Robust Optical Picometrology Through Data Diversity

CHENG-HUNG CHI,[1] ERIC PLUM,[1] NIKOLAY I. ZHELUDEV,[1,2] AND KEVIN F. MACDONALD[1,*]

[1] *Optoelectronics Research Centre, University of Southampton, Southampton, SO17 1BJ, UK*
[2] *Centre for Disruptive Photonic Technologies, School of Physical and Mathematical Sciences and The Photonics Institute, Nanyang Technological University, Singapore; 637378, Singapore*
**kfm@orc.soton.ac.uk*

Abstract: Topologically structured light contains deeply subwavelength features, such as phase singularities, and the scattering of such light can therefore be sensitive to the geometry or movement of scattering objects at such scales. Indeed, it has been shown recently that single-shot optical measurements can yield positional precision better than 100 pm (less than one five-thousandth of the wavelength λ) via a deep-learning analysis of scattering patterns. Measurement performance, and the extent to which it can be sustained are constrained by the stability of experimental apparatus and the quality and depth of neural network training data. Here we show that a neural network can be trained, through exposure to an extended envelope of instrumental/ambient noise conditions, to robustly quantify picometric displacements of a target against orders-of-magnitude larger background fluctuations: maintaining precision and accuracy of 100-150 pm in optical measurements (at λ = 488 nm) of nanowire positional change. This capability opens up a range of application opportunities, for example in the optical study of nanostructural dynamics, stiction, material fatigue, and phase transitions.

1. Introduction

In a range of non-contact, label-free optical measurement and imaging techniques, machine learning has emerged as a powerful tool for retrieving object information from far-field scattering patterns, through its aptitude for effectively solving the inverse scattering problem – i.e. to determine characteristics of an unknown scatterer from measurements of a scattered light field - an inverse problem that can be reduced to the Fredholm integral equation [1-3]. For example, it has been shown that a trained neural network can retrieve positional and dimensional characteristics of sub-wavelength apertures under plane wave illumination from their transmission scattering/diffraction patterns with single-shot precision down to λ/130 (where λ is the wavelength of light) [4, 5]; and more recently that with a combination of topologically structured illumination – an incident light field containing deeply subwavelength features, such as phase singularities [6] – and '*in-situ*' neural network training data that is perfectly congruent with the object of study, picometric precision down to λ/5300 can be achieved in single-shot localization measurements of the position of a nanowire [7].

However, in any optical measurement or imaging application, there are constraints placed on performance by the stability of the instrumental platform. Even a robust optical microscope in a typical temperature-controlled laboratory environment will be subject to short-term random fluctuations and long-term drift at nano- to micrometric length scales over time scales of minutes to hours [8], for example through ambient acoustic noise or differential thermal expansion of components (especially those intended to provide motion, e.g. sample stages, objective turrets). The power, wavelength and polarization of a laser illumination source may likewise be subject to fluctuations and drift over similar time scales. In systems dependent on machine learning, these noise contributions prevent generalization or application to 'out-of-distribution' data - meaning data containing artefacts and noise features different from, or outside the range of such features encountered during training. Training data diversity is essential to addressing this limitation [9, 10]. Here, we show how increased diversity, in the form of just a few iterative training cycles, each subject to random time-dependent noise, can enable a neural network to continue delivering nanowire localization measurements with precision and accuracy orders of magnitude smaller than the scale of concomitant instrumental fluctuations, for minutes beyond the training interval (over a duration $>10^9$ times longer than the nanowire's natural oscillation period), where previously (Ref. [7]) it could only do so within the training interval. The neural network learns to identify changes in the transmission scattering pattern that relate specifically to changes in the position of the nanowire relative to the rest of the sample, while disregarding other, extraneous changes in the scattered light field.

2. Method

The experimental sample and microscopic apparatus for this study were as described in Ref [7]: The sample comprises a 17 µm long, 200 nm wide nanowire, cut by focused ion beam milling from a 50-nm-thick Si_3N_4 membrane coated with 65 nm of gold, with ~100 nm gaps on either side (Fig. 1). The lateral position of the nanowire, i.e. distance to the parallel edges of the surrounding membrane, is electrostatically controlled over a few-nanometer range through the application of a DC bias between the nanowire and one side of the adjacent membrane. The sample was illuminated at the mid-point of the nanowire length, by a coherent (λ = 488 nm) superoscillatory wavefront generated by a pair of spatial light modulators (see Supplementary Information). Transmission scattering patterns from an image plane at a distance ~λ from the membrane were recorded via an objective with a numerical aperture of 0.9 on a 16-bit image sensor. The experiment was operated remotely to exclude perturbations due to the presence of people (as a source of vibrational disturbance, heat, moisture, etc.) in the laboratory during data collection.

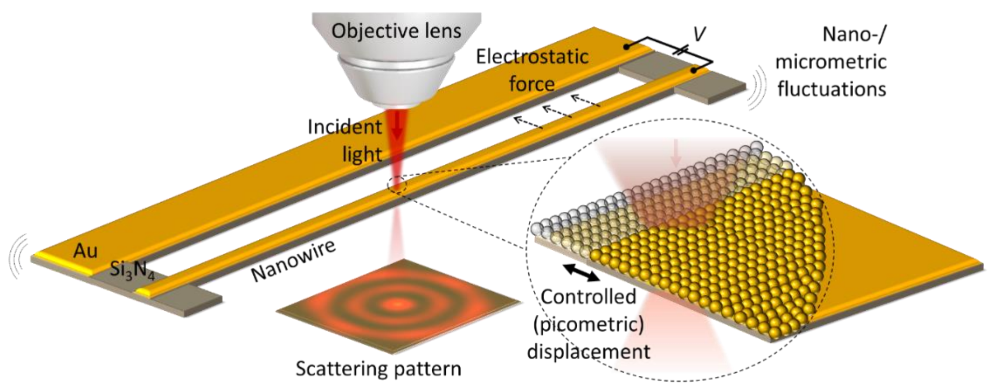

Fig. 1. Measuring picometric nanowire displacement in the presence of nano- to micrometric noise. Topologically structured light scattered from the nanowire is imaged in transmission via a high-numerical-aperture microscope objective (not shown). Lateral displacements of the wire, controlled by application of a DC bias, *V*, between the wire and the adjacent edge of the supporting membrane, are quantified via a deep-learning-enabled analysis of single-shot scattering patterns. With sufficiently 'diverse' training, a neural network can determine nanowire position with high precision and accuracy, even while the whole sample and/or apparatus are subject to much larger positional fluctuations.

We recorded 19 consecutive sets of single-shot scattering patterns for 201 different nanowire positions over a 360-4000 pm range of displacements from the zero-bias position. These were recorded in a random sequence (repeated in each set) to exclude the possibility of neural network learning from any correlation of position and time within a set, and the applied bias (displacement) was reset to zero between each recorded position to eliminate any effect of stress history in the nanowire. Scattering patterns were recorded with a 7.7λ × 7.7λ (301×301 pixel) field of view and an integration time of 100 ms. Each 201-image set was recorded over 2.5 min, giving a total span of 47.5 min between the beginning of the first and the end of the last. The final set was reserved for use as a delayed measurement set, i.e. never used in neural network training, except to establish benchmarks for self-referenced training and testing within a single dataset (as per Ref. [7]). Among all sets, the same randomly selected 80% of scattering patterns and corresponding (known) nanowire displacement values were designated for use in neural network training and validation. The other 20% are put aside as unseen images (nominally unknown displacements) for testing (see Fig. 2). This ensures that, regardless to the 'depth' of training, i.e. the number of repeated datasets used in training: (i) all neural networks are trained and validated on the same 160 known nanowire displacements; and (ii) the assessment of metrological performance is always based on the same set of scattering patterns from the measurement dataset, for nanowire positions that have never been seen in training. (Further details of the neural network architecture and training/validation procedure are given in Supplementary Information.)

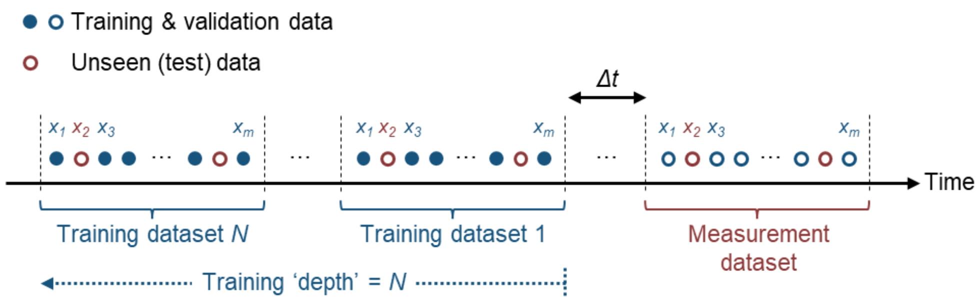

Fig. 2. Neural network training for resilience to noise. Training data with a greater diversity of background noise artefacts is accumulated by iterative repetition of the set of m randomly sequenced nanowire displacement settings (x_1, x_2, x_3, …, x_m), where m = 201 in the present case. The time interval between the end of the last dataset used in neural network training and the subsequently recorded 'measurement dataset' is denoted Δt. 'Training depth' N refers to the number of datasets used in training, counting backwards from the closest in time to the measurement dataset. Within each set, scattering patterns for the same randomly selected 80% of nanowire positions (blue solid circles) are designated for use in training and validation, while the other 20% of positions (red open circles) remain 'unseen' by the network. (Training points within the 'measurement' dataset [blue open circles] are only used for benchmarking against in-distribution measurement performance, i.e. training and measurement within a single dataset.)

3. Results and discussion

Figure 3 shows how metrological performance deteriorates rapidly as the time interval Δt between the end of training data collection and measurement increases, while only single datasets are used for training. We assess neural network performance in the retrieval of optically-measured nanowire displacement in terms of: precision (Fig. 3d) , which describes the reproducibility of a set of nominally identical measurements (cf. the magnitude of the error bars in Figs. 3a-c) – calculated as the average measurement standard deviation over the range of measured displacements; and accuracy, which describes how close the mean value of a set of nominally identical measurements is to the true value (cf. how close the datapoints are to the measured = actual displacement diagonal in Figs. 3a-c) – evaluated as root mean square error over the range of measured displacements. As one would expect, performance in time-delayed measurements is invariably inferior to the self-referenced benchmark case (i.e. of testing on unseen images recorded within the training interval). Precision and accuracy fluctuate wildly as a function of Δt – randomly better or worse from one point to the next as noise artefacts in each training image set are randomly (as a consequence of short-timescale instrumental fluctuations) closer to or further from those present during recording of the measurement dataset. In simple terms, useful measurement capability is lost within <5 min., as illustrated in Fig. 3c.

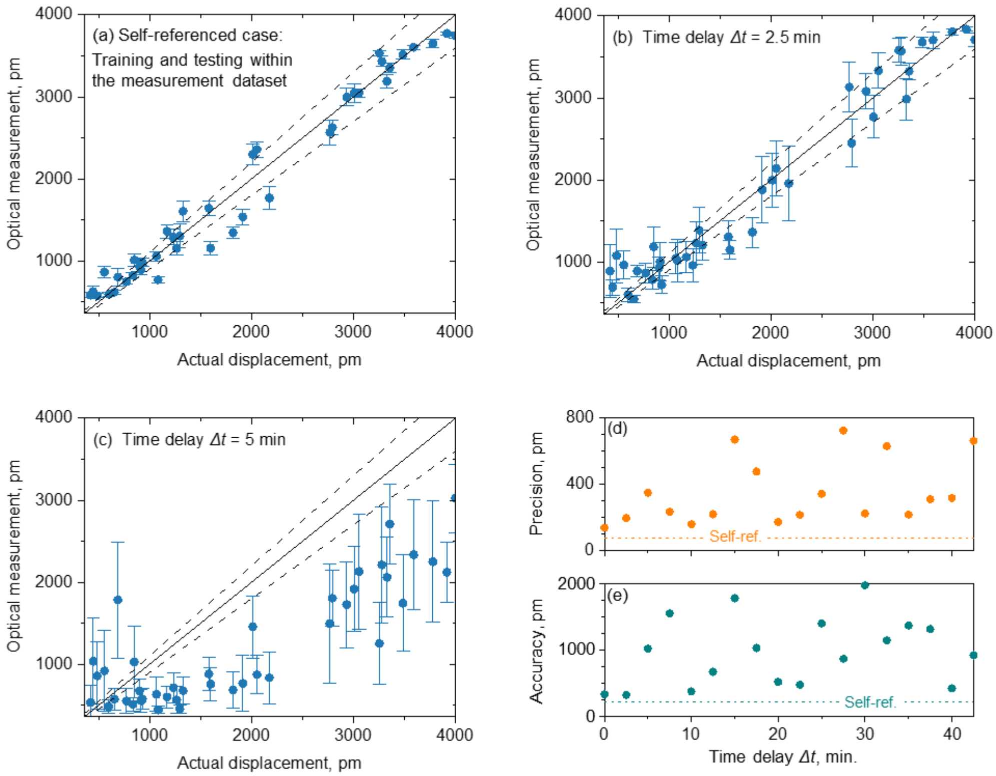

Fig. 3. Optical localization of nanowire position – performance as a function of time delay between training and measurement. (a-c) Optically measured versus actual values of nanowire displacement for neural networks trained: (a) in the self-referenced, benchmark case of training and testing within a single dataset; (b, c) on single datasets with a delay *Δt* [as labelled] between the end of training and start of measurement dataset recording. Points and error bars represent respectively the mean and standard deviation of ten measurements from independently trained networks. Dashed lines above and below the solid ideal correlation diagonals are plotted at ±10% relative error. (d) Measurement precision [average standard deviation] and (e) accuracy [root mean square error] as a function of *Δt*. In (d) and (e), values for the self-referenced benchmark case are shown as dashed lines.

Figure 4 shows how increasing the diversity (or 'depth') of training can improve performance in time-delayed measurements, to the extent of achieving comparable precision and better accuracy than the self-referenced benchmark at values of *Δt* where measurements based on single-dataset training fail. In this case, we fix *Δt* at 5 minutes and vary the training depth *N* – the number of prior datasets used in neural network training. Figures 4a-c show examples for training depths *N* = 1, 6, and 16 (Fig. 4a, for *N* = 1 and *Δt* = 5 min. being identical to Fig. 3c). As can be seen in Figs. 4d,e a training depth >6 is sufficient to recover precision ≤100 pm, approaching the 72 pm single-dataset self-referenced benchmark, and accuracy slightly but consistently better than the 214 pm self-referenced benchmark. This is not a simple manifestation of the idea that the predictive power of a neural network increases with the size of its training dataset [11]: In the present case, increasing *N* does not add any additional information to the training dataset on the measurand (the nanowire position) – it adds repetitions of the same 160 known nanowire positions. This should not and does not improve measurement precision (Fig. 4d). That it marginally improves measurement accuracy (Fig. 4e) is an interesting consequence (discussed below) rather than the intended purpose of the iterative training regime, which is to sustain performance over time after training. Through training on several scattering patterns for each nanowire position, recorded over time (here, at intervals of ~2.5 min) such that each presents different random noise artefacts related to instrumental fluctuations on such timescales, the neural network learns to distinguish between changes in the patterns which relate specifically to movement of the nanowire relative to the nearby edges of the membrane (i.e. electrostatically controlled variations in the gap sizes on either side of the wire) and other changes in the scattering patterns, related to instrumental perturbations that may be orders of magnitude larger.

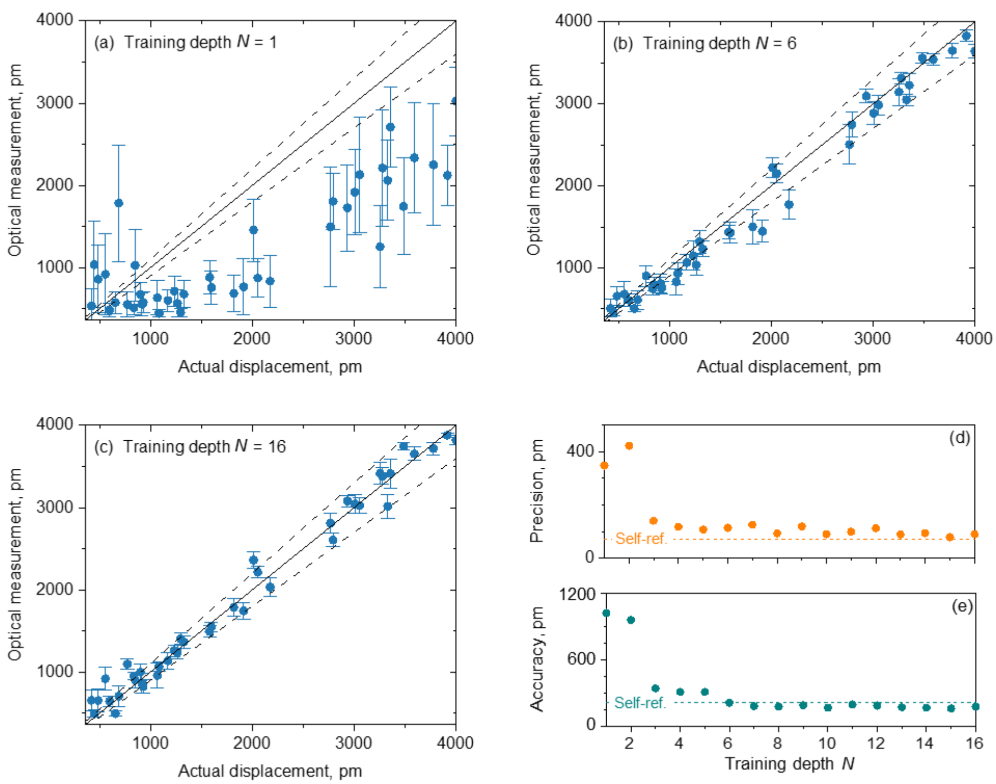

Fig. 4. Optical localization of nanowire position – performance as a function of training depth at Δt = 5 min. (a-c) Optically measured versus actual values of nanowire displacement for neural networks trained over N [as labelled] datasets, with a fixed interval Δt = 5 min between the end of training and the start of measurement dataset recording. (d) Measurement precision [average standard deviation] and (e) accuracy [root mean square error] as a function of N.

The same behavior – recovery of precision and accuracy to levels respectively just above and just below the self-referenced benchmark at N>6; little or no further improvement in performance for N >8, is seen at Δt = 10 min. (see supplementary Fig. S2). This speaks to the fact that in general, the depth of training required to maximize metrological performance and the length of time over which performance can then be maintained will depend on the nature of the measurement being made, and the nature, magnitude and characteristic timescale of instrumental and ambient noise and/or drift, i.e. on the physical and environmental stability of the experiment. In the present case (an optical microscope with a closed-loop piezoelectric sample stage, on a vibration-isolated optical table), as few as six iterations of each nanowire position over a total training period of ~15 min. are sufficient for the neural network to map the parameter space of noise subsequently encountered over at least another 12.5 min. (i.e. Δt = 10 min. plus the 2.5 min. duration of measurement dataset collection). Performance will be maintained for as long as experimental conditions remain within the trained envelope of measurement range and noise; it will be lost (the neural network's model will fail) at whatever point the apparatus exceed that envelope, e.g. due to long-term drift of a given parameter with laboratory temperature over many hours.

Ordinarily (Figs. 3, 4) neural networks are tested only on previously-unseen nanowire positions (the red points within the measurement dataset, as illustrated in Fig. 2). However, it is informative to compare measurement performance for previously-seen (i.e. in training and validation) versus -unseen positions, both subject (in the time-delayed measurement dataset) to previously-unseen noise. At Δt = 5 min., there is no difference between results for these two cases (Fig. 5). This indicates firstly that noise in the scattering patterns at Δt = 5 min. is sufficiently large that networks with shallow (low-*N*) training cannot distinguish between nominally known and unknown positions (i.e. all lie outside the envelope of trained noise parameter space). At high *N*, the fact that precision (Fig. 5a) for previously-seen positions tends to a self-referenced benchmark identical to that for previously-unseen positions (rather than a lower level, or even zero) indicates that this is a limit imposed by experimental uncertainty, i.e. principally in knowledge and reproducibility of actual displacement values. Accuracy (Fig. 5b) at high *N*, for both previously-seen and -unseen positions, tends to the self-referenced benchmark for previously-seen positions, which surpasses the level for previously-unseen positions because it is derived from the data specifically used to optimize accuracy (minimize error) as the singular objective of network training and validation.

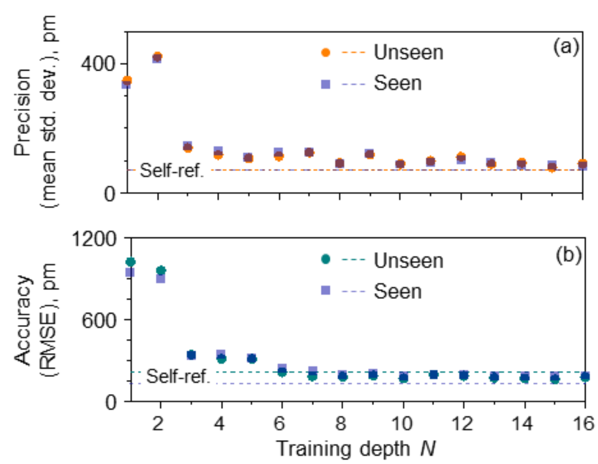

Fig. 5. Comparison of optical measurement performance as a function of training depth at Δt = 5 min. between nanowire positions unseen by neural networks during training [datapoints as in Figs. 4d,e] and nominally known [previously-seen] positions used in training. (a) Measurement precision [average standard deviation] and (b) accuracy [root mean square error] as a function of *N*.

4. Conclusion

In summary, we have shown experimentally that picometrically precise single-shot optical measurements enabled by machine learning can be made robust against instrumental fluctuations at orders of magnitude larger scale. With suitably diverse (iterative depth of) *in-*

situ training, a neural network retrieving (sub)nanometric positional/dimensional information on an object from scattering patterns can learn to distinguish meaningful changes in the patterns from extraneous artefacts related to random instrumental noise. Precise and accurate measurements can thus be maintained for some time beyond the training window – in the present case, for at least 12.5 minutes after only six 2.5-min. training data acquisition cycles. In the context of studying the motion of nanoscale objects, this interval should be compared with the nanowire's ~1.6 MHz in-plane natural mechanical resonance frequency: it corresponds to $>10^9$ oscillation periods. This sustained measurement capability opens up a range of interesting applications for optical picometrology, for example in the study of nanostructural dynamics and the action of forces and fields on nano-objects, in systems providing for a neural network to be trained under a regime of controlled (quasi)static positioning, for subsequent observation of free or externally driven motion.

Funding

This work was supported by the UK Engineering and Physical Sciences Research Council (grant EP/T02643X/1) and the National Research Foundation Singapore (NRF-CRP23-2019-0006).

Disclosures

The authors declare no conflicts of interest.

Data availability

The data from this paper can be obtained from the University of Southampton ePrints research repository: https://doi.org/10.5258/SOTON/D3200.

Supplemental document

See Supplement 1 for supporting content.

References

[1] V. Vemuri, and G.-S. Jang, "Inversion of Fredholm integral equations of the first kind with fully connected neural networks," J. Frankl. Inst. **329**, 241-257 (1992).
[2] Y. Li, Y. Xue, and L. Tian, "Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media," Optica **5**, 1181-1190 (2018).
[3] T. Li, A. Chen, L. Fan, M. Zheng, J. Wang, G. Lu, M. Zhao, X. Cheng, W. Li, X. Liu, H. Yin, L. Shi, and J. Zi, "Photonic-dispersion neural networks for inverse scattering problems," Light Sci. Appl. **10**, 154 (2021).
[4] C. Rendón-Barraza, E. A. Chan, G. Yuan, G. Adamo, T. Pu, and N. I. Zheludev, "Deeply sub-wavelength non contact optical metrology of sub-wavelength objects," APL Photon. **6**, 066107 (2021).
[5] T. Pu, J. Y. Ou, N. Papasimakis, and N. I. Zheludev, "Label-free deeply subwavelength optical microscopy," Appl. Phys. Lett. **116** (2020).
[6] N. I. Zheludev, and G. Yuan, "Optical superoscillation technologies beyond the diffraction limit," Nat. Rev. Phys. **4**, 16-32 (2022).
[7] T. Liu, J. Y. Ou, J. Xu, E. A. Chan, K. F. MacDonald, and N. I. Zheludev, "Picophotonic localization metrology beyond thermal fluctuations," Nat. Mater. **22**, 844-847 (2023).
[8] G. Grover, W. Mohrman, and R. Piestun, "Real-time adaptive drift correction for super-resolution localization microscopy," Opt. Express **23**, 23887-23898 (2015).
[9] S. Madan, T. Henry, J. Dozier, H. Ho, N. Bhandari, T. Sasaki, F. Durand, H. Pfister, and X. Boix, "When and how convolutional neural networks generalize to out-of-distribution category–viewpoint combinations," Nature Machine Intelligence **4**, 146-153 (2022).
[10] D. Hendrycks, S. Basart, N. Mu, S. Kadavath, F. Wang, E. Dorundo, R. Desai, T. Zhu, S. Parajuli, M. Guo, D. Song, J. Steinhardt, and J. Gilmer, "The Many Faces of Robustness: A Critical Analysis of Out-of-Distribution Generalization," in *2021 IEEE/CVF International Conference on Computer Vision*, (2021), pp. 8320-8329.
[11] B. D. Ripley, *Pattern Recognition and Neural Networks* (Cambridge University Press, Cambridge, 1996).