

Neural-network-based multimode fiber imaging and position sensing under thermal perturbations

Tristan Kremp^a, Nicholas Bagley^b, Erin S. Lamb^a, Paul S. Westbrook^a, and David J. DiGiovanni^a

^aOFS Laboratories, Somerset, NJ 08873, USA

^bDepartment of Mathematics, Southern Methodist University, Dallas, TX 75275, USA

ABSTRACT

Multimode fibers (MMFs) have a very large number of propagating modes per unit area and therefore allow for imaging with a very large number of pixels relative to their diameter. This makes MMFs perfect candidates for ultrathin endoscopes in applications such as deep brain imaging. However, the accuracy of the input-output relation that is needed, e.g., for distal spot scanning without moving parts, requires a new calibration after the fiber position or temperature has been significantly altered.

While neural networks have been used before to attempt to solve these challenges, we present an MMF-based imaging method that tolerates and classifies different fiber positions, using two single-layer fully-connected neural networks that only require the optical intensity without measuring the optical phase. One network learns the nonlinear relation between the input and output intensities and allows for image reconstruction in the presence of position changes, while the other network classifies that position change for different images. We show that our method is superior to memory-effect-based position sensing, both for small position changes where the relation between position change and output specklegram rotation angle is linear, as well as for larger position changes where this linearity and uniqueness break down. We also show that the position classification results are robust to temperature and polarization perturbations, and that our position classifier is able to effectively generalize. Likewise, we show that our imaging network also is robust to 30°C perturbations in temperature and 10° in polarization.

Keywords: Multimode fiber imaging, optical fiber sensing, memory effect, artificial intelligence, neural networks

1. INTRODUCTION

Multimode fibers (MMFs) have been used since the 1970s in applications such as telecommunications, sensing, and high power lasers. Due to their larger cores, MMFs enable reduced nonlinear effects in comparison to single-mode fibers and support the propagation of a large number of modes. The interference of these modes at the fiber output gives rise to intensity speckle patterns that may look random, but in fact represent a wealth of deterministic information about the light propagation and mode-coupling dynamics throughout the fiber. Fiber specklegrams, therefore, have been utilized for numerous sensing applications by statistically relating changes in the speckle pattern to changes in the parameter under study¹⁻⁴.

Additionally, various reports have characterized MMFs through complete measurements of the fiber's transmission matrix^{5,6}, which can then be used to control the electric field at the distal end of the fiber for applications such as medical endoscopy⁷⁻¹⁰. However, the transmission matrix measurement depends critically on the bend configuration of the fiber, necessitating a method of continuously updating it during any real-world application involving movement of the fiber^{11,12}.

In addition to MMF fiber imaging applications, many fiber-based sensors of bend, position, twist and displacement have been demonstrated. In the most general case, it has been shown that the entire three-dimensional shape of a fiber can be reconstructed using twisted multicore fibers and optical frequency domain reflectometry

E-mail: tkremp@ofsoptics.com

(OFDR)^{13–15}. Such fiber shape reconstruction can be used in sensors and encoders for three-dimensional motion control applications in fields ranging from robotics to medical procedures. Moreover, investigation of the changes to specklegrams induced by alteration of the bend state of the fiber have provided evidence of the speckle memory effect, relating the new bend state of the fiber to translational and rotational shifts present in the resulting specklegrams^{16,17}. Such memory-effect correlations have been used to perform imaging at the distal end of an MMF^{18,19}.

As an alternative to these approaches, artificial neural networks (ANNs) have been in use for decades. Their structure of connected layers of nodes is based on the neural pathways in the human brain^{20,21} and is utilized for the “learning” of nonlinear mappings. Numerous demonstrations^{22–34} have shown that various ANNs are well-suited to learning connections between the input and output of optical fibers, in particular MMFs. While transmission matrix methods rely on a linear relation between input and output and therefore require the knowledge of phase and intensity (or real and imaginary part of the electric field) at both ends, ANNs can learn nonlinear mappings between the input and output of the MMF, such as the dependency of the output intensity on the input phase²⁷ or on the input intensity^{24,32,33}, negating the need for a complex interferometric measurement of the electric field. Additionally, compared to the transmission matrix method, ANNs exhibit a higher degree of robustness to fiber shape and temperature changes if they are trained on data taken in varying configurations³⁵. Recent work has also shown success of ANNs to perform imaging in the case of images with multiple spectral channels, and images with multiple linear polarization channels^{36,37}.

Image classification at the distal end of MMFs^{25,38} or through scattering plates³⁹ is a popular application of machine learning systems. Separate studies have shown that ANNs are capable of detecting different positions and bend states at which a given specklegram is launched, thereby allowing for specklegram-based sensing^{40,41}, which can be extended to higher-dimensional displacement sensing⁴².

While the focus of most of these ANN studies has been on either imaging *or* position sensing, many applications would benefit from simultaneously performing *both* functions. Thus, it would be of interest to perform fiber shape sensing and imaging through the same fiber with only a single dataset. While some studies have previously considered image reconstruction in the presence of different fiber shapes³⁵, there has been comparatively little consideration of a fiber system in which both the output image *and* a given fiber configuration may be reconstructed in situations where another perturbation such as a temperature and/or polarization variation is imposed on the fiber. Our consideration of these additional perturbations is a step towards applying ANN-based fiber systems in real-world conditions where environmental changes during routine operation are common.

In this context, we recently showed⁴³ that by training an ANN through different launched images while the fiber moves through a prescribed set of positions, both the image and the fiber position can be simultaneously recovered even in the presence of a fixed and temporally varying thermal perturbation of 30°C on the fiber during the training. Furthermore, we analyzed the dependence of the specklegram perturbations on fiber position and temperature and showed that the accuracy of the ANN is maintained over the 48-hour time period used to test the long-term accuracy.

In the present work, we first compare the position-sensing performance of our ANN approach to that of a classical memory-effect-based MMF position sensor. We show that the ANN approach is substantially more accurate and robust not only for large position changes at which the memory effect becomes nonlinear or even ambiguous (non-injective mapping), but even for small position increments where the memory-effect-based specklegram rotation is an almost perfectly linear function of the fiber position change. Finally, we investigate how our previous results⁴³ generalize to 10° and 20° perturbations of the input polarization, as well as to a change of the type of input dataset (MNIST to Fashion-MNIST) after the training.

2. EXPERIMENTAL SETUP

The experimental set-up is depicted in Fig. 1(a). A continuous-wave (CW) laser at 532 nm (CrystalLaser CL532-100-S) is used to illuminate a phase-only spatial light modulator (SLM, Meadowlark Optics P1920) that displays a grayscale image from the MNIST database of handwritten digits. While the SLM modulates the digits in a single linear polarization, a half-wave plate prior to the MMF allows for an arbitrary linear polarization to be launched into the fiber. The input facet on the proximal end of the graded-index MMF (50 cm of OFS LaserWave

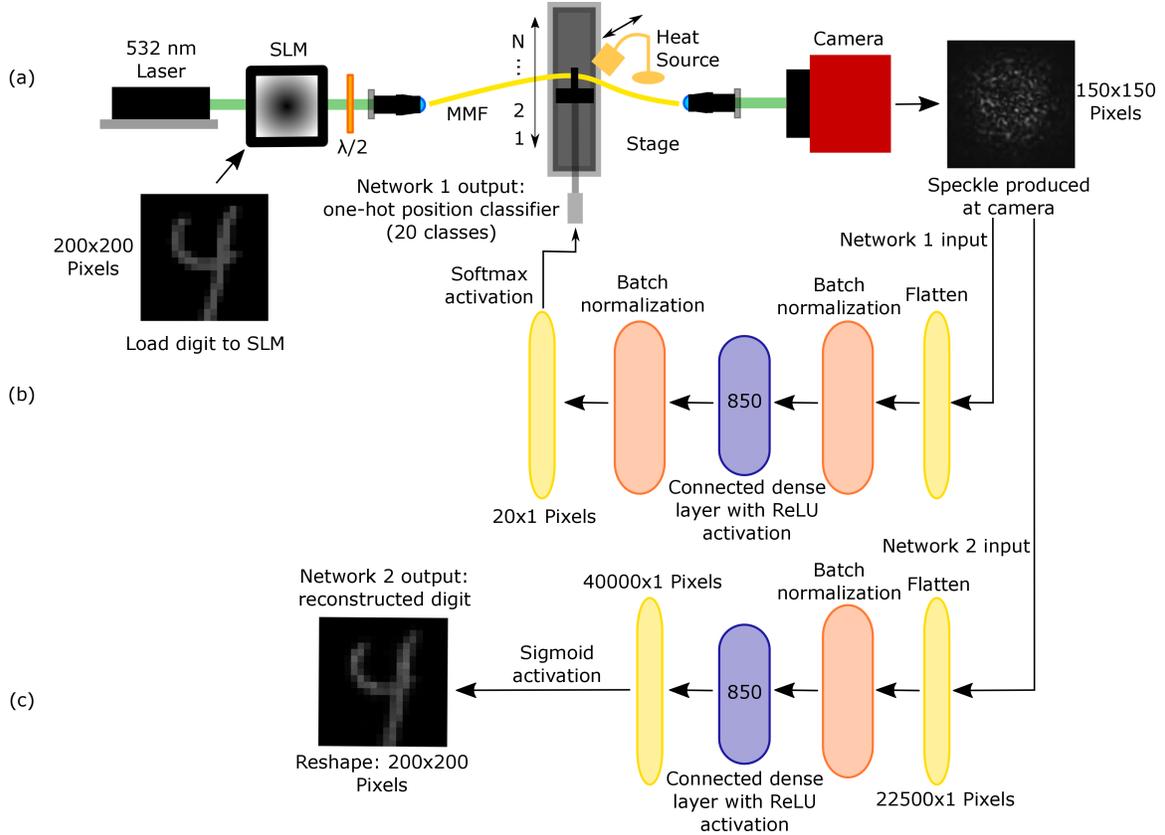


Figure 1. (a) Experimental schematic. (b) Position-sensing neural network architecture. (c) Imaging neural network architecture.

Wideband OM5) is placed in the image plane of the SLM using a $20\times$ microscope objective. The center of the length of the fiber is taped to a motorized linear translation stage (Newport UTM150CC1HL) as indicated in Fig. 1(a). A heat source (halogen lamp) is placed above the portion of the fiber that is affixed to the translation stage. The output speckle patterns from the MMF are imaged to a complementary metal oxide semiconductor (CMOS) camera (Photonfocus MV1-D2048x1088-96-G2-10) that represents the measured intensity at each pixel by an 8-bit unsigned integer, thus allowing relative intensity values between 0 and $2^8 - 1 = 255$.

The linear translation stage is moved in $100\ \mu\text{m}$ increments (Fig. 3) or $50\ \mu\text{m}$ increments (all other figures) during position-sensing experiments. The output speckle patterns from the first 10 MNIST digits are recorded for the first stage position, then for the second stage position, and so on. Then, the stage is returned to the starting position, and the speckle patterns for the next 10 MNIST digits are recorded for each position, and so forth until all of the desired MNIST digits have been launched for all stage positions. This intermixing of the data acquisition over position is done to ensure that the network is indeed classifying based on position changes, rather than simply identifying speckle patterns based on drifts in the experiment over time.

Temperature gradients are induced in the fiber through the heat source centered on the portion of the fiber attached to the stage. The heat source (lamp) is positioned in near-contact with the fiber and perturbs the air temperature over a few-centimeter length of the fiber. This perturbation ranges from room temperature (lamp off, ambient temperature around 25°C with no gradient) to a high-temperature perturbation of approximately 55°C with the lamp on, inducing a spatial gradient along the fiber as well as potentially a slow temporal drift of the temperature. In addition to the ambient room temperature data and this data with a static heat source, one dataset is also acquired with the heat source moving randomly closer and farther away at a time scale of seconds to tens of seconds to change the induced gradient. The specklegram data are recorded every 25 ms, so the temperature can safely be assumed to be stable for each set of 10 launched MNIST digits.

3. ANALYSIS OF SPECKLEGRAM SENSITIVITY AND MEMORY-EFFECT

To confirm that the changes we have made to the fiber position and temperature are significant, we measure⁴³ the output specklegrams and compute their correlation as these variables are changed. Due to the circular geometry of the MMF, there is a roughly circular region of interest (ROI) that contains the actual specklegram for all launched digits; see the example on the right of Fig. 1(a). To minimize computation time, memory, and the impact of the dark corners, each specklegram is cropped to this common ROI, which contains about 10,000 pixels. We quantify the correlation between any two specklegrams using Pearson’s r-coefficient⁴³, which ranges from 1 (identical images) to -1 (perfectly anti-correlated images). After averaging at each position over 500 MNIST digits, the correlation is shown in Fig. 2(a-c) for different heat source settings. In a *static* heat source setting (lamp on or off), Figs. 2(a) and 2(b) show a high correlation > 0.9 between specklegrams across all positions, due to the relatively small range of the position perturbation ($20 \times 50 \mu\text{m} = 1\text{mm}$). As a consequence, most specklegrams taken at different positions are almost indistinguishable. This is similar to the high correlations at a constant fiber position for different launched MNIST digits as explained, e.g., in²⁴ by the dominance of the DC component of the light from the SLM. The fall-off in correlation with increasing difference in position is similar to that seen in previous work¹ using specklegram correlations for position sensing without the use of ANNs.

In⁴⁴, it was found that an 8°C change in temperature significantly decorrelates the speckle patterns generated through 1 m of MMF at a wavelength of 1550 nm. Since we measured the above-mentioned approximate 30°C temperature change over a length of a few centimeters with the heat source on, we are expecting that the correlation of specklegrams at different heat source settings is lower than in Fig. 2(a,b) where only specklegrams at identical heat source settings were compared. Indeed, correlations between specklegrams at different heat source settings (lamp on or off) are much lower, with an average (across all 500 MNIST digits) of only about 0.34, see Fig. 2(c). In this case, the impact of the position changes on the correlation is even less pronounced (in comparison to the same-setting experiments from Figs. 2(a,b)), with average correlations varying only from 0.332 to 0.351. As demonstrated below, at such low correlation values, it is crucial that our imaging network can generalize across these conditions to produce robust and accurate image reconstructions and position sensing.

For large changes of the fiber position, the resulting perturbations of the specklegrams can in general become very complex. For small changes, however, it makes sense to investigate the dominant geometrical causes for the reduction of the correlation of the specklegrams. We therefore fit the specklegram recorded at the fixed position 1 to the specklegrams at all the other positions 2 to 20 to determine any translational and rotational changes between them, and we then compute the averages and standard deviations of these shifts and angles for each position and heat source setting. When translating¹⁶ or rotating¹⁷ the input light field, a resulting translation or rotation of the output field has been referred to as a shift or rotational memory effect^{18,19}. In contrast, in our experiment, the input field is constant for any given MNIST launch, and instead the fiber shape (position) is changed. The translational shift between specklegrams is shown in Fig. 2(d-f), separately for the horizontal x -direction and vertical y -direction. This shift is independent of the position difference, but differs strongly between the different heat source settings. Vice versa, the average rotation between specklegrams shown in Fig. 2(g-i) is an almost perfectly linear function of the small position changes, but almost independent of the heat source setting at which these two specklegrams were recorded. In particular, this shows that the standard deviation of the rotation (being about the same in Fig. 2(g-i)) is basically independent of the standard deviation of the shift (vastly different in Fig. 2(d/e) vs. 2(f)). Thus, this experiment is operating in the speckle memory regime with respect to both a fiber-position-based memory effect and a temperature-based memory effect. Even more importantly, the average rotational increments that correspond to neighboring positions are smaller than the standard deviation, i.e., smaller than the angle difference when using different launched MNIST digits. This is the case not only when comparing different heat source settings (right column of Fig. 2), but also for an unchanged heat source setting (left and middle column of Fig. 2).

Hence, even if the position range is so small that the specklegram rotation is approximately linear, a pure measurement of the shift and rotation is not sufficient to reliably classify the position for all launched MNIST digits. This prediction is quantitatively confirmed by Figs. 2(j/k/l) that show histograms of the position-sensing error when using the linear slope from Figs. 2(g/h/i) to convert measured rotation to estimated position as a classical cross-correlation-based memory-effect method, resulting in average absolute errors of more than one position. In contrast, the ANN learns these nontrivial perturbations of the specklegrams to perform highly

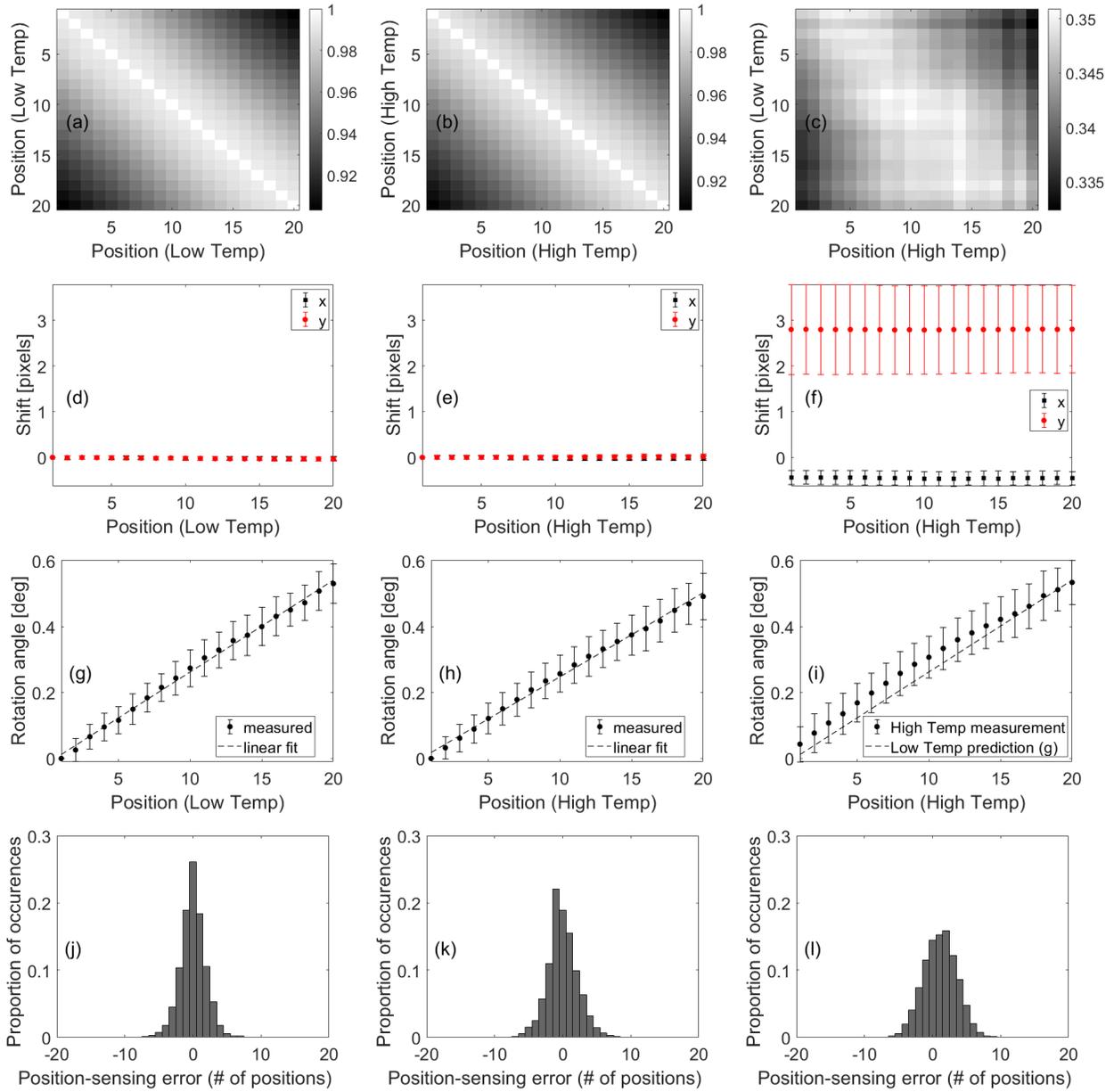


Figure 2. Analysis of output specklegrams from 500 launched MNIST digits recorded at each of 20 measured positions spaced $50\mu\text{m}$ apart. Top row (a,b,c): Correlations between specklegrams recorded at different positions. Second row (d,e,f): Detected shift of the specklegrams relative to the specklegram at position 1. Third row (g,h,i): Detected rotation angle of the specklegrams relative to the specklegram at position 1. In the second and third row, the error bars indicate standard deviations across the 500 specklegrams at each position. Bottom row (j,k,l): Position-sensing error histograms when using rotation angle (memory effect). Left column (a,d,g,j): Ambient (low) temperature, i.e., heat source off. Middle column (b,e,h,k): Heat source on. Right column (c,f,i,l): Data taken with heat source on versus data taken with heat source off (only position 1 in (f,i)). E.g., (f) shows the shift between the specklegram pairs with heat source on vs. the single specklegram at position 1 with heat source off, and this shift is removed before computing the corresponding rotation in (i). Similarly, (l) shows the position-sensing error from specklegrams with heat source on when using the rotation-vs-position slope that was measured (trained) with heat source off (dashed line in (g)). In (j,k,l), the average position-sensing errors are 1.39, 1.66 and 2.14 positions, corresponding to relative errors of 0.0697, 0.0829 and 0.107 compared to the total range of 1mm.

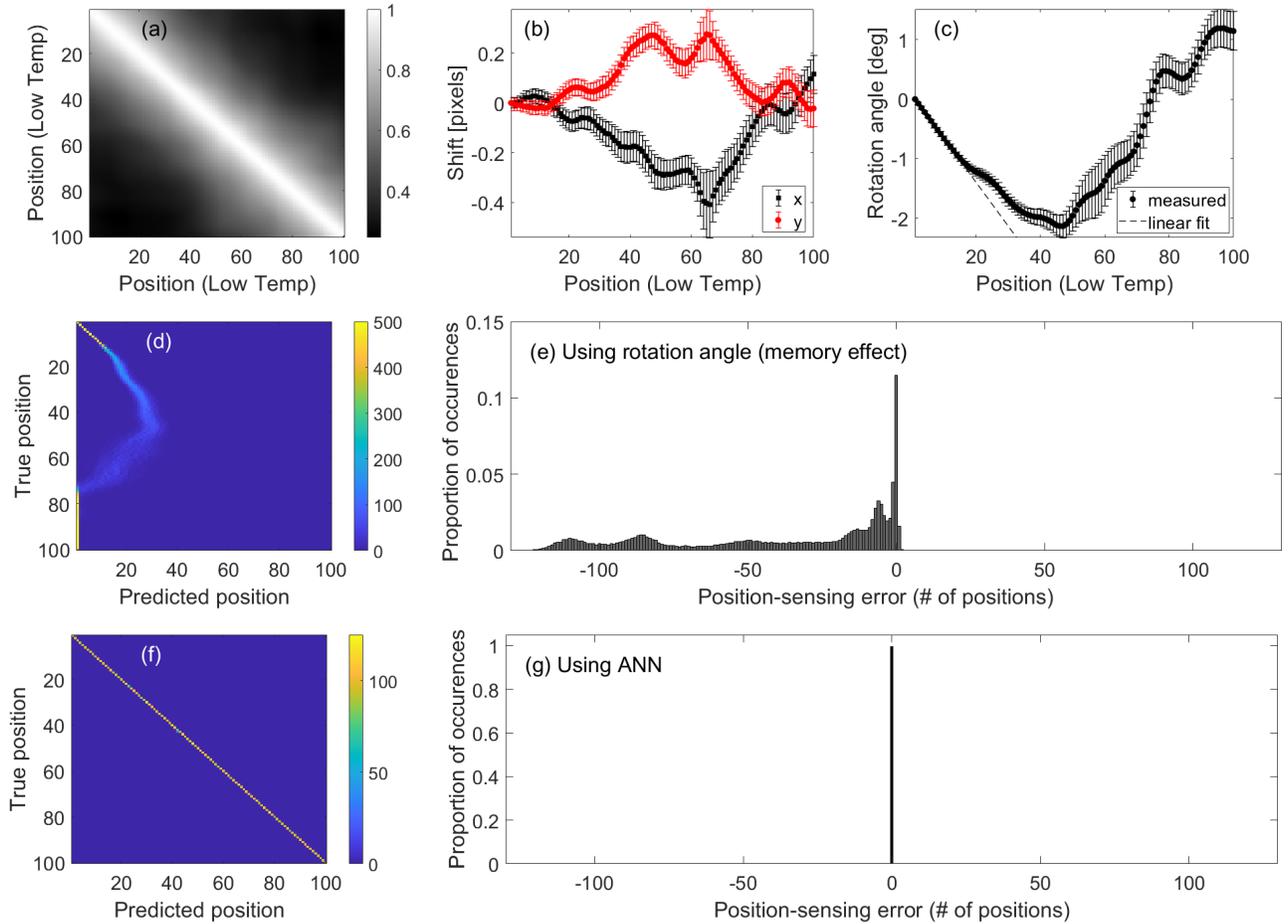


Figure 3. Memory effect vs. ANN for 500 launched MNIST digits recorded at each of 100 measured positions spaced 100 μm apart at room temperature (heat source off). (a) Correlations between specklegrams recorded at different positions. (b) Detected shift of the specklegrams relative to the specklegram at position 1. (c) Detected averages of rotation angles of the specklegrams relative to the specklegram at position 1. The error bars indicate standard deviations across the 500 specklegrams at each position. (d) Confusion matrix of predicted vs. true position, based on memory-effect-based computation. (e) Position-sensing error histograms when using memory effect via detected rotation angle, with average position error of 39.11 positions = 3.911 mm, corresponding to a relative error of 0.3911 compared to the total range of 10mm. (f) Confusion matrix of predicted vs. true position, based on position-sensing network from Fig. 1(b). (g) Position-sensing error histograms when using the position-sensing network from Fig. 1(b), with average position error of 0.00273 positions = 0.273 μm , corresponding to a relative error of $2.73 \cdot 10^{-5}$ (too small to be resolved in the plotted histogram) compared to the total range of 10mm.

accurate position sensing, as one can see in Fig. 5, even though the similarity of the specklegrams exacerbates the distinction of individual positions in comparison to larger position increments, as in the case of Fig. 3 with the fivefold number of positions (100) and doubled position increment (100 μm). Due to this tenfold range of motion (10mm), there is a more pronounced variation between specklegrams taken at the first and last positions, resulting in a larger variation of the correlation in Fig. 3(a) in comparison to Fig. 2(a). Similarly, the shift in Fig. 3(b) and rotation angle in Fig. 3(c) also span a wider range than in Fig. 2(d/g). More importantly, the rotation angle starts to become ambiguous (non-injective mapping) for position index differences > 46 , in the case of Fig. 3(c) bending back towards small rotation angles. Hence, using the memory-effect-induced slope (dashed line in Fig. 3(c)) to quantify the position, we would dramatically misinterpret the rotation angle, leading to the hugely negative position-sensing errors in Fig. 3(e) and an average error of 3.911mm. The ANN, on the other hand, reliably learns the position-induced changes of the specklegrams, regardless if the rotation angle is

an injective function of the position or not, and classifies the position with a negligible average error of only 273nm in Fig. 3(g).

In the following, we describe our ANN in more detail.

4. NEURAL NETWORK ARCHITECTURE AND TRAINING

The ability of neural networks to unscramble MMF specklegrams and recover proximal inputs has been demonstrated in^{24,26,27,35} and others. Single-layer dense ANNs have been shown to be capable of this task³², and in some cases may even outperform more complex convolutional networks, as shown in³¹. We therefore use the networks shown in Fig. 1(b/c), which are described in detail in⁴³. To obtain the training data for both the position and imaging networks, we launch 10,000 input SLM patterns and record the resulting output specklegrams. As indicated in Sect. 2, the position of the fiber is changed by 50 μm after every 10 images and is cycled through all (20) fiber positions 50 times. In the case of Fig. 2 in the previous Sect. 3 only, we used five times as many (100) position increments and SLM patterns (50,000) at double the position increment (100 μm) to investigate a larger position range. The grayscale output specklegrams are recorded on a 150 \times 150 pixel region of the camera. The ground truth input SLM phase images are 200 \times 200 pixel grayscale images as described in Sect. 2.

In the case of the position-sensing network in Fig. 1(b), the training data consists of the same set of images, but the outputs are now mapped onto a ground truth of 10,000 one-hot (20 component, binary) vectors representing the 20 positions. The training labels thus consist of a vector with a single unity value and zeros for all other (19) vector elements. This one-hot representation⁴⁵ encodes a given specklegram’s discretized fiber position in units of the 50 μm displacement steps mentioned in Sect. 2. The training data in this case includes all images that were launched for that position. The softmax final activation layer in Fig. 1(b) captures the “class probabilities”, i.e., the probabilities that the fiber is at the different discrete positions for a given specklegram. There are, in total $850(150^2 + 20) = 19,142,000$ trainable weights in the position-sensing network in Fig. 1(b). The predicted label is the class with the highest predicted class probability using categorical-crossentropy loss⁴⁵.

The imaging network in Fig. 1(c) has a sigmoid final activation layer to limit the range of possible outputs. It is trained on the input SLM images of size 200 \times 200, so there are in total $850(150^2 + 200^2) = 53,125,000$ trainable weights in the imaging network in Fig. 1(c). After the sigmoid activation is applied, the flat output of the deep layer is reshaped into a square image (200 \times 200 pixels) and then fed to the loss function. We use a batch size of 300 for position sensing (Fig. 1(b)) and 30 for imaging (Fig. 1(c)). Every epoch uses every data point exactly once. We train for 200 epochs for position sensing, and 40 epochs for imaging. Altogether, the training process takes only a few hours on a typical workstation, and the inference, i.e., prediction of the fiber position and SLM input phase from an unseen test specklegram, can then be obtained in less than a second⁴³.

After our first successful demonstration of position sensing at a single unperturbed state in Fig. 3(f/g), we now investigate the performance of our position-sensing network when we generalize over perturbations in temperature, the proximal input dataset type, and polarization.

5. GENERALIZATION OF THE POSITION SENSOR

5.1 Position Sensing: Generalization over temperature

Our first example of generalization (from⁴³, included here for completeness) is the usage of an MMF-based position sensor at a heat-source state that it has not been trained on. To achieve this task, we train the position-sensing classifier from Fig. 1(b) with variations in the position of the fiber and the resulting specklegrams at the output of the MMF. The data is gathered by training on 10,000 images from the 20 different positions (again spaced 50 μm apart) per heat-source state. We then test the position classifier on 2,000 images from either the same or a different heat source state, while sampling from all positions uniformly in both testing and training. Finally, we collect the found positions in a matrix (see Fig. 4) that shows the average position-classification accuracy across all 20 positions for all possible combinations of training and testing thermal perturbations.

The experimental results shown on the main diagonal of Fig. 4 demonstrate that when tested on the same temperature, this ANN is able to classify unseen images, i.e., to generalize, for 20 different positions with nearly 100% accuracy, similar to^{40,41}. Combined with the fact that a position-sensing accuracy above 6.46% is

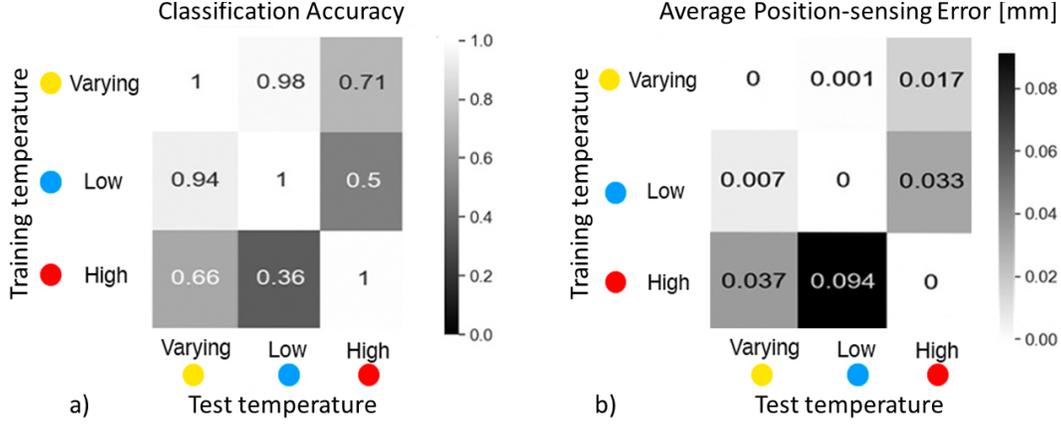


Figure 4. Generalization across different temperature conditions for training and testing. (a) Position-sensing classification accuracy. The position classifier achieved statistically significant generalization in all cases. (b) Average absolute position-sensing error in mm, equivalent to the relative error in comparison to the total range of 1mm. From⁴³.

statistically significant⁴³, this demonstrates that the ANN learned a sufficient amount of fine-scale differences between specklegrams to classify the position much more accurately than the memory-effect-based approach from Fig. 2(j,k,l). We note that the nearly 100% position classification accuracy of the ANN persists even if we are not operating in the speckle memory regime, see Fig. 3(f,g). According to the top left elements (varying temperature) in Figs. 4(a/b), this high accuracy is true not only for a static low temperature and a static heat source with spatial temperature gradient, but also in the combined case of classifying temperature while some specklegrams in the training set are taken at each temperature state (low *and* high). Hence, the trained ANN can reliably generalize for unseen SLM launch images and distinguish between all discrete displacement steps, even though the differences between specklegrams can be very small according to the correlation coefficients in Fig. 2(a/b), especially for directly adjacent positions.

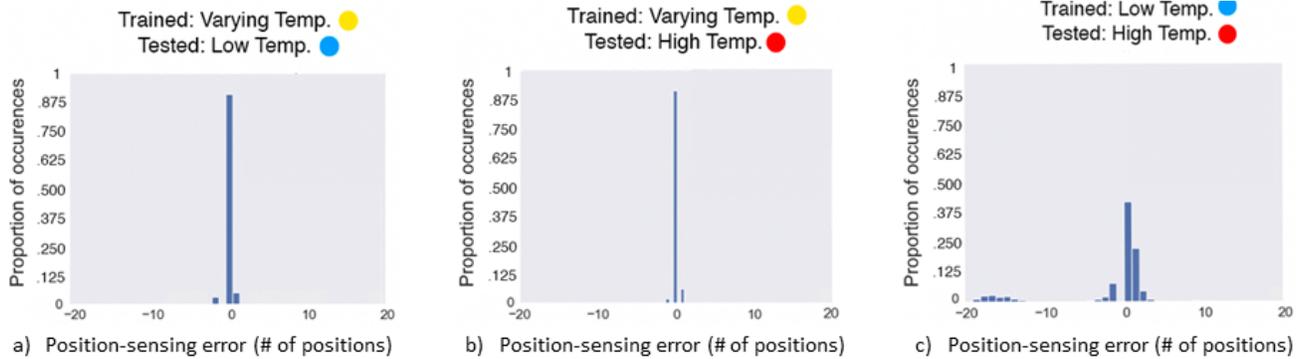


Figure 5. Histograms of absolute position-sensing classification error after generalizing across temperature perturbation. Since only the images from the test set can be used for such quantifications, the total number of images in Fig. 5 is only 20% of those in Fig. 2. (a) Varying \rightarrow Low. (b) Varying \rightarrow High. (c) Low \rightarrow High. From⁴³.

The off-diagonal elements in Fig. 4(a) show the generalization by displaying the position-classification accuracy for different test-training heat source combinations. When the generalization accuracy in Fig. 4(a) is high, most of the network’s mistakes are off by only one or two positions (see Fig. 5), leading to low average position error values in the corresponding entry of Fig. 4(b). Comparing these values in Fig. 4(b) to the average position error values given in the caption of Fig. 2, we see that our ANN-based position classifier has a much smaller position-sensing error than the classical memory-effect-based classifier from Figs. 2(j,k,l) for all temperature combinations, confirming the accuracy and robustness of our approach.

We note that, while these results are consistent when our system is run on a single dataset, Figs. 4, 5, and 9

were all obtained from datasets taken on different days and include natural variation in the results due to experimental noise and the randomness inherent in training a neural network via a stochastic-gradient-descent-based optimizer. For the task of training on varying temperatures and testing on high temperatures, for instance, Fig. 4 represents 71.0% classification accuracy and Fig. 5 represents 91.5% classification accuracy. Likewise, for the task of training on low temperature and testing on high temperature, Fig. 4 represents 50.0% classification accuracy and the confusion matrix in Fig. 9(b) represents 60.5% classification accuracy. In all of these cases, testing on varying conditions performed the best, and the generalization from low to high temperature conditions (or vice versa) was much more challenging.

5.2 Position Sensing: Generalization over the proximal input dataset (MNIST to Fashion-MNIST) and temperature

In this second example of generalization, we demonstrate our classifier’s ability to generalize to specklegrams produced by SLM input phases coming from a different database (“Fashion-MNIST”, which contains images of pieces of clothing instead of handwritten digits), and also in presence of different temperature conditions. The ability of the imaging network to generalize from MNIST to Fashion-MNIST at constant temperature has already been established by³⁵, which is what we also find in Fig. 6 (second row, second column). The training heat conditions “varying”, “off” and “on” are represented by the top, middle and bottom rows in Fig. 6. The three columns in Fig. 6 test with Fashion-MNIST (first column: “varying” heat, second column: “off” heat, third column: “on” heat). While generalization errors are slightly higher now that testing is done on the Fashion-MNIST dataset, which was unseen in training, evidence of robustness of the position sensing is still present, especially in the case where the training conditions (“varying”) contain data from MNIST digit patterns launched in various temperature states across the 25-55°C range.

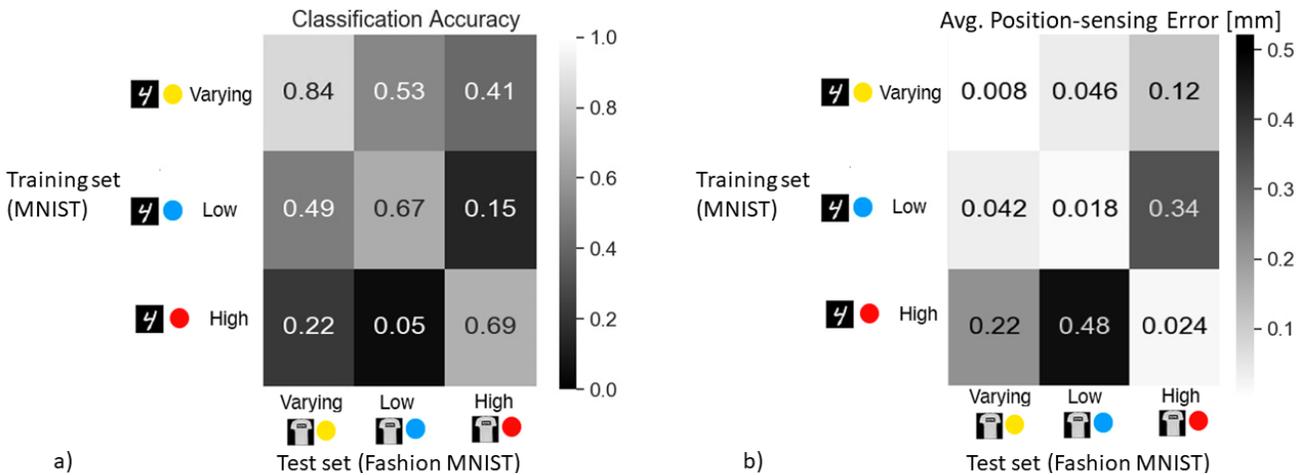


Figure 6. Generalization across temperature and proximal input dataset (MNIST vs. Fashion-MNIST). (a) Position-sensing classification accuracy. (b) Average position-sensing classification error in mm.

5.3 Position Sensing: Generalization over input polarization state

Our third example of generalization is the usage of varying input polarization states. When trained and tested on the same polarization state, the neural network achieves perfect position classification accuracy (zero average position error) as shown on the main diagonal of Fig. 7. As expected, the position classification accuracy decreases (increasing average position error) with an increasing difference of the input polarization state between training and testing, as quantified by the off-diagonal elements in Fig. 7. However, the fact that the first off-diagonal entries in Fig. 7(b) are small shows that our position sensor can operate not only in the presence of temperature changes that affect the MMF as shown in Sect. 5.1, but also in the presence of environmental changes that affect the input polarization between the test and training datasets to a modest degree.

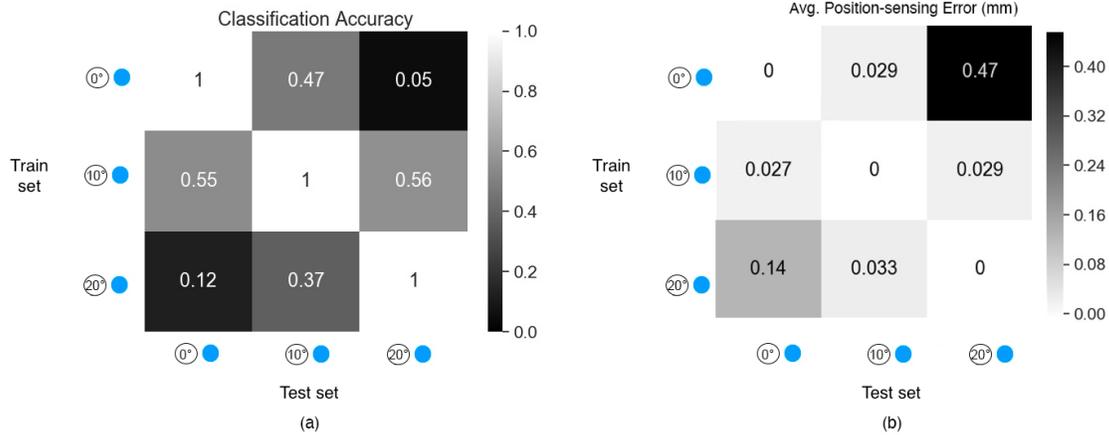


Figure 7. Generalization across input polarization. (a) Position-sensing classification accuracy. (b) Average position-sensing classification error in mm.

5.4 Position Sensing: Generalization over temperature and polarization

Our fourth example of generalization is the combined usage of varying input polarization states and varying temperature conditions. In Fig. 8, the different training sets are all taken at the same polarization (0°), but at the different temperature conditions “varying” and “on”, represented by the top and bottom rows. The test sets use different polarization states (0° and 10°) and the temperature condition “off” and generate the first and second column. In agreement with the findings from Fig. 7 in Sect. 5.3, the errors are small when the test and training sets differ by only 10° in polarization, even in the presence of some additional temperature variations (top right element in Fig. 8). Only in the case of training at “high” and testing at “off” heat, in combination with a 10° polarization angle difference (bottom right element in Fig. 7), do the errors become significant.

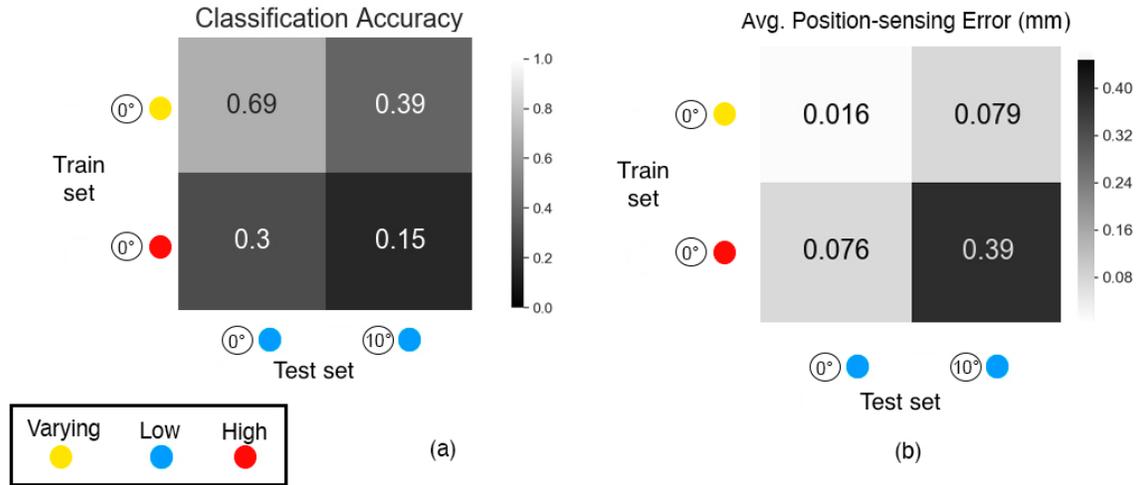


Figure 8. Generalization across polarization and temperature. (a) Position-sensing classification accuracy. (b) Average position-sensing classification error in mm.

6. GENERALIZATION OF THE IMAGING SYSTEM

6.1 Imaging: Generalization over temperature

Having investigated the generalization abilities across position and temperature for the position sensor network from Fig. 1(b) in Sect. 5.1, we do the same for the imaging network from Fig. 1(c) as in⁴³, included here again for

completeness. As illustrated in Fig. 9(a), we find that training on the randomly sampled temperature training set and testing on the set with a static heat source (55°C at the location of the heat lamp as mentioned in Sect. 2) gives an excellent image reconstruction quality with an average structural similarity index measure (SSIM)⁴⁶ of 0.802. When tested on the low temperature data set (25°C), the result is similar. In the case of generalization from heat source on to off (and similarly for heat source off to on), the presence or absence of the induced temperature perturbation results in a lower average SSIM of 0.311 in Fig. 9(b), but some reconstructed digits are still recognizable, and the position classifier is still reliable. In this sense, sampling a variety of temperature perturbations, of course, maximizes robustness, but the predictions are still useful even in the case of a large temperature difference between calibration (training) and testing conditions where generalization is being achieved in a stricter sense.

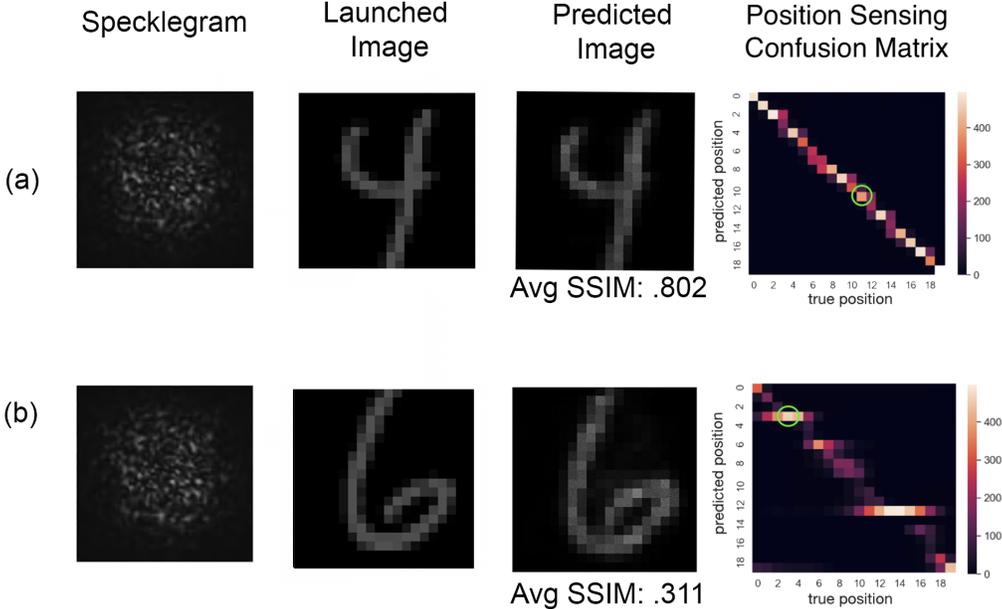


Figure 9. Average SSIM and classification accuracy refer to averages taken over the entire test set (10,000 images). The circled entries in the confusion matrices shown represent where an example image (the launched digit shown) lies on the confusion matrix. (a) Trained on varying temperature perturbation, tested on high temperature perturbation (55°C at the location of the heat lamp). (b) Generalization: Trained on low temperature (25°C), tested on high temperature perturbation. From⁴³.

Since the heat source profile is constant for the “low temperature” training and “high temperature” testing datasets, this combination used in Fig. 9(b) is another example of successful generalization. In Fig. 9(a), the fiber position and launched MNIST digit are different for different training samples as well, so this is also an example of generalization, making the imaging system more robust to different bend (position) states. As such, accurate position predictions can be obtained for the specklegrams corresponding to the images in Fig. 9. Therefore, the same experimental data and code can be used to produce both robust sensing and imaging predictions.

6.2 Imaging: Generalization over input polarization state

Our second example of generalization of the imaging network is the usage of varying input polarization states, shown in Fig. 10. While the results are not as high-quality as in the case of temperature perturbations, some digits are still recognizable in the case of a 10°-perturbation to the input polarization, resulting in an SSIM of 0.378 when averaged across the entire 2000-datapoint test set. For this dataset, the classification accuracy on the (10°-polarization-perturbed) test set was 0.776, demonstrating the ability of the system to simultaneously perform both imaging and position tasks with success.

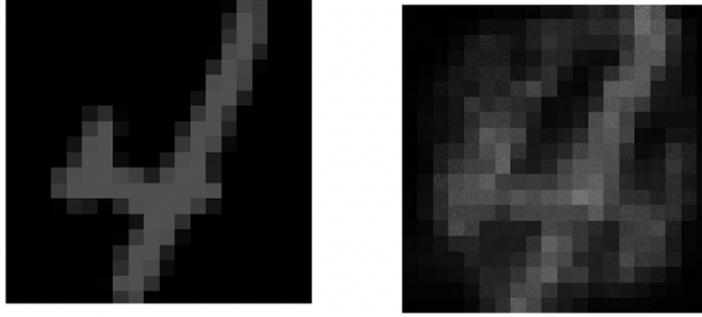


Figure 10. (Left) Example SLM pattern launched. (Right) Corresponding predicted pattern when input polarization is perturbed by 10° compared to the training set.

7. CONCLUSION

We have demonstrated a dense neural network that is capable of simultaneous image reconstruction and position classification through a multimode fiber. Our position classifier is much more accurate than a classical approach that utilizes the memory-effect-induced quasi-linear dependence of the specklegram rotation as a function of the fiber position in the regime of small position ranges. Further, we have shown generalization of the image reconstruction and position sensing from data taken at one input polarization, thermal condition and proximal input database type (MNIST) to data taken at a different input polarization, thermal condition and/or input database type (Fashion-MNIST). This represents progress towards a calibrated MMF endoscope that can withstand perturbations to its shape and temperature, which is a property that is desirable for real-world settings in, for instance, medical endoscopy. Our system is able to quickly and efficiently calibrate its sensing and imaging capabilities and produce a model that can provide reliable predictions in practical situations where the temperature and position of the multimode fiber endoscope is either unknown or difficult to control.

REFERENCES

- [1] Yu, F. T. S., Wen, M., Yin, S., and Uang, C.-M., “Submicrometer displacement sensing using inner-product multimode fiber speckle fields,” *Appl. Opt.* **32**(25), 4685–4689 (1993).
- [2] Efendioglu, H. S., “A review of fiber-optic modal modulated sensors: specklegram and modal power distribution sensing,” *IEEE Sens. J.* **17**(7), 2055–2064 (2017).
- [3] Leal-Junior, A. G., Frizera, A., Marques, C., and Pontes, M. J., “Optical fiber specklegram sensors for mechanical measurements: A review,” *IEEE Sens. J.* **20**(2), 569–576 (2019).
- [4] Liew, S. F., Redding, B., Choma, M. A., Tagare, H. D., and Cao, H., “Broadband multimode fiber spectrometer,” *Opt. Lett.* **41**(9), 2029–2032 (2016).
- [5] Čižmár, T. and Dholakia, K., “Shaping the light transmission through a multimode optical fibre: complex transformation analysis and applications in biophotonics,” *Opt. Express* **19**(20), 18871–18884 (2011).
- [6] Bianchi, S. and Di Leonardo, R., “A multi-mode fiber probe for holographic micromanipulation and microscopy,” *Lab on a Chip* **12**(3), 635–639 (2012).
- [7] Papadopoulos, I. N., Farahi, S., Moser, C., and Psaltis, D., “High-resolution, lensless endoscope based on digital scanning through a multimode optical fiber,” *Biomed. Opt. Express* **4**(2), 260–270 (2013).
- [8] Loterie, D., Farahi, S., Papadopoulos, I., Goy, A., Psaltis, D., and Moser, C., “Digital confocal microscopy through a multimode fiber,” *Opt. Express* **23**(18), 23845–23858 (2015).
- [9] Morales-Delgado, E. E., Psaltis, D., and Moser, C., “Two-photon imaging through a multimode fiber,” *Opt. Express* **23**(25), 32158–32170 (2015).
- [10] Trägårdh, J., Pikálek, T., Šerý, M., Meyer, T., Popp, J., and Čižmár, T., “Label-free cars microscopy through a multimode fiber endoscope,” *Opt. Express* **27**(21), 30055–30066 (2019).
- [11] Farahi, S., Ziegler, D., Papadopoulos, I. N., Psaltis, D., and Moser, C., “Dynamic bending compensation while focusing through a multimode fiber,” *Opt. Express* **21**(19), 22504–22514 (2013).

- [12] Plöschner, M., Tyc, T., and Čížmár, T., “Seeing through chaos in multimode fibres,” *Nat. Photonics* **9**(8), 529–535 (2015).
- [13] Duncan, R. G., Froggatt, M. E., Kreger, S. T., Seeley, R. J., Gifford, D. K., Sang, A. K., and Wolfe, M. S., “High-accuracy fiber-optic shape sensing,” in [*Sensor Systems and Networks: Phenomena, Technology, and Applications for NDE and Health Monitoring 2007*], **6530**, 65301S, SPIE (2007).
- [14] Moore, J. P. and Rogge, M. D., “Shape sensing using multi-core fiber optic cable and parametric curve solutions,” *Opt. Express* **20**(3), 2967–2973 (2012).
- [15] Ahmad, R., Westbrook, P. S., Ko, W., and Feder, K. S., “Probing micron-scale distributed contortions via a twisted multicore optical fiber,” *APL Photonics* **4**(6), 066101 (2019).
- [16] Judkewitz, B., Horstmeyer, R., Vellekoop, I. M., Papadopoulos, I. N., and Yang, C., “Translation correlations in anisotropically scattering media,” *Nat. Phys.* **11**, 684–689 (2015).
- [17] Amitonova, L. V., Mosk, A. P., and Pinkse, P. W. H., “Rotational memory effect of a multimode fiber,” *Opt. Express* **23**(16), 20569–20575 (2015).
- [18] Stasio, N., Conkey, D. B., Moser, C., and Psaltis, D., “Light control in a multicore fiber using the memory effect,” *Opt. Express* **23**(23), 30532–30544 (2015).
- [19] Li, S., Simon A. R. H., Tyc, T., Čížmár, T., and Phillips, D. B., “Memory effect assisted imaging through multimode optical fibres,” *Nat. Commun.* **12**, 3751 (2021).
- [20] McCulloch, W. and Pitts, W., “A logical calculus of the ideas immanent in nervous activity,” *Bull. Math. Biol.* **5**, 115–133 (1943).
- [21] Ivakhnenko, A. G., “Polynomial theory of complex systems,” *IEEE Trans. Syst. Man Cybern.* **SMC-1**(4), 364–378 (1971).
- [22] Marusarz, R. K. and Sayeh, M. R., “Neural network-based multimode fiber-optic information transmission,” *Appl. Opt.* **40**(2), 219–227 (2001).
- [23] Borhani, N., Kakkava, E., Moser, C., and Psaltis, D., “Seeing through multimode fibers with deep learning,” in [*Imaging and Applied Optics*], paper CTH1B.4, Optica Publishing Group (2018).
- [24] Borhani, N., Kakkava, E., Moser, C., and Psaltis, D., “Learning to see through multimode fibers,” *Optica* **5**(8), 960–966 (2018).
- [25] Wang, P. and Di, J., “Deep learning-based object classification through multimode fiber via a CNN-architecture specklenet,” *Appl. Opt.* **57**(28), 8258–8263 (2018).
- [26] Rahmani, B., Loterie, D., Konstantinou, G., Psaltis, D., and Moser, C., “Multimode optical fiber transmission with a deep learning network,” *Light Sci. Appl.* **7**(1), 1–11 (2018).
- [27] Moran, O., Caramazza, P., Faccio, D., and Murray-Smith, R., “Deep, complex, invertible networks for inversion of transmission effects in multimode optical fibres,” in [*Advances in Neural Information Processing Systems (NeurIPS)*], 3280–3291 (2018).
- [28] Fan, P., Deng, L., and Su, L., “Light propagation prediction through multimode optical fibers with a deep neural network,” in [*IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*], 1080–1084, IEEE (2018).
- [29] Kakkava, E., Rahmani, B., Borhani, N., Teĝin, U., Loterie, D., Konstantinou, G., Moser, C., and Psaltis, D., “Imaging through multimode fibers using deep learning: The effects of intensity versus holographic recording of the speckle pattern,” *Opt. Fiber Technol.* **52**, 101985 (2019).
- [30] Rahmani, B., Loterie, D., Kakkava, E., Borhani, N., Teĝin, U., Psaltis, D., and Moser, C., “Competing neural networks for robust control of nonlinear systems,” *arXiv*, 1907.00126 (2019).
- [31] Zhu, C., Chan, E. A., Wang, Y., Peng, W., Guo, R., Zhang, B., Soci, C., and Chong, Y., “Image reconstruction through a multimode fiber with a simple neural network architecture,” *Sci. Rep.* **11**(1), 1–10 (2021).
- [32] Turpin, A., Vishniakou, I., and Seelig, J. D., “Light scattering control in transmission and reflection with neural networks,” *Opt. Express* **26**(23), 30911–30929 (2018).
- [33] Shabairou, N., Cohen, E., Wagner, O., Malka, D., and Zalevsky, Z., “Color image identification and reconstruction using artificial neural networks on multimode fiber images: Towards an all-optical design,” *Opt. Lett.* **43**(22), 5603–5606 (2018).

- [34] Aisawa, S., Noguchi, K., and Matsumoto, T., “Remote image classification through multimode optical fiber using a neural network,” *Opt. Lett.* **16**(9) (1991).
- [35] Fan, P., Zhao, T., and Su, L., “Deep learning the high variability and randomness inside multimode fibers,” *Opt. Express* **27**(15), 20241–20258 (2019).
- [36] Zhu, R., Feng, H., and Xu, F., “Deep learning-based multimode fiber imaging in multispectral and multi-polarimetric channels,” *Opt. Lasers Eng.* **161**, 107386 (2023).
- [37] Zhu, R., Feng, H., Xiong, Y., Zhan, L., and Xu, F., “All-fiber reflective single-pixel imaging with long working distance,” *Opt. Laser Technol.* **158**, 108909 (2023).
- [38] Aisawa, S., Noguchi, K., and Matsumoto, T., “Remote image classification through multimode optical fiber using a neural network,” *Opt. Lett.* **16**(9), 645–647 (1991).
- [39] Ando, T., Horisaki, R., and Tanida, J., “Speckle-learning-based object recognition through scattering media,” *Opt. Express* **23**(26), 33902–33910 (2015).
- [40] Merritt, J. H., Sayeh, M., and Mostafavi, M. T., “Intelligent shape sensor (a novel positional sensor for 3d vascular reconstruction),” *Southern Illinois University, Carbondale, IL, USA* (2007).
- [41] Fontana, M., *Using Machine Learning to Turn Optical Fiber Specklegram Sensor into a Spatially Resolved Sensing System*, Master’s thesis, University of Padua (2018).
- [42] Yang, Q. and Butler, C., “Sensor signal processing using neural networks for a 3-D fibre-optic position sensor,” *Sensors and Actuators A: Physical* **41**(1-3), 102–109 (1994).
- [43] Bagley, N., Kremp, T., Lamb, E. S., and Westbrook, P. S., “Transfer learning and generalization of a neural-network-based multimode fiber position and imaging sensor under thermal perturbations,” *Opt. Fiber Technol.* **70**, 102855 (2022).
- [44] Redding, B., Popoff, S. M., and Cao, H., “All-fiber spectrometer based on speckle pattern reconstruction,” *Opt. Express* **21**(5), 6584–6600 (2013).
- [45] Aggarwal, C., [*Neural Networks and Deep Learning: A Textbook*], Springer International Publishing (2018).
- [46] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P., “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. Image Process.* **13**(4), 600–612 (2004).