ACS APPLIED MATERIALS & INTERFACES

www.acsami.org

Research Article

Toward Machine-Learning-Accelerated Design of All-Dielectric Magnetophotonic Nanostructures

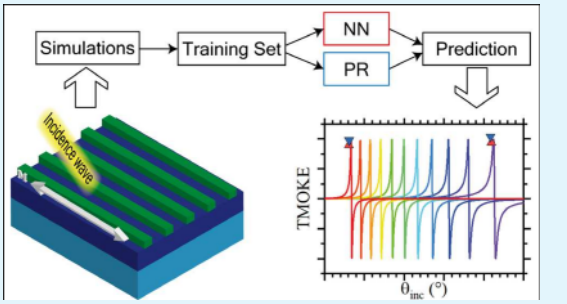
William O. F. Carvalho,* Marcio Tulio Aiex Taier Filho, Osvaldo N. Oliveira, Jr., Jorge Ricardo Mejía-Salazar, and Felipe Augusto Pereira de Figueiredo

Cite This: ACS Appl. Mater. Interfaces 2024, 16, 42828−42834

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: All-dielectric magnetophotonic nanostructures are promising for integrated nanophotonic devices with high resolution and sensitivity, but their design requires computationally demanding electromagnetic simulations evaluated through trial and error. In this paper, we propose a machine-learning approach to accelerate the design of these nanostructures. Using a data set of 12 170 samples containing four geometric parameters of the nanostructure and the incidence wavelength, trained neural network and polynomial regression algorithms were capable of predicting the amplitude of the transverse magneto-optical Kerr effect (TMOKE) within a time frame of 10^{-3} s and mean square error below 4.2%. With this approach, one can readily identify nanostructures suitable for sensing at ultralow analyte concentrations in aqueous solutions. As a proof of principle, we used the machine-learning models to determine the sensitivity ($S = |\Delta\theta^{res}/\Delta n_a|$) of a nanophotonic grating, which is competitive with state-of-the-art systems and exhibits a figure of merit of 672 RIU^{-1}. Furthermore, researchers can use the predictions of TMOKE peaks generated by the algorithms to assess the suitability for experimental setups, adding a layer of utility to the machine-learning methodology.



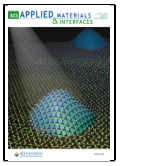
KEYWORDS: *all-dielectric, machine-learning-accelerated design, magnetophotonic, neural networks, polynomial regression, TMOKE sensing*

INTRODUCTION

Developing a new host of active optical devices depends upon the ability to design nanostructures whose optical properties can be modulated dynamically. In recent years, the optical response of magnetic materials has been tuned by exploiting magneto-optical (MO) effects,[1] which has led to breakthroughs in (bio)sensing,[1] routers/switches,[2] modulators,[3,4] circulators and isolators,[5−7] and buffering.[8] Because MO effects are weak in the nanoscale, enhancement has been sought through plasmonic resonance incorporating metallic building components. MO activity is increased in these magnetoplasmonic (MO metal) nanostructures that have enhanced, localized plasmonic fields (at the metal surface) distributed inside an adjacent MO layer.[9] However, energy efficiency is limited by the intrinsic joule heating in these magnetoplasmonic systems. This has motivated the use of all-dielectric magnetophotonic nanostructures,[10−12] which feature low levels of losses as a result of the lack of joule heating. The main challenge in designing these all-dielectric nanophotonic platforms is to identify the optical resonances with high MO amplitudes. Another difficulty is related to the confinement of optical fields within the scatterers with the highest refractive index (hindering interaction with nearby MO materials), in contrast to magnetoplasmonic approaches, where the field is confined at metal/dielectric surfaces.[9] Owing to such problems, current methodologies for designing all-dielectric MO nanostructures require time-consuming electromagnetic simulations, mostly with trial-and-error strategies. Additionally, the design of each nanostructure for a given magnetization state demands a meticulous analysis to attain the desired MO features, making the process laborious and iterative. Strategies are therefore needed to overcome these drawbacks when designing all-dielectric nanostructures with tailored MO properties.

One possible solution is to employ machine-learning (ML) algorithms that could identify nanostructures with optimized values by training with a limited number of electromagnetic simulations. In fact, neural networks (NNs) and other ML methods have proven effective for the design of core−shell nanoparticles[13] and metasurfaces,[14] in addition to other examples of material design and discovery that reduced the reliance on labor-intensive experiments and simulations.[15,16] While some prior works have employed ML techniques for

Received: April 24, 2024
Revised: July 8, 2024
Accepted: July 15, 2024
Published: July 30, 2024



optimizing MO devices, such as MO traps for cold atoms[17] and MO imaging devices,[18] this has not been the case for magnetophotonic nanostructures. In this paper, we report on a ML-based approach for the rapid and efficient design of all-dielectric MO nanostructures. Two ML models were used for comparison: NNs and polynomial regression (PR). These models were trained and validated using a data set of 15 213 (comprising 12 170 for training and 3043 for validation) samples to predict the amplitude of the transverse magneto-optical Kerr effect (TMOKE), which can be used in sensing and biosensing. In particular, the algorithms were not specifically trained to deliver designs with targeted TMOKE values. Instead, they were developed to provide researchers with the maximum achievable TMOKE amplitude and sensitivity for a given set of geometrical parameters in a nanostructure. Therefore, our methodology can be applied to various materials and geometries, allowing for broad applications and generalization in nanophotonic-based MO applications.

METHODS AND MATERIALS

The aim is to automate the design of nanostructures, such as the grating coupler in Figure 1, which comprises a one-dimensional

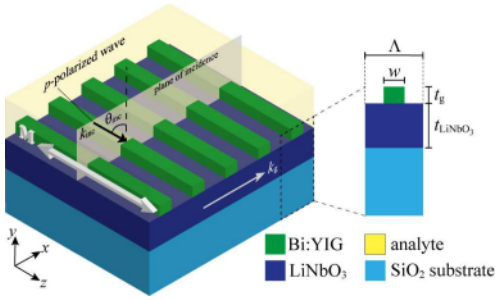


Figure 1. Schematic illustration of all-dielectric magnetophotonic gratings, with the geometric parameters in the inset.

periodic arrangement of dielectric MO ribs placed on a high refractive index (HRI) guiding layer, grown on a SiO_2 substrate. For a realistic simulation, we considered materials used experimentally, with MO ribs made of bismuth-substituted yttrium iron garnet (Bi:YIG)[19,20] and the HRI layer made of lithium niobate ($LiNbO_3$).[21] The working wavelength (λ) was varied from 570 to 670 nm, compatible with commercial setups. The corresponding permittivities (as functions of λ) ε_{SiO_2},[22] ε_{LiNbO_3},[21] and $\tilde{\varepsilon}_{Bi:YIG}$[23]

$$\tilde{\varepsilon}_{Bi:YIG} = \begin{pmatrix} \varepsilon_{Bi:YIG} & im\varepsilon_{xy} & 0 \\ -im\varepsilon_{yx} & \varepsilon_{Bi:YIG} & 0 \\ 0 & 0 & \varepsilon_{Bi:YIG} \end{pmatrix} \quad (1)$$

for SiO_2, $LiNbO_3$, and Bi:YIG, respectively, were used from experimental reports. Because Bi:YIG is an anisotropic material, its permittivity is represented by the tensor $\tilde{\varepsilon}_{Bi:YIG}$[23] in eq 1, where $\varepsilon_{Bi:YIG}$ and $\varepsilon_{xy} = \varepsilon_{yx}$ are the diagonal and off-diagonal tensor components in the transverse configuration ($\mathbf{M}\|z$, in relation to Figure 1).

Among the different MO configurations (polar, longitudinal, and transverse), we focused on the TMOKE as a result of its unique ability to preserve the polarization of the incident light, because only the amplitude of reflected/transmitted intensity is modulated.[24] Moreover, the sharp Fano-like curves of TMOKE are used for improved resolution in magnetoplasmonic biosensing and magnetometry,[25,26] which we also exploit here to illustrate the applicability of our concept. In the transmission mode, TMOKE is defined as[24]

$$\text{TMOKE} = 2\frac{T_p(+\mathbf{M}) - T_p(-\mathbf{M})}{T_p(+\mathbf{M}) + T_p(-\mathbf{M})} \quad (2)$$

where $T_p(\pm\mathbf{M})$ is the transmittance amplitude associated with $\mathbf{M}$ pointing along the $\pm z$ axis. The superstrate region, referred to as the analyte region in Figure 1, is considered to possess a refractive index n_a varying between 1.3 and 1.4 to illustrate the potential application in all-dielectric magnetophotonic biosensing platforms. The geometric parameters of the structure are illustrated in Figure 1, where Λ, w, and t_g are the period length, width, and thickness of the grating ribs, whereas t_{LiNbO_3} represents the thickness of the HRI layer. The full-wave (FW) numerical simulations were made using the finite element method (FEM) within the commercial software COMSOL Multiphysics. Floquet periodic boundary conditions were set on the lateral boundaries to consider an infinite periodic system along the x axis, while perfectly matched layers (PMLs) were used along the y boundaries to avoid undesired numerical reflections.

We compared two ML models for the automated design of the nanostructures, namely, NN and PR. The training and evaluation metric chosen to assess these ML models was the mean square error (MSE). The input data for the ML algorithms consists of a set of geometrical parameters (illustrated in the inset of Figure 1) and the incident wavelength. For computational purposes, the data set has undergone two preprocessing steps. The first step eliminated incoherent and divergent geometrical parameters, also known in the ML context as features. The second step normalized the features to enhance the performance of the ML models. Specifically, the features were normalized to the range $[-1, 1]$. In contrast, the output of the algorithms corresponds to the maximum TMOKE amplitudes (|TMOKE|) for the refractive indexes 1.3 and 1.4 and their corresponding angles, respectively. Given the input data, supervised training algorithms iteratively minimize the MSE between the expected values, e.g., TMOKE amplitudes (ranging from $[0, 2]$) and the angles (ranging from $[0°, 90°]$) corresponding to the selected refractive indexes, and the values predicted by the models. We first used the neural architecture search (NAS) technique to automate searching for the best NN architecture for this specific task.[27,28] Unlike the tedious and time-consuming trial-and-error process of finding a proper NN architecture, the NAS technique utilizes an optimization algorithm, such as greedy, Bayesian, and hyperband algorithm, to do so.[29] The evaluation metric chosen at the beginning of the NAS process is MSE. Then, a cyclic optimization algorithm generates candidate NN architectures, which are subsequently trained and assessed on the basis of the evaluation metric.[30] The latter process is illustrated in the left panel of Figure 2. The search process continues until a termination criterion, such as a trial limit, is reached. The final architecture is the one that demonstrates the best performance according to the evaluation metric. At the end of the

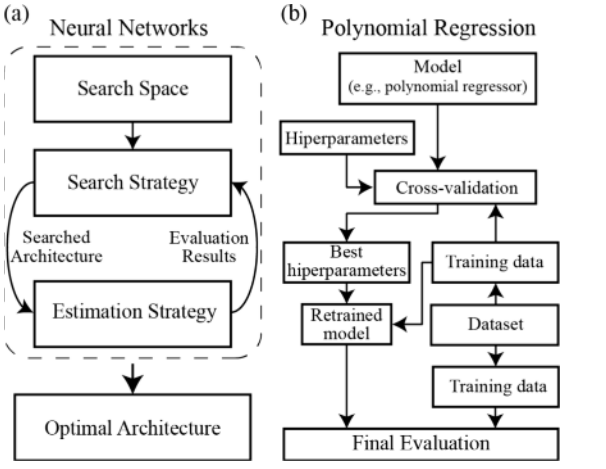


Figure 2. Working flowcharts of (a) NN and (b) PR algorithms.

https://doi.org/10.1021/acsami.4c06740
ACS Appl. Mater. Interfaces 2024, 16, 42828−42834

search, the optimized NN architecture is evaluated using a set of unseen samples, obtained through FW FEM simulations, to verify its generalization performance.[31] We used AutoKeras, an open-source NAS library.[31]

The other ML method, PR, is one of the oldest and most efficient ML techniques. It involves adjusting a polynomial function to the data samples, which requires finding the optimal polynomial degree. Cross-validation approaches can be used to discover the best values for hyperparameters, such as the degree of a polynomial function that can be applied to the problem being addressed.[32] They assess the performance of a model accurately by splitting the data set into multiple training and testing sets, allowing the model to be trained and tested multiple times. Cross-validation avoids the model overfitting to the training data set.[33] The procedure to find the best PR model is illustrated on the right panel of Figure 2; it employs the k-fold approach, which is a simple, efficient cross-validation method.[34] In the k-fold method, the data set is divided into k subsets or folds where $k-1$ subsets are used for model training and 1 subset is used for validation. The method iteratively trains models with different polynomial degrees on k training sets created out of the $k-1$ subsets left for training. After each training iteration, the MSE of the model is calculated using the current validation set. This procedure is repeated k times, resulting in k MSE values. These values are then averaged and used to estimate the variance of the MSE. The averaged MSE and its variance are the metrics used to compare the performance of this model against others. The previous procedure is repeated for each polynomial degree. Then, the MSE and variance values of all models are compared, and the degree resulting in the smallest values is selected as the best one for the PR model. Finally, the best PR model is also evaluated with the set of unseen samples used previously to assess its generalization capacity. This cross-validation procedure ensures a robust model with a suitable trade-off between underfitting and overfitting.[35]

RESULTS AND DISCUSSION

Optimization of all-dielectric magnetophotonic devices is a key requirement to reach responses suitable for applications, such as sensing. Trial-and-error strategies for experimental fabrication of the devices are not an option owing to the number of geometric parameters and materials involved. Computer simulations are then normally employed, but these are costly in terms of computational resources and time. We, therefore, exploit here an accelerated design optimization by combining ML methods and simulations. The data set used for training the ML algorithms comprises 12 170 samples, which are simulated properties of grating couplers as in Figure 1. Each sample is structured as a 1×5 array, with four geometrical parameters illustrated in the inset of Figure 1, and the operating wavelength λ. Thus, an individual sample, labeled "data instance", is represented as data = $[\Lambda w t_g t_{LiNbO_3} \lambda]$, where the corresponding values in the array consist of pseudorandom combinations of values for $215 \le \Lambda \le 255$ nm, $70 \le w \le 110$ nm, $43 \le t_g \le 65$ nm, $130 \le t_{LiNbO_3} \le 175$ nm, and $570 \le \lambda \le 670$ nm. Because the resonant angle/wavelength in magnetophotonic gratings normally varies linearly with n_a,[11] calculations of TMOKE were only made for $n_a = 1.3$ and 1.4 (i.e., the extreme values in the range). Hence, for each sample, there are two TMOKE values, namely, $TMOKE_1$ (for $n_a = 1.3$) and $TMOKE_2$ (for $n_a = 1.4$), and the resonant angle θ^{res} associated with TMOKE peaks ($TMOKE^{max}$), which were obtained using FW electromagnetic simulations by sweeping the incident angle in the range of $0° \le \theta_{inc} < 90°$. The ML models are then trained to predict these four target values, $|TMOKE_1^{max}|$, $|TMOKE_2^{max}|$, θ_1^{res}, and θ_2^{res}. The algorithms provide the values for $|TMOKE_1^{max}|$ and $|TMOKE_2^{max}|$, which range from 0 to 2 (according to eq 2). The values of θ_1^{res} and θ_2^{res}, ranging from 0° to 90°, are intricately associated with each $TMOKE_i^{max}$ through the resonant physical properties of the respective nanostructures. It is worth noting that the concept presented here is not limited to a particular set of materials, thus enabling its generalization for various combinations of materials and/or geometries.

Figure 3 depicts the Pearson correlation coefficient between all pairs of input features and targets. The correlation

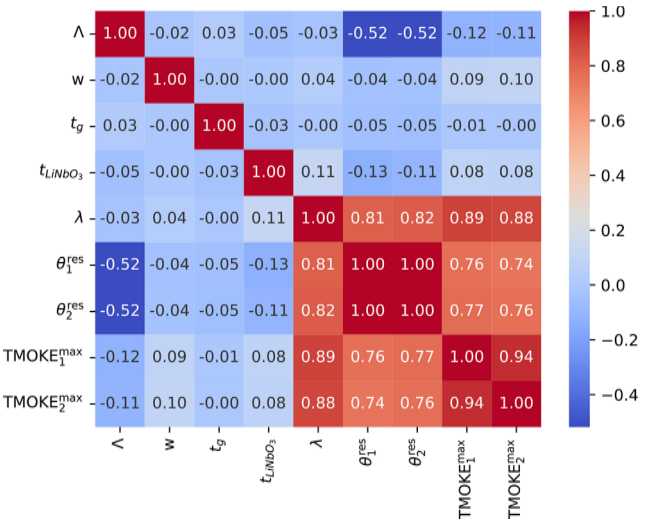


Figure 3. Correlation matrix between each of the input features and target values.

coefficient refers to the degree to which a pair of variables are linearly related. As seen, there is a high positive correlation (values equal or greater than 0.81) between the λ feature and all of the four targets, $|TMOKE_1^{max}|$, $|TMOKE_2^{max}|$, θ_1^{res}, and θ_2^{res}. A positive correlation indicates that, as the attribute value increases, the target values tend to increase. There is also a medium negative correlation of −0.52 between the Λ feature and targets θ_1^{res} and θ_2^{res}, indicating that, as the attribute value increases, the target value tends to decrease. Apart from that, features w, t_g, and t_{LiNbO_3} exhibit small to very small correlations with the four targets. However, as some experiments have shown, removing them from the training set has a negative impact on the MSE, increasing it when compared to the scenario where all of them are employed. Therefore, although they do not have a high correlation with the targets, these features might present nonlinear relationships with them that are not accounted for by the (linear) correlation coefficient. If the relationship between two variables is not linear, the correlation coefficient will not fully characterize their relationship. Additionally, because NN and PR models are nonlinear in the sense that they combine input features internally and, in the case of NNs, have nonlinear activation functions, they can explore complex relationships between them and the targets. The figure also shows that the input features are non-co-linear (i.e., the correlation between features is low). Therefore, it is not necessary to remove or create/transform features. Thus, we treat all features equally and employ all of them as input to the PR and NN models.

The most efficient NN and PR models identified had their performance validated using 3043 random samples out of 15 213 of the whole data set. For visualization purposes, Figure

https://doi.org/10.1021/acsami.4c06740
ACS Appl. Mater. Interfaces 2024, 16, 42828−42834

4a shows results only for the first 30 samples in the validation set, as indicated on the horizontal axis. The ML results appear

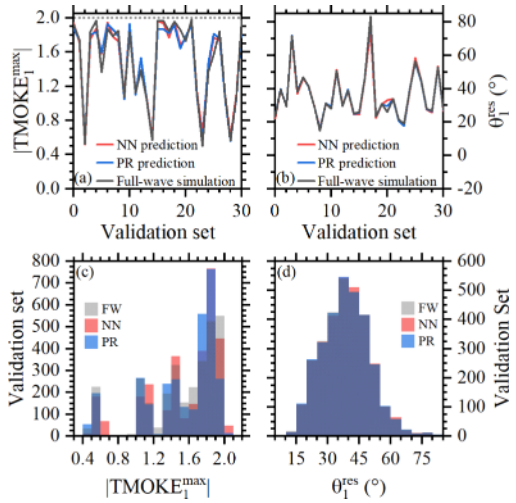


Figure 4. Validation of algorithms: (a) $TMOKE_1^{max}$ and (b) θ_1^{res} comparison among FW simulations (black) and NN (red) and PR (blue) predictions and validation histograms for FW simulations (gray), NN (red), and PR (blue) for (c) $TMOKE_1^{max}$ and (d) θ_1^{res}, respectively.

in red and blue for NN and PR, respectively, while the FW numerical simulations are shown in black. Because similar results were obtained for $|TMOKE_1^{max}|$ and $|TMOKE_2^{max}|$ and their corresponding resonant angles θ_1^{res} and θ_2^{res}, we only show results for the validation samples associated with $|TMOKE_1^{max}|$ and θ_1^{res} in panels a and b of Figure 4. Small differences for $TMOKE_1^{max}$ and θ_1^{res} are observed only in a few validation samples, as can be noted, indicating the feasibility of our approach. Panels c and d of Figure 4 present histograms offering a general overview comparison using all 3043 random validation samples. These histograms compare the distribution of the target values (i.e., results from the numerical simulations) against the distribution of the values predicted by the PR and NN models. As observed, the distributions of both predictions, PR and NN, for θ_1^{res} almost perfectly match the target distribution. Conversely, the distributions of the predictions for $TMOKE_1^{max}$ deviate somewhat from the target distribution, with the prediction distribution of the NN model being the closest. This difference might be explained by the fact that the $TMOKE_1^{max}$ target is highly correlated only with the λ feature, as seen in Figure 3, and the multimodal nature of this distribution.

Validation was also carried out using MSE, i.e., the mean of squares of the differences between FW simulations (Y_{FW}) and predicted (Y_p) values [$MSE = (Y_{FW} - Y_p)^2$], for NN (MSE_{NN}) and PR (MSE_{PR}) in panels a and b of Figure 5. The results indicate low MSE values, as corroborated by the larger number (2978) of data with $0 \leq MSE \leq 0.1$ in comparison to the PR model (2736). The MSE values calculated over the entire validation data set for the PR and NN models are equal to 0.1044 and 0.009688, respectively. Nevertheless, the PR model exhibits faster prediction times than the NN model, as seen from histograms in panels c and d of Figure 5. This is due to their different computational complexities, with NNs being more computationally demanding than PR models.

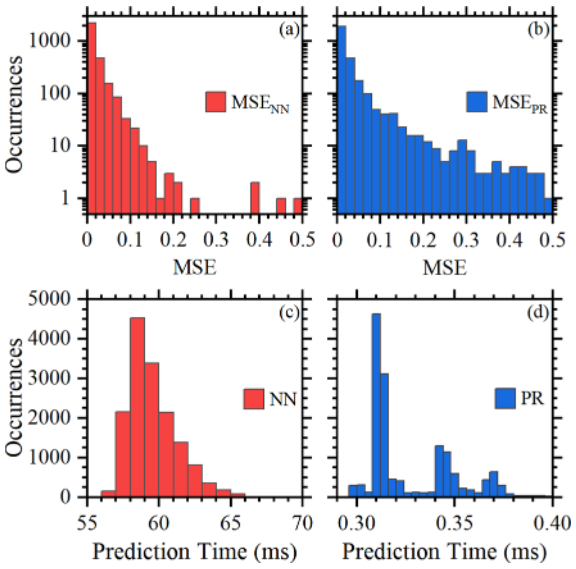


Figure 5. MSE histograms for the validation set using (a) NN and (b) PR, respectively, and inference time histograms for all available samples (15 213), i.e., for the training + validation sets, using the (c) NN and (d) PR methods.

The optimal NN found with NAS and used in this study consists of 3 hidden dense layers with 256, 128, and 64 nodes per layer, respectively, each utilizing ReLU activation functions. The output layer is a dense layer with 4 nodes and employs linear activation. The Adam optimizer is employed with a learning rate of 0.0001 and a batch size of 32 samples. The time complexity of the NN can be approximately represented as $O(\sum_{l=1}^{L} n_l^2)$, where L is the number of layers and n_l denotes the number of nodes in the lth layer.

On the other hand, the PR model is a seventh-order polynomial, which has a complexity of $O(n^d)$, where n is the number of input parameters and d is the polynomial order. Despite the differences in the prediction times and efficiency, both methods deliver suitable results in intervals of milliseconds, which are several orders of magnitude shorter than the hours or days spent on conventional design/simulation tasks.

As mentioned before, the PR model uses a cross-validation approach to select the optimal polynomial degree. With the data set divided into multiple folds and the model iteratively trained and validated, cross-validation ensures that the performance of the model is evaluated on different subsets of data. This technique helps prevent overfitting by ensuring that the model generalizes well across various data splits. The MSE on the validation set for the PR model with cross-validation is equal to 0.1044. For the NN, the NAS technique was employed to find the optimal architecture. NAS systematically explored different architectures and selected the architecture with the best performance based on the evaluation metric, which, in our study, is the minimization of the MSE on the validation data set. Minimizing the MSE on the validation set mitigates the risk of overfitting. This method ensures that the chosen NN architecture is well-suited for the task without being overly complex. The MSE on the validation set for the NN with the NAS technique is equal to 0.009688. Therefore, it is reasonable to say that both models have the necessary complexity to capture the underlying behavior behind the

https://doi.org/10.1021/acsami.4c06740
ACS Appl. Mater. Interfaces 2024, 16, 42828−42834

samples and generalize effectively. Among them, the NN model demonstrates superior generalization capacity, as evidenced by its lower MSE value.

With regard to the NAS trajectory, Tables S1 and S2 of the Supporting Information present all tested configurations. Each trial trained a NN with the respective configurations for 500 epochs. The trials carried out using the NAS technique highlighted several key findings: normalization significantly improved MSE minimization; deeper and wider networks with a decreasing number of units per layer generally performed better; batch normalization and dropout showed no improvement; the Adam optimizer often outperformed SGD and was generally the best optimizer; all attempts to use the Adam optimizer with weight decay failed; and models trained with a learning rate of 0.001 generally achieved lower MSE values, indicating better performance compared to those with significantly higher or lower learning rates. The best hyperparameters for the optimal model can be seen in trial 103 of Table S2 of the Supporting Information.

To demonstrate the suitability of the accelerated design, we used trained NN and PR models to determine the sensitivity ($S = |\Delta\theta^{res}/\Delta n_a|$) of a grating design not previously encountered in the training or validation data sets. We manually chose a nanostructure with the following geometrical parameters, which we anticipate could lead to high TMOKE responses: Λ = 215 nm, w = 70 nm, t_g = 65 nm, t_{LiNbO_3} = 130 nm, and λ = 650 nm. In further work, one may conceive combining our approach with an optimization process to select high-performance nanostructures. The transmittances $T_p(\pm\mathbf{M})$ and TMOKE amplitudes plotted as a function of the incident angle in Figure 6a for the analyte refractive index n_a = 1.33

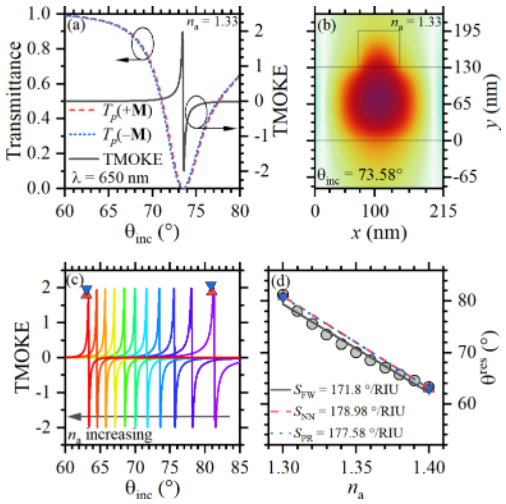


Figure 6. (a) Transmittances and TMOKE curves using n_a = 1.33, (b) H_z field coupled mode in the $LiNbO_3$ waveguide, (c) TMOKE responses with corresponding n_a, where the triangles are the NN (red) and PR (blue) predictions, and (d) sensitivity for FW (gray solid), NN (red dashed), and PR (blue dotted) lines.

(conventionally used for aqueous media) and λ = 650 nm coincide with the prediction using the ML approach. The near-maximum (~2) TMOKE amplitudes are attributed to the excitation of lateral leaky Bloch modes coupled with the guiding sublayer and the ribs of the MO grating.[11] Figure 6b shows a highly concentrated near field within the MO ribs, which is responsible for enhancing the MO activity. To quantify sensitivity, we varied the analyte refractive index and calculated the corresponding resonance shifts of the TMOKE peaks, as illustrated in Figure 6c. The TMOKE curves corresponding to n_a ranging from 1.30 to 1.40 are displayed from right to left. The normal and inverted triangles represent predicted values from the NN and PR methods, respectively, at the extreme n_a values. These predictions are consistent with full-wave electromagnetic simulation results, as noticed. The figure of merit (FoM) ranged from 672 RIU^{-1} for n_a = 1.30 to 2126 RIU^{-1} for n_a = 1.40. This variation highlights not only the high sensitivity of the designed nanostructures but also their spectral resolution capability. Such attributes are significant for biosensing at ultralow analyte concentrations. Finally, the sensitivity S_{FW} is determined through linear regression of TMOKE peak variations with n_a, as shown by the solid gray line in Figure 6d, which is compared to sensitivities predicted from the NN (S_{NN}) and PR (S_{PR}) methods in the same figure. The generalization ability of the ML-accelerated design is confirmed with the results in Table 1. A maximum relative

Table 1. Comparison of the TMOKE^{max}, θ^{res}, and RIU Values Obtained in Electromagnetic Simulations (FW) to Those Predicted with the ML Models NN and PR

	$TMOKE_1^{max}$	$TMOKE_2^{max}$	θ_1^{res} (deg)	θ_2^{res} (deg)	S (deg/RIU)
FW	1.9905	1.9775	81.20	63.32	171.8
NN	1.8812	1.7998	80.93	63.03	178.98
PR	2.0870	1.9637	80.82	63.06	177.58
error (NN) (%)	−5.49	−8.99	−0.33	−0.46	+4.18
error (PR) (%)	+4.84	−0.7	−0.47	−0.41	+3.36

error of −8.99% was obtained in NN and PR predictions for the TMOKE peaks $TMOKE_i^{max}$ and resonant angles θ_i^{res} for n_a = 1.30 (associated with i = 1) and n_a = 1.40 (associated with i = 2), along with their corresponding sensitivities.

CONCLUSION

We have devised a ML-aided approach for the accelerated design of all-dielectric magnetophotonic nanostructures. The robustness of the approach was confirmed by obtaining similar results with two distinct ML algorithms, whose predictions coincided with the results from simulations. With these ML-driven methodologies, one may circumvent the time-intensive and computationally demanding electromagnetic simulations. The ML algorithms were trained using a database of 12 170 samples containing five geometric parameters that define the nanostructure and the incident wavelength. The output was the $|TMOKE^{max}|$ and θ^{res} for two refractive indexes of the incident medium, i.e., n_a = 1.3 and 1.4, typical in biosensing in aqueous environments. It was also possible to assess sensitivity $S = |\Delta\theta_{res}/\Delta n_a|$, whose high value for some of the nanostructures justifies the accelerated design of all-dielectric magnetophotonic-based biosensors. It is significant that properties are predicted within milliseconds with ML algorithms, to be compared to 12 min for each full-wave electromagnetic simulation using the same computer settings. Furthermore, both training sets and algorithms can be further refined to identify nanostructure designs with optimized responses for specific working wavelengths, incidence angles, and materials tailored to the requirements of manufacturers. This versatility renders our approach valuable to experimen-

https://doi.org/10.1021/acsami.4c06740
ACS Appl. Mater. Interfaces 2024, 16, 42828−42834

talists who wish to fabricate magnetophotonic nanostructures. Moreover, we should remark that the ML models presented in this work generate results based on the geometry of grating nanostructures, independent of any specific set of materials. This flexibility means that the trained algorithms can be applied to geometries that fall within the range of parameters used in the training database. However, a limitation exists: ML-based regression models typically struggle to extrapolate beyond the scope of their training data. Consequently, they do not produce reliable predictions for structures with parameters outside this predefined range. Furthermore, while the algorithms are not tied to specific materials, the resonant features of each nanograting are inherently linked to the optical properties of the materials used. This relationship suggests that generalizing the approach is feasible. By incorporation of further training, the models can learn the behavior of resonant features for different sets of materials or geometries beyond those initially considered in this study.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsami.4c06740.

Training experimental results with the NAS technique (PDF)

■ AUTHOR INFORMATION

Corresponding Author

William O. F. Carvalho − *Sao Carlos Institute of Physics, University of Sao Paulo, São Carlos, São Paulo 13560-970, Brazil;* orcid.org/0000-0001-6178-0061; Email: williamofcarvalho@gmail.com

Authors

Marcio Tulio Aiex Taier Filho − *National Institute of Telecommunications (Inatel), Santa Rita do Sapucaí, Minas Gerais 37540-000, Brazil*

Osvaldo N. Oliveira, Jr. − *Sao Carlos Institute of Physics, University of Sao Paulo, São Carlos, São Paulo 13560-970, Brazil;* orcid.org/0000-0002-5399-5860

Jorge Ricardo Mejía-Salazar − *National Institute of Telecommunications (Inatel), Santa Rita do Sapucaí, Minas Gerais 37540-000, Brazil;* orcid.org/0000-0003-1742-9957

Felipe Augusto Pereira de Figueiredo − *National Institute of Telecommunications (Inatel), Santa Rita do Sapucaí, Minas Gerais 37540-000, Brazil*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsami.4c06740

Funding

The Article Processing Charge for the publication of this research was funded by the Coordination for the Improvement of Higher Education Personnel - CAPES (ROR identifier: 00x0ma614).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work has been partially funded by the projects XGM-AFCCT-2024-3-1-1 and XGM-AFCCT-2024-2-5-1 supported by the xGMobile−EMBRAPII−Inatel Competence Center on 5G and 6G Networks, with financial resources from the PPI IoT/Manufatura 4.0 from MCTI Grant 052/2023, signed with EMBRAPII. The authors also acknowledge financial support from the Brazilian agencies National Council for Scientific and Technological Development (CNPq, 403612/2020-9, 403827/2021-3, and 314671/2021-8), FAPESP (2023/08999-9 and 2018/22214-6), FAPEMIG (2070.01.0004709/2021-28, 2070.01.0007442/2023-48, APQ-03162-24, and PPE-00124-23), and RNP, with resources from MCTIC, Grant 01245.020548/2021-07, under the Brazil 6G project of the Radiocommunication Reference Center [Centro de Referência em Radiocomunicações (CRR)] of the National Institute of Telecommunications [Instituto Nacional de Telecomunicações (Inatel)], Brazil.

■ REFERENCES

(1) Rizal, C.; Manera, M. G.; Ignatyeva, D. O.; Mejía-Salazar, J. R.; Rella, R.; Belotelov, V. I.; Pineider, F.; Maccaferri, N. Magneto-photonics for sensing and magnetometry toward industrial applications. *J. Appl. Phys.* **2021**, *130*, 230901.

(2) Carvalho, W. O. F.; Mejía-Salazar, J. R. Magneto-optical micro-ring resonators for dynamic tuning of add/drop channels in dense wavelength division multiplexing applications. *Opt. Lett.* **2021**, *46*, 2396−2399.

(3) Neto, J. V. S.; Carvalho, W. O. F.; Mejía-Salazar, J. R. Magnetically Tunable Micro-Ring Resonators for Massive Magneto-Optical Modulation in Dense Wavelength Division Multiplexing Systems. *Sensors* **2022**, *22*, 8163.

(4) Nayak, B. K.; Buks, E. Polarization-selective magneto-optical modulation. *J. Appl. Phys.* **2022**, *132*, 193905.

(5) Yan, W.; Yang, Y.; Liu, S.; Zhang, Y.; Xia, S.; Kang, T.; Yang, W.; Qin, J.; Deng, L.; Bi, L. Waveguide-integrated high-performance magneto-optical isolators and circulators on silicon nitride platforms. *Optica* **2020**, *7*, 1555−1562.

(6) Portela, G.; Levy, M.; Hernandez-Figueroa, H. E. Magnetless optical circulator based on an iron garnet with reduced magnetization saturation. *Molecules* **2021**, *26*, 4692.

(7) Portela, G.; Levy, M.; Hernandez-Figueroa, H. E. Novel compact magnetless isolator based on a magneto-optical garnet material. *Opt. Laser Technol.* **2023**, *157*, 108638.

(8) Wang, H.; Wang, Z.; Fu, Z.; Jiang, R.; Liu, L.; Wang, J.; Wu, C. Silicon-on-Insulator Optical Buffer Based on Magneto-Optical 1 × 3 Micro-Rings Array Coupled Sagnac Ring. *Photonics* **2022**, *9*, 366.

(9) Maccaferri, N.; Zubritskaya, I.; Razdolski, I.; Chioar, I.-A.; Belotelov, V.; Kapaklis, V.; Oppeneer, P. M.; Dmitriev, A. Nanoscale magnetophotonics. *J. Appl. Phys.* **2020**, *127*, 080903.

(10) Carvalho, W. O. F.; Moncada-Villa, E.; Oliveira Jr, O. N.; Mejía-Salazar, J. R. Beyond plasmonic enhancement of the transverse magneto-optical Kerr effect with low-loss high-refractive-index nanostructures. *Phys. Rev. B* **2021**, *103*, 075412.

(11) Carvalho, W. O. F.; Mejía-Salazar, J. R. All-dielectric magnetophotonic gratings for maximum TMOKE enhancement. *Phys. Chem. Chem. Phys.* **2022**, *24*, 5431−5436.

(12) Li, J.; Liu, Y.; Li, X. Enhancement of Transverse Magneto-Optic Kerr effect in subwavelength all-dielectric guided mode gratings with a buffer layer. *Results Phys.* **2024**, *59*, 107597.

(13) So, S.; Mun, J.; Rho, J. Simultaneous Inverse Design of Materials and Structures via Deep Learning: Demonstration of Dipole Resonance Engineering Using Core−Shell Nanoparticles. *ACS Appl. Mater. Interfaces* **2019**, *11*, 24264−24268.

(14) Zhang, J.; Wang, G.; Wang, T.; Li, F. Genetic Algorithms to Automate the Design of Metasurfaces for Absorption Bandwidth Broadening. *ACS Appl. Mater. Interfaces* **2021**, *13*, 7792−7800.

(15) Brown, K. A.; Brittman, S.; Maccaferri, N.; Jariwala, D.; Celano, U. Machine Learning in Nanoscience: Big Data at Small Scales. *Nano Lett.* **2020**, *20*, 2−10.

(16) Kuznetsova, V.; Coogan, A.; Botov, D.; Gromova, Y.; Ushakova, E. V.; Gun'ko, Y. K. Expanding the Horizons of Machine Learning in Nanomaterials to Chiral Nanostructures. *Adv. Mater.* **2024**, *36*, 2308912.
(17) Tranter, A. D.; Slatyer, H. J.; Hush, M. R.; Leung, A. C.; Everett, J. L.; Paul, K. V.; Vernaz-Gris, P.; Lam, P. K.; Buchler, B. C.; Campbell, G. T. Multiparameter optimization of a magneto-optical trap using deep learning. *Nat. Commun.* **2018**, *9*, 4360.
(18) Fujita, T.; Sakaguchi, H.; Zhang, J.; Nonaka, H.; Sumi, S.; Awano, H.; Ishibashi, T. Magneto-optical diffractive deep neural network. *Opt. Express* **2022**, *30*, 36889−36899.
(19) Fakhrul, T.; Tazlaru, S.; Beran, L.; Zhang, Y.; Veis, M.; Ross, C. A. Magneto-Optical Bi:YIG Films with High Figure of Merit for Nonreciprocal Photonics. *Adv. Opt. Mater.* **2019**, *7*, 1900056.
(20) Fakhrul, T.; Tazlaru, S.; Khurana, B.; Beran, L.; Bauer, J.; Vančík, M.; Marchese, A.; Tsotsos, E.; Kučera, M.; Zhang, Y.; Veis, M.; Ross, C. A. High Figure of Merit Magneto-Optical Ce- and Bi-Substituted Terbium Iron Garnet Films Integrated on Si. *Adv. Opt. Mater.* **2021**, *9*, 2100512.
(21) Zelmon, D. E.; Small, D. L.; Jundt, D. Infrared corrected Sellmeier coefficients for congruently grown lithium niobate and 5 mol.% magnesium oxide−doped lithium niobate. *J. Opt. Soc. Am. B* **1997**, *14*, 3319−3322.
(22) Malitson, I. H. Interspecimen comparison of the refractive index of fused silica. *J. Opt. Soc. Am.* **1965**, *55*, 1205−1209.
(23) Jesenska, E.; Yoshida, T.; Shinozaki, K.; Ishibashi, T.; Beran, L.; Zahradnik, M.; Antos, R.; Kučera, M.; Veis, M. Optical and magneto-optical properties of Bi substituted yttrium iron garnets prepared by metal organic decomposition. *Opt. Mater. Express* **2016**, *6*, 1986−1997.
(24) Kreilkamp, L. E.; Belotelov, V. I.; Chin, J. Y.; Neutzner, S.; Dregely, D.; Wehlus, T.; Akimov, I. A.; Bayer, M.; Stritzker, B.; Giessen, H. Waveguide-Plasmon Polaritons Enhance Transverse Magneto-Optical Kerr Effect. *Phys. Rev. X* **2013**, *3*, 041019.
(25) Voronov, A. A.; Karki, D.; Ignatyeva, D. O.; Kozhaev, M. A.; Levy, M.; Belotelov, V. I. Magneto-optics of subwavelength all-dielectric gratings. *Opt. Express* **2020**, *28*, 17988−17996.
(26) Ignatyeva, D. O.; Krichevsky, D. M.; Belotelov, V. I.; Royer, F.; Dash, S.; Levy, M. All-dielectric magneto-photonic metasurfaces. *J. Appl. Phys.* **2022**, *132*, 100902.
(27) Ren, P.; Xiao, Y.; Chang, X.; Huang, P.-Y.; Li, Z.; Chen, X.; Wang, X. A comprehensive survey of neural architecture search: Challenges and solutions. *ACM Computing Surveys (CSUR)* **2022**, *54*, 1−34.
(28) Liu, Y.; Sun, Y.; Xue, B.; Zhang, M.; Yen, G. G.; Tan, K. C. A survey on evolutionary neural architecture search. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, *34*, 550−570.
(29) Wu, J.; Chen, X.-Y.; Zhang, H.; Xiong, L.-D.; Lei, H.; Deng, S.-H. Hyperparameter optimization for machine learning models based on Bayesian optimization. *J. Electron. Sci. Technol.* **2019**, *17*, 26−40.
(30) Weisberg, S. *Applied Linear Regression*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, 2005; DOI: 10.1002/0471704091.
(31) Jin, H.; Chollet, F.; Song, Q.; Hu, X. AutoKeras: An AutoML Library for Deep Learning. *J. Mach. Learn. Res.* **2023**, *24*, 169−174.
(32) Browne, M. W. Cross-validation methods. *J. Math. Psychol.* **2000**, *44*, 108−132.
(33) Picard, R. R.; Cook, R. D. Cross-validation of regression models. *J. Am. Stat. Assoc.* **1984**, *79*, 575−583.
(34) Anguita, D.; Ghelardoni, L.; Ghio, A.; Oneto, L.; Ridella, S. The 'K' in K-fold Cross Validation. *Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)*; Bruges, Belgium, April 25−27, 2012; pp 441−446.
(35) Koehrsen, W. Overfitting vs. underfitting: A complete example. *Towards Data Science* **2018**, 405.

https://doi.org/10.1021/acsami.4c06740
ACS Appl. Mater. Interfaces 2024, 16, 42828−42834