Machine-Learning-Enabled Vectorial Opto-Magnetization Orientation

Weichao Yan, Zhongquan Nie,* Xunwen Zeng, Guohong Dai, Mengqiang Cai, Yun Shen, and Xiaohua Deng

Manipulation of light-induced magnetization has become a fundamentally hot topic with a potentially high impact for atom trapping, confocal and magnetic resonance microscopy, and data storage. The control of the magnetization orientation mainly relies on the direct methods composed of amplitude, phase, and polarization modulations of the incident light under the tight focusing condition, leaving the achievement of arbitrary desirable 3D magnetization orientation complicated, inflexible, and inefficient. Here, a facile approach called machine learning inverse design to achieve expected vectorial opto-magnetization orientation is proposed. This pathway is time-efficient and accurate to produce the demanded incident beam for arbitrary prescribed 3D magnetization orientation. It is highlighted that the machine learning method is not only applied for magnetization orientations, but also widely used in the control of magnetization structures.

1. Introduction

Discovery of the giant magnetoresistance\(^{[1]}\) greatly facilitated the longitudinal magnetization, enabling to accelerate the development of high-capacity magnetic storage devices. In the 1960s, the inverse Faraday effect (IFE)\(^{[2]}\) was firstly proposed to describe an optically induced magnetization in a nonabsorbing material. In this seminal work, the optically induced magnetization in nonabsorbing material Eu\(^{12+}\)-CaF\(_2\), a number of diamagnetic glasses, and several organic and inorganic liquids is observed and demonstrated by the IFE in both the theory and the experiment. The ultrafast non-thermal control of magnetization created by the IFE in an antiferromagnetic 3D-metal oxide DyFeO\(_3\) is also revealed experimentally.\(^{[3]}\) Apart from these pioneering works, a new physical phenomenon called all-optical helicity-dependent switching (AO-HDS) was experimentally discovered.\(^{[4]}\) Later, the behavior of AO-HDS is found that this phenomenon is not a trivial one and related to paramagnetic materials\(^{[2]}\) or nonmagnetic metals and antiferromagnetic metals\(^{[5]}\) or specific laser fluence in ferrimagnetic materials\(^{[6–8]}\). Owing to the high density, high-energy efficiency, and erasable feature, all-optical magnetic recording (AOMR), has emerged as a key component for the next generation storage technology of an ultra-high areal density.\(^{[8,9]}\) To improve the storage density, the circularly polarized beam was focused into a half-wavelength region under a high numerical aperture (NA) lens.\(^{[10–12]}\) For the purpose of potential practical application of the longitudinal magnetization, the magnetization needle,\(^{[13]}\) magnetization chain,\(^{[14]}\) and magnetization spot arrays\(^{[15,16]}\) were also produced by the control of the phase, amplitude, and polarization distributions of the incident beam.

The above achieved magnetization structures are confined to the longitudinal magnetization components, which limit the capacity of magnetization storage. Recently, a uniformly oriented in-plane magnetization was produced through focusing two counter-propagating vector beams.\(^{[17]}\) Furthermore, arbitrary 3D orientations were achieved by diverse methods, such as the reversing electric dipoles,\(^{[18,19]}\) raytracing models,\(^{[20,21]}\) and the axially symmetrical destruction.\(^{[22]}\) Although the inverse designing methods can be implemented by the reversing electric dipoles and the axially symmetrical destruction, the achieved magnetization orientation is a little not coincident with these means only in the tight focusing system of an ideal 4\(\pi\) apparatus. One significant key is needed to find the corresponding relation between the magnetization orientation and spin orientation in realistic tight focusing condition. And the method of the raytracing models is not restricted to the ideal tight focusing system, the corresponding relation is still needed to be revealed. Apart from
2. Theoretical Illustration for Achieving 3D Orientated Magnetization

The principle of vectorial magnetization orientation is described in Figure 1.

Figure 1a illustrates a feasible strategy for the generation of the magnetization spot with 3D orientation with the help of ray-tracing model. The incident beam consists of four kinds of beams, as written by the first column vector. The first kind of beam is an azimuthally polarized beam imposed with both $\pi/2$ phase delay and $\pi$-phase-step filter along the $x$-axis, as given by $j E_{ax} e^{j\phi} e_x$. Here, $j$ denotes the $\pi/2$ phase delay, and $e_x$ is the azimuthal base vector. The focal field in the focus for this kind of beam can be written by $j E_{ax} e_x$ in the first row of the first column vector, where $E_{ax}$ is the amplitude factor and $e_x$ is the Cartesian coordinate base vector along the $x$-axis. The second and third kinds of beams are azimuthally polarized beam with $\pi$-phase-step filter along the $y$-axis, which are expressed as $j E_{ay1} e^{j\phi} e_y$ and $j E_{ay2} e^{j\phi} e_y$ in the second row of the first column vector, Here, $E_{ay1}$ and $E_{ay2}$ are the corresponding amplitude factors, and $e_y$ is the Cartesian coordinate base vector along the $y$-axis. The last kind of beam is a purely radial beam, written as $E_{\rho} e^{j\phi} e_{\rho}$ in the third row of the first column vector, where $E_{\rho}$ and $e_{\rho}$ are the radial amplitude weight and the radial base vector, respectively. The focal field in the focus for the last kind of beam is a longitudinally polarized field written by $E_{\rho} e_{\rho}$ in the third row of the first column vector, where $E_{\rho}$ and $e_{\rho}$ are the amplitude factor and Cartesian coordinate base vector along the $z$-axis, respectively.

this factor, it is required to repeatedly search the corresponding relations for realizing numerous different magnetization orientations. On the whole, the reported methods to obtain 3D magnetization orientations lack of considerable flexibilities and high efficiency. To achieve the prescribed 3D magnetization orientation, what the required incident beam is like. This is an inverse question and a challenging problem, which is rarely considered in previous magnetization shapings. Nowadays, machine learning has become a powerful tool in handling various problems, for example, the space weather forecast and image completion. Specially, machine learning inverse design has been demonstrated to be an accurate and time-efficient approach for solving complicated inverse problems in information storage with high density, metasurface imager and recognizer, and 3D vectorial holography. Here, we will demonstrate the generation of 3D vectorial magnetization orientation through exploiting machine learning inverse design based on multilayer perceptron artificial neural networks (MANN).
According to the general definition of the IFE proportional to the vector product between two complex conjugate fields, the combination of the focal fields such as $jE_1 e_1$ and $E_2 e_2$ can induce magnetization component along the $z$-axis ($M_z$). Similarly, the combination of other focal fields can induce magnetization components along the $x$-axis and $y$-axis ($M_x$ and $M_y$). Via adjusting the weights of the mentioned four kinds of beams, the ratios between the magnetization components ($M_x$, $M_y$, and $M_z$) can be controlled, thereby achieving a magnetization spot with 3D dynamic orientation. It is pointed out that the four weights meet the vector product between two complex conjugate fields, the $\mathbf{M}$ercury component along the $z$-axis ($M_z$) is the incident beam composed of four sorts of beams represented by independent amplitude parameters ($E_{x0}$, $E_{y1}$, and $E_{y2}$), where $E_x$ is set as 1. The key point is to train the MANN to statistically learn the relationship between a given 3D vectorial magnetization orientation ($\mathbf{I}$) on the nonabsorbing magnetic-optic (MO) film and the 2D vector field distribution in the pupil plane ($\mathbf{O}$).

In the light of the above schematic to produce vectorial magnetization orientation, the input and output datasets are required. According to the Richards and Wolf diffraction theory,

$\text{19,20}$ the Cartesian components of the arbitrarily electric field near focus can be written as

$$
\mathbf{E}(\rho, \phi, z) = A \int_0^\infty \int_0^{2\pi} \sin \theta \sqrt{\cos \theta} T(\theta, \phi) \mathbf{E}_m(\theta, \phi) \times \exp[jkz \cos \theta + jk\rho \sin \theta \cos(\varphi - \phi)] d\theta d\phi
$$

(1)

with

$$
T = \begin{bmatrix}
1 + (\cos \theta - 1) \cos^2 \varphi & (\cos \theta - 1) \cos \varphi \sin \varphi & - \sin \theta \cos \varphi & \cos \theta

\cos \theta - 1 & \cos \varphi \sin \varphi & 1 + (\cos \theta - 1) \sin^2 \varphi & - \sin \theta \cos \varphi

\sin \theta \cos \varphi & \sin \varphi & \cos \theta & \sin \theta \sin \varphi
\end{bmatrix}
$$

(2)

In Equation (1), $\mathbf{E}_m(\theta, \varphi)$ denotes the incident electric field and $T(\theta, \varphi)$ implies the polarization transformation matrix due to the depolarization effect under the tight focusing condition. Besides, $\theta$ and $\varphi$ are the spherical coordinates in the pupil space, and $\rho$, $\phi$, and $z$ are the cylindrical coordinates in the focal region. When the focal fields impinge on the nonabsorbing and isotropic MO film, the light-induced magnetization on the MO film can be described by the IFE

$$
\mathbf{M} = j\gamma \mathbf{E} \times \mathbf{E}'
$$

(3)

In Equation (3), $\gamma$ is magneto-optical constant. $\mathbf{E}$ and $\mathbf{E}'$ indicate the focal electric field and its complex conjugate, respectively. Substituting Equation (1) into Equation (3), the magnetization induced by tightly focusing arbitrary incident light can be garnered.

3. Simulated Results Based on Machine Learning

As we all know, we can directly calculate the magnetization for arbitrary incident light. But alike some complex mathematic functions, the inverse function may be difficult to obtain or even nonexistent. To achieve the prescribed magnetization distribution, it is not easy to find the proper incident light. In order to solve this problem, the machine learning inverse design is adopted. The preacquired magnetizations and incident vector beams are regarded as input and output datasets, respectively. The input and output datasets are trained by our MANN model. In the MANN model, the general architecture is composed of an input layer with three input elements, an output layer with three output elements, and four hidden layers with 500 hidden neurons in each layer. The input datasets are normalized (minimal to maximal scaling) before being fed into the training model. A rectified linear unit is used as the activation function. The uniform distribution to initialize all weights for neurons in each layer is adopted. For this typical regression problem, the mean square error (MSE) is chosen as the loss function. In order to get the optimal weights for the MANN model, the Gradient Descent is selected as the optimizer.

According to the vector diffraction theory and the IFE on the MO film, we calculate the magnetizations for 250,000 incident beam samples. These data samples are split into two subdatasets: training (200,000) and testing (50,000). It is noted that the testing datasets are not used in the training, which is completely blind. In order to estimate the testing data, the hyperparameters such as the batch size, learning rate, and iteration are needed to be adjusted during the training. When the batch size is chosen as a larger value, the required number of iterations rapidly decreases and the speed of convergence is vastly accelerated. Moreover, the convergence tendency is along the dropping direction and the corresponding oscillations decrease. When the learning rate is selected as a larger value, the best state may be missed. On the contrary, if the learning rate is enough small, the best state can be reached at the expense of a large amount of time. Therefore, the learning rate should be selected as a moderate value. The convergence result is improved with the increase of the iterations. Fortunately, these hyperparameters are easily tuned. By choosing the optimal hyperparameters, better performance can be achieved.

In our optimizing MANN model, the proper hyperparameters batch size, learning rate, and the number of iterations are chosen as 1000, 0.001, and 5000, respectively. By making use of the trained machine learning model, the predicted values for the testing datasets are given. The predicted values of the amplitude parameters for the random 100 incident beam samples are presented in Figure 2.

In Figure 2a–c, both true values and predicted values for the amplitude parameters $E_{x0}$, $E_{y1}$, and $E_{y2}$ are plotted. In our calculation, the absolute values of $E_{x0}$, $E_{y1}$, and $E_{y2}$ are all confined to $[0, \sqrt{2}/2]$. From Figure 2, it is easily found that predicted values are close to the true values. The training and testing loss values represented by the MSE are plotted in Figure 3, where the number of iterations increases. It is seen from Figure 3 that the loss values on the training and testing datasets show the same variation trend as the iteration increases. First, the loss values decline rapidly when the iterations are the range of $[0, 100]$. Subsequently,
Figure 2. Comparison of the predicted values and the true values in the 100 data randomly extracted from the testing dataset. (a), (b), and (c) are for $E_{ax}$, $E_{ay1}$, and $E_{ay2}$, respectively.

Figure 3. Loss values represented by the mean squared error (MSE) in the training (red) and testing (blue) datasets as a function of the iteration indicated by epoch.

the loss values decrease in a quite low speed. At last, loss values do not decrease and converge in a stable small value (about 0.2%).

Combining with the results in Figures 2 and 3, the reasonability and availability of our MANN model are sufficiently proved.

To give a direct and vivid demonstration of the power in our MANN model, we give an example. Here, we choose one sample data. The electric amplitude factors are selected as $E_{ax} = 0.398$, $E_{ay1} = -0.611$, and $E_{ay2} = 0.467$. Meanwhile, the estimated electric amplitude factors are $E_{ax} = 0.385$, $E_{ay1} = -0.601$, and $E_{ay2} = 0.464$. These two groups of values are nearly close. According to these two groups of electric amplitude factors, the projections of magnetization intensities and magnetization orientations along orthogonal planes as well as the 3D isosurface with the 80 percentages of the maximum magnetization intensity and the corresponding surrounding 3D magnetization orientations are depicted in Figure 4. It is easily found from Figure 4 that the orientations of the magnetization are distributed in three orthogonal planes. And as seen from the 3D isosurface

Figure 4. a,b) The projections of magnetization intensity (color) and magnetization orientation (arrows) along orthogonal planes as well as the 3D isosurface (color) with the 80 percentages of the maximum magnetization intensity and the corresponding surrounding 3D magnetization orientations (arrows) for the testing values in one sample and its corresponding estimated values.
magnetization intensity surrounded by magnetization orientation, the magnetization orientations near the focus are along 3D direction.

More importantly, the magnetization distributions of intensities and orientations in Figure 4a,b are almost the same except for some trivial side lobes.

The 3D magnetization orientations in Figure 4a,b are vectors. Mathematically, the similarities and discrepancies between two vectors can be estimated by \( S_m = \cos(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a} \cdot \mathbf{b}}{||\mathbf{a}|| \cdot ||\mathbf{b}||} \). Here, \( \mathbf{a}, \mathbf{b} \) denotes the angle between two vectors (\( \mathbf{a} \) and \( \mathbf{b} \)). When the angle between the two vectors is equal to 0, the cosine of the angle is 1. In this case, the directions of the two vectors are identical. When the angle between the two vectors is equal to 90 degrees, the cosine of the angle is 0. In other words, the two vectors are perpendicular. Therefore, \( S_m \) ranging from 0 to 1 can be applied to estimate the similarities between two different vectors. The true magnetization orientation (\( M_x, M_y, M_z \)) in the focus is \((-0.559, -0.479, -0.560)\). And the estimated magnetization orientation is \((M_{xe}, M_{ye}, M_{ze})=(-0.570, -0.475, -0.546)\). The cosine of the angle between the true and estimated magnetization orientations can be calculated by our definition \( S_m \). And the calculated \( S_m \) is equal to 0.9998. This value is in close proximity to 1, which means that the true and estimated magnetization orientations are nearly identical.

To further perform a reliable examination of the quality about the trained neural network, some different input configurations which are not contained in the training and testing datasets are used as input values. The corresponding electric fields are predicted by our well trained neural network. According to the IFE in the MO film, the final estimated magnetization orientations are calculated. The desired magnetization orientations (\( M_x, M_y, M_z \)) and the estimated magnetization orientations (\( M_{xe}, M_{ye}, M_{ze} \)) are shown in Table 1. Besides, the similarities (\( S_m \)) between the desired magnetization orientations and the estimated magnetization orientations are also calculated. As seen from this table, most values of \( S_m \) reach above 0.96, which are very close to 1. Only one \( S_m \) is equal to about 0.85. It is a common phenomenon for a small number of worse samples to arise. This is because the machine learning guarantees the overall estimated effect for all training samples not only for a small number of samples. As a whole, the estimated and desired magnetization orientations are almost identical. This result again proves that the predictions based on our MANN model are quite accurate.

By adopting our proposed method, the required incident beam can be predicted for an expected magnetization orientation. In the experiment, the complex incident beam can be produced by vectorial optical field generator. Through the modulation of the incident beam, the expected magnetization orientation can be realized under the tight focusing condition. Magnetization spot arrays with controllable 3D orientations are produced by point-to-point scanning. Each spot in these magnetization distributions has independent magnetization orientation and can be controlled arbitrarily. The controllable magnetization orientation based on our machine learning algorithm greatly facilitates the formation of magnetic holography, which strengthens ultrahigh information security. The magnetic holography can be used as identity verification for a credit card with magnetic stripe. On account of the freedom of storage of the 3D magnetization orientation, the realized multistate could advance the development in multiplexing magnetization storage and information processing technologies, and it may also find applications in the detection of magnetic particles.

4. Conclusion

In conclusion, we put up with the machine learning inverse design to achieve vectorial magnetization orientation. According to the RMUs under tightly focusing condition and the IFE on the nonabsorbing MO film, a structured incident beam composed of four kinds of beams is focused to four polarized focal fields, allowing to produce a magnetization spot with 3D orientation. Based on the vector diffraction theory and the IFE, two groups of datasets comprised of the magnetizations and incident beams are garnered. These two groups of datasets are fed into the MANN model and trained by optimizing hyperparameters. After training, it is revealed that the predicted values are almost close to the testing values for the majority of samples. The loss value decreases with the increase of the iteration and finally converges to the stable value (0.2%). Importantly, the calculated magnetization distributions based on the testing and estimated values of one incident beam sample are almost the same. Besides, some prospects on how to experimentally realize the 3D magnetization and its interesting applications are discussed. It is highly expected that the proposed machine learning inverse design can open up broad potentials in magnetization shapings containing structures and orientations.

**Acknowledgements**

This work was supported by the National Natural Science Foundation of China (12004155, 11974258, 11904152, 11604236, 61575139, 61865009, 61927813, 60504052); Key Research and Development (R&D) Projects of Shanxi Province, China (201903D121217); Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi (2019L0151).

**Conflict of Interest**

The authors declare no conflict of interest.
Data Availability Statement

Research data are not shared.

Keywords

magnetization, machine learning, polarization, tight focusing

Received: June 16, 2021
Revised: October 21, 2021
Published online: December 2, 2021

[1] M. Baibich, J. Broto, A. Fert, V. D. F. Nguyen, J. Chazelas, Phys. Rev. Lett. 1988, 61, 2472.
[2] J. P. van der Ziel, P. S. Pershan, L. D. Malmstrom, Phys. Rev. Lett. 1965, 15, 190.
[3] A. V. Kimel, A. Kirilyuk, P. A. Usachev, R. V. Pisarev, A. M. Balbashov, Th. Rasing, Nature 2005, 435, 655.
[4] C. D. Stanciu, F. Hansteen, A. V. Kimel, A. Kirilyuk, A. Tsukamoto, A. Itoh, T. Rasing, Phys. Rev. Lett. 2007, 99, 047601.
[5] M. Berritta, R. Mondal, K. Carva, P. M. Oppeneer, Phys. Rev. Lett. 2016, 117, 137203.
[6] S. Wang, C. Wei, Y. Feng, Y. Cao, H. Wang, W. Cheng, C. Xie, A. Tsukamoto, A. Kirilyuk, T. Rasing, A. V. Kimel, X. Li, Appl. Phys. Lett. 2018, 113, 171108.
[7] A. Khorsand, M. Savoini, A. Kirilyuk, A. Kimel, A. Tsukamoto, A. Itoh, T. Rasing, Phys. Rev. Lett. 2012, 108, 127205.
[8] S. Mangin, M. Gottwald, C. Lambert, D. Steil, V. Uhl r, L. Pang, M. Hehn, S. Alebrand, M. Cinchetti, G. Malinowski, Y. Fainman, M. Aeschlimann, E. E. Fullerton, Nat. Mater. 2014, 13, 286.
[9] C.-H. Lambert, S. Mangin, B. C. S. Varaprasad, Y. Takahashi, M. Hehn, M. Cinchetti, G. Malinowski, K. Hono, Y. Fainman, M. Aeschlimann, E. E. Fullerton, Science 2014, 345, 1337.
[10] L. Helseth, Opt. Commun. 2008, 281, 5671.
[11] Y. Zhang, J. Bai, Phys. Lett. A 2008, 372, 6294.
[12] Y. Zhang, J. Bai, J. Opt. Soc. Am. B 2009, 26, 176.
[13] S. Wang, X. Li, J. Zhou, M. Gu, Opt. Lett. 2014, 39, 5022.
[14] Z. Nie, W. Ding, G. Shi, D. Li, X. Zhang, Y. Wang, Y. Song, Opt. Express 2015, 23, 21296.
[15] Z. Q. Nie, H. Lin, X. F. Liu, A. P. Zhai, Y. T. Tian, W. J. Wang, D. Y. Li, W. Q. Ding, X. R. Zhang, Y. L. Song, B. Jia, Light. Sci. Appl. 2017, 6, e17032.
[16] C. Hao, Z. Nie, H. Ye, H. Li, Y. Luo, R. Feng, X. Yu, F. Wen, Y. Zhang, C. Yu, J. Teng, B. Lukyanchuk, C.-W. Qiu, Sci. Adv. 2017, 3, e1701398.
[17] S. Wang, Y. Cao, X. Li, Opt. Lett. 2017, 42, 5050.
[18] S. Wang, J. Luo, Z. Zhu, Y. Cao, H. Wang, C. Xie, X. Li, Opt. Lett. 2018, 43, 5551.
[19] J. Luo, H. Zhang, S. Wang, S. Liu, Z. Zhu, B. Gu, X. Wang, X. Li, Opt. Lett. 2019, 44, 727.
[20] W. Yan, Z. Nie, X. Liu, X. Zhang, Y. Wang, Y. Song, Opt. Lett. 2018, 43, 3826.
[21] W. Yan, S. Lin, H. Lin, Y. Shen, Z. Nie, B. Jia, X. Deng, Opt. Exp. 2021, 29, 961.
[22] W. Yan, Z. Nie, X. Liu, X. Zhang, Y. Wang, Y. Song, APL Photonics 2018, 3, 116101.
[23] R. Tang, F. Zeng, Z. Chen, J.-S. Wang, C.-M. Huang, Z. Wu, Atmosphere 2020, 11, 316.
[24] Z. Chen, M. Jin, Y. Deng, J.-S. Wang, H. Huang, X. Deng, C.-M. Huang, J. Geophys. Res. Space Phys. 2019, 124, 790.
[25] P. R. Wiecha, A. Lecestre, N. Mallet, G. Larrieu, Nat. Nanotech. 2019, 14, 237.
[26] L. Li, Y. Shuang, Q. Ma, H. Li, H. Zhao, M. Wei, C. Liu, C. Hao, C.-W. Qiu, T. J. Cui, Light. Sci. Appl. 2019, 8, 1.
[27] H. Ren, W. Shao, Y. Li, F. Salim, M. Gu, Sci. Adv. 2020, 6, eaaz4261.
[28] H. Ren, X. Li, M. Gu, Opt. Lett. 2014, 39, 6771.
[29] B. Richards, E. Wolf, Proc. Roy. Soc. A 1959, 253, 358.
[30] K. Youngworth, T. Brown, Opt. Exp. 2000, 7, 77.
[31] W. Han, Y. Yang, W. Cheng, Q. Zhan, Opt. Exp. 2013, 21, 20692.