



HAL
open science

Deep learning for nano-photonic materials -The solution to everything!?

Peter Wiecha

► **To cite this version:**

Peter Wiecha. Deep learning for nano-photonic materials -The solution to everything!?. *Current Opinion in Solid State and Materials Science*, 2023, 28, pp.101129. 10.1016/j.cossms.2023.101129 . hal-04760142

HAL Id: hal-04760142

<https://laas.hal.science/hal-04760142v1>

Submitted on 30 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Deep learning for nano-photonic materials - The solution to everything!?

Peter R. Wiecha^{1,*}

¹LAAS, Université de Toulouse, CNRS, Toulouse, France

Deep learning is currently being hyped as an almost magical tool for solving all kinds of difficult problems that computers have not been able to solve in the past. Particularly in the fields of computer vision and natural language processing, spectacular results have been achieved. The hype has now infiltrated several scientific communities. In (nano-)photonics, researchers are trying to apply deep learning to all kinds of forward and inverse problems. A particularly challenging problem is for instance the rational design of nanophotonic materials and devices. In this opinion article, I will first discuss the public expectations of deep learning and give an overview of the quite different scales at which actors from industry and research are operating their deep learning models. I then examine the weaknesses and dangers associated with deep learning. Finally, I'll discuss the key strengths that make this new set of statistical methods so attractive, and review a personal selection of opportunities that shouldn't be missed in the current developments.

In the past year, large language models (LLMs) like “chat-GPT”, “GPT-4” or “LLM200” [1–3] and text-to-image generators like “Midjourney” or “stable diffusion” [4, 5] have demonstrated the literally breathtaking capabilities of large deep learning models. As a result, Big Tech companies currently engage in a literal battle over the integration of artificial intelligence (AI) into their products [6]. Supposedly for marketing reasons, the term “AI” is thereby often used synonymously for “deep learning”, which boils down to, a little disrespectfully stated, fitting large mathematical functions in a statistical approach to gigantic amounts of data [7]. Although, in contrast to conventional statistical methods, deep learning models are typically much larger in terms of internal degrees of freedom and the number of input / output channels. So large actually, that it becomes extremely hard or even impossible to understand, how a model comes up with its predictions [8]. Nevertheless, DL has proven to perform almost “unreasonably effective” [9] on problems encompassing an abundant variety of applications. Already for more than a decade, deep learning methods are breaking records in computer vision challenges [10, 11] or enable previously unimaginable applications in natural language processing [12, 13]. Other examples to illustrate the remarkable effectiveness of deep learning, may be protein folding predictions [14], astrophysical problems like the identification and analysis of galaxy merger events [15], or even games like “Go”, which, before the era of deep learning, was assumed to be unsolvable by computer algorithms [16].

I. EXPECTATIONS

In the early stages of applying deep learning to the study of photonic nano-structures and materials, researchers reported successful applications [17–21]. These findings seemed to validate the prospect of a transformative shift in our computational capacity to address nano-photonics problems. In the meanwhile however, several challenges have turned out to

be more difficult to overcome than anticipated, and other, unexpected problems have emerged. A representative example from the field around nano-photonic materials may be meta-material design. For several years already, intense research is taking place with the goal to create deep learning algorithms that increase the accuracy in the conception of photonic metasurfaces [22–28]. However, and despite frequent claims of unprecedented design accuracy, works demonstrating deep learning methods are quite generally at the level of proofs of concept. Application oriented and large-scale metasurfaces are still almost exclusively designed via the traditional method based on pre-simulated lookup tables [29–31]. As another example, while indeed providing impressive results, also the performance of the initially mentioned protein folding prediction model (“alphaFold2”) has turned out to be less stable and the results less useful, than initially hoped and claimed [32].

Big Tech has its own examples of overly raised expectations, a prominent one is autonomous driving. Enabling self driving cars by machine learning is in fact an ongoing promise for almost forty years [33–38]. Yet, regardless the tremendous research and development efforts over this considerable time-span, progress remains slow, autonomous vehicles still underlie extremely rigid regulations and their market share is insignificant. A recently shut down San Francisco self-driving taxi company employed 1.5 persons per autonomous car, supervising the fleet remotely and intervening every few miles (keep in mind that a conventional taxi requires exactly one driver) [39]. All this despite enormous global invests, in particular during the last decade [40].

In summary, the expectations with regards to the power and capabilities of deep learning are tremendous, yet sometimes illusory. It is probably safe to claim that these expectations are across the board exaggerated. Consequently I believe that it is important to clarify some myths about the capabilities and weaknesses of deep learning methods, and in this light revisit its main limitations, but also strengths and key potentials.

II. YET ANOTHER DEEP LEARNING PAPER?

To start, I want to briefly discuss a rarely mentioned problem that is in some ways also a consequence of exaggerated

* e-mail : pwiecha@laas.fr

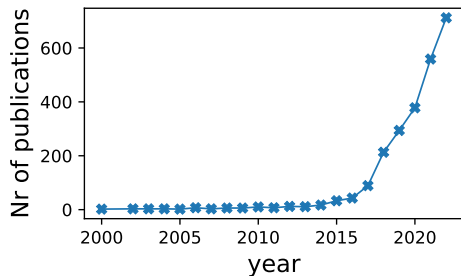


Figure 1. Cumulated number of publications per year from 2000 to 2022, combining the topics “photonics” and one of: “deep learning”, “machine learning” or “artificial intelligence”. Citation Report graphic is derived from Clarivate Web of Science, Copyright Clarivate 2023. All rights reserved.

expectations: The flood of publications that can be referred to as “*yet another deep learning paper*”. Because of its medial attention and promises, many researchers try currently to apply deep learning to their problems. Naturally, first tests are often done on simple problems, like design challenges with only a few free parameters, or with very constrained geometries, merely in the perturbation regime. Such problems however, are typically better solved with conventional methods [41]. The main reason is the tremendous computational overhead of deep learning. Moreover, being a statistical approach, it also does not make much sense to apply it to problems with only weak variations. However, scientific publication culture has strongly accelerated in the last decades [42]. Researchers experience high pressure for publication, especially in an exploding research field, where a year of hesitation can make all the difference. As a consequence, currently an overwhelming number of publications around deep learning applications for photonic materials and nano-optics problems is thrown on the audience (figure 1). A significant fraction of these may be cautiously called “incremental results”. On the other hand, there *are* various highly relevant results that merit the community’s full attention. The ever increasing inflation of publications unfortunately dilutes the important works and renders them more and more difficult to spot. For photonics materials researchers that are new to deep learning and think about applying it to their research, I recommend to get an overview of the more significant results through recent review articles, such as [41, 43, 44].

III. CONSIDERATIONS ABOUT SCALES

Another problem associated with overwhelming expectations is the difference between the scale of deep learning in the products of Big Tech companies that we hear about in the mass media, and the scale of deep learning that is accessible to researchers in physics or materials science.

a. Big Tech The threshold of networks with more than a billion parameters has been overcome already many years ago [45]. Today, the challenge that Big Tech has set, are mod-

els with trillions of free parameters [46]. Some commercial products have in fact already reached this order of magnitude. Google’s PaLM for instance has 540 billion parameters [47], and GPT-4 is considered to have almost 2 trillions of trained parameters [48]. Concerning the dataset sizes we face similar orders of magnitude. According to a leaked document, GPT-4 is said to be trained on more than 12 trillion (10^{13}) text samples [48], which probably amounts to a significant fraction of the internet’s entire accessible information. OpenAI’s competitors are not significantly behind this data scale. Meta’s Llama for example was trained on 2 Trillion text samples [49].

An incredible amount of computing power is required to configure these massive degrees of freedom and process the vast amount of data. GPT-4 is the current leading example. It has been trained for several months, running in parallel on about 25,000 (25,000!) Nvidia H100 GPUs. A single H100 GPU integrates 80 billion transistors and is capable to perform 2 petaFLOPS (2×10^{15} 16 bit floating point operations per second), while running on 700 Watts of electric power. This signifies more than 17 Megawatts of continuous power consumption and several tens of Gigawatt hours of electric energy spent for just a single training run. According to OpenAI’s former and once again CEO Sam Altman, the cost for the training of GPT-4 alone amounts to more than 100 million US dollars [50]. Yet, this development is continuing to accelerate. Only a few years back, the second-last generation of large language models (GPT-2, BERT, etc...) required per-training energy in the order of hundreds of Megawatt hours, orders of magnitude less than the latest generation of models [51]. In the meanwhile, also the hardware costs are immense. As of mid 2023, one H100 GPU alone costs around 40,000\$ [52].

b. Physics labs In comparison to Big Tech, the size of typical deep learning models, their training costs, and the amount of data processed are of a completely different order of magnitude in physics, especially in the fields of nano-optics and photonic devices and materials. The number of free parameters are typically in the order of millions to tens of millions ($10^6 - 10^7$), the training is often done on single consumer grade GPUs, in general no longer than during a few days. In most cases, the amount of available training data is particularly limited, since their generation is usually expensive. Common training sets comprise in the order of ten thousand training samples [53–58]. Only in rare cases, studies are based on hundreds of thousands (10^5) or even more samples [59, 60].

A comparison of the orders of magnitude of the relevant

Table I. Orders of magnitude of important metrics in commercial (Big Tech) vs. physics lab scale deep learning.

	physics lab	Big Tech
N_{gpu}	1	10^4
$N_{\text{train time}}$ (GPU hours)	10	10^7
$N_{\text{train cost}}$ (\$)	100	10^8
$N_{\text{NN params}}$	10^7	10^{12}
$N_{\text{train samples}}$	10^6	10^{13}

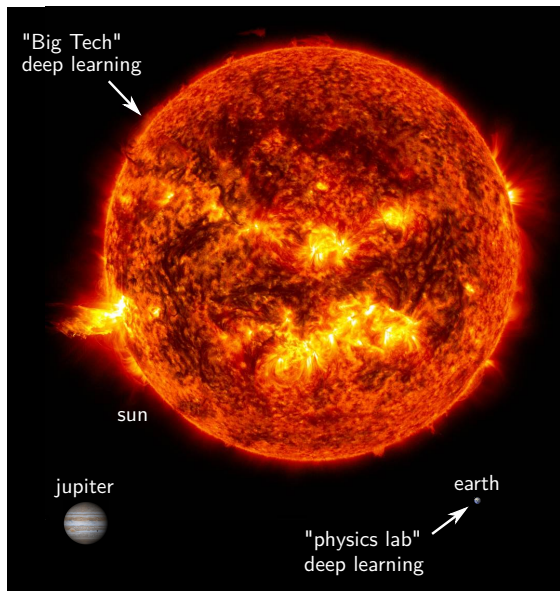


Figure 2. Deep learning scale comparison. In almost every metric, deep learning in physics labs is 6 or more orders of magnitude smaller compared to Big Tech applications like “chat-GPT” (c.f. table I). This difference compares like the size of the sun to that of earth, the latter fitting around 1.3 million times in the star at our solar system’s center. Images by NASA, arranged with permission through NASA’s open access policy.

quantities is given in table I. In fact, Big Tech’s neural network models compare to physics lab deep learning in a remarkably consistent way, like comparing the size of our Sun to the size of the planet Earth, as shown at scale in figure 2. The Earth is about 1.3 million times (10^6) smaller in volume than the Sun.

IV. INTERPOLATION AND EXTRAPOLATION

The main assumption that underlies deep learning is that a mathematical function of large enough degree of freedom can approximate any other function to arbitrary precision [64]. The large mathematical model is the artificial neural network. The target function to be approximated is only implicitly defined by a (large) set of data samples. Practically, the parameters of the network are then configured in a fitting process (the training) to match the given set of data. In this process, the hope is that the neural network will develop some kind of global model that describes the data and is able to generalize. However, such generalization happens very rarely, in most cases extrapolation outside the training data region fails quickly [41]. This is illustrated in figure 3 by the example of a multilayer perceptron model, trained on datapoints sampled from a random polynomial function. While interpolation works well, extrapolation fails immediately outside of the training data range.

In nano-optics and photonics materials applications, similar trends are observed. Limiting trained deep learning models to a valid parameter range is therefore an important secondary

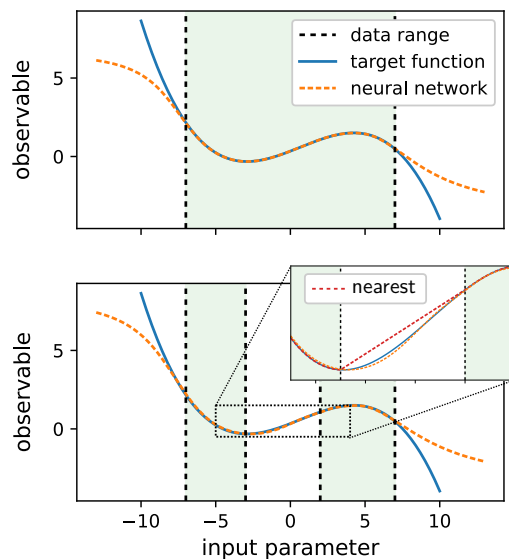


Figure 3. Illustration of a neural network’s performance at interpolation and extrapolation tasks. Two simple 3-layer networks are trained on data, describing a polynomial function. In one case, a larger area of connected data is given (top panel). In a second experiment, the data in a limited parameter region is removed (bottom panel). In the interpolation regime where training data is available (inside the green shaded ranges), the accuracy is high. Outside of the data range (left and right sides), extrapolation fails immediately. Yet, interpolating between two adjacent zones of available data (white center area in the bottom panel), works better than the trivial nearest neighbor linear interpolation (red dotted line in the inset).

problem [65, 66]. Especially, because in high-dimensional problems it is often not trivial to specify the interpolation regime [67]. Figure 5c shows an example of extrapolation to a photonic light router design with around 150 perturbations, whereas the dataset was limited to cases with no more than 20 perturbations [68]. A possible approach to constrain surrogate models to the interpolation regime of the input space is to use auxiliary networks like variational autoencoders or generative adversarial networks, that learn a compact, regularized representation of the input parameters [41, 69–71]. It can also help to provide physics through additional loss functions, that test whether a solution is compatible with known physics laws (for example in form of partial differential equations or causality) [72–74]. But these steps to improve extrapolation require significant efforts, and there is no guarantee to what extent such techniques will work. Simply speaking, one must always expect “to get out what has been put in”, and careful testing of the results is necessary when operating in extrapolation.

V. WEAKNESSES AND DANGERS

In addition to the limited extrapolation capabilities discussed before, various further weaknesses and possible pitfalls come with deep learning. A selection is illustrated in figure 4a. Some limitations can be countervailed with large

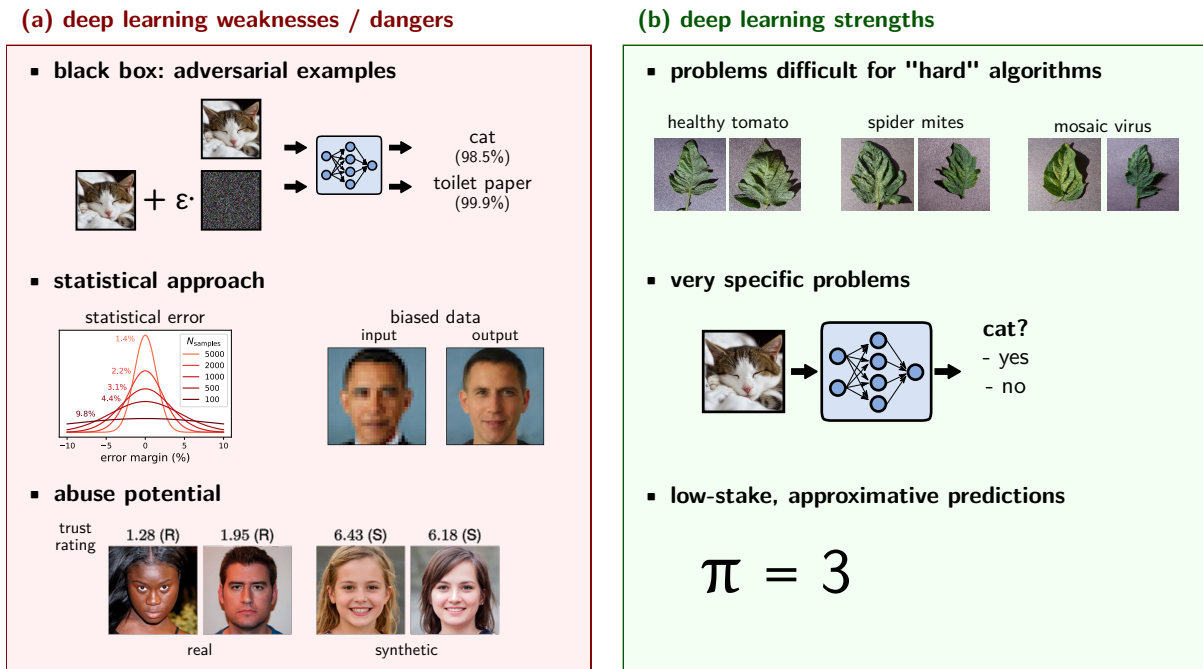


Figure 4. Deep learning weaknesses and strengths. On the weakness side (a) we have: *Black-box*: Deep learning models are black boxes, untrustworthy and prone to adversarial attacks [61]. *Statistical approach*: On smaller datasets, deep learning will always add a significant statistical error. For example with 5000 samples, the margin of error for a 2σ confidence interval is still around $\pm 1.5\%$. Furthermore, a neural network will reproduce biases from the datasets, therefore deep learning stands and falls with the data quality. *Abuse potential*: Deep learning can be used for generative tasks beyond detectability. On the strength side (b) we have: *Hard algorithmic problems*: Problems that are very difficult to solve by conventional algorithms can often be solved with little effort by deep learning. Very detailed image analysis is a good example. *Specific problems*: The more specific a problem can be defined, the easier it will be for deep learning. *Low stake approximations*: Problems where rough approximations are good enough but speed is all that matters. Cat image (c) 2016 by Alexandru Zdrobău, reproduced with permission of the author. Obama’s white skin upsampling result reproduced from twitter post on 20th June 2020 by user “chicken3egg”. Generated vs real portraits, reproduced with permission from [62]. Tomato leaf dataset images reprinted from [63], copyright 2020, with permission from Elsevier.

enough datasets. Others however are inherent to the method.

a. Black Box A fundamental problem of deep learning as such, is the black box character of basically all artificial neural networks. In general it is not possible to understand microscopically how a trained neural network makes its predictions. An illustrative symptom of this issue are so-called “adversarial examples”, that can be found for basically all deep learning image classification models [61, 75, 76]. An adversarial example is a weak, but specifically designed noise pattern, that maximally activates a target pathway through the deep learning model and eventually leads to a complete misclassification of the input image. Also in various regression models, similar singularities can be found at which predictions totally fail [53]. This is illustrated in figure 5a, where, for not understood reasons, a nanophotonics regression model shows prediction failure with very high error in around 5% of cases. This renders neural network predictions untrustworthy in a quite general sense, imposing severe limitations for security-relevant applications such as the above mentioned autonomous driving [77]. For instance, only a few years back it was possible by very simple means to fool Tesla’s car assistant system to drive on the wrong side of the street or to ignore speed limits [78]. Another example for a currently

emerging security gap are so-called “indirect prompt injection” attacks against large language models like AI-assistants (e.g. Microsoft “co-pilot” or Google’s “Bard”). These LLMs are known to be vulnerable to “engineered” prompts, which hackers may hide in specifically tailored websites. A model that parses such website and which is equipped with advanced privileges like access to personal data and the internet, can then be forced to execute malicious tasks [79]. Despite these known shortcomings, all major Big Tech companies are, as of 2023, about to grant LLM-based personal assistant tools access to the internet and to personal customer data such as e-mails.

Please note that there are more interpretable machine learning models than deep neural networks [80], and there are considerable ongoing efforts to render deep learning more explainable [81, 82].

b. Data quantity and quality Other problems relate to the data quality or quantity, and boil down to the statistical nature of deep learning. Naturally, a sufficient amount of data is necessary as the training process will statistically evaluate correlations between the samples. If little data is available, it is therefore important to keep in mind that the predictions of a neural network model will be necessarily worse than the

accuracy of the training data. A statistical error will always be added on top. This can be circumvented to some degree using data augmentation methods, yet excessive use of such techniques bears the risk of inducing bias in the model [83, 84].

Data quality on the other hand, can pose problems that could also be denominated with the term “systematic errors”. A prominent problem with data quality is biased statistics, that is, data sets that are not representative of the entire parameter space, or that contain an excessive number of samples from a small subspace of the problem. Biased datasets behind natural language processing models or computer vision tools have been widely covered by the media. Both type of applications have reportedly learned racism and sexism, among other social or ethnical biases [85–87]. In photonics materials design, a typical bias could be the resonant or non-resonant nature of samples in the training data, as illustrated in figure 5b, where the dataset consisted of mostly non-resonant samples [53]. A neural network may also statistically learn to ignore the possibility of optical losses if trained on a dataset of highly transmissive devices [68]. Other systematic problems can be caused by erroneous data (e.g. non-converged numerical simulations) or datasets containing many outliers (e.g. noisy measurements) that break the statistical assumptions of the learning process [88, 89]. Since high quality data is the most essential ingredient for deep learning, in Appendix A, I discuss some of the most frequent problems with data quality and possible solutions.

c. Deep fakes Finally, a very dangerous element of deep learning is actually related to the unquestionable strength of large models to generate seemingly real content. They genuinely reproduce the characteristics of the training data in now, often undetectable ways. These generative models are nowadays widely used to disseminate deep fakes of images, videos or voices [90–92]. In fact, a recent survey that compared real photographs and generated portraits, found that deep learning generated faces were believed to be real significantly more often than photos of actual persons [62]. The authors of the study assume that the neural network learned to combine key features of faces in a statistically perfectly averaged way, whereas real persons sometimes feature “outlier” characteristics (like a slightly crooked nose or asymmetric eyes or ears). Such characteristics of real faces were associated by the test-persons with generation errors of the neural network, whereas the used face generation network was sophisticated at a level to produce seemingly perfect images [92]. As another example, deep voice fakes were reportedly used by criminals to access the bank account of the CEO of a British energy company, relieving him by around a quarter million dollars [93].

Just as waves of AI generated fake news are washed ashore social media platforms [94], also in science these technologies make it increasingly easier to generate fraudulent results and flood scientific publication channels with fake research [95, 96]. This imminent danger can be compared in some ways with distributed denial of service (DDoS) attacks in computer networks. The fraudulent potential can already be impressively illustrated by the two-generations old, openly accessible language model GPT-2 [97]. It was specif-

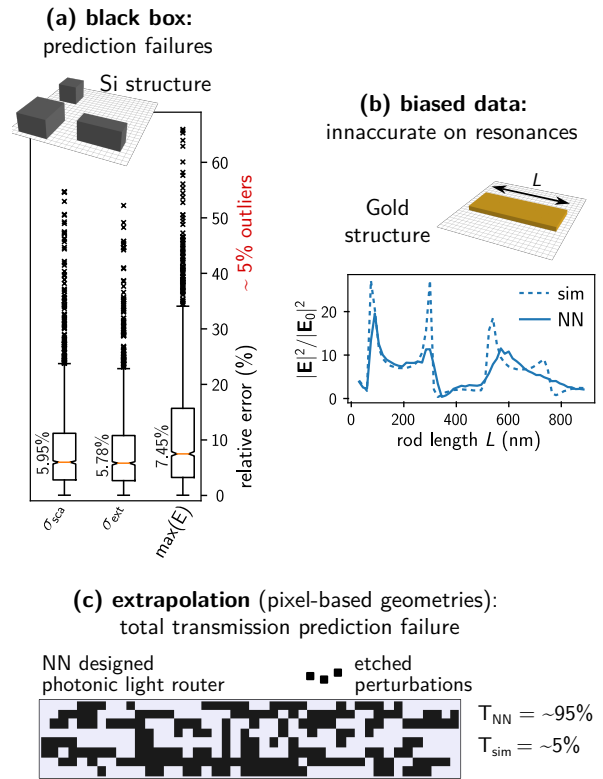


Figure 5. Illustrations of the potential impact of deep learning weaknesses on applications in nano-photonics. (a) In regression tasks, a certain percentage of predictions is often significantly worse than the statistical average error. Adapted with permission from [53]. Copyright 2020 American Chemical Society. (b) Infrequent physical phenomena, for example resonances, can easily be under represented in a dataset. Predictions will be inaccurate in such cases. Adapted with permission from [53]. Copyright 2020 American Chemical Society. (c) Extrapolation to parameter regimes outside the dataset range very generally leads to total prediction failures. Here the network learned photonic structures with up to 20 perturbations that typically feature very high optical transmission. The design network extrapolated in the shown example to a device with > 150 perturbations, leading in reality to almost zero transmission of the device. Adapted with permission from [68]. Copyright 2021 American Chemical Society.

ically fine-tuned on arxiv pre-print abstracts, rendering it capable to produce perfectly convincing, yet fake research descriptions. I invite the readers to have a look for themselves on <https://thispaperdoesnotexist.netlify.app/>, keeping in mind that the latest generations of LLMs are incomparably more capable in their reasoning proficiency and processing capacity. Modern models are able to process and generate tens to hundreds of pages of self-consistent, high quality text [3].

VI. STRENGTHS

Apart from quite a number of risks and dangers that a responsible user should be aware of, there is a reason for the current hype around deep learning. The approach does offer a number of very interesting strengths. A selection is illustrated

in figure 4b.

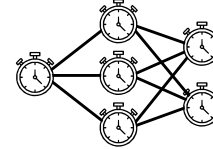
a. Problems that are hard for conventional algorithms Deep learning is often very effective at solving problems that are extremely hard to approach with conventional algorithms. In fact, today’s AI hype started with an application on such a type of problem: Image classification. Since 2012, no other algorithm could beat deep learning methods in this problem category [10, 98]. Countless further examples could be found alone in computer vision. This ranges from classification tasks like tomato plant disease recognition from images of leaves [63], over painting style extraction and its transfer to other images [99] to entirely generative tasks like bicycle design [100].

b. Very specific problems Similarly, yet to be distinguished from the former point, deep learning performs exceptionally well and is typically very easy to apply on problems that can be defined very specifically. This point is particularly true in the case of small and moderate dataset sizes as the amount of required data scales non-linearly with the complexity of a problem. Network models that generalize well usually require gigantic datasets and large efforts in hyper-parameter optimization.

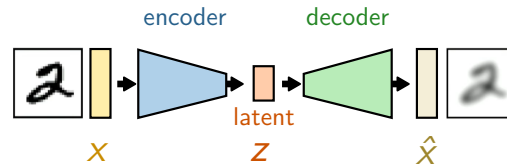
“Fine-tuning” or “transfer learning” are common strategies that exploit the fact that very specific problems are easy to learn. Both of these terms describe similar approaches in which models that have previously been pre-trained on a very large, generic data set are, in a second step, trained on a much smaller data set that contains entirely new samples. The goal is to learn a new, very specific task, either with the same input and output dimensions (fine tuning), or with data of different dimensions that are still of a similar type of problem (transfer learning). This is used for example in computer vision, where pre-trained models learn to interpret images in general [101], and fine-tuning is used subsequently to learn identification of specific, new objects [102, 103]. In natural language processing, fine-tuning is used on models that were pre-trained on a large, yet unspecific data corpus, to learn performing very specific tasks, for instance following explicit, complex instructions, instead of predicting just the next word in a text [104]. In physics and photonics, fine-tuning or transfer learning can be used for example to first teach the general physics to a neural network model through “cheap” simulated or analytically calculated data, and fine-tune it subsequently on specific experimental results that are “expensive” to generate [105]. Another application of transfer learning is to migrate concepts between different applications of similar physics, for example from multi-layered planar to multi-shell spherical geometries [106].

c. Low stake, rapid approximations A great strength of deep learning is its capability to deliver rough approximations, fast. In time critical applications where stakes are low, this can be an outstanding advantage. An everyday application, that many of us have grown used to, are smartphone tools like automatic spelling correcting or next word prediction [107]. It is very useful to get such propositions fast, and a wrongly suggested next word is of no big harm. In photonics, equivalent applications can be tools for the rapid estimation of optical performances [108–110], to obtain a rough first suggestion

(a) inference speed



(b) the latent space



(c) differentiability

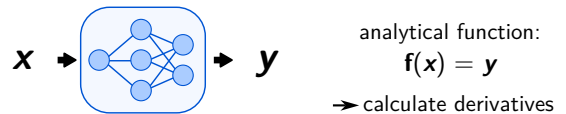


Figure 6. The key capabilities and concepts of deep learning in the opinion of the author. (a) the high inference speed of a trained model. (b) the concept of the latent space together with various available regularization schemes. (c) the analytical character of deep learning models, enabling for example physics informed neural networks or to learn differentiable empirical models even from experimental data. Picture of number “2” from the MNIST database of handwritten digits [116].

for photonic device designs [111–113], or up-scaling techniques for cheap microscopy equipment [114, 115]. But just like the predicted next word on our smartphone app is, such deep learning predictions should always be carefully double-checked.

VII. KEY CAPABILITIES

Beyond the above list of assets, I personally believe that three specific key attributes of deep learning models, are the most important strengths, that make deep learning stand out with respect to other (statistical) methods. Those points are summarized in figure 6.

a. Inference speed There’s little to add on this fundamental aspect of trained deep learning models: Their evaluation is in general very fast. This is probably the main key strength that is exploited in almost every application.

b. The latent space One of the, if not the, most crucial concept in deep learning is the idea of the latent space. A latent description of some thing or concept is a condensed, representative depiction of the object. For instance, the word “fish”, a fish-symbol or also the according Chinese letter are each in fact some latent description for this kind of animal. The appealing capacity of deep learning is, to automatically find latent descriptions via statistical analysis of a large dataset. This can even be done on unlabeled, raw data, us-

ing unsupervised learning approaches like variational autoencoders (VAEs) or generative adversarial networks (GANs) [117, 118], often used to tackle inverse problems in photonic materials design [119, 120]. Such approaches are similar to classical principle component analysis [121], yet the nonlinear, hierarchical character of deep learning models renders them potentially far more powerful. Using appropriate regularization techniques, it is possible to extract from the latent space actually meaningful information, like identifying specific properties of persons on portrait pictures (like age, the amount, length or color of hair, the gender, ...) [122–124]. The latent space is also the key concept behind generative deep learning. It can be regularized such that every point in a learned latent space corresponds to meaningful generated content or enable smooth interpolation between distant samples in latent space [4, 118, 125].

In physics, learning latent representations can be used for example to identify meaningful coordinate systems. For centuries people believed that the earth is the center of the solar system, yes even of the universe. The debate whether it wouldn't rather be the sun came up, when describing the trajectories of the other planets turned out to be a nightmare in the geocentric coordinate system. With their trajectories forming almost perfect circles in a heliocentric model, the search for adequate coordinate systems has proven to be of crucial importance. Training a network with low dimensional latent space can be helpful to identify such coordinate systems, in which better description of high dimensional, raw observations is possible [126], or it allows dimensionality reduction in other problems [127–129]. Such reduced representation then often allows an easier identification of correlations in a large dataset. One can also include a further, small neural network, into the latent space of a larger mode. This can for instance allow to learn mappings between different latent descriptions, for example to understand system dynamics (by mapping static to dynamic representations) or to discover linear approximations for non-linear physics phenomena [130, 131].

A latent space can be learned to categorize data, otherwise difficult to parametrize. In nanophotonics for example, low-dimensional parametrizations can be learned from top-view images of planar nanostructures or of meta-atoms for photonic metasurfaces [41, 70, 71]. Furthermore, thanks to the analytical nature of the entire deep learning pipeline (which is a requirement for training via backpropagation), it is possible to include operations like geometric transformations in a model or to work on dynamic discretizations like non isotropic meshes (e.g. from finite element method simulations). Such mesh description can either be learned, or implemented algorithmically, for example via graph neural networks or via hard-coded coordinate transforms [132–135]

The latent space does not necessarily need to describe a single property or a single physical object. So-called multi-modal latent models can be built, that map multiple observables, representations or objects into the same latent space [136]. Multi modal physics models can for instance combine different optical properties like far-field and near-field in a single latent space to facilitate according mappings. It

prompt

Formulate a results paragraph for a nano-photonics paper discussing the following:

- single crystalline gold, chemical synthesized flakes
- helium ion milling created gold nanostructures
- individually positioned using AFM tip
- designed a 5x5 micrometers metasurface
- demonstrate to focus incident 600nm light with NA 0.5

chat-GPT output (GPT 3.5)

In this study, we present the successful fabrication and characterization of a novel metasurface based on single crystalline gold flakes synthesized via a chemical route. These gold flakes were precisely shaped into intricate nanostructures using helium ion milling, allowing for the controlled manipulation of their geometry. Remarkably, the individual positioning of these gold nanostructures was achieved with unparalleled precision using an Atomic Force Microscopy (AFM) tip, enabling the creation of a metasurface spanning 5x5 micrometers. Our results indicate that this metasurface demonstrates the unique capability to effectively focus incident 600nm light with a numerical aperture (NA) of 0.5. **These findings underscore the potential of our approach for advancing nano-photonics applications and open new avenues for the design and engineering of metasurfaces with superior light manipulation capabilities.**

Figure 7. Case study to test the capabilities of chat GPT as an editing assistant in scientific research article writing. Top: Bullet point list summarizing a fictitious photonics materials research is used as input together with the instruction to formulate a results paragraph for a paper in nano-photonics. Bottom: The text generated by chat-GPT (freely accessible version, running on GPT-3.5). Remarkably, chat-GPT inserted novelty claims and exaggerated praising (italic, underlined) as well as a meaningless babble concluding sentence (bold), not given in the brief summary. Experiment performed in October 2023.

could also be possible to combine optical properties of photonic devices with other physical properties like mechanical or thermodynamical attributes. Related approaches have been recently used to infer the optical response in a specific frequency range from data of a different frequency regime [137], or to understand physics across multiple length scales in one model [138, 139].

c. Differentiability The third key property of deep learning is the analytic nature of the approach. Remember that an artificial neural network is nothing else than a gigantic mathematical function. In consequence, a deep learning model is analytically differentiable. This is the key behind so-called physics informed neural networks (PINNs) [72, 140] and neural operators [141–143], which aim at learning an approximating to the solution of partial differential equations (PDEs). Using the analytical derivatives of the network predictions, these models can test the validity of a solution within a PDE

to arbitrary accuracy, without the need of pre-calculated data. In optics, this concept can be used to find approximate solutions to Maxwell's equations [144], the Helmholtz or wave equation [145, 146], but also for physics-based regularization of data-based models [74].

Also aside from PINNs and neural operator networks, the differentiability is a very appealing property. Actually, if enough experimental data is available, but the observations cannot be described with a conventional model, deep learning offers the possibility to train an analytical, empirical model based on actual experimental observation. Such differentiable model allows more sophisticated applications than numerical simulations, for example derivatives can help to assess the robustness of solutions against small perturbations [147]. It also enables gradient based inverse design [41, 71, 148]. And since the entire calculation pipeline is necessarily analytical and differentiable, it is possible to include other kinds of mathematical operations in the data processing, for example geometry transformations, to render problem descriptions or parametrizations more compact and efficient [132].

d. Writing and editorial tools Last but not least, I believe that aside from direct applications in physics or research on photonic materials, today's deep learning tools can be extremely useful for the daily routine in research. Especially LLMs like chat-GPT can be powerful writing tools, in particular for non-english native speakers, that struggle with the language barrier for the dissemination of their research – provided they are used in an honest way [149].

An example to illustrate both the potential as well as the dangers associated with such application is shown in figure 7. Chat-GPT was instructed to write a paper paragraph given a bullet point list, that describes a fictitious research project about a plasmonic metasurface. The generated text is pertinent and contains all provided information in an eloquently written English. However, the writing style follows more what would be expected for an abstract and not for a results paragraph. Still, this may be due to the lack of sufficient information. It is far more alarming though, that the model inserted various exaggerated novelty and performance assertions that were not given in the prompt (highlighted by underlining). Finally, chat-GPT even added meaningless prattle to conclude its paragraph (indicated in boldface) – quite apparently the language model was trained on a corpus of very typical research articles.

VIII. CONCLUSIONS

Since the invention of the computer, numerical tools like statistical data processing or physics simulations have enabled unimaginable research insights and opened access to new knowledge [150, 151]. However, computers and technology in general have accelerated the world drastically. This has a strong impact on all aspects of our lives [152]. Driven by technology and digitalization, also research is accelerating fast, the number of published articles is growing exponentially every year [42]. Today, deep learning starts to be used in more and more fields of research – and deep learning is clearly a

tool, *designed to accelerate*. This development is likely to put additional pressure on every participant in the system, leading to even more self-amplification of these accelerations. Moreover, methods of increased efficiency are usually accompanied by a high level of abstraction. This often implies a certain detachment from reality, which in turn carries serious ethical risks [153].

To conclude I would therefore like to raise some questions: What impact will the increased use of accelerating tools like deep learning have on the thoroughness, the integrity and eventually on the quality of research? What will be the impact on the meaningfulness of average scientific publications? Will scientists be soon overtaken by the operation speed of their numerical tools? Or will big-data driven research rather lead to spectacular new discoveries, comparable to achievements enabled by the era of computer-based simulations?

In order to find positive answers to these these questions, I believe it is advisable to be prudent. Especially regarding the dissemination of effusive claims about the capabilities of deep learning methods in science. I am convinced that deep learning is not the solution to everything. It is rather merely a tool, yet unquestionable a powerful one, with tremendous potential for numerous applications. Particularly in an age of abundant available data, it bears the potential to reveal hidden correlations, for example through latent space methods, or to allow developing differentiable empirical models from raw observations. In the meanwhile it is essential to keep double-checking results, since deep learning models are black boxes, impossible to be categorically trusted. As a last comment, I would like to remind that deep learning cannot create knowledge from nowhere, but relies on hidden information in large datasets. Ideally, its capability to correlate, classify and interpolate intricate data will inspire researchers in their critical thinking and as a result guide the human scientists to the discovery of new phenomena.

ACKNOWLEDGMENTS

I thank Arnaud Arbouet, Christian Girard, Otto L. Muskens and Aurélien Cuche for fruitful discussions and the Toulouse HPC CALMIP for their support through grant p20010. I acknowledge funding from the French Agence Nationale de la Recherche (ANR) under the grants ANR-22-CE24-0002 (project NAINOS) and ANR-22-CE42-0021 (project VERDICT).

Appendix A: Data quality

As discussed above, high quality data is the foundation of any good deep learning model, and problems with data can have a huge impact on model performance. Furthermore, as data problems stand in the very beginning of the process, they can be very hard to identify in the final model [154]. Therefore I want to briefly discuss common problems with the data and possible approaches to alleviate them:

Missing parameter regions / Biased data If parameter regions are uncovered by the training data, the model will extrapolate in the inference task. This leads to poor, faulty results. Similarly, an over-representation of certain regions in parameter space can lead to reproduction of these biases. In cases where it is hard to estimate if such problems exist, machine learning techniques can be used to assess whether a model extrapolates (e.g. comparing latent space projections of predictions and representative training set samples).

Redundancies: If some samples are occurring multiple times, it distorts the statistics, making these samples appear more important than they actually may be. There are more or less sophisticated approaches to identify and remove duplicates, depending for instance on whether duplicates are identical copies or have slightly different signatures [155].

Errors / Outliers: Too many such samples break the statistical assumption of the training. Again, latent space embeddings can be used to remove outliers. Data that clusters in latent space are kept, while samples between clusters are removed [156, 157]. In such cases it is crucial to make sure that the data between clusters are actually outliers and not in fact relevant for the problem. Similarity learning or metric learning can help to assess data with difficult statistical properties [158].

Noise: With noisy data, the predictions may become noisy as well, reducing the accuracy of the results. But in case of random noise, the statistics are still valid, so generating more (noisy) data will improve the situation. Enough training data will then eventually even result in smooth pre-

dictions. It may also be possible to apply denoising before training.

Inconsistent data: These could be for example incorrect labels, non-converged simulations, a systematic temporal drift in a measurement setup, imperceptible random components of the data source (e.g. in stock market prizes [159]), data that partly comes from other generative models [160] etc... It is important to identify such problems since they potentially break the correlations between input and output features. Statistical correlation tests can give a hint of such problems [161]. Also machine learning methods such as confident learning exist, that can be used to assess data label quality [162].

Irrelevant information: Samples that are not related with the actual task can reduce the accuracy on the target problem. This can be particularly problematic with small models and little data. Beginning with a thorough definition of the task can help avoiding such problems and is actually crucial in low-data scenarios.

Quite generally, a few preventive recommendations can be given that will help avoiding the most frequent problems with data quality.

- Define the problem well, including a clear data format
- Think about data verification early in the process
- Perform (conventional) statistical analysis of your data
- Verify and document your data sources

Extensive further discussions on data problems and countermeasures can be found in literature [163].

-
- [1] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, in *Advances in Neural Information Processing Systems*, Vol. 300 (2020) pp. 1877–1901, [arxiv:2005.14165](https://arxiv.org/abs/2005.14165).
- [2] N. Team, M. R. Costa-jussà, J. Cross, O. Çelebi, M. Elbayad, K. Heafield, K. Heffernan, E. Kalbassi, J. Lam, D. Licht, J. Maillard, A. Sun, S. Wang, G. Wenzek, A. Youngblood, B. Akula, L. Barrault, G. M. Gonzalez, P. Hansanti, J. Hoffman, S. Jarrett, K. R. Sadagopan, D. Rowe, S. Spruit, C. Tran, P. Andrews, N. F. Ayan, S. Bhosale, S. Edunov, A. Fan, C. Gao, V. Goswami, F. Guzmán, P. Koehn, A. Mourachko, C. Ropers, S. Saleem, H. Schwenk, and J. Wang, “No Language Left Behind: Scaling Human-Centered Machine Translation,” (2022), [arxiv:2207.04672](https://arxiv.org/abs/2207.04672) [cs].
- [3] S. Bubeck, V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz, E. Kamar, P. Lee, Y. T. Lee, Y. Li, S. Lundberg, H. Nori, H. Palangi, M. T. Ribeiro, and Y. Zhang, “Sparks of Artificial General Intelligence: Early experiments with GPT-4,” (2023), [arxiv:2303.12712](https://arxiv.org/abs/2303.12712) [cs].
- [4] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, “High-Resolution Image Synthesis with Latent Diffusion Models,” (2022), [arxiv:2112.10752](https://arxiv.org/abs/2112.10752) [cs].
- [5] D. Podell, Z. English, K. Lacey, A. Blattmann, T. Dockhorn, J. Müller, J. Penna, and R. Rombach, “SDXL: Improving Latent Diffusion Models for High-Resolution Image Synthesis,” (2023), [arxiv:2307.01952](https://arxiv.org/abs/2307.01952) [cs].
- [6] S. Sanyal, A. Khandekar, and S. Sanyal, Reuters (2023).
- [7] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, 2016).
- [8] S. Chakraborty, R. Tomsett, R. Raghavendra, D. Harborne, M. Alzantot, F. Cerutti, M. Srivastava, A. Preece, S. Julier, R. M. Rao, T. D. Kelley, D. Braines, M. Sensoy, C. J. Willis, and P. Gurram, in *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI)* (2017) pp. 1–6.
- [9] T. J. Sejnowski, *Proceedings of the National Academy of Sciences* **117**, 30033 (2020).
- [10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, *Advances in Neural Information Processing Systems* **25**, 1097 (2012).
- [11] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* (2016) pp. 4278–4284, [arxiv:1602.07261](https://arxiv.org/abs/1602.07261).
- [12] I. Sutskever, O. Vinyals, and Q. V. Le, in *Advances in Neural Information Processing Systems*, Vol. 27, edited by Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger (Curran Associates, Inc., 2014) [arxiv:1409.3215](https://arxiv.org/abs/1409.3215) [cs].

- [13] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention Is All You Need,” (2017), [arxiv:1706.03762 \[cs\]](https://arxiv.org/abs/1706.03762).
- [14] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis, *Nature* **596**, 583 (2021).
- [15] L. Ferreira, C. J. Conselice, K. Duncan, T.-Y. Cheng, A. Griffiths, and A. Whitney, *The Astrophysical Journal* **895**, 115 (2020).
- [16] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, *Nature* **529**, 484 (2016).
- [17] I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, *Light: Science & Applications* **7**, 60 (2018).
- [18] M. Ziatdinov, O. Dyck, A. Maksov, X. Li, X. Sang, K. Xiao, R. R. Unocic, R. Vasudevan, S. Jesse, and S. V. Kalinin, *ACS Nano* **11**, 12742 (2017).
- [19] D. Liu, Y. Tan, E. Khoram, and Z. Yu, *ACS Photonics* **5**, 1365 (2018).
- [20] W. Ma, F. Cheng, and Y. Liu, *ACS Nano* **12**, 6326 (2018).
- [21] P. R. Wiecha, A. Lecestre, N. Mallet, and G. Larrieu, *Nature Nanotechnology* **14**, 237 (2019).
- [22] Z. Liu, D. Zhu, S. P. Rodrigues, K.-T. Lee, and W. Cai, *Nano Letters* **18**, 6570 (2018).
- [23] S. An, C. Fowler, B. Zheng, M. Y. Shalaginov, H. Tang, H. Li, L. Zhou, J. Ding, A. M. Agarwal, C. Rivero-Baleine, K. A. Richardson, T. Gu, J. Hu, and H. Zhang, *ACS Photonics* **6**, 3196 (2019), [arxiv:1906.03387](https://arxiv.org/abs/1906.03387).
- [24] F. Wen, J. Jiang, and J. A. Fan, *ACS Photonics* **7**, 2098 (2020), [arxiv:1911.13029](https://arxiv.org/abs/1911.13029).
- [25] L. Lu, R. Pestourie, W. Yao, Z. Wang, F. Verdugo, and S. G. Johnson, *SIAM Journal on Scientific Computing* **43**, B1105 (2021).
- [26] S. An, B. Zheng, M. Y. Shalaginov, H. Tang, H. Li, L. Zhou, Y. Dong, M. Haerinia, A. M. Agarwal, C. Rivero-Baleine, M. Kang, K. A. Richardson, T. Gu, J. Hu, C. Fowler, and H. Zhang, *Advanced Optical Materials* , 2102113 (2021), [arxiv:2102.01761](https://arxiv.org/abs/2102.01761).
- [27] T. Gahlmann and P. Tassin, *Physical Review B* **106**, 085408 (2022).
- [28] S. Gladyshev, T. D. Karamanos, L. Kuhn, D. Beutel, T. Weiss, C. Rockstuhl, and A. Bogdanov, *Nanophotonics* **12**, 3767 (2023).
- [29] H.-X. Xu, G. Hu, X. Kong, Y. Shao, P. Genevet, and C.-W. Qiu, *Light: Science & Applications* **12**, 78 (2023).
- [30] J.-S. Park, S. W. D. Lim, A. M. Amirzhan, H. Kang, K. Karrfalt, D. Kim, J. Leger, A. M. Urbas, M. Ossiander, Z. Li, and F. Capasso, “All-glass 100 mm Diameter Visible Metalens for Imaging the Cosmos,” (2023), [arxiv:2307.08186 \[astro-ph, physics:physics\]](https://arxiv.org/abs/2307.08186).
- [31] G. Palermo, A. Lininger, A. Guglielmelli, L. Ricciardi, G. Nicoletta, A. De Luca, J.-S. Park, S. W. D. Lim, M. L. Meretska, F. Capasso, and G. Strangi, *ACS Nano* **16**, 16539 (2022).
- [32] D. T. Jones and J. M. Thornton, *Nature Methods* **19**, 15 (2022).
- [33] T. Kanade, C. Thorpe, and W. Whittaker, in *Proceedings of the 1986 ACM Fourteenth Annual Conference on Computer Science*, CSC '86 (Association for Computing Machinery, New York, NY, USA, 1986) pp. 71–80.
- [34] C. Thorpe, M. Herbert, T. Kanade, and S. Shafer, *IEEE Expert* **6**, 31 (1991).
- [35] T. M. Mitchell, *AI Magazine* **18**, 11 (1997).
- [36] J. Michels, A. Saxena, and A. Y. Ng, in *Proceedings of the 22nd International Conference on Machine Learning*, ICML '05 (Association for Computing Machinery, New York, NY, USA, 2005) pp. 593–600.
- [37] K. Bengler, K. Dietmayer, B. Farber, M. Maurer, C. Stiller, and H. Winner, *IEEE Intelligent Transportation Systems Magazine* **6**, 6 (2014).
- [38] S. Grigorescu, B. Trasnea, T. Cocias, and G. Macesanu, *Journal of Field Robotics* **37**, 362 (2020).
- [39] T. Mickle, C. Metz, and Y. Lu, *The New York Times* (2023).
- [40] T. Higgins, *Wall Street Journal* (2022).
- [41] A. Khairah-Walieh, D. Langevin, P. Bennet, O. Teytaud, A. Moreau, and P. R. Wiecha, *Nanophotonics* **aop** (2023), [10.1515/nanoph-2023-0527](https://arxiv.org/abs/2307.08618), [arxiv:2307.08618 \[physics\]](https://arxiv.org/abs/2307.08618).
- [42] L. Bornmann and R. Mutz, *Journal of the Association for Information Science and Technology* **66**, 2215 (2015).
- [43] J. Jiang, M. Chen, and J. A. Fan, *Nature Reviews Materials* **6**, 679 (2021), [arxiv:2007.00084](https://arxiv.org/abs/2007.00084).
- [44] M. K. Chen, X. Liu, Y. Sun, and D. P. Tsai, *Chemical Reviews* (2022), [10.1021/acs.chemrev.2c00012](https://arxiv.org/abs/2022.10.1021/acs.chemrev.2c00012).
- [45] P. Covington, J. Adams, and E. Sargin, in *Proceedings of the 10th ACM Conference on Recommender Systems*, RecSys '16 (Association for Computing Machinery, New York, NY, USA, 2016) pp. 191–198.
- [46] W. Fedus, B. Zoph, and N. Shazeer, *The Journal of Machine Learning Research* **23**, 120:5232 (2022).
- [47] A. Chowdhery, S. Narang, J. Devlin, M. Bosma, G. Mishra, A. Roberts, P. Barham, H. W. Chung, C. Sutton, S. Gehrmann, P. Schuh, K. Shi, S. Tsvyashchenko, J. Maynez, A. Rao, P. Barnes, Y. Tay, N. Shazeer, V. Prabhakaran, E. Reif, N. Du, B. Hutchinson, R. Pope, J. Bradbury, J. Austin, M. Isard, G. Gur-Ari, P. Yin, T. Duke, A. Levskaya, S. Ghemawat, S. Dev, H. Michalewski, X. Garcia, V. Misra, K. Robinson, L. Fedus, D. Zhou, D. Ippolito, D. Luan, H. Lim, B. Zoph, A. Spiridonov, R. Sepassi, D. Dohan, S. Agrawal, M. Omernick, A. M. Dai, T. S. Pillai, M. Pellat, A. Lewkowycz, E. Moreira, R. Child, O. Polozov, K. Lee, Z. Zhou, X. Wang, B. Saeta, M. Diaz, O. Firat, M. Catasta, J. Wei, K. Meier-Hellstern, D. Eck, J. Dean, S. Petrov, and N. Fiedel, “PaLM: Scaling Language Modeling with Pathways,” (2022), [arxiv:2204.02311 \[cs\]](https://arxiv.org/abs/2204.02311).
- [48] D. Yalalov, “GPT-4’s Leaked Details Shed Light on its Massive Scale and Impressive Architecture,” (2023).
- [49] H. Touvron, L. Martin, K. Stone, P. Albert, A. Almahairi, Y. Babaei, N. Bashlykov, S. Batra, P. Bhargava, S. Bhosale, D. Bikel, L. Blecher, C. C. Ferrer, M. Chen, G. Cucurull, D. Esiobu, J. Fernandes, J. Fu, W. Fu, B. Fuller, C. Gao, V. Goswami, N. Goyal, A. Hartshorn, S. Hosseini, R. Hou, H. Inan, M. Kardas, V. Kerkez, M. Khabsa, I. Kloumann, A. Korenev, P. S. Koura, M.-A. Lachaux, T. Lavril, J. Lee, D. Liskovich, Y. Lu, Y. Mao, X. Martinet, T. Mihaylov, P. Mishra, I. Molybog, Y. Nie, A. Poulton, J. Reizenstein, R. Rungta, K. Saladi, A. Schelten, R. Silva, E. M. Smith, R. Subramanian, X. E. Tan, B. Tang, R. Taylor, A. Williams, J. X. Kuan, P. Xu, Z. Yan, I. Zarov, Y. Zhang, A. Fan, M. Kambadur, S. Narang, A. Rodriguez, R. Stojnic, S. Edunov, and T. Scialom, “Llama 2: Open Foundation and Fine-Tuned Chat

- Models,” (2023), [arxiv:2307.09288 \[cs\]](#).
- [50] W. Knight, *Wired* (2023).
- [51] E. Strubell, A. Ganesh, and A. McCallum, “Energy and Policy Considerations for Deep Learning in NLP,” (2019), [arxiv:1906.02243 \[cs\]](#).
- [52] R. Waters and T. Bradshaw, *Financial Times* (2023).
- [53] P. R. Wiecha and O. L. Muskens, *Nano Letters* **20**, 329 (2020), [arxiv:1909.12056](#).
- [54] J. Noh, Y.-H. Nam, S.-G. Lee, I.-G. Lee, Y. Kim, J.-H. Lee, and J. Rho, *Photonics and Nanostructures - Fundamentals and Applications* **52**, 101071 (2022).
- [55] T. Ma, M. Tobah, H. Wang, and L. J. Guo, *Opto-Electronic Science* **1**, 210012 (2022).
- [56] A. Khairah-Walieh, A. Arnoult, S. Plissard, and P. R. Wiecha, *Crystal Growth & Design* **23**, 892 (2023).
- [57] G.-X. Liu, J.-F. Liu, W.-J. Zhou, L.-Y. Li, C.-L. You, C.-W. Qiu, and L. Wu, *Nanophotonics* **12**, 1943 (2023).
- [58] D. Gostimirovic, Y. Grinberg, D.-X. Xu, and O. Liboiron-Ladouceur, *ACS Photonics* **10**, 1953 (2023).
- [59] T. Ma, H. Wang, and L. J. Guo, “OptoGPT: A Foundation Model for Inverse Design in Optical Multilayer Thin Film Structures,” (2023), [arxiv:2304.10294 \[physics\]](#).
- [60] Z. Zhang, C. Yang, Y. Qin, H. Feng, J. Feng, and H. Li, *Nanophotonics* **12**, 3871 (2023).
- [61] I. J. Goodfellow, J. Shlens, and C. Szegedy, “Explaining and Harnessing Adversarial Examples,” (2015), [arxiv:1412.6572 \[cs, stat\]](#).
- [62] S. J. Nightingale and H. Farid, *Proceedings of the National Academy of Sciences* **119**, e2120481119 (2022).
- [63] M. Agarwal, A. Singh, S. Arjaria, A. Sinha, and S. Gupta, *Procedia Computer Science International Conference on Computational Intelligence and Data Science*, **167**, 293 (2020).
- [64] H. Robbins and S. Monro, *The Annals of Mathematical Statistics* **22**, 400 (1951).
- [65] Y. Deng, S. Ren, K. Fan, J. M. Malof, and W. J. Padilla, *Optics Express* **29**, 7526 (2021).
- [66] Y. Deng, S. Ren, J. Malof, and W. J. Padilla, *Photonics and Nanostructures - Fundamentals and Applications* **52**, 101070 (2022).
- [67] R. Balestriero, J. Pesenti, and Y. LeCun, “Learning in High Dimension Always Amounts to Extrapolation,” (2021), [arxiv:2110.09485](#).
- [68] N. J. Dinsdale, P. R. Wiecha, M. Delaney, J. Reynolds, M. Ebert, I. Zempakis, D. J. Thomson, G. T. Reed, P. Lalanne, K. Vynck, and O. L. Muskens, *ACS Photonics* **8**, 283 (2021), [arxiv:2009.11810](#).
- [69] D. Melati, Y. Grinberg, M. Kamandar Dezfouli, S. Janz, P. Cheben, J. H. Schmid, A. Sánchez-Postigo, and D.-X. Xu, *Nature Communications* **10**, 4775 (2019).
- [70] Z. Liu, L. Raju, D. Zhu, and W. Cai, *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* **10**, 126 (2020), [arxiv:1902.02293](#).
- [71] Y. Augenstein, T. Repän, and C. Rockstuhl, *ACS Photonics* **10**, 1547 (2023), [arxiv:2302.01934 \[physics\]](#).
- [72] M. Raissi, P. Perdikaris, and G. E. Karniadakis, *Journal of Computational Physics* **378**, 686 (2019).
- [73] A.-P. Blanchard-Dionne and O. J. F. Martin, *Optics Letters* **45**, 2922 (2020).
- [74] M. Chen, R. Lupoiu, C. Mao, D.-H. Huang, J. Jiang, P. Lalanne, and J. A. Fan, *ACS Photonics* **9**, 3110 (2022).
- [75] J. Su, D. V. Vargas, and S. Kouichi, *IEEE Transactions on Evolutionary Computation* **23**, 828 (2019), [arxiv:1710.08864](#).
- [76] N. Liu, M. Du, R. Guo, H. Liu, and X. Hu, *ACM SIGKDD Explorations Newsletter* **23**, 86 (2021).
- [77] A. Qayyum, M. Usama, J. Qadir, and A. Al-Fuqaha, *IEEE Communications Surveys & Tutorials* **22**, 998 (2020).
- [78] B. Nassi, J. Shams, R. B. Netanel, and Y. Elovici, “bAdvertisement: Attacking Advanced Driver-Assistance Systems Using Print Advertisements,” (2022), [arxiv:2202.10080 \[cs\]](#).
- [79] K. Greshake, S. Abdelnabi, S. Mishra, C. Endres, T. Holz, and M. Fritz, “Not what you’ve signed up for: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection,” (2023), [arxiv:2302.12173 \[cs\]](#).
- [80] C. Molnar, *Interpretable Machine Learning. A Guide for Making Black Box Models Explainable*, 2nd ed. (2022).
- [81] X. Li, H. Xiong, X. Li, X. Wu, X. Zhang, J. Liu, J. Bian, and D. Dou, *Knowledge and Information Systems* **64**, 3197 (2022).
- [82] G. Ras, N. Xie, M. van Gerven, and D. Doran, *Journal of Artificial Intelligence Research* **73**, 329 (2022).
- [83] C. Shorten and T. M. Khoshgoftaar, *Journal of Big Data* **6**, 60 (2019).
- [84] R. Balestriero, L. Bottou, and Y. LeCun, “The Effects of Regularization and Data Augmentation are Class Dependent,” (2022), [arxiv:2204.03632 \[cs, stat\]](#).
- [85] M. T. Ribeiro, S. Singh, and C. Guestrin, ““Why Should I Trust You?”: Explaining the Predictions of Any Classifier,” (2016), [arxiv:1602.04938 \[cs, stat\]](#).
- [86] T. Gebru, J. Krause, Y. Wang, D. Chen, J. Deng, E. L. Aiden, and L. Fei-Fei, *Proceedings of the National Academy of Sciences* **114**, 13108 (2017).
- [87] J. Zou and L. Schiebinger, *Nature* **559**, 324 (2018).
- [88] Y. Xia, X. Cao, F. Wen, G. Hua, and J. Sun, in *Proceedings of the IEEE International Conference on Computer Vision* (2015) pp. 1511–1519.
- [89] D. Hendrycks, M. Mazeika, and T. Dietterich, “Deep Anomaly Detection with Outlier Exposure,” (2019), [arxiv:1812.04606 \[cs, stat\]](#).
- [90] T. Karras, T. Aila, S. Laine, and J. Lehtinen, [arXiv:1710.10196 \[cs, stat\]](#) (2017), [arxiv:1710.10196 \[cs, stat\]](#).
- [91] E. Zakharov, A. Shysheya, E. Burkov, and V. Lempitsky, “Few-Shot Adversarial Learning of Realistic Neural Talking Head Models,” (2019), [arxiv:1905.08233 \[cs\]](#).
- [92] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, “Analyzing and Improving the Image Quality of StyleGAN,” (2020), [arxiv:1912.04958 \[cs, eess, stat\]](#).
- [93] C. Stupp, *Wall Street Journal* (2019).
- [94] P. Wang, R. Angarita, and I. Renna, in *Companion Proceedings of the The Web Conference 2018*, WWW ’18 (International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 2018) pp. 1557–1561.
- [95] M. Májovský, M. Černý, M. Kasal, M. Komarc, and D. Netuka, *Journal of Medical Internet Research* **25**, e46924 (2023).
- [96] F. R. Elali and L. N. Rachid, *Patterns* **4**, 100706 (2023).
- [97] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever, *et al.*, *OpenAI blog* **1**, 9 (2019).
- [98] N. O. Mahony, S. Campbell, A. Carvalho, S. Harapanahalli, G. Velasco-Hernandez, L. Krpalkova, D. Riordan, and J. Walsh, “Deep Learning vs. Traditional Computer Vision,” (2019), [arxiv:1910.13796](#).
- [99] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, [arXiv:1703.10593 \[cs\]](#) (2017), [arxiv:1703.10593 \[cs\]](#).
- [100] L. Regenwetter, B. Curry, and F. Ahmed, *Journal of Mechanical Design* **144** (2021), 10.1115/1.4052585.
- [101] R. Bachmann, D. Mizrahi, A. Atanov, and A. Zamir, in *Computer Vision – ECCV 2022*, Lecture Notes in Computer

- Science, edited by S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, and T. Hassner (Springer Nature Switzerland, Cham, 2022) pp. 348–367.
- [102] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, *Computational Intelligence and Neuroscience* **2016**, e3289801 (2016).
- [103] H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, *BMC Medical Imaging* **22**, 69 (2022).
- [104] Y. Wang, Y. Kordi, S. Mishra, A. Liu, N. A. Smith, D. Khashabi, and H. Hajishirzi, “Self-Instruct: Aligning Language Models with Self-Generated Instructions,” (2022), [arxiv:2212.10560](https://arxiv.org/abs/2212.10560).
- [105] A. Ivanov and I. Agapov, *Physical Review Accelerators and Beams* **23**, 074601 (2020).
- [106] Y. Qu, L. Jing, Y. Shen, M. Qiu, and M. Soljačić, *ACS Photonics* **6**, 1168 (2019).
- [107] A. Hard, C. M. Kiddon, D. Ramage, F. Beaufays, H. Eichner, K. Rao, R. Mathews, and S. Augenstein, “Federated Learning for Mobile Keyboard Prediction,” (2018).
- [108] C. C. Nadell, B. Huang, J. M. Malof, and W. J. Padilla, *Optics Express* **27**, 27523 (2019).
- [109] R. S. Hegde, *IEEE Journal of Selected Topics in Quantum Electronics* **26**, 1 (2020).
- [110] C. Majorel, C. Girard, A. Arbouet, O. L. Muskens, and P. R. Wiecha, *ACS Photonics* **9**, 575 (2022), [arxiv:2110.02109](https://arxiv.org/abs/2110.02109).
- [111] A. Estrada-Real, A. Khairah-Walieh, B. Urbaszek, and P. R. Wiecha, *Photonics and Nanostructures - Fundamentals and Applications* **52**, 101066 (2022), [arxiv:2207.03431](https://arxiv.org/abs/2207.03431) [physics].
- [112] Y. Ha, Y. Luo, M. Pu, F. Zhang, Q. He, J. Jin, M. Xu, Y. Guo, X. Li, X. Li, X. Ma, and X. Luo, “Physics-data-driven intelligent optimization for large-scale meta-devices,” (2023), [arxiv:2306.01978](https://arxiv.org/abs/2306.01978) [physics].
- [113] A. Luce, A. Mahdavi, H. Wankerl, and F. Marquardt, *Machine Learning: Science and Technology* **4**, 015014 (2023).
- [114] Y. Rivenson, H. Ceylan Koydemir, H. Wang, Z. Wei, Z. Ren, H. Günaydin, Y. Zhang, Z. Göröcs, K. Liang, D. Tseng, and A. Ozcan, *ACS Photonics* **5**, 2354 (2018).
- [115] P. Sadeghli Dizaji, H. Habibiyan, and H. Arabalibeik, *Photonics and Nanostructures - Fundamentals and Applications* **52**, 101057 (2022).
- [116] L. Deng, *IEEE Signal Processing Magazine* **29**, 141 (2012).
- [117] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, [arXiv:1406.2661](https://arxiv.org/abs/1406.2661) [cs, stat] (2014), [arxiv:1406.2661](https://arxiv.org/abs/1406.2661) [cs, stat].
- [118] D. P. Kingma and M. Welling, *Foundations and Trends in Machine Learning* **12**, 307 (2019), [arxiv:1906.02691](https://arxiv.org/abs/1906.02691).
- [119] S. So and J. Rho, *Nanophotonics* **8**, 1255 (2019), [arxiv:1903.08432](https://arxiv.org/abs/1903.08432) [Nanophotonics].
- [120] J. Jiang, D. Sell, S. Hoyer, J. Hickey, J. Yang, and J. A. Fan, *ACS Nano* **13**, 8872 (2019).
- [121] I. T. Jolliffe and J. Cadima, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **374**, 20150202 (2016).
- [122] T. Karras, S. Laine, and T. Aila, “A Style-Based Generator Architecture for Generative Adversarial Networks,” (2019), [arxiv:1812.04948](https://arxiv.org/abs/1812.04948) [cs, stat].
- [123] R. Abdal, Y. Qin, and P. Wonka, in *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2019) pp. 4432–4441.
- [124] R. Abdal, Y. Qin, and P. Wonka, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2020) pp. 8296–8305.
- [125] T. Sainburg, M. Thielk, B. Theilman, B. Migliori, and T. Gerner, “Generative adversarial interpolative autoencoding: Adversarial training on latent space interpolations encourage convex latent distributions,” (2019), [arxiv:1807.06650](https://arxiv.org/abs/1807.06650) [cs, stat].
- [126] K. Champion, B. Lusch, J. N. Kutz, and S. L. Brunton, *Proceedings of the National Academy of Sciences* **116**, 22445 (2019).
- [127] Y. Kiarashinejad, M. Zandehshahvar, S. Abdollahramezani, O. Hemmatyar, R. Pourabolghasem, and A. Adibi, *Advanced Intelligent Systems* **2**, 1900132 (2020), [arxiv:1909.07330](https://arxiv.org/abs/1909.07330).
- [128] R. Iten, T. Metger, H. Wilming, L. del Rio, and R. Renner, *Physical Review Letters* **124**, 010508 (2020).
- [129] M. Zandehshahvar, Y. Kiarashinejad, M. Zhu, H. Maleki, T. Brown, and A. Adibi, *ACS Photonics* **9**, 714 (2022), [arxiv:2102.04454](https://arxiv.org/abs/2102.04454).
- [130] B. Lusch, J. N. Kutz, and S. L. Brunton, *Nature Communications* **9**, 4950 (2018).
- [131] U. Fasel, J. N. Kutz, B. W. Brunton, and S. L. Brunton, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **478**, 20210904 (2022).
- [132] H. Gao, L. Sun, and J.-X. Wang, *Journal of Computational Physics* **428**, 110079 (2020).
- [133] L. Kuhn, T. Repán, and C. Rockstuhl, *APL Photonics* **8**, 036109 (2023).
- [134] E. Khoram, Z. Wu, Y. Qu, M. Zhou, and Z. Yu, *ACS Photonics* **10**, 892 (2023).
- [135] S. Deshpande, S. P. A. Bordas, and J. Lengiewicz, “MAGNET: A Graph U-Net Architecture for Mesh-Based Simulations,” (2023), [arxiv:2211.00713](https://arxiv.org/abs/2211.00713) [cs].
- [136] W. Guo, J. Wang, and S. Wang, *IEEE Access* **7**, 63373 (2019).
- [137] J. Chen, C. Qian, J. Zhang, Y. Jia, and H. Chen, *Nature Communications* **14**, 4872 (2023).
- [138] S. Wan, C. Gong, P. Zhong, B. Du, L. Zhang, and J. Yang, “Multi-scale Dynamic Graph Convolutional Network for Hyperspectral Image Classification,” (2019), [arxiv:1905.06133](https://arxiv.org/abs/1905.06133) [cs, eess, stat].
- [139] L. Lu, R. Pestourie, S. G. Johnson, and G. Romano, *Physical Review Research* **4**, 023210 (2022).
- [140] M. Raissi, A. Yazdani, and G. E. Karniadakis, *Science* **367**, 1026 (2020).
- [141] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar, “Fourier Neural Operator for Parametric Partial Differential Equations,” (2021), [arxiv:2010.08895](https://arxiv.org/abs/2010.08895) [cs, math].
- [142] Z. Li, H. Zheng, N. Kovachki, D. Jin, H. Chen, B. Liu, K. Azizzadenesheli, and A. Anandkumar, “Physics-Informed Neural Operator for Learning Partial Differential Equations,” (2023), [arxiv:2111.03794](https://arxiv.org/abs/2111.03794) [cs, math].
- [143] N. Kovachki, Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, and A. Anandkumar, “Neural Operator: Learning Maps Between Function Spaces,” (2023), [arxiv:2108.08481](https://arxiv.org/abs/2108.08481) [cs, math].
- [144] Z. Fang and J. Zhan, *IEEE Access* **8**, 24506 (2020).
- [145] B. Moseley, A. Markham, and T. Nissen-Meyer, [arXiv:2006.11894](https://arxiv.org/abs/2006.11894) [physics] (2020), [arxiv:2006.11894](https://arxiv.org/abs/2006.11894) [physics].
- [146] M. Zhelyeznyakov, J. Fröch, A. Wirth-Singh, J. Noh, J. Rho, S. Brunton, and A. Majumdar, *Communications Engineering* **2**, 1 (2023).
- [147] C. Yeung, J.-M. Tsai, B. King, Y. Kawagoe, D. Ho, M. Knight, and A. P. Raman, *ACS Photonics* **7**, 2309 (2020), [arxiv:2003.06075](https://arxiv.org/abs/2003.06075).
- [148] S. Ren, A. Mahendra, O. Khatib, Y. Deng, W. J. Padilla, and J. M. Malof, *Nanoscale* **14**, 3958 (2022), [arxiv:2009.12919](https://arxiv.org/abs/2009.12919).
- [149] C. Stokel-Walker and R. Van Noorden, *Nature* **614**, 214

- (2023).
- [150] P. Denning, D. Comer, D. Gries, M. Mulder, A. Tucker, A. Turner, and P. Young, *Computer* **22**, 63 (1989).
- [151] J. Thijssen, *Computational Physics* (Cambridge University Press, 2007).
- [152] S. Turkle, *Life on the Screen* (Simon and Schuster, 2011).
- [153] S. Turkle, *Simulation and Its Discontents* (MIT Press, 2009).
- [154] L. Budach, M. Feuerpfeil, N. Ihde, A. Nathansen, N. Noack, H. Patzlaff, F. Naumann, and H. Harmouch, “The Effects of Data Quality on Machine Learning Performance,” (2022), arxiv:2207.14529 [cs].
- [155] I. Koumarelas, L. Jiang, and F. Naumann, *Journal of Data and Information Quality* **12**, 15:1 (2020).
- [156] X. Z. Fern and W. Lin, *Statistical Analysis and Data Mining: The ASA Data Science Journal* **1**, 128 (2008).
- [157] T. Kieu, B. Yang, and C. S. Jensen, in *2018 19th IEEE International Conference on Mobile Data Management (MDM)* (2018) pp. 125–134.
- [158] A. Bellet, A. Habrard, and M. Sebban, “A Survey on Metric Learning for Feature Vectors and Structured Data,” (2014), arxiv:1306.6709 [cs, stat].
- [159] E. F. Fama, *Financial Analysts Journal* **21**, 55 (1965).
- [160] I. Shumailov, Z. Shumaylov, Y. Zhao, Y. Gal, N. Papernot, and R. Anderson, “The Curse of Recursion: Training on Generated Data Makes Models Forget,” (2023), arxiv:2305.17493 [cs].
- [161] P. M. Robinson, *Journal of Econometrics Econometric Modelling in Finance and Risk Management: An Overview*, **147**, 5 (2008).
- [162] C. G. Northcutt, L. Jiang, and I. L. Chuang, “Confident Learning: Estimating Uncertainty in Dataset Labels,” (2022), arxiv:1911.00068 [cs, stat].
- [163] D. Becker, T. D. King, and B. McMullen, in *2015 IEEE International Conference on Big Data (Big Data)* (2015) pp. 2644–2653.