (2020)

[19] Rothe, S., Zhang, Q., Koukourakis, N., Czarske, J.: Intensity-only mode decomposition on multimode fibers using a densely connected convolutional network. Journal of Lightwave Technology **39**(6), 1672–1679 (2021)

[20] Zhang, Q., Rothe, S., Koukourakis, N., Czarske, J.: Learning the matrix of few-mode fibers for high-fidelity spatial mode transmission. APL Photonics **7**(6), 066104 (2022)

[21] An, Y., Li, J., Huang, L., Li, L., Leng, J., Yang, L., Zhou, P.: Numerical mode decomposition for multimode fiber: From multi-variable optimization to deep learning. Optical Fiber Technology **52**, 101960 (2019)

[22] Situ, G.: Deep holography. Light: Advanced Manufacturing **3**(2), 278–300 (2022)

[23] Saba, A., Gigli, C., Ayoub, A.B., Psaltis, D.: Physics-informed neural networks for diffraction tomography. Advanced Photonics **4**(6), 066001–066001 (2022)

[24] Wang, F., Bian, Y., Wang, H., Lyu, M., Pedrini, G., Osten, W., Barbastathis, G., Situ, G.: Phase imaging with an untrained neural network. Light: Science & Applications **9**(1), 77 (2020)

[25] Lim, J., Psaltis, D.: Maxwellnet: Physics-driven deep neural network training based on maxwell's equations. Apl Photonics **7**(1) (2022)

[26] Zang, Y., Yu, Z., Xu, K., Lan, X., Chen, M., Yang, S., Chen, H.: Principle-driven fiber transmission model based on pinn neural network. Journal of Lightwave Technology **40**(2), 404–414 (2021)

[27] Tian, Z., Pei, L., Wang, J., Hu, K., Xu, W., Zheng, J., Li, J., Ning, T.: High-performance mode decomposition using physics-and data-driven deep learning. Optics Express **30**(22), 39932–39945 (2022)

[28] Baddoo, P.J., Herrmann, B., McKeon, B.J., Nathan Kutz, J., Brunton, S.L.: Physics-informed dynamic mode decomposition. Proceedings of the Royal Society A **479**(2271), 20220576 (2023)

[29] Tian, Z., Pei, L., Wang, J., Hu, K., Zheng, J., Li, J., Ning, T.: Physics-driven mode decomposition based on a training-free shallow neural network. Optical Fiber Technology **76**, 103239 (2023)

[30] Wu, X., Wu, Z., Shanmugavel, S.C., Yu, H.Z., Zhu, Y.: Physics-informed neural network for phase imaging based on transport of intensity equation. Optics Express **30**(24), 43398–43416 (2022)

[31] Gigli, C., Saba, A., Ayoub, A.B., Psaltis, D.: Predicting nonlinear optical scattering with physics-driven neural networks. Apl Photonics **8**(2) (2023)

[32] Wang, F., Wang, C., Deng, C., Han, S., Situ, G.: Single-pixel imaging using physics enhanced deep learning. Photonics Research **10**(1), 104–110 (2021)

[33] Tang, Z., Wang, F., Fu, Z., Zheng, S., Jin, Y., Situ, G.: Deepsci: scalable speckle correlation imaging using physics-enhanced deep learning. Optics Letters **48**(9), 2285–2288 (2023)

[34] Li, R., Pedrini, G., Huang, Z., Reichelt, S., Cao, L.: Physics-enhanced neural network for phase retrieval from two diffraction patterns. Optics Express **30**(18), 32680–32692 (2022)

[35] Wang, F., Wang, C., Chen, M., Gong, W., Zhang, Y., Han, S., Situ, G.: Far-field super-resolution ghost imaging with a deep neural network constraint. Light: Science & Applications **11**(1), 1 (2022)

[36] Snyder, A.W., Love, J.D., *et al.*: Optical Waveguide Theory vol. 175. Chapman and hall London, ??? (1983)

[37] Gloge, D.: Weakly guiding fibers. Applied optics **10**(10), 2252–2258 (1971)

[38] Cao, H., Čižmár, T., Turtaev, S., Tyc, T., Rotter, S.: Controlling light propagation in multimode fibers for imaging, spectroscopy, and beyond. Advances in Optics

and Photonics **15**(2), 524–612 (2023)

[39] Popescu, M.-C., Balas, V.E., Perescu-Popescu, L., Mastorakis, N.: Multilayer perceptron and neural networks. WSEAS Transactions on Circuits and Systems **8**(7), 579–588 (2009)

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

[41] An, Y., Huang, L., Li, J., Leng, J., Yang, L., Zhou, P.: Deep learning-based real-time mode decomposition for multimode fibers. IEEE Journal of Selected Topics in Quantum Electronics **26**(4), 1–6 (2020)

[42] Rothe, S., Daferner, P., Heide, S., Krause, D., Schmieder, F., Koukourakis, N., Czarske, J.W.: Benchmarking analysis of computer generated holograms for complex wavefront shaping using pixelated phase modulators. Optics Express **29**(23), 37602–37616 (2021)

[43] Gu, R.Y., Mahalati, R.N., Kahn, J.M.: Design of flexible multi-mode fiber endoscope. Optics express **23**(21), 26905–26918 (2015)

[44] Borhani, N., Kakkava, E., Moser, C., Psaltis, D.: Learning to see through multimode fibers. Optica **5**(8), 960–966 (2018)

[45] Zhou, Y., Braverman, B., Fyffe, A., Zhang, R., Zhao, J., Willner, A.E., Shi, Z., Boyd, R.W.: High-fidelity spatial mode transmission through a 1-km-long multimode fiber via vectorial time reversal. Nature communications **12**(1), 1866 (2021)

[46] Rothe, S., Besser, K.-L., Krause, D., Kuschmierz, R., Koukourakis, N., Jorswieck, E., Czarske, J.W.: Securing data in multimode fibers by exploiting mode-dependent light propagation effects. Research **6**, 0065 (2023)

[47] Dosovitskiy, A.: An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929 (2020)

[48] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)

[49] Czarske, J.W., Haufe, D., Koukourakis, N., Büttner, L.: Transmission of independent signals through a multimode fiber using digital optical phase conjugation. Optics express **24**(13), 15128–15136 (2016)

[50] Büttner, L., Thümmler, M., Czarske, J.: Velocity measurements with structured light transmitted through a multimode optical fiber using digital optical phase

conjugation. Optics Express **28**(6), 8064–8075 (2020)

[51] Haufe, D., Koukourakis, N., Büttner, L., Czarske, J.W.: Transmission of multiple signals through an optical fiber using wavefront shaping. JoVE (Journal of Visualized Experiments) (121), 55407 (2017)

[52] Aziz, A., Medina-Sánchez, M., Koukourakis, N., Wang, J., Kuschmierz, R., Radner, H., Czarske, J.W., Schmidt, O.G.: Real-time ir tracking of single reflective micromotors through scattering tissues. Advanced Functional Materials **29**(51), 1905272 (2019)

[53] Koukourakis, N., Wagner, F., Rothe, S., Karl, M.O., Czarske, J.W.: Investigation of human organoid retina with digital holographic transmission matrix measurements. Light: Advanced Manufacturing **3**(2), 211–225 (2022)

[54] Lich, J., Glosemeyer, T., Czarske, J., Kuschmierz, R.: Single-shot 3d incoherent imaging with diffuser endoscopy. Light: Advanced Manufacturing **5**(1), 1–11 (2024)

[55] Sun, J., Zhao, B., Wang, D., Wang, Z., Zhang, J., Koukourakis, N., Czarske, J.W., Li, X.: Calibration-free quantitative phase imaging in multi-core fiber endoscopes using end-to-end deep learning. Optics Letters **49**(2), 342–345 (2024)

[56] Wang, T., Dremel, J., Richter, S., Polanski, W., Uckermann, O., Eyüpoglu, I., Czarske, J.W., Kuschmierz, R.: Resolution-enhanced multi-core fiber imaging learned on a digital twin for cancer diagnosis. Neurophotonics **11**(S1), 11505–11505 (2024)

[57] Wu, J., Wang, T., Uckermann, O., Galli, R., Schackert, G., Cao, L., Czarske, J., Kuschmierz, R.: Learned end-to-end high-resolution lensless fiber imaging towards real-time cancer diagnosis. Scientific Reports **12**(1), 18846 (2022)

[58] Sun, J., Kuschmierz, R., Katz, O., Koukourakis, N., Czarske, J.W.: Lensless fiber endomicroscopy in biomedicine. PhotoniX **5**(1), 18 (2024)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- PhyMoNetsupplement20240917.pdf