



Light scattering control in transmission and reflection with neural networks

ALEX TURPIN,^{1,2} IVAN VISHNIAKOU,^{1,2} AND JOHANNES D. SEELIG^{1,*}

¹Center of Advanced European Studies and Research, 53175 Bonn, Germany

²These authors contributed equally

*johannes.seelig@caesar.de

Abstract: Scattering often limits the controlled delivery of light in applications such as biomedical imaging, optogenetics, optical trapping, and fiber-optic communication or imaging. Such scattering can be controlled by appropriately shaping the light wavefront entering the material. Here, we develop a machine-learning approach for light control. Using pairs of binary intensity patterns and intensity measurements we train neural networks (NNs) to provide the wavefront corrections necessary to shape the beam after the scatterer. Additionally, we demonstrate that NNs can be used to find a functional relationship between transmitted and reflected speckle patterns. Establishing the validity of this relationship, we focus and scan in transmission through opaque media using reflected light. Our approach shows the versatility of NNs for light shaping, for efficiently and flexibly correcting for scattering, and in particular the feasibility of transmission control based on reflected light.

© 2018 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

When light propagates through a non-homogeneous and non-isotropic material its wavefront becomes distorted due to aberrations and scattering, resulting in an apparently random interference pattern of granular speckles [1, 2]. Such scattering conditions hamper the controlled delivery of light and the engineering of the PSF, which is a basic requirement for many applications [3–6]. To counteract this effect, methods based on shaping the light wavefront entering the scattering material have been developed. Wavefront shaping is typically achieved by using spatial light modulators (SLMs) [7, 8] which, with their millions of degrees of freedom (pixels), allow focusing through diffusers [9–11], multimode fibers [3, 12, 13], and biological tissue [14–17]. Different techniques have been developed to determine the appropriate wavefront corrections to be displayed on the SLM. The first demonstration of scattering control took advantage of iterative wavefront optimization [9, 18, 19], which approaches the targeted light distribution, typically a single or multiple focal spots, by updating the wavefront depending on the result after each optimization step [4, 9, 18, 20, 21]. These feedback-based algorithms calculate the wavefront correction separately for each focal position or shape and can be optimized to achieve very fast focusing times [11, 22]. A second approach, also typically used to control a single focus, is digital optical phase conjugation which uses interferometry to measure the scattered light field and reverses it with an SLM [16, 23–26]. This technique has the advantage of achieving update rates approaching the millisecond range needed for imaging in dynamic biological tissue [27], while however requiring a focus or other guidestar to measure the appropriate correction. A third group of methods aims for describing and controlling the scattering process simultaneously across an entire field of view which was first achieved with the help of a transmission matrix [10, 28, 29]. For obtaining the transmission matrix, one needs to measure light phase, which, similar to digital optical phase conjugation, requires technically more demanding interferometric approaches. To simplify such experiments, computational methods for estimating incompletely measured information have been implemented which for example can infer the light phase from intensity measurements [30, 31].

Another set of computational techniques that, thanks to the development of programming frameworks together with the computational power of GPUs, is increasingly being applied in imaging and microscopy relies on machine learning (ML) [32–35] and in particular on NNs [36–40]. The usefulness of these techniques [41–50] has been demonstrated in the context of light scattering for image analysis [37, 45, 46, 48], where the goal lies in the classification of an object across a scattering layer, or image reconstruction based on a predefined data set [49]. In astronomy, NNs have been applied for the correction of weak scattering encountered when imaging through the atmosphere, for example for the control of multi mirror telescopes [50]. For light control, genetic algorithms, a class of iterative optimization algorithms, have been used for optimizing focusing across scattering materials [41, 42, 44]. Single- and multi-focus single-shot control (after training) over a 5×5 pixel area has been achieved using support vector regression [47], but the reported small field of view, low signal to noise ratio, and long training times (97 min) are limiting for high-resolution PSF engineering.

While the methods outlined above allow focusing through scatterers in transmission, an additional set of challenges arises when the focal plane lies hidden behind or inside the scatterer, remote from direct optical access. For applications such as sensing, imaging or communication this is the more relevant configuration [51]. For example in biological imaging, fluorescent signals or guidestars can be used to monitor excitation intensity inside or behind a scatterer ([2, 15, 19, 21]). Particularly in the presence of strong scattering, however, these signals are often dim and generally have a limited photon budget [52]. Alternatively, back-scattered excitation light can provide feedback about the beam [53–55]. These signals need to be additionally filtered to remove out-of-focus light and using various combinations of temporal, frequency, or spatial gating [53–59] one aims for extracting photons that are scattered little and therefore retain image information. Since such weakly scattered photons disappear exponentially with depth, they in turn limit imaging depth.

However, even under strongly scattering conditions reflected (or backscattered) photons carry information about transmitted light [51, 60, 61]. Mutual information between speckle patterns generated in these two opposite scattering directions indicates that reflected light might potentially be used to control transmitted light [51, 60, 61]. This would require that a functionally explicit relationship between these two scattering signals can be found and that the available information is sufficient for controlling one signal through the other. So far, reflected light has been used to maximize the energy sent into a sample [62, 63], but without control over the resulting light distribution, or required an embedded highly scattering target to achieve a localized light distribution [64]. Other schemes to take advantage of backscattered light have been suggested in theoretical work [65, 66], but these concepts have so far not been implemented experimentally.

Here, we discuss how neural networks can be used to image through materials with different scattering characteristics such as glass diffusers, multi-mode fibers, or paper. First, we show that single-layer NNs (SLNNs) and multi-layer convolutional NNs (CNNs) can be trained to control the light distribution behind scattering materials with high accuracy. Second, we show that NNs can be used to find a functional relationship between transmitted and reflected light, i.e., they can predict transmitted speckle patterns from reflected speckle patterns with sufficient accuracy for light control through opaque materials. Taking advantage of this relationship we then show that NNs can be used for focusing in transmission using reflected light.

2. Neural network approach for scattering control

We here first outline the underlying approach of using NNs for light control through a scatterer, which is also sketched in Fig. 1. In an initial step, we generate a dataset consisting of pairs of binary illumination patterns displayed on the SLM and corresponding speckle patterns recorded with a CCD camera after transmission through the scatterer (64×64 macropixels for illumination patterns and 96×96 pixels for the CCD camera). These pairs of illumination and speckle patterns

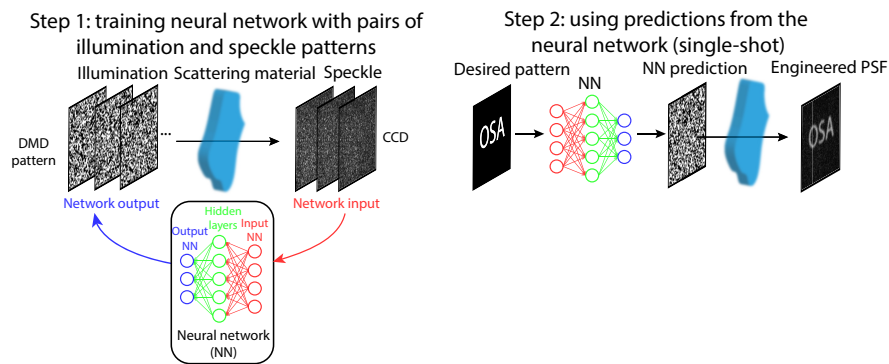


Fig. 1. Approach for light control through scattering media with NNs. A NN is trained with pairs of illumination and speckle patterns (illustration, see figures below for examples of actual illumination and speckle patterns), using the speckle patterns as input of the network and the illumination as output. Once the NN is trained, it is used to predict the illumination necessary to generate a target pattern after the scattering material. The predicted illumination is subsequently sent through the material resulting in the desired light pattern.

(typically on the order of 10000, but see below for training with a reduced number of patterns) are used to train the NNs as detailed below and in the Appendix, with the goal of inferring the relationship between the resulting scattered light distributions and the illumination patterns. We then feed the desired distribution into the trained NNs to predict the corresponding illumination pattern. This pattern is finally displayed on the SLM and the resulting light pattern is recorded with the camera. Each pattern, $\mathbf{C}(\mathbf{k})$, can be considered as a combination of plane waves with different wave vectors \mathbf{k} . This distribution of plane waves is modified by the scattering material through a function $F[\mathbf{C}(\mathbf{k})]$ in a deterministic way and results in the speckle pattern, \mathbf{S} , i.e. $\mathbf{S} = F[\mathbf{C}(\mathbf{k})]$. Through training, the NN learns an approximation of the function F needed to generate any light distribution after the scatterer.

3. Experimental set-up

The experimental setup is schematically shown in Fig. 2. A laser beam ($\lambda = 640$ nm, with an intensity of up to $P = 100$ mW; iBeamSmart, Toptica) is expanded with a telescope ($f_1 = 15$ mm, $f_2 = 150$ mm) and sent to the SLM. (For the experiment shown in the Appendix we additionally included an optical isolator to minimize reflections into the laser (Thorlabs, IO-3D-633-VLP)). The SLM is a high-speed digital micromirror device (DMD, 768×1024 pixels, pixel size = $13.7 \mu\text{m}^2$; model V-7000 from Vialux) allowing binary amplitude modulation at a maximum frame rate of 22.7 kHz and is used to display the illumination patterns (we have tested our system both with pseudo-random checkerboard-like patterns and with patterns obtained from Hadamard matrices) with typically 64×64 macropixels extending over the central 768×768 pixels of the DMD (12×12 micromirrors per macropixel). Two additional lenses ($f_3 = 200$ mm, $f_4 = 50$ mm) combined with a pinhole are used after the DMD to filter the maximum-intensity diffraction order mode and to demagnify and image the DMD onto the back aperture of the microscope objective (10X, 0.25 NA, or 40X, 0.6 NA, WD=2.7-4.0 for one of the experiments shown in the Appendix, both from Olympus). The objective focuses the light beam through the scatterer (a glass diffuser, Thorlabs DG20-120, a step-index multimode fiber optic patch cable, Thorlabs M38L02; and a piece of white paper of $100 \mu\text{m}$ thickness) and a second identical microscope objective is used to collect the scattered light. Finally, a pair of lenses ($f_5 = 100$ mm, $f_6 = 75$ mm) in $2f$ configuration (or only a single lens with $f = 60$ mm for one of the experiments shown in the Appendix) images the back aperture of the second microscope objective (or the sample for one of

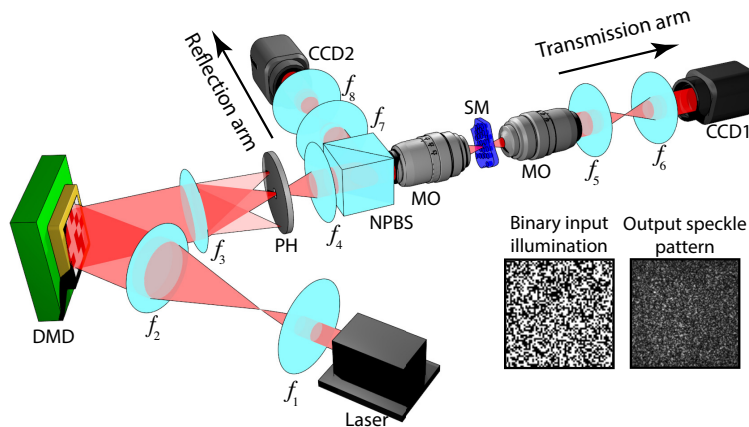


Fig. 2. Sketch of the experimental setup (see text for details). A DMD generates pseudo-random binary illumination patterns that are projected onto the back aperture of a microscope objective. After passing through the scattering material, light is projected onto a CCD camera by using a second identical microscope objective. A beam splitter placed before the first objective and a CCD camera are used to retrieve speckle patterns reflected by the sample for experiments with combined transmission and reflection. Bottom-right insets depict illustrations of illumination and speckle patterns. (PH: pinhole, NPBS: non-polarizing beam splitter, MO: microscope objective, SM: scattering material.)

the experiments shown in the Appendix) onto the CCD camera (acA640-750um, Basler), with a frame rate of 500 fps at full resolution of 480×640 pixels (pixel size $4.8 \mu\text{m}^2$). Both microscope objectives and the scattering material are mounted on XYZ stages (omitted in Fig. 2) for aligning the system and moving the sample to different positions, as well as for displacing the image plane axially. In our experiments typically 10000 checkerboard patterns are uploaded to the internal memory of the DMD. Then, the projection of a pattern on the DMD triggers the frame capture of the CCD camera (transmission arm). The maximum frame rate of the DMD is 22.7 kHz and the maximum frame rate of the CCD camera is about 1000 fps at a resolution of 96×96 pixels, which allowed us to record the whole sequence in about 10 s. We also note that our approach is valid for larger fields of view than those shown in the main figures ($20 \times 20 \mu\text{m}^2$, (see Appendix for details). For experiments with reflected light, a non-polarizing beam splitter redirects the backscattered speckles towards a pair of lenses ($f_7 = 50$ mm, $f_8 = 25$ mm) in $2f$ configuration that image the back aperture of the first microscope objective onto a second CCD camera identical to and synchronized with the one used to capture the transmitted speckles (reflection arm). We used a computer with a Linux-Ubuntu operating system, an Intel Xeon CPU E5-1620 v4 @ 3.50 GHz, 32Gb of DDR5 RAM memory, and a Nvidia Titan XP GPU possessing 3840 CUDA cores running at 1.60 GHz and with 12GB of GDDR5X memory running at over 11 Gbps.

4. Neural networks for light control through scattering materials

In Fig. 3(a) we demonstrate the ability of SLNNs to generate diffraction-limited Gaussian foci through a glass diffuser (as used for example in [17]) at different positions within the field of view. Top images schematically illustrate the NN architecture and training process, as detailed in the Appendix. Briefly, the SLNN connects all input to output channels through a single, so-called fully connected layer. Below that, the first rows show the intensity distribution captured with the CCD camera, while the second and third rows display horizontal and vertical cross sections through the center of the focus. Insets and red-dashed lines show the position and shape of the target distribution that is fed into the neural network and for which the network then calculates

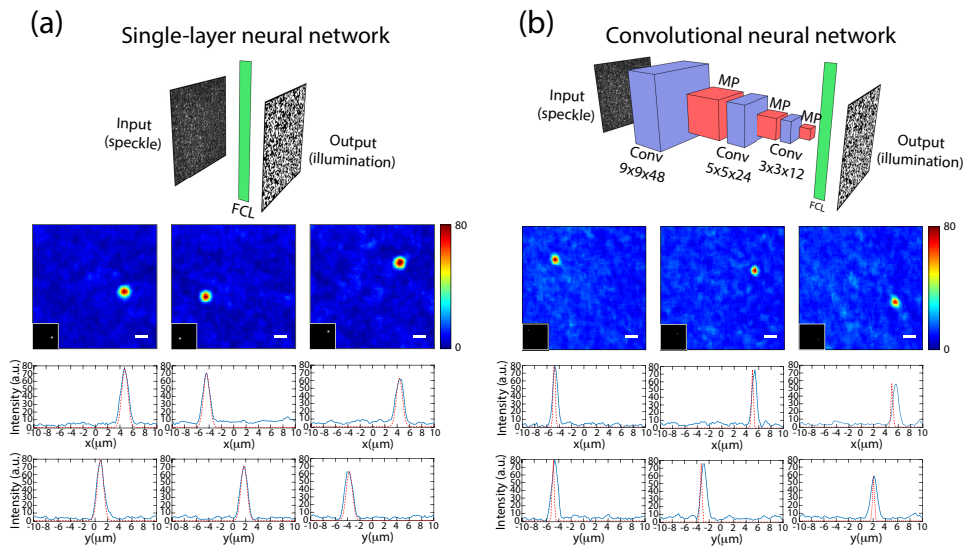


Fig. 3. Focusing with neural networks. Top: Illustration of SLNN (a) or CNN (b) with speckle and illumination pattern. Bottom: Intensity distributions (first row) and intensity profiles through the foci along horizontal (second row) and vertical (third row) directions of Gaussian beams obtained at different positions after training (a) the SLNN and (b) the CNN. Red-dashed lines (--) are the targeted intensity distributions that enter the NN, normalized to the corresponding experimental result. Scale bars = $2 \mu\text{m}$. FCL: fully-connected layer; Conv $n \times m \times p$: convolutional layer of p kernels with dimensions $n \times m$; MP: max pooling operation reducing the previous element size. Color bars: intensity (a.u.).

the appropriate SLM mask (the target distribution is displayed normalized to the experimentally recorded intensity). The quality of the generated foci is analyzed with an automated procedure that generates spots at different positions placed in a grid throughout the whole field of view and measures the enhancement, defined as: $\eta \equiv I_{\text{focus}} / \langle I_{\text{speckle}} \rangle$, where I_{focus} is the intensity at the generated foci and $\langle I_{\text{speckle}} \rangle$ is the mean value of the background speckle [67].

The images show an excellent agreement between the desired and recorded patterns (see also the Appendix for quantification of the match between the target and actual light distribution) with a signal-to-noise ratio > 10 and an enhancement $\eta = 32 \pm 5$ for the 10X objective (see Appendix for scanning across the entire field of view) and an enhancement $\eta = 81 \pm 18$ for the 40X objective (see Appendix). The time to achieve light control depends on the number of recorded frames and the training time. For the typical datasets of 10000 frames (with a resolution of 64×64 macropixels on the DMD and 96×96 pixels on the CCD, recorded at 1000 Hz) training on a single GPU required 34 s, and could be reduced down to 18 s while keeping an enhancement $\eta > 10$ (for the 10X objective see Appendix).

While SLNNs are easy to implement and train, the underlying linearity limits their performance for many tasks [68]. A plethora of other network architectures have therefore been developed with the goal to improve over the performance of SLNNs. The most straightforward generalization of SLNNs combines multiple NN layers with connections between all neurons, resulting in a densely connected network. While such densely connected networks are not limited by linearity, the increased number of parameters also makes them more challenging to train, particularly for large data sets such as stacks of high resolution images. Network architectures were therefore developed to take into account the structure of the underlying data and convolutional neural networks (CNNs) have emerged as one of the most successful solutions for image processing [68]. The typical architecture of a CNN consists of multiple convolutional layers that extract features

across an entire field of view, interspersed with pooling layers that down-sample the image, and fully connected layers. While a large number of different networks are applied for different tasks, with a few to a few hundred convolutional layers [32, 33, 68], we here found that a three-layer CNN (see Fig. 3(b) and Appendix for details) could be used for scattering control through a glass diffuser. To circumvent the difficulties of training nonlinear networks we pretrained the network with an autoencoder [68], a network that compresses and then uncompresses the data into a close approximation of the input. (See also Appendix for focusing through paper with a different CNN architecture and training procedure.) The part of the network that was used for compression then served as the initial CNN for scattering control.

In Fig. 3(b) we demonstrate the ability of CNNs to generate diffraction-limited Gaussian foci through a glass diffuser (see Appendix for scanning across the field of view). The images again show a good agreement between the target pattern that was fed into the CNN (red dashed lines and inset) and the recorded patterns (see the Appendix for quantification) with a signal-to-noise ratio > 10 and an enhancement $\eta = 3.6 \pm 0.9$ (measured over an ensemble of 25 different focus positions). For this particular application, CNNs reduced the number of network parameters by 80% compared to SLNNs at the cost of lower enhancement with similar number of training samples. A larger enhancement of $\eta = 10 \pm 5$ was obtained with the 40X objective, as shown in the Appendix for focusing across the more strongly scattering paper.

4.1. SLNNs for point spread function engineering

Since the SLNN is a linear network, we reasoned that after training it should be able to take advantage of the linearity of light scattering in non-absorbing media to generate arbitrary light distributions. To demonstrate the validity of our approach for controlling the light intensity distribution after the scatter we generated in Fig. 4 a variety of non-trivial shapes using SLNNs. Again, there is an excellent correspondence between the target distribution that enters that network (insets) and the recorded patterns. We note that thanks to the high frame rate of the DMD (22.7 kHz), alternatively, any shape can be generated with high fidelity by painting it spot by spot, e. g. similar to approaches for trapping ultra cold atoms [69] or optogenetics [70], as shown in [Visualization 1](#), [Visualization 2](#), [Visualization 3](#), [Visualization 4](#), [Visualization 5](#), and [Visualization 6](#).

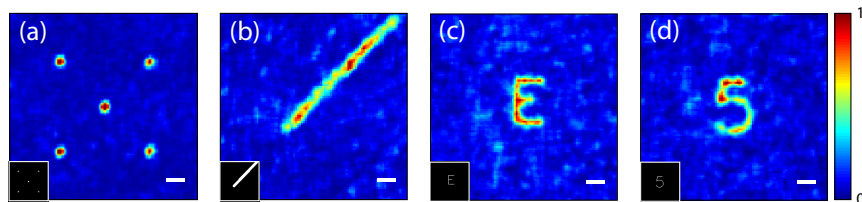


Fig. 4. Light control through a glass diffuser with a SLNN. Normalized intensity patterns obtained after the glass diffuser with the SLNN: (a) five Gaussian foci; (b) a line at 45° ; (c) the letter "E"; and (d) the number "5"; Insets show the desired light distribution. Scale bars = $2 \mu\text{m}$. Color bar: intensity (a.u.) normalized for each image.

4.2. SLNNs for light control through optical fibers

Our system is suited well to correct for scattering in materials with slow dynamics (on the order of a few tens of seconds, see Appendix), such as optical fibers [12, 71–73]. In particular, multimode optical fibers are ideal for applications in imaging and optogenetics, but modal dispersion and cross-talk distribute light into an apparently random speckle pattern. We therefore tested the performance of SLNNs for controlled light delivery through multimode fibers. In Fig. 5 and

Visualization 7, Visualization 8, and Visualization 9 a single focus is scanned ($\eta = 10 \pm 3$) with different paths across the field of view of the fiber (including a circle, a square, and a 5×5 array of points), demonstrating that SLNNs are able to precisely control light through optical fibers.

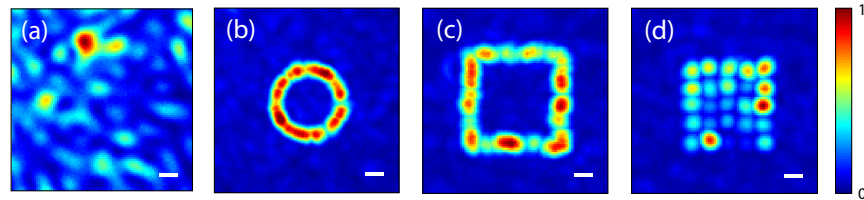


Fig. 5. Neural networks focus light through multimode fibers. Normalized transverse maximum intensity projection of the light field (see Visualization 7, Visualization 8, and Visualization 9 for further information) after a multimode fiber when (a) no correction is applied and when a single focus is scanned (maximum intensity projection) along (b) a circle, (c) a square, and (d) an array of 5×5 points. Scale bars = $2.3 \mu\text{m}$. Color bar: intensity (a.u.) normalized for each image.

5. Neural networks find functional relationships between transmitted and reflected speckle patterns

While most methods for focusing light through strongly scattering media rely on measuring transmitted light (as in the experiments described so far), many applications could benefit from using reflected light. Towards that goal we tested whether neural networks (NNs) can take advantage of mutual information between transmission and reflection images [51, 60, 61] for light control. In the following we show that with the help of NNs it is indeed possible to find a functional relationship between reflected and transmitted speckle patterns to control transmitted light using reflected light.

For this experiment we simultaneously recorded transmitted and reflected light by adding a non-polarizing beam splitter, a pair of imaging lenses and a CCD camera to the setup, as shown in Fig. 2. To achieve good signal to noise ratio of transmitted as well as reflected speckle patterns, we used paper as scattering material, which was more strongly scattering than the glass diffuser [57] and led to an increased amount of backscattered light. Sets of simultaneously recorded transmitted and reflected speckle patterns (with size of 128×128 pixels) were then generated by illuminating the sample with a series of checkerboard projections (64×64). Once the speckle patterns were recorded, we trained a SLNN (SLNN1) to find the relationship between transmitted and reflected light. To quantify the performance of this network we used the Pearson correlation coefficient as a similarity measure [74] between transmission speckle patterns predicted by the network and measured transmission speckle patterns. Fig. 6(a) shows the histogram of these correlation coefficients and the correlation coefficient between transmitted and reflected speckle patterns for comparison. Figs. 6(b) and (c) show, respectively, an example of measured and predicted speckle patterns with median correlation ($\rho = 0.50$), while Figs. 6(d) and (e) show an example of measured reflected speckle pattern when a corresponding focus is generated in transmission. Note that when focusing in transmission the intensity of the speckle pattern and the number of speckle grains in reflection decrease.

To take advantage of this relationship between transmitted and reflected light for light control, we trained a second independent network (SLNN2 in Fig. 7(a)) to infer the relation between reflected speckles and illumination patterns, similar to the training of the SLNN in the transmission configuration in the previous sections of the article (that is, with the reflected speckles as input of the SLNN and the illumination patterns as output, see Appendix for details). Combining these SLNNs, as shown in Fig. 7(a), allowed us to form transmission foci by only taking

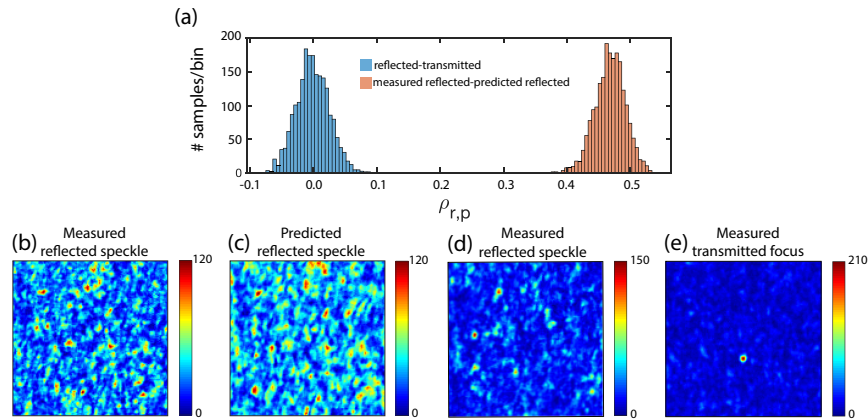


Fig. 6. Neural networks find functional relationships between transmitted and reflected speckle patterns. (a) Histogram of the Pearson correlation coefficient $\rho_{r,p}$ between the measured reflected and transmitted speckle patterns (in blue) and the measured and predicted reflected speckle patterns (in orange). Number of bins: 30. Total number of pairs of samples used: 2000. Bottom: Examples of the normalized transverse intensity distributions of the light field of the (b) measured reflected speckle pattern, (c) predicted reflected speckle pattern, and (d) measured reflected speckle pattern while focusing in transmission (e).

advantage of reflected light, based on SLNN1 relating reflected to transmitted speckle patterns. In Figs. 7(b)-(d) we show, respectively, that we can scan a circle, a square, and a grid, demonstrating full control of transmitted modes using reflected modes over the entire field of view. This additionally demonstrates that the predicted speckle patterns (Fig. 6) are sufficiently accurate for high-resolution light control. The measured enhancement in this case was $\eta = 12 \pm 4$ (measured over an ensemble of 25 different focus positions). Note also that even though paper is more strongly scattering [57] than the glass diffuser, this does not hinder the SLNN from light control. Focusing through paper with a SLNN as reported in the previous sections is shown in the Appendix and [Visualization 10](#), [Visualization 11](#), and [Visualization 12](#).

6. Discussion and conclusions

In summary, we showed that NNs can be used to efficiently shape light through a variety of media with different scattering characteristics (Figs. 3 – 5, 8–11). Once the NNs are trained, we achieve real-time, single-shot light control through the scattering material with high fidelity, in a fashion similar to transmission matrix approaches [10, 29, 75]. Specifically, we demonstrated the ability of SLNNs to focus and scan light through glass diffusers, multimode fibers, and paper, and to generate arbitrary light distributions through glass diffusers. We further showed that nonlinear networks, specifically CNNs, can focus light through a glass diffuser and paper.

In a second set of experiments, we demonstrated that with the help of two networks, one establishing an explicit functional relationship between light that is transmitted through a scatterer and light that is reflected, and one relating reflected light to illumination patterns, we can control transmission using reflection at diffraction limited resolution. SLNNs therefore prove to be well suited to take advantage of a recently described mutual information between transmitted and backscattered light for light control [51, 60, 61].

To compare the performance of the NN method for focusing in transmission with other schemes, we quantified the enhancement as in [67] and obtained values similar to those reported for intensity-only modulation for the SLNN (see Fig. 9 in the Appendix for scanning through paper with the SLNN) [30, 42, 67] and lower values for the CNNs (but still with a sufficient SNR

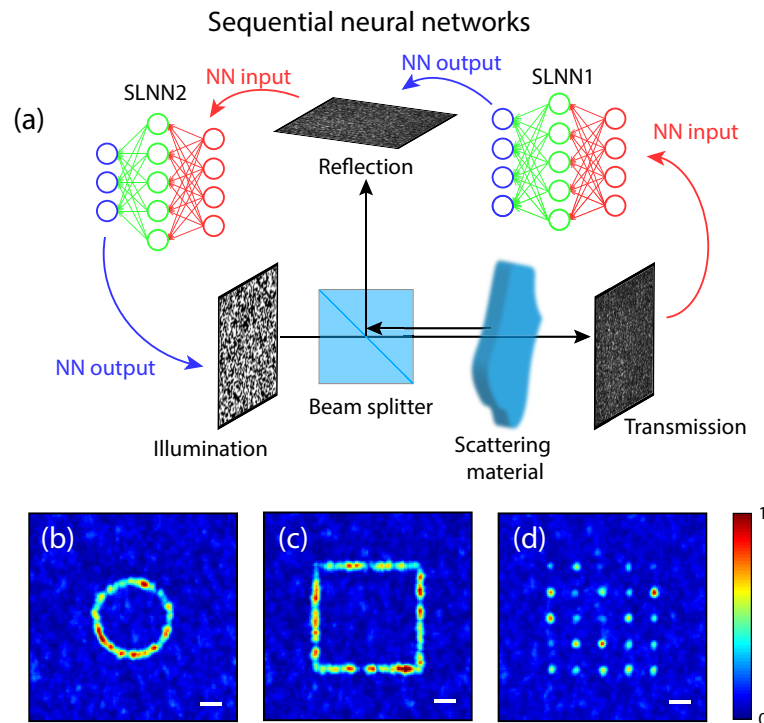


Fig. 7. Focusing and scanning in transmission using reflected light. Illustration of network approach to control transmitted light using reflected light. A SLNN is trained to learn the relationship between simultaneously recorded transmitted and reflected speckle patterns. After training of the network, it is sufficient to train a second SLNN to relate reflection to illumination for controlling light through the combined network. Bottom: Normalized transverse maximum intensity projection of the light field after a sheet of paper when a single focus is scanned (maximum intensity projection) along (a) a circle, (b) a square, and (c) an array of 5×5 points. Scale bars = $2.7 \mu\text{m}$. Color bar: intensity (a.u.) normalized for each image.

and enhancement for imaging applications, see also Fig. 11 in the Appendix for an example of scanning through paper with CNNs), with the caveat that a direct quantitative comparison needs to take into account the specific combination of scatterer, optical setup (we used lower N.A. objectives than many of the reports with higher enhancement, and as expected also measured lower enhancement with the 10X objectives as compared to the 40X objectives), and the number of controlled modes. The maximum number of controlled modes in our experiments was 4096 (64×64). While the SLM has 768×1024 pixels, the number of controllable modes depends on the memory of the GPU of 12 GB which needs to accommodate the NN model and a single batch of training data, in our case typically 150 frames. This limit could however be overcome by using multiple GPUs.

To increase the enhancement and light control one could in addition modulate phase (see p. 42 of [1] for the effect on enhancement and signal-to-noise ratio) and the NN approach could be extended to any combination of stimulus-response pairs, including phase or polarization on either the detection or projection side, or both. An advantage of only using binary intensity modulation and intensity measurements [30], is that it simplifies the setup compared to approaches that also rely on phase information. Additionally, although we used monochromatic light, our approach could also be used with pulsed light [76].

For applications, the time it takes to achieve light control is critical. For the transmission matrix approach as well as the NN approach this time can be broken down into two parts, the time for acquiring the data and the time to compute the wavefront correction. The acquisition time is ultimately limited by the number of required frames. Typical numbers for the transmission matrix approach in recent reports range from 4000 [29] to 12000 [72] which is similar to the number of frames used in our experiments, which was typically 10000 or less, see Figs. 12 and 13 (see also [73] for measuring the transmission matrix using 5000 sample pairs). The time-limiting factor in our experiments was training of the NNs. For the largest data sets the time required for training was less than 35 seconds for the SLNNs and less than 50 seconds for fine-tuning the CNNs or 150 seconds for training the second CNN architecture (Fig. 11). To accelerate the process we tested training with a reduced amount of data (Fig. 13) which sped up training at the cost of lower enhancement. The shortest training time on a single GPU with the SLNN that lead to a focus with significant enhancement ($\eta > 10$ for the 10X objectives) was obtained with 5000 frames in 18 seconds. For comparison, the time required for calculating the transmission matrix varies for different techniques, from a simple Hermitian conjugation operation, to computationally more demanding approaches which require 15 seconds matrix multiplication on a GPU [72]. While some methods that optimize a single mode can be very fast (for example 33.8 ms in [11]), this still results in a comparable correction time for a full field of view (of about 5 minutes for 96×96 pixels in this example).

Further improvements of the NN approach could be achieved by optimizing the network architectures. Here, we compared two basic networks, SLNNs and CNNs. SLNNs take advantage of the linearity of scattering (as does the transmission matrix approach) and therefore can generalize from speckle patterns to arbitrary light distributions. Multi-layer NNs in contrast need to be specifically designed and trained to generate a desired type of light distribution. That CNNs are not constrained to the linearity of the underlying scattering process also might explain their worse performance in our experiments which could potentially be remedied with a larger training data set. At the same time, that multi-layer NNs are independent of assumptions about the underlying physical model (such as linearity of scattering) and can efficiently reduce the dimensionality of the images through convolutional layers as well as lower the number of parameters required for training (by 80% compared to the SLNNs in our case), will likely prove advantageous for applications for example in nonlinear situations [77].

For many applications of light control through scattering media, such as imaging, sensing or communication, it will be necessary to develop methods that can work with reflected light [51]. For example in biological microscopy, fluorescence signals can serve as feedback for scattering correction [2, 19], but they require labeling of the sample and are often dim, particularly before wavefront correction. Other schemes for light control in tissue resort to the assistance of acoustics waves [5, 78] but do not achieve diffraction limited optical resolution [19]. The most broadly applicable implementation for wavefront correction takes advantage of backscattered light as for example in optical coherence tomography or related approaches [53–59, 79]. However, ultimately the availability of weakly scattered photons is limiting the imaging depth of these methods and ways to take advantage of strongly scattered light are therefore needed. Strategies for light control using strongly scattered, reflected light have indeed been developed [62–64] for maximizing the energy delivered into the material [62, 63] or an embedded strongly scattering target [64] without, however, exerting full independent control over the transmitted modes. We here took advantage of mutual information between transmitted and reflected speckle patterns [51, 60, 61] and used NNs to show that it is indeed possible to control transmitted light with reflected light with sufficient accuracy for high-resolution focusing and scanning (Fig. 7). We achieved this by using NNs to establish an explicit functional relation between transmitted and reflected speckle patterns (Fig. 6). That such a relationship can be established (with a linear network) could not necessarily be expected based on the mutual information relationship in [51].

The limitation of the current approach for applications is that it first requires characterizing the transmission and reflection properties of the scatterer for the specific field of view, which still requires unobstructed access to the focal plane behind the scatterer. How could this limitation be overcome? One of the distinctions of neural networks is their ability to generalize. A potential avenue would therefore be to train appropriate NN models on sufficiently broad training sets and to adapt these models to the specific sample or field of view [80], e.g. using backscattered light. For example, CNNs are the building blocks for many of the more advanced network techniques that analyze novel visual scenes based on previously learned data sets [32, 33, 68], and such methods might also be harnessed for light scattering.

Independent of this, the simplicity, effectiveness and flexibility of the method presented here makes it suitable for scattering control in transmission or in reflection as well as for the further analysis of the relationship between transmission and reflection in scattering materials.

Appendix

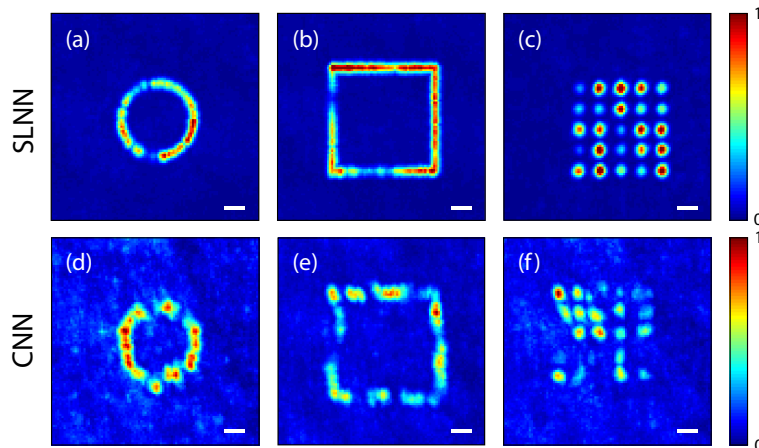


Fig. 8. Single-focus scanning allows time-averaged pattern projection. Patterns obtained when a single focus is scanned following (a/d) a circle (128/96 scanning points), (b/e) a square (256/256 scanning points), and (c/f) a grid of 5×5 points with the SLNN (first row) and the CNN (second row). Color bars: intensity (a.u.) normalized for each image. Scale bars: $2.3 \mu\text{m}$

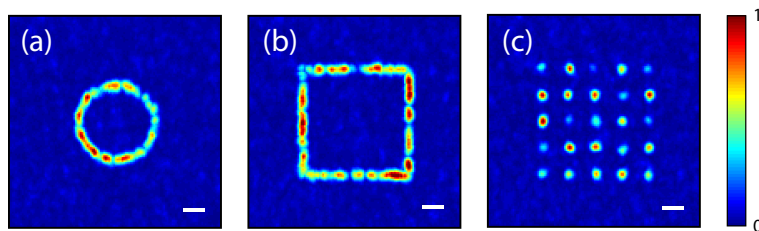


Fig. 9. Single-focus scanning allows time-averaged pattern projection through paper. Maximum intensity projections of patterns obtained when a single focus is scanned following (a) a circle (128 scanning points), (b) a square (256 scanning points), and (c) a grid of 5×5 points with the SLNN and paper as scattering material. Color bars: intensity (a.u.) normalized for each image. Scale bars: $2.7 \mu\text{m}$

Light control over different fields of view

In the main text we showed the ability of NNs to shape light through disordered media within a field of view of $20 \times 20 \mu\text{m}^2$ imaged onto 96×96 pixels on the CCD camera. However, the presented method performs similarly for other fields of view, as shown in Fig. 10 for the case of the glass diffuser, with 64×64 macropixels displayed on the DMD controlled with the SLNN. In all cases, the experimental setup is the same as the one shown in the main text and the only difference is the resolution of the images captured with the CCD camera and used to train the SLNN. Note that the illumination resolution additionally impacts the quality of light shaping through the diffuser (in agreement with previous reports, see e.g. Ref. [9]).

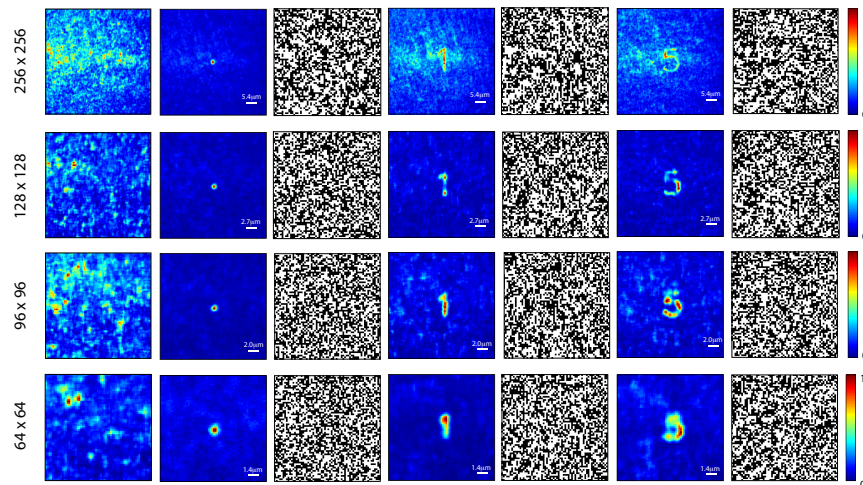


Fig. 10. Light control over different fields of view. Normalized transverse intensity distributions of an example of speckle pattern (first column), a single focus (second column), the number “1” (fourth column), and the number “5” (6th column) for different fields of view: 256×256 pixels (first row), 128×128 pixels (second row), 96×96 pixels (third row), and 64×64 pixels (last row). Columns number 3, 5, and 7 are the actual DMD patterns used to generate the light distributions from columns number 2, 4, and 6, respectively. Color bars: intensity (a.u.) normalized for each image.

Pattern generation by scanning a focus at high frequency

Since neural networks can generate single foci with high fidelity, one can take advantage of the fast operation of the DMD (22.7 kHz) to obtain any transverse intensity distribution after the scatter by scanning a single spot (or multiple spots) at high speed. This is shown in Visualization 1, Visualization 2, Visualization 3, Visualization 4, Visualization 5, Visualization 6, Visualization 7, Visualization 8, Visualization 9, Visualization 10, Visualization 11, and Visualization 12 for a single focus tracing out a circle (Visualization 1, Visualization 4, Visualization 7, and Visualization 10), a square (Visualization 2, Visualization 5, Visualization 8, Visualization 3, Visualization 6, Visualization 9, and Visualization 12) consisting of 128/96, 256/256, and 25/25 scanning positions, respectively, through a glass diffuser both for the SLNN (Visualization 1, Visualization 2, and Visualization 3) and the CNN (Visualization 4, Visualization 5, and Visualization 6); a multimode optical fiber (Visualization 7, Visualization 8, and Visualization 9), and a piece of paper (Visualization 10, Visualization 11, and Visualization 12). These sequences were recorded at a speed of 500 Hz (the fastest allowed by our CCD camera). Note that for the DMD operating at full speed, i.e. without restrictions imposed by the camera, a sequence

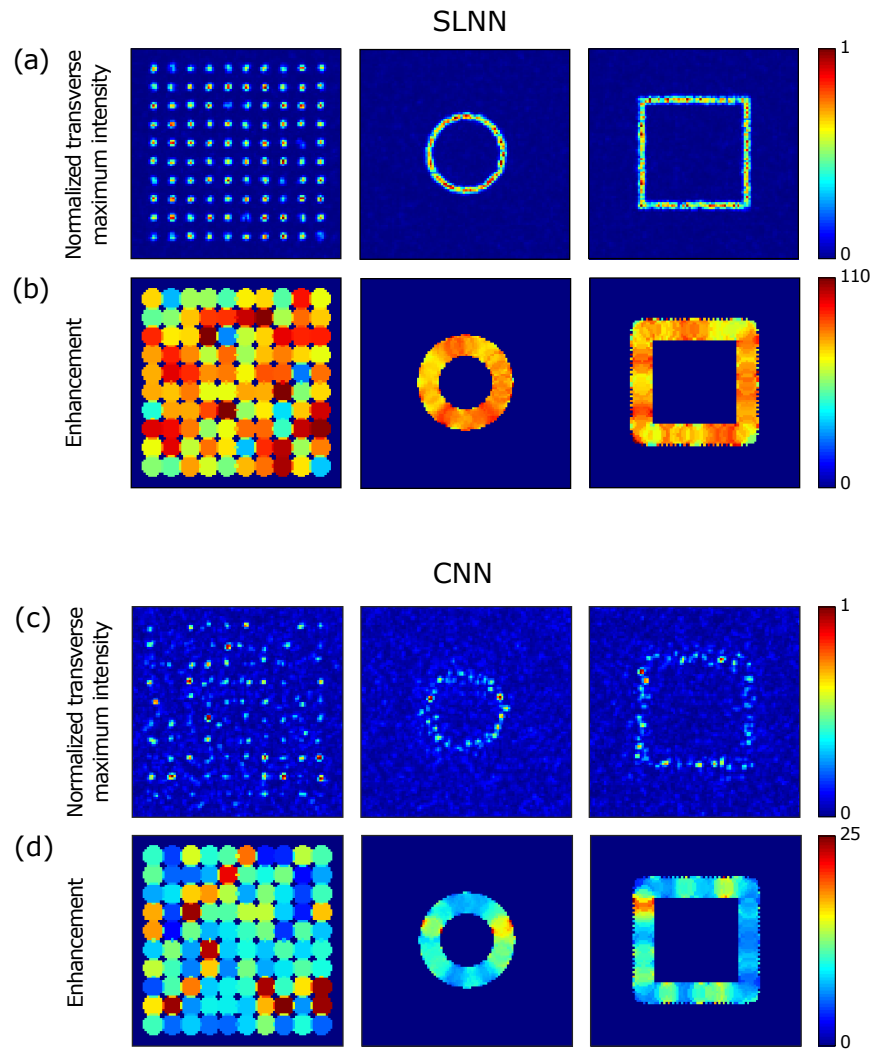


Fig. 11. Scanning and corresponding enhancement for focusing across paper with the SLNN (a, b) and the CNN (c, d) with the 40X objective. (a) shows different scan patterns across the entire field of view (grid, circle and square) for the SLNN, and (b) shows the corresponding enhancement for each point in (a). (c) Same as (a) for the CNN. (d) same as (b) for the CNN. The width and height of the full field of view is $51 \mu\text{m}$, the single pixel width and height is $0.53 \mu\text{m}$. Note the different scale bars for the enhancement for the SLNN and CNN.

composed of 96 focus positions projecting a certain pattern will result in a pattern projection frame rate higher than 200 Hz.

To quantify the focusing accuracy of the different network architectures presented in Fig. 11 we calculated the average distance between the target focus position and the measured focus position (which was measured by finding the focus centroid by fitting it with a Gaussian distributions). The values obtained for the SLNN were 0.63 pixels (standard deviation 0.20) and for the CNN 0.99 pixels, (standard deviation 0.60), with a pixel size of $0.53 \mu\text{m}$.

Neural network design and performance

We use the Keras library [81] with the TensorFlow [82] back-end for GPU-accelerated neural network training. The networks are trained to map grayscale speckle images to the corresponding binary illumination patterns with a subset of the total dataset of image pairs (8000 pairs in our case) and tested on the previously not introduced data (the remaining 2000 pairs). Once the network is trained, we input the desired PSF and the output binary map is uploaded to the DMD for light control through the diffuser or fiber.

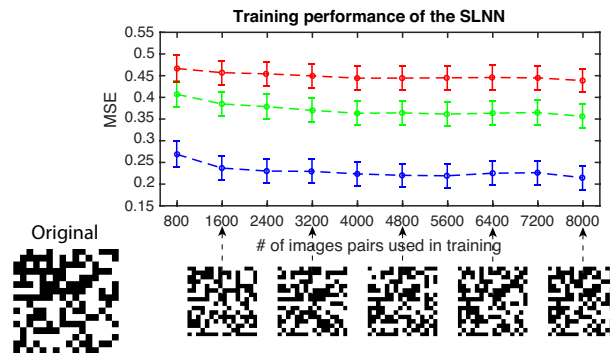


Fig. 12. Training performance of SLNN. Mean-square error (MSE) between the predicted and original illumination patterns for the single-layer neural network for different sizes of the dataset. Red-, green-, and blue-dashed curves correspond to illumination sizes of 64×64 , 32×32 , and 16×16 , respectively. Insets below show the the predicted illumination at different stages of the training for the 16×16 case.

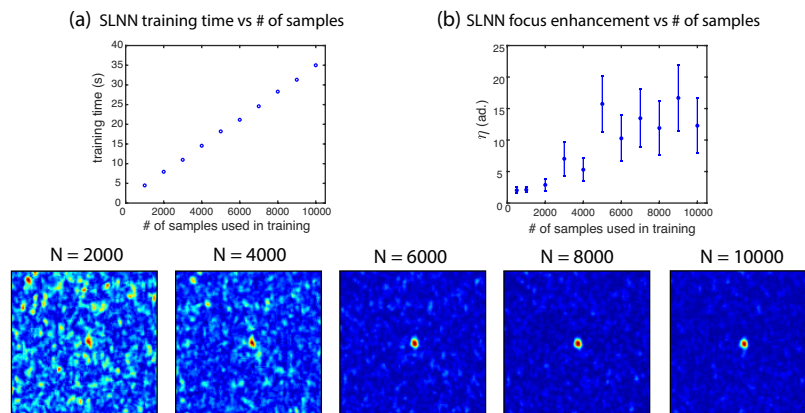


Fig. 13. Speed of SLNN training and quality of focus. (a) Training time of the SLNN and (b) enhancement η of the generated foci with paper as scattering sample for different values of the number of image pairs used during training. The enhancement is defined as $\eta \equiv I_{\text{focus}} / \langle I_{\text{speckle}} \rangle$, where I_{focus} is the intensity at the generated foci and $\langle I_{\text{speckle}} \rangle$ is the mean value of the background speckle [67]. Image in the lower part of the figure are examples of focusing with N samples used for training (with the number N of samples used for training at the top of each image).

The SLNN we used is a single-layer perceptron, which is a network consisting of one fully connected layer followed by a non-linear activation function bounding the output to the 0-1 range. In principle, it can be represented as a matrix dot product, with bias addition and a

sigmoid function applied element-wise to the resulting vector. We found that with the activation function applied per each individual element the model is prone to over-fitting and does not make good generalizations. As a solution, we replaced the nonlinear activation function with a binarization function with a threshold common for the whole predicted pattern (mean value of the prediction) which results in a more robust model with better focus enhancement and faster training. The training time depends on the number of images used (8000 in our case), the batch size (number of images taken for each iteration of the training algorithm, 150 in our case), and the number of epochs (up to 20 for the results presented here, see Figs. 12 and 12 for SLNN training performance). With these parameters the single-layer perceptron requires less than 35 seconds for training, while the predicted patterns take about 1 s to be calculated. However, we have verified that lower training times with acceptable enhancement can be obtained by reducing the number of image pairs used in the training. For an analysis on how the training times and enhancement of the focus depend with the number of images pairs used in the training, see Fig. 13.

For exploring functional relations between transmitted and reflected speckle patterns we concatenated two SLNNs similar to the one described above. The first SLNN (SLNN1) uses the transmitted speckle as input and connects it to the reflected speckle pattern (output) through a single fully-connected layer (without binarization). The second SLNN (SLNN2) connects the reflected speckle (input) to the illumination patterns (output) also through a fully connected layer including the mean threshold binarization. Both sets of speckle patterns can be generated with independent checkerboard illumination patterns. For training of these two SLNNs we used the same number of images, batch size, and number of epochs as for the SLNN discussed above, with similar performance. The desired illumination in transmission is finally fed into SLNN1, which predicts a speckle pattern as output, which in turn serves as input for SLNN2. The output of SLNN2 is a binary illumination pattern that is sent to the DMD in order to experimentally obtain the desired illumination.

In perceptron-like models a single fully-connected layer contains a large number of parameters (the product of input and output vector dimensions) which makes these models more demanding to train as the resolution of the illumination and speckle images and the memory demand increase. CNNs can efficiently reduce the number of trainable parameters and we first used a model with three convolutional layers with 48 (9×9), 24 (5×5) and 12 (3×3) filters respectively, each succeeded by rectified linear unit (ReLU) activation and (2×2) max pooling operation, followed by a fully connected layer with 0.25 dropout rate (see below and Fig. 11 for a different CNN architecture). This configuration achieves a performance similar to the SLNN in controlling a single focal spot while having 20% of the SLNN's number of parameters.

As any deeper network, the CNN requires longer training time and a more extensive dataset. A workaround is offered by the fine tuning technique: the convolutional layers are pretrained separately on a dataset of 40000 speckle images in an autoencoder. An autoencoder is a network trained to map its input to itself, however it contains a bottleneck - a lower dimensional middle layer (latent space) where a compressed representation of the data is learned. Our autoencoder has three convolutional layers as needed for the proposed CNN model and a symmetrical deconvolutional decoder. The training time largely varies with dataset size and speckle image resolution, and it is best to provide as much data as possible. Good training results were achieved after 20 minutes of training with 40000 samples sized 256×256 .

For the results presented in Fig. 11, a different CNN structure was used: two convolutional layers of 48 (7×7), and 24 (5×5) filters, both succeeded by rectified linear unit (ReLU) activation, (2×2) max pooling and batch normalization and a 0.3 dropout used in the training. The subsequent fully-connected layer used sigmoid activation. This network could be trained from scratch in 150 seconds (25 epochs of 10000 samples dataset, with batch size of 100). Additionally, the input scaling used in the illumination prediction was adapted. The CNN is trained on images produced by random illumination patterns, which contain speckles equally

distributed across the whole field of view. The image of the focus which serves as the input in the prediction stage is however significantly different: it has zeros in most parts except for a Gaussian peak. This input results in lower activation values in the intermediate layers compared to random speckle input images, and as a result the output illumination is not binary, but rather continuous-valued, because the sigmoid activation units of the output layer do not receive input of sufficient magnitude to saturate to either 0 or 1 values. To overcome this, the focus input is scaled by empirically adjusting the value of 10^5 to force the CNN to output binary patterns.

Training performance of SLNN

Although neural networks are capable of predicting the illumination patterns necessary for light control through scattering media, there are multiple variables affecting the efficiency of the training. Here, we discuss the impact of the size of the dataset used for training. In Fig. 12 we plot the mean-square error (MSE) between the predicted illuminations after training the single-layer neural network and a set of 100 original illuminations that were not included in the training. The analysis is performed for different sizes of the dataset (ranging from 800 pairs to 8000 pairs) and different illumination sizes: 64×64 (red), 32×32 (green), and 16×16 (blue). As an illustration of the training performance, we have included an example of original illumination and corresponding predictions obtained when 1600, 3200, 4800, 6200, and 8000 pairs are used (16×16 case). As expected, the lower the size of the illumination pattern, the higher the MSE.

In Fig. 13 we further explore the training performance of the SLNN as the size of the dataset used in the training varies, in this case in terms of Fig. 13(a) training time, and Fig. 13(b) enhancement of the generated foci. While the training time shows a linearly increasing tendency, the enhancement tends to saturate beyond a certain number of samples (typically for $N > 5000$).

Funding

Max Planck Society; Center of Advanced European Studies and Research (caesar); Nvidia (Titan Xp GPU).

Acknowledgements

We thank Andres Flores for Python support and Bernd Scheiding for electronics support.

References

1. S. Rotter and S. Gigan, "Light fields in complex media: Mesoscopic scattering meets wave control," *Rev. Mod. Phys.* **89**, 015005 (2017).
2. N. Ji, "Adaptive optical fluorescence microscopy," *Nat. Methods* **14**, 374–380 (2017).
3. T. Cizmar and K. Dholakia, "Exploiting multimode waveguides for pure fibre-based imaging," *Nat. Commun.* **3**, 1027 (2012).
4. T. Cizmar, M. Mazilu, and K. Dholakia, "In situ wavefront correction and its application to micromanipulation," *Nat. Photonics* **3**, 388–394 (2010).
5. H. W. Ruan, J. Brake, J. E. Robinson, Y. Liu, M. Jang, C. Xiao, C. Y. Zhou, Y. Gradinaru, and C. H. Yang, "Deep tissue optical focusing and optogenetic modulation with time-reversed ultrasonically encoded light," *Sci. Adv.* **3**, eaao5520 (2017).
6. 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).
7. A. Forbes, A. Dudley, and M. McLaren, "Creation and detection of optical modes with spatial light modulators," *Adv. Opt. Photon.* **8**, 200–227 (2016).
8. S. Turtaev, I. T. Leite, K. J. Mitchell, M. J. Padgett, D. B. Phillips, and T. Cizmar, "Comparison of nematic liquid-crystal and dmd based spatial light modulation in complex photonics," *Opt. Express* **25**(24), 29874–29884 (2017).
9. I. M. Vellekoop and A. P. Mosk, "Focusing coherent light through opaque strongly scattering media," *Opt. Lett.* **32**, 2309–2311 (2007).
10. S. M. Popoff, G. Lerosey, R. Carminati, M. Fink, A. C. Boccarda, and S. Gigan, "Measuring the transmission matrix in optics: An approach to the study and control of light propagation in disordered media," *Phys. Rev. Lett.* **104**, 100601 (2010).

11. D. B. Conkey, A. M. Caravaca-Aguirre, and R. Piestun, "High-speed scattering medium characterization with application to focusing light through turbid media," *Opt. Express* **20**, 1733–1740 (2012).
12. R. D. Leonardo and S. Bianchi, "Hologram transmission through multi-mode optical fibers," *Opt. Express* **19**, 247–254 (2011).
13. 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. Express* **39**, 1291–1294 (2014).
14. H. Ruan, J. Brake, J. E. Robinson, Y. Liu, M. Jang, C. Xiao, C. Zhou, V. Gradinaru, and C. Yang, "Deep tissue optical focusing and optogenetic modulation with time-reversed ultrasonically encoded light," *Sci. Adv.* **3**, eaao5520 (2017).
15. J. Yoon, M. Lee, K. Lee, N. Kim, J. M. Kim, J. Park, H. Yu, C. Choi, W. D. Heo, and Y. Park, "Optogenetic control of cell signaling pathway through scattering skull using wavefront shaping," *Sci. Reports* **3**, 13289 (2015).
16. D. Wang, E. Zhou, J. Brake, H. Ruan, M. Jang, and C. Yang, "Focusing through dynamic tissue with millisecond digital optical phase conjugation," *Optica* **2**, 728–735 (2015).
17. Y. Liu, P. Lai, C. Ma, X. Xu, and A. A. Grabar, and L. V. Wang, "Optical focusing deep inside dynamic scattering media with near infrared time-reversed ultrasonically encoded (true) light," *Nat. Commun.* **6**, 5904 (2015).
18. I. M. Vellekoop, "Feedback-based wavefront shaping," *Opt. Express* **23**, 12189–12206 (2015).
19. R. Horstmeyer, H. Ruan, and C. Yang, "Guidestar-assisted wavefront-shaping methods for focusing light into biological tissue," *Nat. Photonics* **9**, 563–571 (2015).
20. A. P. Mosk, A. Lagendijk, G. Leroosey, and M. Fink, "Controlling waves in space and time for imaging and focusing in complex media," *Nat. Photonics* **6**, 283–292 (2012).
21. J. Tanga, R. N. Germain, , and M. Cui, "Superpenetration optical microscopy by iterative multiphoton adaptive compensation technique," *Proc. Natl. Acad. Sci.* **129**, 8434–8439 (2012).
22. B. Blochet, L. Bourdieu, and S. Gigan, "Focusing light through dynamical samples using fast continuous wavefront optimization," *Opt. Lett.* **42**, 4994–4997 (2017).
23. M. Cui and C. Yang, "Implementation of a digital optical phase conjugation system and its application to study the robustness of turbidity suppression by phase conjugation," *Opt. Express* **18**, 3444–3455 (2010).
24. C.-L. Hsieh, Y. Pu, R. Grange, G. Laporte, and D. Psaltis, "Imaging through turbid layers by scanning the phase conjugated second harmonic radiation from a nanoparticle," *Opt. Express* **18**, 20723–20731 (2010).
25. Y. M. Wang, B. Judkewitz, C. A. DiMarzio, and C. Yang, "Deep-tissue focal fluorescence imaging with digitally time-reversed ultrasound-encoded light," *Nat. Commun.* **3**, 928 (2012).
26. T. R. Hillman, T. Yamauchi, W. Choi, R. R. Dasari, M. S. Feld, Y. Park, and Z. Yaqoob, "Digital optical phase conjugation for delivering two-dimensional images through turbid media," *Sci. Reports* **3**, 1909 (2013).
27. Y. Liu, C. Ma, Y. Shen, J. Shi, and L. V. Wang, "Focusing light inside dynamic scattering media with millisecond digital optical phase conjugation," *Optica* **4**, 280–288 (2017).
28. S. Popoff, G. Leroosey, M. Fink, A. C. Boccara, and S. Gigan, "Image transmission through an opaque material," *Nat. Commun.* **1**, 81 (2010).
29. A. Boniface, M. Mounaix, B. Blochet, R. Piestun, and S. Gigan, "Transmission-matrix-based point-spread-function engineering through a complex medium," *Optica* **4**, 54–59 (2017).
30. A. Dreameau, A. Liutkus, D. Martina, O. Katz, C. Schülke, F. Krzakala, S. Gigan, and L. Daudet, "Reference-less measurement of the transmission matrix of a highly scattering material using a dmd and phase retrieval techniques," *Opt. Express* **23**, 11898–11911 (2015).
31. C. A. Metzler, M. K. Sharma, S. Nagesh, R. G. Baraniuk, O. Cossairt, and A. Veeraraghavan, "Coherent inverse scattering via transmission matrices: Efficient phase retrieval algorithms and a public dataset," in *2017 IEEE International Conference on Computational Photography (ICCP)* (2017), pp. 1–16.
32. M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science* **349**, 255–260 (2015).
33. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
34. L. Waller and L. Tian, "Machine learning for 3d microscopy," *Nature* **523**, 416–417 (2015).
35. C. Deans, L. D. Griffin, L. Marmugi, , and F. Renzoni, "Machine learning based localization and classification with atomic magnetometers," *Phys. Rev. Lett.* **120**, 033204 (2018).
36. 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**, 517–522 (2015).
37. A. Sinha, J. L. S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica* **4**, 1117–1125 (2017).
38. Y. Rivenson, Z. Göröcs, H. Günaydin, Y. Zhang, H. Wang, and A. Ozcan, "Deep learning microscopy," *Optica* **4**, 1437–1443 (2017).
39. E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Deep-storm: Super resolution single molecule microscopy by deep learning," *Optica* **5**, 458–464 (2018).
40. N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, "Learning to see through multimode fibers," *Optica* **5**, 960–966 (2018).
41. D. B. Conkey, A. N. Brown, A. M. Caravaca-Aguirre, and R. Piestun, "Genetic algorithm optimization for focusing through turbid media in noisy environments," *J. Opt.* **20**, 4840–4849 (2012).
42. X. Zhang and P. Kner, "Binary wavefront optimization using a genetic algorithm," *J. Opt.* **17**, 125704 (2014).
43. K. F. Tehrani, J. Xu, Y. Zhang, P. Shen, and P. Kner, "Adaptive optics stochastic optical reconstruction microscopy

- (ao-storm) using a genetic algorithm,” *Opt. Express* **23**, 13677–13692 (2015).
44. B. Zhang, Z. Zhang, Q. Feng, Z. Liu, C. Lin, and Y. Ding, “Focusing light through strongly scattering media using genetic algorithm with sbr discriminant,” *J. Opt.* **20**, 025601 (2017).
 45. T. Ando, R. Horisaki, and J. Tanida, “Speckle learning-based object recognition through scattering media,” *Opt. Express* **23**, 33902–33910 (2015).
 46. R. Horisaki, R. Takagi, and J. Tanida, “Learning-based imaging through scattering media,” *Opt. Express* **24**, 13738–13743 (2016).
 47. R. Horisaki, R. Takagi, and J. Tanida, “Learning-based focusing through scattering media,” *Appl. Opt.* **56**, 4358–4362 (2017).
 48. S. Li, M. Deng, J. Lee, A. Sinha, and G. Barbastathis, “Imaging through glass diffusers using densely connected convolutional networks,” *Optica* **5**, 803–813 (2018).
 49. M. Lyu, H. Wang, G. Li, and G. Situ, “Exploit imaging through opaque wall via deep learning,” arXiv preprint [arXiv:1711.06810v1](https://arxiv.org/abs/1711.06810v1) (2017).
 50. J. R. P. Angel, P. Wizinowich, M. Lloyd-Hart, and D. Sandler, “Adaptive optics for array telescopes using neural-network techniques,” *Nature* **348**, 221–224 (1990).
 51. N. Fayard, A. Goetschy, R. Pierrat, and R. Carminati, “Mutual information between reflected and transmitted speckle images,” *Phys. Rev. Lett.* **120**, 073901 (2018).
 52. Z. Liu, L. D. Lavis, and E. Betzig, “Imaging live-cell dynamics and structure at the single-molecule level,” *Mol. Cell* **58**, 644–659 (2015).
 53. M. Feierabend, M. Ruckel, and W. Denk, “Coherence-gated wave-front sensing in strongly scattering samples,” *Opt. Lett.* **29**, 2255–2257 (2004).
 54. R. Fiolka, K. Si, and M. Cui, “Parallel wavefront measurements in ultrasound pulse guided digital phase conjugation,” *Opt. Express* **20**, 24827–24834 (2012).
 55. W. Drexler and J. G. Fujimoto, *Optical Coherence Tomography* (Springer International Publishing, 2015).
 56. S. Kang, S. Jeong, W. Choi, H. Ko, T. D. Yang, J. H. Joo, J. S. Lee, Y. S. Lim, Q. H. Park, and W. Choi, “Imaging deep within a scattering medium using collective accumulation of single-scattered waves,” *Nat. Photonics* **9**, 253–258 (2015).
 57. A. Badon, D. Y. Li, G. Lerosey, A. C. Boccarda, M. Fink, and A. Aubry, “Smart optical coherence tomography for ultra-deep imaging through highly scattering media,” *Sci. Adv.* **2**, e1600370 (2016).
 58. S. Kang, P. Kang, S. Jeong, Y. Kwon, T. D. Yang, J. H. Hong, M. Kim, K. D. Song, J. H. Park, J. H. Lee, M. J. Kim, K. H. Kim, and W. Choi, “High-resolution adaptive optical imaging within thick scattering media using closed-loop accumulation of single scattering,” *Nat. Commun.* **8**, 2157 (2017).
 59. M. Kadobianskyi, I. N. Papadopoulos, T. Chaigne, R. Horstmeyer, and B. Judkewitz, “Scattering correlations of time-gated light,” *Optica* **5**, 389–394 (2018).
 60. N. Fayard, A. Caze, R. Pierrat, and R. Carminati, “Intensity correlations between reflected and transmitted speckle patterns,” *Phys. Rev. A* **92**, 033827 (2015).
 61. I. Starshynov, A. M. Paniagua-Diaz, N. Fayard, A. Goetschy, R. Pierrat, R. Carminati, and J. Bertolotti, “Non-gaussian correlations between reflected and transmitted intensity patterns emerging from opaque disordered media,” *Phys. Rev. X* **8**, 021041 (2018).
 62. Y. Choi, T. R. Hillman, W. Choi, N. Lue, R. R. Dasari, P. T. C. So, W. Choi, and Z. Yaqoob, “Measurement of the time-resolved reflection matrix for enhancing light energy delivery into a scattering medium,” *Phys. Rev. Lett.* **111**, 243901 (2013).
 63. H. Yu, J. H. Park, and Y. Park, “Measuring large optical reflection matrices of turbid media,” *Opt. Commun.* **352**, 33–38 (2015).
 64. S. Jeong, Y.-R. Lee, W. Choi, S. Kang, J. H. Hong, J.-S. Park, Y.-S. Lim, H.-G. Park, and W. Choi, “Focusing of light energy inside a scattering medium by controlling the time-gated multiple light scattering,” *Nat. Photonics* **12**, 277–283 (2018).
 65. C. Jin, R. R. Nadakuditi, E. Michielssen, and S. C. Rand, “Iterative, backscatter-analysis algorithms for increasing transmission and focusing light through highly scattering random media,” *J. Opt. Soc. Am. A* **30**, 1592–1602 (2013).
 66. C. Jin, R. R. Nadakuditi, E. Michielssen, and S. C. Rand, “Backscatter analysis based algorithms for increasing transmission through highly scattering random media using phase-only-modulated wavefronts,” *J. Opt. Soc. Am. A* **31**, 1788–1800 (2014).
 67. D. Akbulut, T. J. Huisman, E. G. van Putten, W. L. Vos, and A. P. Mosk, “Focusing light through random photonic media by binary amplitude modulation,” *Opt. Express* **19**, 4017–4029 (2011).
 68. J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural Networks* **61**, 85–117 (2014).
 69. K. Henderson, C. Ryu, C. MacCormick, and M. G. Boshier, “Experimental demonstration of painting arbitrary and dynamic potentials for Bose–Einstein condensates,” *New J. Phys.* **11**, 043030 (2009).
 70. J. P. Rickgauer and D. W. Tank, “Two-photon excitation of channelrhodopsin-2 at saturation,” *PNAS* **106**, 15025–15030 (2009).
 71. A. M. Caravaca-Aguirre and R. Piestun, “Single multimode fiber endoscope,” *Opt. Express* **25**, 1656–1665 (2017).
 72. S. Ohayon, A. M. Caravaca-Aguirre, R. Piestun, and J. J. DiCarlo, “Deep brain fluorescence imaging with minimally invasive ultra-thin optical fibers,” arxiv preprint [arXiv:1703.07633](https://arxiv.org/abs/1703.07633) (2017).
 73. T. Zhao, L. Deng, W. Wang, D. S. Elson, and L. Su, “Bayes’ theorem-based binary algorithm for fast reference-less

- calibration of a multimode fiber,” *Opt. Express* **26**, 20368–20378 (2018).
74. A. A. Goshtasby, *Image Registration* (Springer, 2012).
 75. D. P. C. M. Damien Loterie, Salma Farahi, “Complex pattern projection through a multimode fiber,” *Proc. SPIE* 9335, **93350I** (2015).
 76. B. Sun, P. S. Salter, C. Roeder, A. Jesacher, J. Strauss, J. Heberle, M. Schmidt, and M. J. Booth, “Four-dimensional light shaping: manipulating ultrafast spatiotemporal foci in space and time,” *Light. Sci. Appl.* **7**, 17117 (2018).
 77. H. Frostig, E. Small, A. Daniel, P. Oulevey, S. Derevyanko and Y. Silberberg Mosk, “Focusing light by wavefront shaping through disorder and nonlinearity,” *Optica* **4**, 1073–1079 (2017).
 78. X. A. Xu, H. L. Liu, and L. V. Wang, “Time-reversed ultrasonically encoded optical focusing into scattering media,” *Nat. Photonics* **5**, 154–157 (2011).
 79. R. Kuschmierz, E. Scharf, N. Koukourakis, and J. W. Czarske, “Self-calibration of lensless holographic endoscope using programmable guide stars,” *Opt. Lett.* **43**, 2997–3000 (2018).
 80. L. T. Yunzhe Li, Yujia Xue, “Deep speckle correlation: a deep learning approach towards scalable imaging through scattering media,” arxiv preprint [arXiv:1806.04139v1](https://arxiv.org/abs/1806.04139v1) (2018).
 81. Chollet, François *et al*, “Keras” GitHub <https://github.com/keras-team/keras>, (2015).
 82. Martín Abadi “TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems,” <http://tensorflow.org/>, (2015).