PHOTONICS Research

Multitask deep-learning-based design of chiral plasmonic metamaterials

Eric Ashalley,[1] Kingsley Acheampong,[2] Lucas V. Besteiro,[1,3] Peng Yu,[1] Arup Neogi,[4] Alexander O. Govorov,[1,5] and Zhiming M. Wang[1,*]

[1]Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu 610054, China
[2]School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China
[3]Centre Énergie Matériaux et Télécommunications, Institut National de la Recherche Scientifique, Varennes QC J3X 1S2, Canada
[4]Department of Physics, University of North Texas, Denton, Texas 76203, USA
[5]Department of Physics and Astronomy, Ohio University, Athens, Ohio 45701, USA
*Corresponding author: zhmwang@uestc.edu.cn

Received 13 January 2020; revised 5 May 2020; accepted 26 May 2020; posted 26 May 2020 (Doc. ID 388253); published 1 July 2020

The field of chiral plasmonics has registered considerable progress with machine-learning (ML)-mediated metamaterial prototyping, drawing from the success of ML frameworks in other applications such as pattern and image recognition. Here, we present an end-to-end functional bidirectional deep-learning (DL) model for three-dimensional chiral metamaterial design and optimization. This ML model utilizes multitask joint learning features to recognize, generalize, and explore in detail the nontrivial relationship between the metamaterials' geometry and their chiroptical response, eliminating the need for auxiliary networks or equivalent approaches to stabilize the physically relevant output. Our model efficiently realizes both forward and inverse retrieval tasks with great precision, offering a promising tool for iterative computational design tasks in complex physical systems. Finally, we explore the behavior of a sample ML-optimized structure in a practical application, assisting the sensing of biomolecular enantiomers. Other potential applications of our metastructure include photodetectors, polarization-resolved imaging, and circular dichroism (CD) spectroscopy, with our ML framework being applicable to a wider range of physical problems.

https://doi.org/10.1364/PRJ.388253

1. INTRODUCTION

Chiral nanostructures are non-superimposable to their mirror image, and produce a different optical response for left circularly polarized (LCP) and right circularly polarized (RCP) light [1–3]. This difference can be quantified by measuring chiroptical responses such as circular dichroism (CD) and chiral anisotropy factor (gCD), providing metrics for the chiral asymmetry of these structures. Hindered by their small chiral asymmetry and electromagnetic interaction volume, naturally occurring chiral structures such as amino acids, DNA, and RNA exhibit low CD signals, which imposes limits on their detectability. This limitation can be resolved by coupling these naturally occurring structures to artificially engineered plasmonic nanostructures under resonant excitation [4–6]. Plasmonic metamaterials can strongly enhance a molecular CD signal [6] and even displace it to longer wavelengths due to the interaction of the strong resonant plasmonic near-field with the chiral biomolecules [5]. Chiral plasmonic metamaterials exhibit inherent CD signals whose amplitude depends on the degree of chiral asymmetry of their geometry and thus presents an avenue for designing and manipulating nanoscale chiroptical effects. By changing the geometry of these plasmonic metamaterials (chiral and nonchiral), one can tune a resonant spectrum and target specific applications using bio-assembled as well as nonbiological systems [7,8]. The interaction of biomolecular structures with superchiral fields from chiral plasmonic nanostructures can also induce asymmetric changes in the retardation phase effect of the modes in the nanostructures and provide fingerprints for enantiomer discrimination [9]. The range of applications for plasmonic chiral systems also includes, among others, improved spectroscopy techniques [10,11], photocurrent generation [3], and bolometry [12]. At the same time, while chiral plasmonic metastructures offer very strong CD signals, they also present design and optimization challenges due to the iterative and case-by-case simulations required to solve Maxwell's equations for a given geometry, which can be a time- and resource-intensive process. Given the complex relationship between chiral plasmonic metamaterials and their chiroptical response, a data-driven approach to the design and optimization of the structures can improve its efficiency.

Deep-learning (DL) as a data-driven technique for analysis and prediction has permeated several disciplines such as natural language processing [13], image recognition [14], genetics, and

biology [15,16]. Instead of generating simulation results by running through predefined systems of equations in a given geometry, a DL architecture can be trained to recognize patterns in a given dataset, identify attributes, and predict responses, thanks to its capability to reproduce arbitrarily complex functions. As a type of representation or feature learning, DL brings machine learning (ML) closer to artificial intelligence where human-like and exceedingly challenging tasks can be completed by trained systems. A typical example of the capabilities of this technique is the success of deep neural networks with reinforcement learning in playing games such as shogi, chess, and Go [17–19]. In recent communications, neural network models have been utilized as a fast prototyping tool in the design of metamaterials [20–26]. The generative adversarial network (GAN) [27] and image processing [28] approaches have also been used to link the geometry of metamaterials with their optical response. Hybrid techniques consolidating compositional pattern-producing networks and cooperative coevolution [29] as well as those blending deep generative models with semisupervised learning [30] have also been proposed for the inverse design of metastructures. The prediction power of DL models even extends to the near- and far-field distributions of arbitrary 3D nanostructures [31]. However, DL-based prediction models for metamaterial design suffer from low accuracy owing to the huge mismatch between the dimension of input and output parameter, especially for inverse retrieval designs. Also, the generation of the training set via traditional simulation routes can be time and resource intensive for the study of nanoscale chirality, requiring an efficient use of the generated training set during the ML design and training. In addition, capturing the local optima (plasmonic/CD resonances) of chiroptical responses is challenging for most DL models and sometimes necessitates the setup of an auxiliary network, which adds to the complexity of the model [20].

In this communication, we propose an end-to-end multitask DL (MDL)-based model for the design and optimization of 3D chiral metamaterials. MDL models have gained root in the computational study of semantics [32] and transportation [33,34], as well as pose and action recognition [35]. Multitask learning draws on its implicit data augmentation, eavesdropping, attention focusing, representation bias, and regularization for effective and efficient generalization [36,37], eliminating the need for auxiliary networks or equivalent approaches to stabilize the model's output of physically relevant information.

Our MDL model comprises a single bidirectional neural network solving two tasks: the accurate prediction of the full chiroptical response of a chiral metamaterial from a set of geometric parameters, via a forward prediction path, and the accurate retrieval of the geometric parameters that can produce a given input of a full chiroptical response, by solving the inverse problem via an inverse prediction path. To bridge the mismatch gap and enhance the prediction accuracy for both forward and inverse predictions, especially at CD and plasmonic resonances, a joint-learning feature is incorporated in the model training. This feature ensures the comparison of errors in the learning of tasks, allowing for a well-generalized system. Consequently, the MDL model ensures an efficient use of the training set to achieve faster convergence [38,39] and provides a practical guideline for implementing similar ML systems in a variety of design problems with nanotechnological applications. This work is organized in three major parts. The first concerns the description of the chiral metamaterial absorber modeled after the yin-yang symbol and its optical properties calculated by the finite element method (FEM). The second contains the discussion of the MDL model composed by the forward and inverse design paths (FDP and IDP, respectively). Finally, in the third, we apply a sample MDL-optimized structure in the application of sensing biomolecular enantiomers. We envisage that our work will inspire the use of ML as an effective and efficient data-driven metamaterial design and optimization tool, taking advantage of the multitask technique that we detail herein.

2. CHIRAL METAMATERIAL MODEL AND FORMALISM

Generally, the metamaterial absorber structure used in this study consists of top yin-yang-shaped Au nanoparticles (YNPs), a PMMA layer, a Au backreflector, and a bottom glass layer. Figure 1(a) shows the unit cell of the meta-absorber structure. The YNPs are defined by their radius R_0 and thickness t_1. The period, ρ, is set following $2R_0 + 50$ nm. To eliminate spurious effects due to mathematically sharp edges, the top corners and the tip of YNP are rounded with a minimum radius of 10 nm, corresponding to an R_0 of 100 nm, and scale up with larger values of R_0 to preserve the YNP shape. The dielectric constants of glass and PMMA are 2.13 and 2.25, respectively. The dielectric constants of the Au YNP and backreflector are

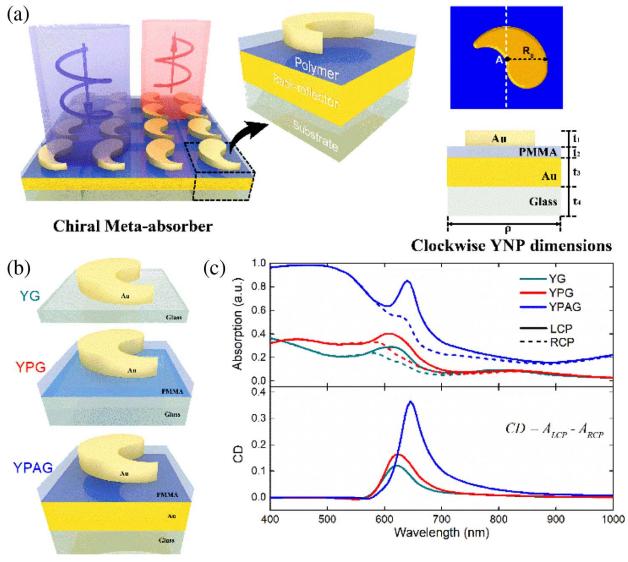

Fig. 1. Schematic of (a) a single YNP chiral meta-absorber array with definition of incident circularly polarized lights and a unit cell with dimensions. (b) Three single YNP metastructure configurations: Au YNP/Glass (YG), Au YNP/PMMA/Glass (YPG), and Au YNP/PMMA/Au/Glass (YPAG). (c) Absorption and CD spectra of the three metastructure configurations (YG, YPG, and YPAG), showing their plasmonic resonances λ_p (λ_p = 620 nm, 625 nm, and 645 nm, respectively) and revealing a strong chiroptical response for the metamaterial absorber case (YPAG). Here, R_0 = 100 nm, t_1 = 40 nm, t_2 = 50 nm, t_3 = 100 nm, and t_4 = 200 nm.

interpolated from the Johnson and Christy dataset [40]. The medium surrounding the metastructure is chosen to be air. The incident beam is normal to the metamaterial. These models were calculated using an FEM-based commercial software package, COMSOL Multiphysics.

To illustrate the relevance of the multilayered structure on the intensity of the fields and chiroptical response under resonance, we consider single YNP under three structures of comparable dimensions: YNP/Glass (YG), YNP/PMMA/Glass (YPG), and YNP/PMMA/Au/Glass (YPAG). Figure 1(b) shows the three metastructure configurations with single YNPs. The YG setup comprises a 40 nm thick Au YNP and a glass substrate with thickness t_4. For the YPG configuration, a PMMA polymer layer with thickness t_2 is introduced on the glass substrate to create a YNP/PMMA interface. Fabrication of the homopolymer layer can be conveniently achieved through spin-coating. The third configuration (YPAG) adds an Au backreflector layer of thickness t_3, between the PMMA and glass layers. Such structural combination with a backreflector largely increases the optical absorption of the system, by allowing the interaction between the nanoantenna and reflected modes [41–45]. Figure 1(c) compares the absorption and CD spectra of the YG, YPG, and YPAG metastructures. The YPAG structure shows a comparatively large and broad absorption peak and large differential absorption, both arising from the interaction between the Au YNP and the Au backreflector. Further comparison of the cross-sectional local field enhancement and surface charge density distributions for the three structures at their plasmonic resonances λ_p reveals sharp differences for LCP and RCP illumination at their respective plasmonic resonances (see Appendix A, Fig. 7). This further illustrates the strong chirality of the YNP-based structures especially in the meta-absorber case where the field is highly enhanced in the polymer spacer layer.

The enhanced chiral field and CD response in the metamaterial absorber case make it ideally competitive for

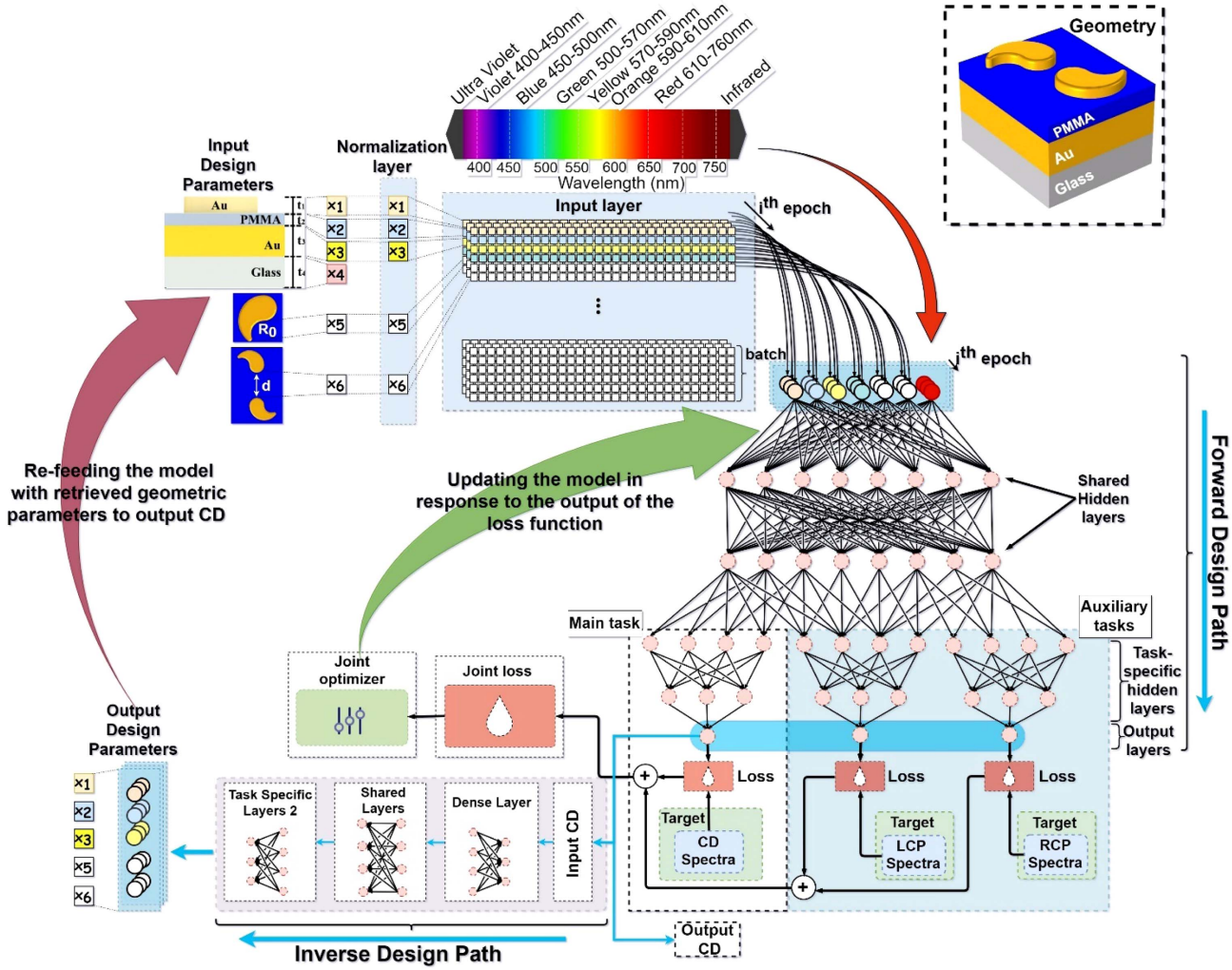

Fig. 2. Schematic of the bidirectional multitask deep-learning model for chiral metamaterial design consisting of forward design path (FDP) and inverse design path (IDP). Each path is composed by shared layers and task specific layers with joint optimization functionality. The model is set up in an end-to-end fashion where the geometric design parameters, CD, and LCP/RCP absorption spectra can be treated as input or output at specific ports. Here, the geometric design parameters are the YNPs thickness, PMMA thickness, YNPs radius, and YNPs (respectively represented as x_i, $i = 1, 2, \ldots, 6$). x_4 has been taken as a constant in the data shown in this study, but is nonetheless included in the model to represent the general parametrization of the system. The inset shows the metamaterial absorber geometry used to exemplify the use of the MDL.

chirality-related applications. The base YNP structure can be modified to generate other complex chiral designs, i.e., dimers and trimers (see Appendix A, Fig. 8). However, the optical response of these structures is highly sensitive to small variations of their geometric parameters. Thus, the optimization of a given design necessitates us to map out the response of many configurations across the design space. Each of these combinations requires solving the full electrodynamic problem under the illumination conditions, which is computationally costly. By combining a relatively small set of full wave simulations with our MDL approach, the design space can be explored in fine detail at a fraction of the computational cost. Experimentally and theoretically, a number of planar chiral geometries have been investigated and demonstrate huge CD signals including spirals [46–49], gammadions [50–54], L-shaped [55], Γ-shaped [12], and z-shaped [3] nanostructures. For the purpose of our demonstration of the MDL model, we adopt a variation of our YNP metastructure defined by double-top YNPs with radius, thickness, and gap distance R_0, t_1, and d, respectively (see inset in Fig. 2). The polymer thickness, t_2, and the Au backreflector thickness, t_3, affect the reflection of the beam and its interaction with the chiral resonance of the YNP, regulating the shift in the CD spectra. Due to the dominant effect of the Au backreflector, we have not explored the variation of thickness of the glass layer, t_4, and assigned a fixed t_4 value of 400 nm. The period of the structures is set following $P_x = 3R_0 + 50$ nm, $P_y = d + 2(R_0 + 25$ nm). Here, P_x is the period along the x direction, and P_y is the period along the y direction. By considering the double YNP structure, we introduce the gap between the nanoresonators as an additional parameter influencing the CD response, which is in a highly nontrivial relationship with the other geometric parameters. We have chosen this structure as a case study for our MDL model because of two sets of characteristics: (1) the continuous curved surfaces and the difference in scale between the whole structure and its finest features, such as its tip, increase the complexity of the numerical optical simulation of the metamaterial, thus representing a desirable candidate for alternative or supporting computational approaches; (2) the complex nonlinear relationship between the different design parameters and the structure's optical response makes it a good example for the power of our MDL approach.

3. MULTITASK DEEP-LEARNING ARCHITECTURE

As depicted in Fig. 2, the model consists of an input layer, tensor layers, a normalization layer, optical and chiroptical task execution layers, and an optimization layer. The inset in the top right shows the schematic of the metamaterial absorber used to train the DL model. This model takes into account the plasmon resonant peaks with a regression learning objective. Our model has unique multitasking and joint learning capabilities that contribute to the efficient generalization of the parameter space [56–58].

A. Forward Design

In the forward prediction path, the varying dataset scales across the input design parameters, absorption, and CD responses would make the model data-scale-dependent if trained directly, resulting in poor generalization. To eliminate the effect of the varying input length scales on the generalization of the model, we employ a normalization layer following the relation

$$\boldsymbol{X}_{\mathrm{norm},(a,b)} = \frac{\boldsymbol{X}_{(a,b)} - \overline{\boldsymbol{X}}_b}{\sigma(\boldsymbol{X}_b)}, \tag{1}$$

where a and b index the row and column, respectively, such that $\boldsymbol{X}_b$ is the bth column of the input parameter matrix. $\overline{\boldsymbol{X}}_b$ and $\sigma(\boldsymbol{X}_b)$ are the mean and standard deviation, respectively.

The normalization provides a well-conditioned dataset for optimization by ensuring that the training is less sensitive to the scale of features to be processed by the shared layer [59]. At the shared layer, we adopt a hard parameter sharing of four hidden layers, each with 1024 nodes. These hidden layers are shared between all the individual task-specific output layers of the network. The shared hidden layers hold computational weights from the task-specific layers. That is, the CD learning leverages the LCP/RCP spectra task-specific learning to enhance accuracy.

We regularize all hidden layers by applying penalties on layer activity during optimization with an $\mathcal{L}_2$ regularizer in order to learn sparse features and internal representations of raw observations [60]. The task-specific layer consists of two independent parts: the main task and auxiliary tasks. The auxiliary tasks are subtasks expected to assist in finding rigorous, rich, and robust representation of the input design parameters to benefit the main task. Learning auxiliary tasks restrict the parameter space during optimization and push for a faster convergence. The main task, which characterizes the desired output response, exploits and jointly learns from the auxiliary tasks via the shared layer. The main and auxiliary tasks correspondingly generate three single-task losses.

To optimize the MDL network, the joint multitask cost function comprising the three single-task losses is minimized. Here, the principal multitask cost function, subject to optimization, is expressed as [61]

$$\prod_j \left(\sum_i (y_{ji} - f_j(\boldsymbol{x}_i; \boldsymbol{w}))^2 \right), \tag{2}$$

where i and j index the training set and the three learning tasks, respectively. y_{ji} refers to the simulated outputs from the three tasks (CD and LCP/RCP absorption signals). The model function f takes as input $\boldsymbol{x}_i$, which is the 1×5 design parameter matrix comprising the YNP radius R_0, the gap distance d, the YNP thickness t_1, the polymer thickness t_2, and the Au backreflector thickness t_3 with weight, $\boldsymbol{w}$. The definition of d is illustrated by the inset in Fig. 3. See Appendix B for derivation.

We adopt the mean squared error (MSE)—the quadratic loss, which is the sum of squared distances between our target variable (simulated CD) and predicted CD values—to average the losses over the output. The Adam moment estimation stochastic optimization approach is used to compute an adaptive learning rate for each of the internal parameters of the model [62]. It is typical of ML models to miss the resonances of datasets with high volatility during prediction. This is because the probability distribution is centered at the off-resonance for each

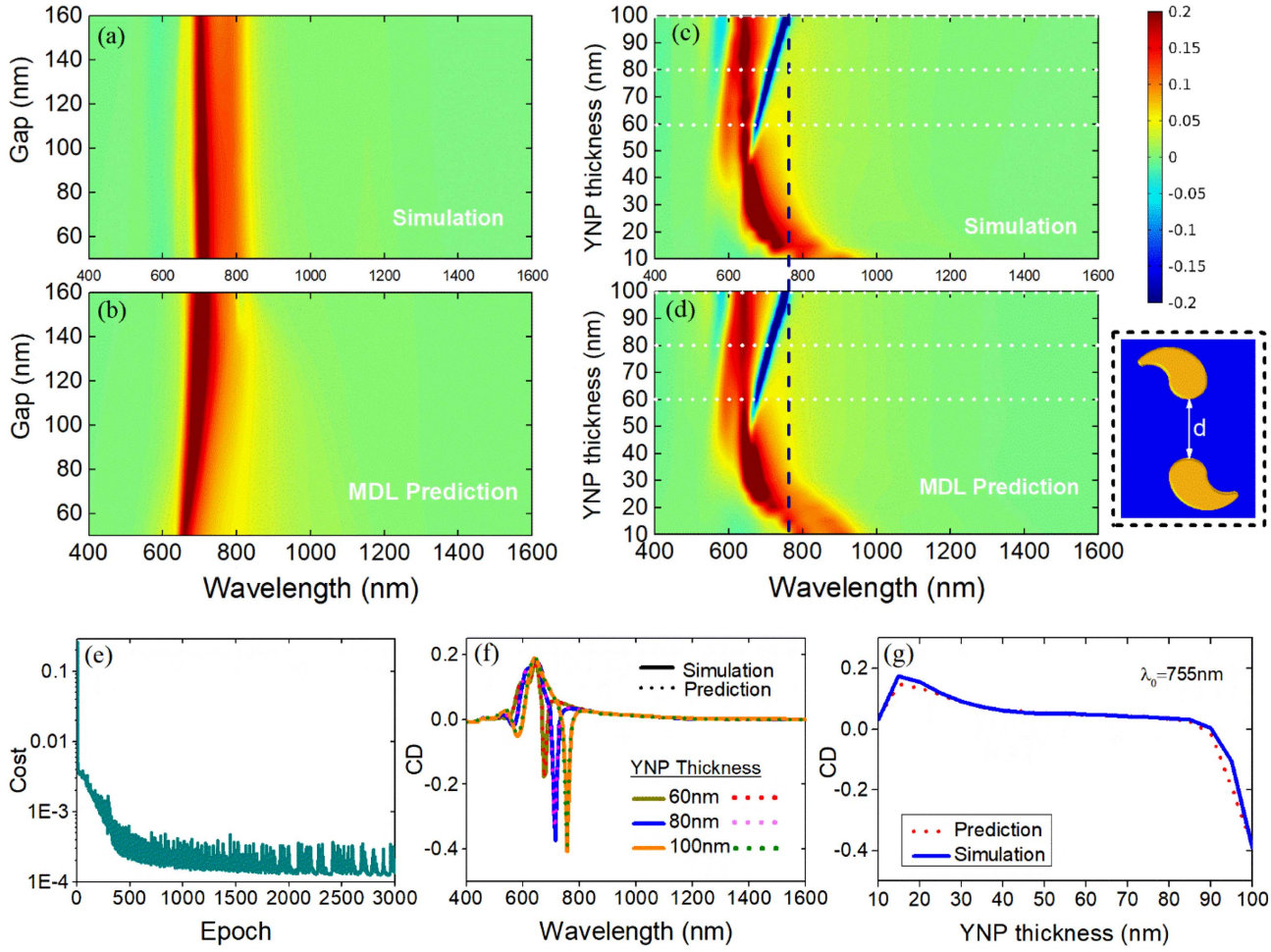

Fig. 3. MDL model performance. (a) Numerical simulation and (b) MDL prediction CD results of the dimer structure at varying gap distance d (50–160 nm), across the visible and near-IR regime. $R_0 = 100$ nm, $t_1 = 40$ nm, $t_2 = 50$ nm, and $t_3 = 100$ nm. (c) Numerical simulation and (d) MDL prediction results of the dimer at varying YNP thickness (t_1). The color legend has been truncated at ± 0.2 for clarity. Inset, definition of the gap d. (e) Learning curve within 3000 epochs. (f) Discretized model performance at selected t_1 = {60 nm, 80 nm, 100 nm} corresponding to the horizontal dots in (c) and (d). (g) Model performance at $\lambda_0 = 755$ nm across varying t_1, corresponding to the vertical short dashes through (c) and (d). Here, $R_0 = 100$ nm, $d = 100$ nm, $t_2 = 50$ nm, and $t_3 = 100$ nm.

neuron in the output layer, neglecting the local optima. The joint loss optimization functionality enables collective error correction for the forward prediction task, allowing an accurate prediction at the local optima. The training set utilized ~80% of the 640 collected samples with the remaining as the validation set. Each sample is constituted by the full 3×241 LCP absorption, RCP absorption, and CD spectra data points. These 241 data points were generated at 5 nm step intervals within the 400–1600 nm wavelengths. The MSE recorded is 0.000441, which is indicative of the model's accuracy. After training, we use the validation dataset, which is unseen throughout the training, to evaluate the model. Within short prediction intervals, the MDL model exhibits prediction results comparable to simulation data. A comparison between simulated [Fig. 3(a)] and MDL-predicted [Fig. 3(b)] CD response at varying dimer gap length d shows a good agreement. The range of wavelength values corresponding to high and low chiroptical activity is essentially identical. Comparing Figs. 3(c) and Fig. 3(d) for varying YNP thickness also shows a good agreement even for ultrathin YNPs. The MDL prediction retains the same wavelength interval as the simulation, providing a full continuous prediction spectrum (241 prediction data points). In this case, essential information within short parameter intervals can be retrieved to enrich the analysis, design, and prototyping process. The learning curve in Fig. 3(e) illustrates the fast convergence of the MDL model. Specifically, the single end-to-end MDL model accurately captures both resonant and off-resonant CD signals of the metastructures for varying geometric parameters as depicted by Figs. 3(f) and 3(g). The training set emphasized larger gap dimensions, which are easier to fabricate consistently, rather than more experimentally challenging dimers with smaller gaps. This bias accounts for the small discrepancies observed in Fig. 3(b) with respect to the simulation values in Fig. 3(a), a deviation that could be further reduced by a more homogeneous sampling. The situation is similar, although less pronounced, for the data in Fig. 3(d).

Given the design parameter space of the structure, vast sets of CD responses from varied parameter configurations

can be retrieved across the visible and near-IR regime ($\lambda_0 = 400$–1600 nm) via the trained MDL model. Figure 4 shows the interdependencies among the design parameters of the chiral metamaterial absorber and the evolution of CD generated by the MDL model with varying geometric parameter dimensions. For each plot, the unrepresented geometric parameters are held constant during the data generation. Figure 4(a) shows the CD values for varying YNP radius at $t_3 = \{100\text{ nm}, 150\text{ nm}, 200\text{ nm}\}$. We observe a slight shift and broadening of the CD peaks within the high chiroptically active wavelength interval (520–800 nm) for R_0 within the range 100–250 nm at varying t_3. Although the influence of the Au backreflector thickness, t_3, on the CD response is significant, its effect is comparatively small. This is clear when we consider that the Au layer indeed serves as a reflector, and we should expect that changes in thickness, when it is already above the penetration depth of the radiation, will not be very impactful to the overall response of the metamaterial.

Figure 4(b) compares the CD response by concurrently varying the YNP radius and t_2 at $t_1 = \{5\text{ nm}, 25\text{ nm}, 50\text{ nm}\}$. This plot is obtained at $\lambda_0 = 780$ nm. We observe the effect of the quasi-linear regions, where the CD remains constant irrespective of the coupling gap between the nanoresonators. Therefore, the YNP and PMMA thicknesses together with the YNP radius control the shift in the CD resonances. Varying these sets of parameters yields highly distinguishable CD map plots as illustrated in Figs. 4(b), 4(c), and 4(d). Figure 4(c) shows CD values at $R_0 = \{100\text{ nm}, 150\text{ nm}, 200\text{ nm}\}$ generated by varying concurrently, t_1 and t_2 at $\lambda_0 = 700$ nm. The contour regions correspond to a CD magnitude of 0.5. Figure 4(d) illustrates the interaction between the YNP radius and YNP thickness at $t_2 = \{10\text{ nm}, 45\text{ nm}, 100\text{ nm}\}$ for $\lambda_0 = 650$ nm. It can be deduced that the CD magnitude is generally lower for larger values of YNP radius. However, with thinner YNPs ($t_1 < 30$ nm), large CD signals can be realized by larger YNP radii ($300\text{ nm} < R_0 < 400$ nm). For thicker YNPs ($t_1 > 45$ nm), large CD signals can be attained for both ultrathin ($t_2 < 20$ nm) and thicker ($65\text{ nm} < t_2 < 100$ nm) PMMAs at larger YNP radius. Appendix D shows variations in CD across the studied wavelengths for varying PMMA thickness.

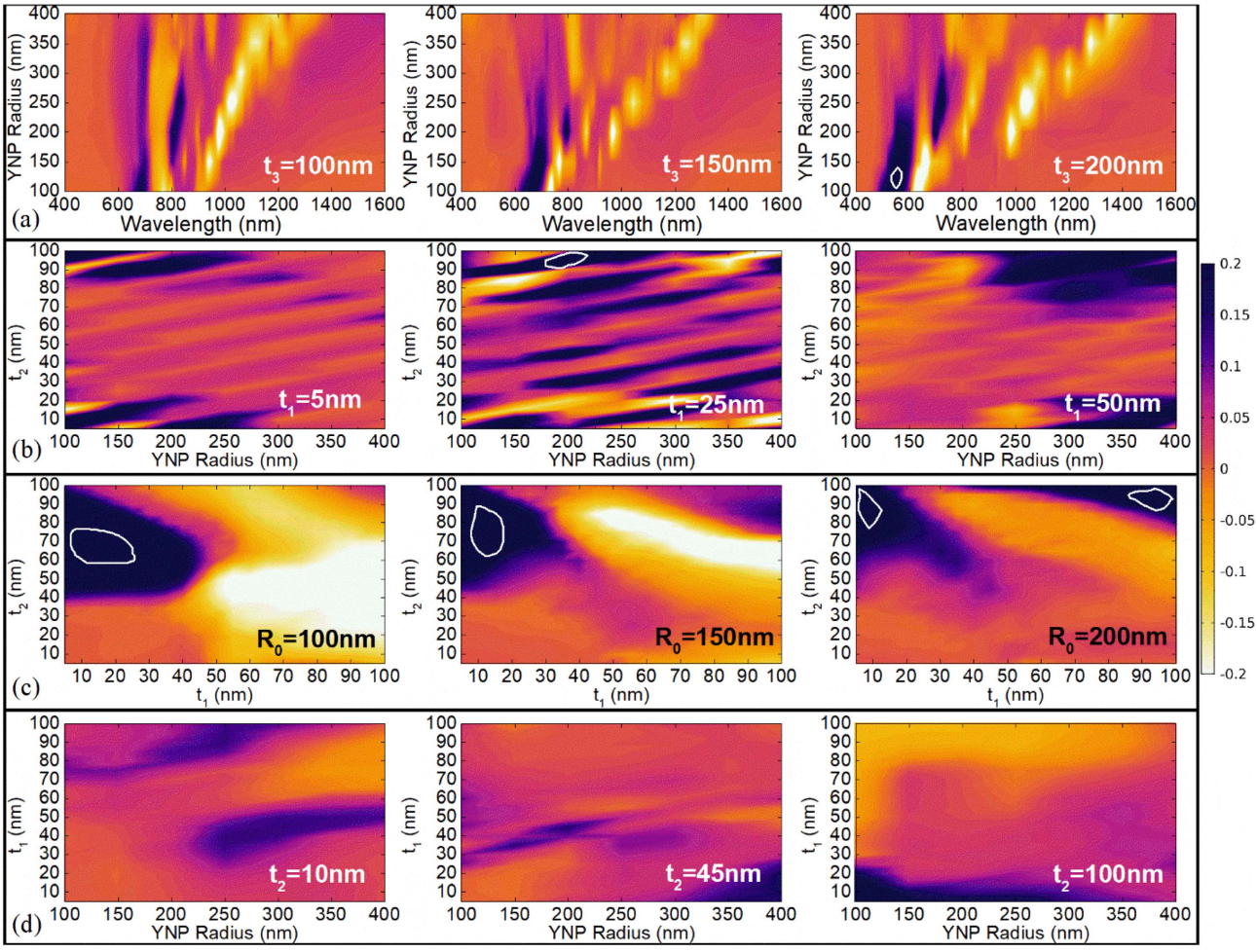

Fig. 4. MDL-predicted CD progressions. (a) CD evolution by varying YNP radius at $t_3 = \{100\text{ nm}, 150\text{ nm}, 200\text{ nm}\}$. CD map plot by varying concurrently, (b) YNP radius and polymer thickness at $t_1 = \{5\text{ nm}, 25\text{ nm}, 50\text{ nm}\}$ for $\lambda_0 = 780$ nm, (c) YNP thickness and polymer thickness at $R_0 = \{100\text{ nm}, 150\text{ nm}, 200\text{ nm}\}$ and $\lambda_0 = 700$ nm, and (d) YNP radius and YNP thickness at $t_2 = \{10\text{ nm}, 45\text{ nm}, 100\text{ nm}\}$ and $\lambda_0 = 650$ nm. The color legend has been truncated at ± 0.2 for clarity, but high-CD areas have been highlighted by adding the contour regions corresponding to CD values of 0.5.

Having the trained network, we can easily extend its output to cover many more geometric parameter combinations, and for different wavelengths, although this sample should suffice for illustrating the method's capabilities, as well as obtaining a qualitative understanding of the physical system.

It is relevant to underscore a point of particular interest of employing such an MDL approach, which is that it requires a much smaller amount of computational resources than producing an equivalently dense dataset through traditional simulation methods. Let us examine a quick estimate for the comparison of both approaches. For each combination of geometric parameters, it takes a typical i7-CPU PC (which, for simplicity, we will assume that consumes the total output of its 330 W power source regardless of the task performed) more than 3 h to compute a full spectrum when using COMSOL with precision levels adequate for our system. This order of magnitude for the duration of the computation is also representative of other simulation packages and frameworks. In our case, the total energy used for creating the model (combining the energy required to generate the training set with COMSOL and actually training the network) is approximately 0.63 MWh spanning 80 days using the above example PC. After training, using the MDL model requires less than a second to produce the response for a given geometric parameter set. On the other hand, for the five geometric parameters and our chosen sampling density, the total number of simulations required to obtain the sample result density that the trained MDL affords would be in the tens of thousands (~28,106 samples), which would require an outrageous number of years (~10,000 years) to produce, with a total energy consumption of approximately 29 GWh using the example PC. Therefore, in this (admittedly extreme) comparison, the MDL approach would outperform a naïve simulation-based approach by 7 orders of magnitude in terms of speed and energy expenditure.

B. Inverse Design

After obtaining a high prediction accuracy in the forward path, we proceed to design an IDP. It is important to realize that this task is very complex and sensitive, because of the large imbalance between the input and output dimensions of the model: five geometric parameters compared with the full CD, LCP, and RCP spectra (1×5 versus 3×241, respectively). While it is plausible to apply up- and down-sampling approaches to resolve this mismatch, essential features may be lost in the process, especially in the case of complex structures like the YNP. Such lost features introduce prediction errors that compound over a training loop, resulting in wide variations in the retrieved geometric parameters for comparatively small changes in the input spectra. Reverse-engineering the multitask forward design path makes use of the entire set of CD data points and maps each output geometric parameter to the desired full input spectra. The IDP takes a simulated CD spectrum as input, with the objective of retrieving the geometric parameters required to produce it. To achieve this objective, we connect the input spectrum to three isolated dense networks with two layers, responsible for mapping the chiroptical response to a shared latent-space representation with four layers, each composed by 2048 neurons. The shared latent space links to five task-specific layers with each layer managing the output of a designated geometric parameter (see Fig. 2).

Running this inverter network gives us a set of geometric parameters, which will be close to the ones generating the target CD spectrum. However, given the interdependency of these parameters and the high sensitivity of the CD response to geometric parameters [see, for instance, Fig. 4(b)], we should not assume that this is sufficient to obtain a satisfactory system prediction. Therefore, assessing the ML-predicted CD spectra will be central to characterizing the efficiency of the inverse retrieval model. For verification, the retrieved geometric parameters are re-fed into the forward prediction path, so that a comparison can be drawn between the target CD from the ground-truth geometric parameters and the ML-predicted CD from the ML-retrieved geometric parameters. Figure 5(a) shows a simulated CD serving as target spectrum (green line), with its corresponding geometric parameters in Fig. 5(c) (green bars). Upon feeding the simulated CD spectrum into the inverse path, the IDP retrieves the geometric parameters in Fig. 5(c) (red bars). Although it is clear that the inverter network found values that are very close to the ground-truth, we proceed to verify the retrieval by inputting these geometric parameters into the forward path and comparing its predicted CD response [red dots in Fig. 5(a)] with the simulated CD spectrum, thus confirming the success of the inverse path.

Figures 5(b) and 5(d) were generated following the same procedure. Both input spectra, in Figs. 5(a) and 5(b), were chosen to illustrate the model acting on two profiles with distinct properties (inflection points, negative and positive resonant CD peaks). Generally, there is a good agreement between simulation and ML-prediction responses, although we can observe slight differences in the spectra that cohere with the small differences between the sets of retrieved and ground-truth geometric parameters. Figure 9(c) in Appendix C compares

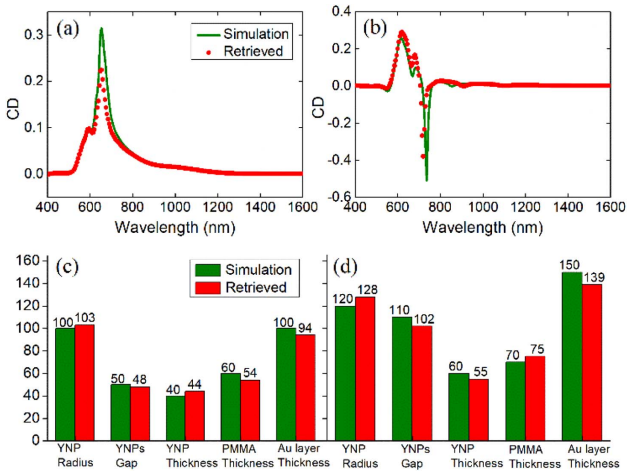

Fig. 5. Inverse design with the MDL model. (a), (b) Simulated (green solid lines) and predicted (red dots) CD spectra. (c), (d) Corresponding simulated (green bars) and retrieved (red bars) geometric parameters. Red dots in (a), (b) are predicted from the MDL model with geometric parameters retrieved [red bar in (c), (d)] for the target simulated CD spectra in (a), (b). [See Figs. 9(a) and 9(b) in Appendix C for absorption spectra comparison.]

different approaches to the training of the model such as training with and without the auxiliary tasks as well as separately training the LCP and RCP spectra. We observe comparatively better performance in the case of our MDL model. The MDL-optimized dimer structure can be adopted for several applications including photodetection of circularly polarized light (CPL), owing to its high consistency and accuracy [see Fig. 10 in Appendix D]. However, we will now explore the use of an example MDL-optimized chiral plasmonic structure in sensing adsorbed molecular enantiomers.

4. CHIRAL BIOSENSING

The strength of the chiroptical response from biomolecules is mostly limited by the small magnitude of their geometric chiral features, relative to the periodicity of CPL, in comparison with chiral plasmonic nanoantennas. Therefore, it is often necessary to use large molecular concentrations of nonracemic mixtures to provide detectable CD signals [5,63,64]. Moreover, chiral biomolecular structures have chiroptical activity in the UV and are thus difficult to detect with common instrumentation.

Figure 6(a) shows the weak CD resonant peak of a molecular enantiomer pair in the UV. The optical activity of chiral molecules can be enhanced by a handedness-preserving Fabry–Perot cavity resonator [65] acting as a metamirror. Such metamirrors selectively reflect one CPL preserving its handedness while absorbing the other [66]. The molecular CD signal can also be enhanced by the interaction with the strong near-field of plasmonic structures, as well as duplicating its CD signal into the visible range, where they are easier to detect [5,67]. These advantages, which greatly enhance the detection capabilities of a chiral biomolecular sensor, are in principle independent from the chirality of the plasmonic metamaterial. Nonchiral nanoparticles can offer an effective and efficient sensing of molecular enantiomers [6,67] by coupling with the chiral molecules. However, local superchiral near-fields can offer additional possibilities for molecular CD detection with chiral plasmonic nanostructures on a metal-biomolecular platform [11]. Superchiral fields are those with a larger chirality than CPL, as characterized through their optical chirality, C [68], and can arise on chiral plasmonic structures illuminated with either linearly or CPL [9]. Such near-fields with enhanced

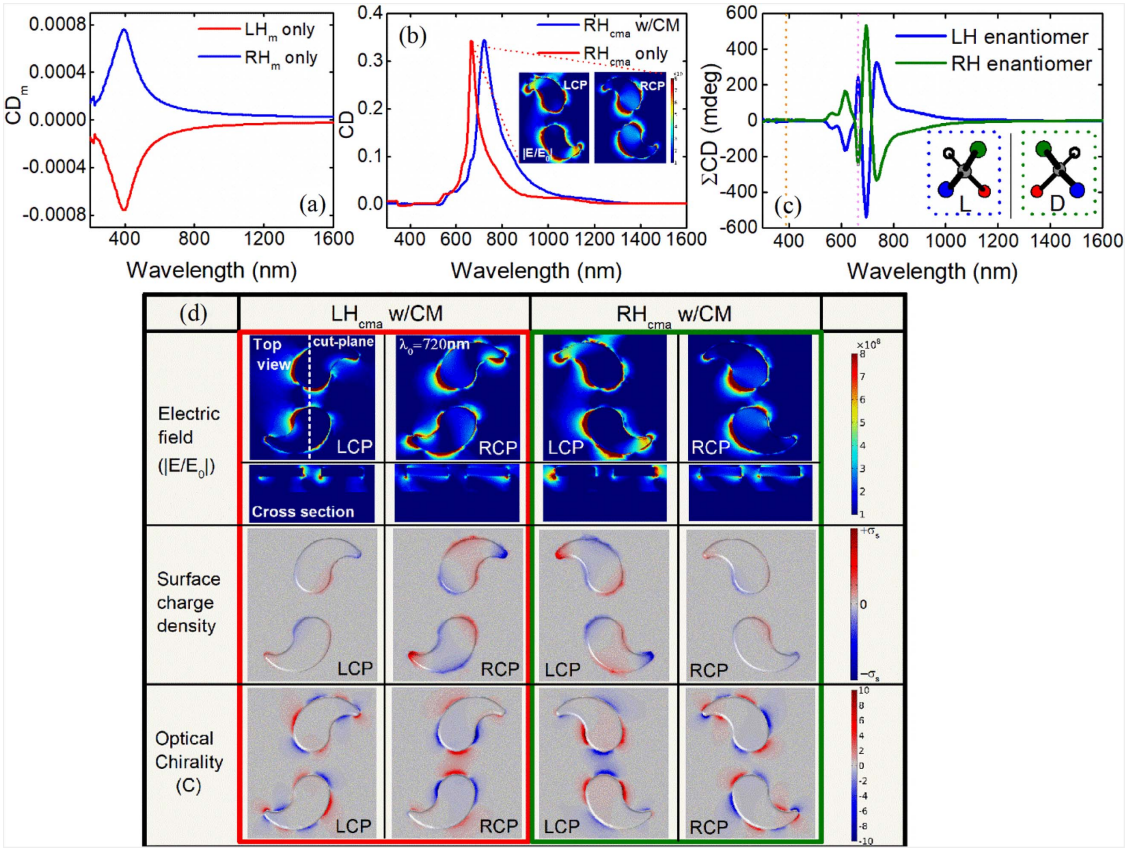

Fig. 6. Enantiomer detection. (a) The CD spectra of (red) left-handed medium (LH$_m$), and (blue) right-handed medium (RH$_m$) with molecular CD resonance ($\lambda_m = 380$ nm) in the UV. (b) CD spectra comparison of the right-handed chiral metamaterial absorber (RH$_{cma}$) with (blue) and without (red) chiral medium (CM). Inset is the electric field at the plasmonic resonance, λ_p ($\lambda_p = 665$ nm), of the bare chiral metamaterial absorber for LCP and RCP light. (c) CD summation to remove metamaterial background CD signal to reveal the LH (blue solid line) and RH (green solid line) enantiomer pair CD signals. λ_m, λ_p, and λ_{mp} represent the resonant wavelengths for the CD of the bare molecules, the plasmonic chiral metamaterial absorber, and the metamaterial covered with chiral media, respectively. Inset, schematic representation of an enantiomeric protein molecular pair (L and D isomers). (d) Electric field, surface charge density, and optical chirality density distributions of the chiral metamaterial absorber covered by chiral media (CM) at $\lambda_{mp} = 720$ nm.

optical chirality cause asymmetric phase differences between the chiral modes in the presence of a chiral dielectric, offering a path for sensing biomolecular chirality [69].

In what follows, we present theoretical results illustrating the behavior of one of the ML-optimized nano-dimer metastructures in the presence of molecular enantiomers, using the excess CD method to quantify the chirality of the molecular sample [70]. The structure was optimized for value of CD of 0.35. The geometric parameters of the metamaterial are $R_0 = 100$ nm, $t_1 = 30$ nm, $t_2 = 50$ nm, $t_3 = 200$ nm, $t_4 = 200$ nm, and $d = 100$ nm. Modeling the chiral molecules, the dimer is covered by a 40 nm thick chiral medium with 1.6 refractive index.

The chiral dielectric medium is modeled following the constitutive equations:

$$\boldsymbol{D} = \varepsilon_0 \varepsilon_r \boldsymbol{E} + i\xi \boldsymbol{B}, \tag{3}$$

$$\boldsymbol{H} = \boldsymbol{B}/\mu_0 \mu_r + i\xi \boldsymbol{E}. \tag{4}$$

Here, ε_0 and ε_r are the permittivity of free space and relative permittivity, respectively. Similarly, μ_0 and μ_r are the permeability of free space and relative permeability, respectively. $\boldsymbol{E}$ and $\boldsymbol{B}$ are the complex electric field and magnetic flux density, respectively. $\boldsymbol{D}$ and $\boldsymbol{H}$ are the electric field displacement and the magnetic field, respectively. ξ is the chirality factor of the molecular sample, which shows very low values for low-density or near-racemic samples. Using a two-state model for the molecules, ξ can be expressed as a function of frequency as

$$\xi = \beta_c \left(\frac{1}{\hbar\omega + \hbar\omega_0 + i\Gamma_{12}} + \frac{1}{\hbar\omega - \hbar\omega_0 + i\Gamma_{12}} \right), \tag{5}$$

where β_c controls the magnitude of the chiral asymmetry, ω is the angular frequency of the radiation, $\omega_0 = 2\pi c_0/\lambda_{\text{mol}}$ at the molecular excitation wavelength, λ_{mol}, and Γ_{12} defines the relaxation rate of the excited molecule, with its indices describing its quantum states. Here, $\lambda_{\text{mol}} = 380$ nm and $\Gamma_{12} = 0.41$ eV. The expression and values for the chirality factor, ξ, are adopted from Govorov *et al.* [5,9] and follow from the quantum equation of motion for the electronic density matrix when assuming a dilute molecular sample.

Now, we proceed with the excess CD method to compute the chiral properties of the molecules, when in interaction with left-handed and right-handed chiral metamaterial (LH_{cma} and RH_{cma}, respectively). First, a baseline CD is calculated from the metamaterial absorber with racemate ($\xi = 0$) molecular coverage. In the process, a distinguishable dielectric medium-induced CD redshift (~55 nm) is observed as illustrated in Fig. 6(b). This redshift arises from the refractive index of the molecular coverage, which is larger than that of the surrounding medium (air). The inset shows the local field enhancement of the bare metamaterial absorber without biomolecules, illustrating its strong and chiral near-field. Then, we can compute the sum of the CD signal of the two YNP metamaterial absorbers, with chiral and racemate molecular coverage, yielding a change in CD response associated with the biomolecules of opposite handedness. The coupling of the enhanced plasmonic near-field to the molecules is critical for biomolecular sensing. In Fig. 6(c), we can see how the CD signal of the chiral molecules shifts its peak to the visible range and increases its magnitude by approximately a factor of 2. The excess CD signal over that of the chiral metamaterial absorber, produced by the coverage of molecular enantiomers, is a detectable magnitude that allows us to characterize the presence and handedness of chiral biomolecules, as illustrated in Fig. 6(c).

Finishing this section, we present additional details on the chiral properties of the example metastructure, showing its near-field enhancement, surface charge density, and optical chirality parameter C. The latter is calculated, for harmonically oscillating fields, as [71]

$$C = -\frac{\varepsilon_0}{2} \omega \, \text{Im}(\boldsymbol{E}^* \cdot \boldsymbol{B}). \tag{6}$$

Figure 6(d) compares these three variables for the LH_{cma} and RH_{cma} systems, under the dielectric chiral medium described above. The first row is the top and cross-sectional view of the electric field enhancement distribution of the LH_{cma} (red border) and RH_{cma} (green border), when illuminated by LCP and RCP light. A *y-z* cross-sectional plane is taken to show the field enhancement inside the layered metamaterial. The surface charge density is illustrated in the second row, and the optical chirality maps in the third row. The electric field, surface charge density, and optical chirality are evaluated at the chiral media-modified plasmonic resonance, $\lambda_{mp} = 720$ nm. Importantly, the optical chirality maps reveal the superchiral near-field regions arising from the curvature and chiral geometry of the nanoantennas.

5. CONCLUSION

We propose a highly portable and functional MDL model to comprehensively study 3D, arbitrarily complex chiral metamaterials, and exemplify its usage with a chiral metamaterial designed after the yin-yang symbol. The model is composed by a single end-to-end bidirectional architecture, capable of performing optimization and inverse retrieval operations, and that takes advantage of the supporting role of two auxiliary tasks to facilitate the learning of the primary task, CD in our case. This feature in particular distinguishes it from other approaches discussed in the literature, and represents a reduction of the complexity of the DL framework while ensuring an efficient use of the training set toward a highly generalized system. As a data-driven approach, this MDL model requires a prior database of results created with methods such as FEM simulations or experimental data, but it helps in avoiding the huge computational cost that would be required to explore the vast design space of the physical system in fine detail. An additional advantage resides in the fact that a trained model is a fast, lightweight, highly transferrable tool that can drastically reduce the computational time used for subsequent studies of the system, both for the designers and for other research groups. Given a set of geometric parameters, the forward design path of the model predicts CD spectra with values virtually identical to the simulations we used as ground-truth. And, vice versa, for input CD spectra, the model retrieves the set of geometric parameters that would produce such input CD spectra by solving the inverse problem. As a result, the trained model can be used to explore the entire design space and thus render a complete account of

the intricate relationship between the metamaterial's geometric parameters and its chiroptical response. This is made possible by the joint-learning feature incorporated in the model. For nanophotonic applications, the design and prototyping process needs to be robust due to the complexity of light–matter interaction with nontrivial geometries. The multitasking DL-based prediction model presented herein can aid in engineering any potential fabrication of the nanophotonic structures for desired optical and chiroptical response toward a variety of applications. Illustrating a potential context for the utilization of this system, we have shown additional chiroptical properties of this chiral metamaterial absorber structure in the context of its interaction with molecular enantiomers. The high efficiency and accuracy of the end-to-end MDL model makes it a valuable tool for the study of complex physical phenomena, particularly for the design and prototyping of nanophotonic structures toward their application as biosensors, as photodetectors, or in polarization-resolved imaging and CD spectroscopy, among others.

APPENDIX A: CD MODEL DESCRIPTION

For chiral structures, the CD and g-factor are the prominent parameters that describe their chiroptical properties. These properties reflect inherent topological characteristics of chiral matter at the nano/microscale. Here, we focus on the CD of the absorption spectra. The CD and g-factor in the absorption are respectively calculated from [12]

$$\mathrm{CD} = A_{\mathrm{LCP}} - A_{\mathrm{RCP}},\ g_{\mathrm{CD}} = \frac{A_{\mathrm{LCP}} - A_{\mathrm{RCP}}}{(A_{\mathrm{LCP}} + A_{\mathrm{RCP}})/2}, \tag{A1}$$

where A_{LCP} and A_{RCP} are the absorbance of the chiral nanostructure illuminated by LCP and RCP beams, respectively.

Figure 7 shows the electric field and surface charge density distributions of the three metastructure configurations: YNP/Glass (YG), YNP/PMMA/Glass (YPG), and YNP/PMMA/Au/Glass (YPAG) revealing comparatively high CD signals in the metamaterial absorber case. Here, $R_0 = 100$ nm, $t_1 = 40$ nm, $t_2 = 50$ nm, $t_3 = 100$ nm, and $t_4 = 100$ nm. Exploring other geometric arrangements of the structure reveals interesting properties (see Fig. 8 for CD and g_{CD} calculations of the meta-absorber and its variations). Figure 8(a) shows the absorption spectra of a single chiral YNP metamaterial absorber at varying polymer thickness upon interaction with LCP and RCP light, with plasmonic peaks in the visible and near-IR range. Figure 8(b) is the corresponding complementary reflection spectra. Due to the effect of the Au backreflector, there is an enhanced differential absorption between LCP and RCP light. There is also an inversion in the CD response of the metamaterial absorber for YNP enantiomer pairs.

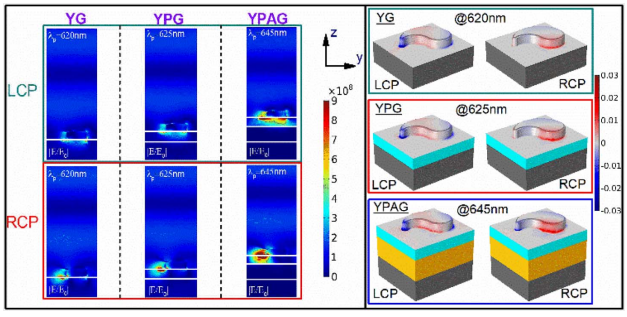

Fig. 7. Electric field and surface charge density distributions of the three single YNP metastructure configurations (YG, YPG, YPAG) at their plasmonic resonances, showing huge enhancement for the metamaterial absorber case. The cut plane is through point A, the unit cell center [Fig. 1(a)].

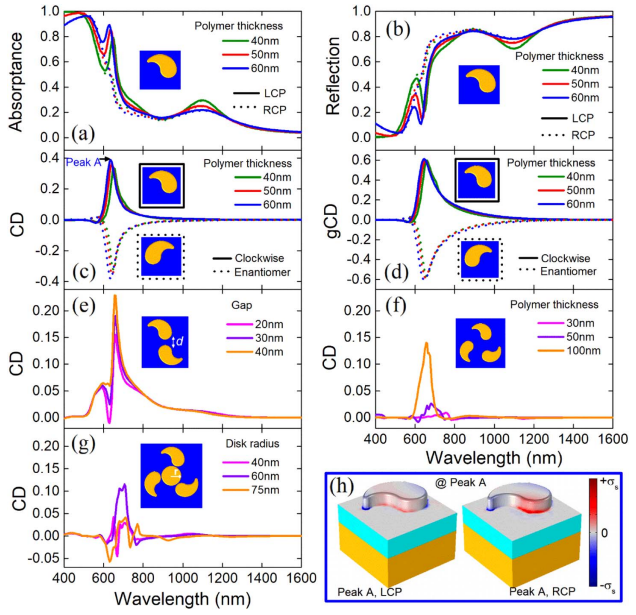

Fig. 8. Optical and chiroptical properties of the chiral metamaterial absorber with 10 nm tip rounding radius. (a) Absorption and (b) reflection spectra of single yin-yang metamaterial absorber at varying polymer thicknesses. Corresponding (c) CD and (d) g_{CD} from (a) exhibiting a strong peak in the visible regime (around 635 nm wavelength). CD response from three gapped configurations of the chiral metamaterial absorber: (e) dimer, varying gap length d, (f) circular, varying PMMA thickness, and (g) circular with central disk, varying central disk radius r. (h) Charge density distribution at the CD maxima (peak A) for the 60 nm PMMA thickness in (c).

Hence, the handedness of the metastructure is dictated by the handedness of the top chiral meta-atom. The CD and g_{CD} response of the single YNP metamaterial absorber is illustrated in Figs. 8(c) and 8(d). In Fig. 8(h) we show the surface charge density distribution of the single YNP meta-absorber corresponding to peak A in Fig. 8(c), for LCP and RCP light. Exploring other geometric arrangements of the structure reveals interesting properties. For a dimer, reducing the separation gap d increases the interaction of the two antennas and separates the CD signal in two distinct peaks, a feature that can be adopted for sensing applications. For a circular arrangement of the YNPs, we observe an interesting dependence of the magnitude of the CD response on the PMMA thickness, with the CD almost vanishing for a PMMA thickness of 100 nm [Fig. 8(f)]. With the introduction of a central disk, however, the plasmonic coupling between the YNP and the central disk can regulate the CD signal [Fig. 8(g)]. That is, at a critical

coupling distance, maximum CD can be achieved, and vice versa, depending on the central disk's radius r.

APPENDIX B: JOINT COST FUNCTION DERIVATION

Let i index the training set and j the dependent variables (the three tasks, that is the LCP, RCP, and CD learning tasks). Assuming these three tasks, with realizations y_{ij}, are independent and conditional to the prediction returned by a model f, with adjustable parameters $\boldsymbol{w}$ (the weight), on input $\boldsymbol{x}_i$, and that the error is normally distributed and zero-mean, with variance σ_j^2, which depends only on j (the task), we can write the log-likelihood function as follows:

$$\sum_{ij} \log\left(\frac{1}{\sqrt{2\pi\sigma_j}} \exp\left(-\frac{(y_{ji} - f_j(\boldsymbol{x}_i; \boldsymbol{w}))^2}{2\sigma_j^2}\right)\right). \tag{B1}$$

Applying the basic properties of the log function, (B1) becomes

$$\sum_{ij}\left(-\log\left(\sqrt{2\pi}\right) - \frac{1}{2}\log\sigma_j^2 - \frac{(y_{ji} - f_j(\boldsymbol{x}_i; \boldsymbol{w}))^2}{2\sigma_j^2}\right). \tag{B2}$$

Looking for stationary points of this loss with respect to the variance σ_j^2, and dropping the constant additive term, we have

$$\frac{\partial}{\partial\sigma_j^2}\sum_i\left(-\frac{1}{2}\log\sigma_j^2 - \frac{(y_{ji} - f_j(\boldsymbol{x}_i; \boldsymbol{w}))^2}{2\sigma_j^2}\right) = 0. \tag{B3}$$

From the linearity of partial derivatives, and simplifying the derivatives for each term, we have

$$\sum_i\left(-\frac{1}{2\sigma_j^2} + \frac{(y_{ji} - f_j(\boldsymbol{x}_i; \boldsymbol{w}))^2}{2\sigma_j^4}\right) = 0. \tag{B4}$$

Extracting the first term from the summation gives

$$-\frac{N}{2\sigma_j^2} + \sum_i\left(\frac{(y_{ji} - f_j(\boldsymbol{x}_i; \boldsymbol{w}))^2}{2\sigma_j^4}\right) = 0. \tag{B5}$$

Here, N is the size of the training set. Further simplification yields

$$-N + \sum_i\left(\frac{(y_{ji} - f_j(\boldsymbol{x}_i; \boldsymbol{w}))^2}{\sigma_j^2}\right) = 0, \tag{B6}$$

$$\sigma_j^2 = \frac{1}{N}\sum_i (y_{ji} - f_j(\boldsymbol{x}_i; \boldsymbol{w}))^2. \tag{B7}$$

Substituting into the log density function and simplifying gives

$$-\sum_j \log\left(\sum_i (y_{ji} - f_j(\boldsymbol{x}_i; \boldsymbol{w}))^2\right). \tag{B8}$$

By changing the signs and exponentiating to get back to the original form of the quadratic loss, we have [61]

$$\prod_j\left(\sum_i (y_{ji} - f_j(\boldsymbol{x}_i; \boldsymbol{w}))^2\right). \tag{B9}$$

APPENDIX C: INVERSE DESIGN PLOTS

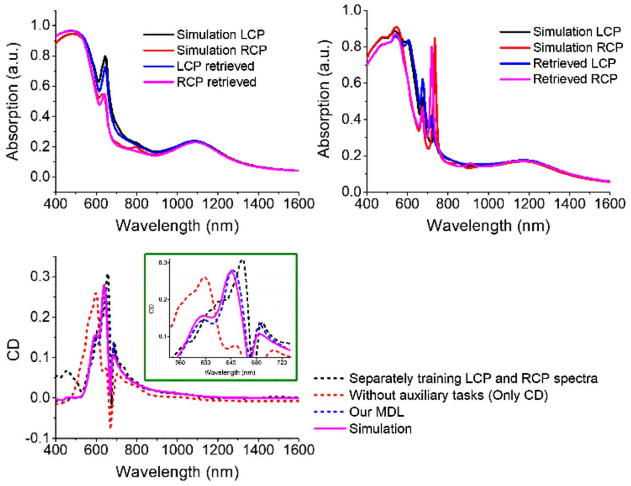

Fig. 9. Inverse design plots. Simulated and MDL-retrieved absorption spectra comparison for the systems providing the data in (a) Fig. 5(a) and (b) Fig. 5(b). (c) Comparison between the CD values obtained with different variations of the ML techniques: separately training LCP and RCP spectra (black dashes), training only CD without auxiliary tasks (red dashes), training by our MDL model (blue dashes), and the ground-truth (pink solid line). The inset is a zoomed-in image at the resonances.

APPENDIX D: ADDITIONAL MDL PREDICTIONS

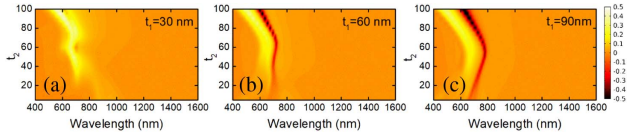

Fig. 10. MDL-predicted CD progressions. CD evolution when varying PMMA thickness at $t_1 = \{30\text{ nm}, 60\text{ nm}, 90\text{ nm}\}$ for $R_0 = 100$ nm. This plot agrees with the Fig. 3, where we observe only CD maxima for $t_1 = 30$ nm but both CD maxima and minima for $t_1 = 60$ nm and $t_1 = 90$ nm, establishing the consistency of the MDL model.

Funding. National Key Research and Development Program of China (2019YFB2203400); China Postdoctoral Science Foundation (2017M622992, 2019M663467, 2019T120820); National Science Foundation Emerging Frontiers & Multidisciplinary Activities (1741677); UESTC Shared Research Facilities of Electromagnetic Wave and Matter Interaction (Y03019012901002O1).

Disclosures. The authors declare no conflicts of interest.

REFERENCES

1. S. Zu, Y. Bao, and Z. Fang, "Planar plasmonic chiral nanostructures," Nanoscale **8**, 3900–3905 (2016).
2. X. Yin, M. Schäferling, A. K. U. Michel, A. Tittl, M. Wuttig, T. Taubner, and H. Giessen, "Active chiral plasmonics," Nano Lett. **15**, 4255–4260 (2015).
3. W. Li, Z. J. Coppens, L. V. Besteiro, W. Wang, A. O. Govorov, and J. Valentine, "Circularly polarized light detection with hot electrons in chiral plasmonic metamaterials," Nat. Commun. **6**, 8379 (2015).
4. J. M. Slocik, A. O. Govorov, and R. R. Naik, "Plasmonic circular dichroism of peptide-functionalized gold nanoparticles," Nano Lett. **11**, 701–705 (2011).
5. A. O. Govorov, Z. Fan, P. Hernandez, J. M. Slocik, and R. R. Naik, "Theory of circular dichroism of nanomaterials comprising chiral molecules and nanocrystals: plasmon enhancement, dipole interactions, and dielectric effects," Nano Lett. **10**, 1374–1382 (2010).
6. B. M. Maoz, Y. Chaikin, A. B. Tesler, O. Bar Elli, Z. Fan, A. O. Govorov, and G. Markovich, "Amplification of chiroptical activity of chiral biomolecules by surface plasmons," Nano Lett. **13**, 1203–1209 (2013).
7. J. R. Mejía-Salazar and O. N. Oliveira, "Plasmonic biosensing," Chem. Rev. **118**, 10617–10625 (2018).
8. X. T. Kong, L. V. Besteiro, Z. Wang, and A. O. Govorov, "Plasmonic chirality and circular dichroism in bioassembled and nonbiological systems: theoretical background and recent progress," Adv. Mater., 1801790 (2018).
9. R. Tullius, G. W. Platt, L. Khosravi Khorashad, N. Gadegaard, A. J. Lapthorn, V. M. Rotello, G. Cooke, L. D. Barron, A. O. Govorov, A. S. Karimullah, and M. Kadodwala, "Superchiral plasmonic phase sensitivity for fingerprinting of protein interface structure," ACS Nano **11**, 12049–12056 (2017).
10. A. Micsonai, F. Wien, L. Kernya, Y. H. Lee, Y. Goto, M. Réfrégiers, and J. Kardos, "Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy," Proc. Natl. Acad. Sci. USA **112**, E3095–E3103 (2015).
11. E. Hendry, T. Carpy, J. Johnston, M. Popland, R. V. Mikhaylovskiy, A. J. Lapthorn, S. M. Kelly, L. D. Barron, N. Gadegaard, and M. Kadodwala, "Ultrasensitive detection and characterization of biomolecules using superchiral fields," Nat. Nanotechnol. **5**, 783–787 (2010).
12. X. T. Kong, L. Khosravi Khorashad, Z. Wang, and A. O. Govorov, "Photothermal circular dichroism induced by plasmon resonances in chiral metamaterial absorbers and bolometers," Nano Lett. **18**, 2001–2008 (2018).
13. J. Hirschberg and C. D. Manning, "Advances in natural language processing," Science **349**, 261–266 (2015).
14. K. Yun and A. Huyen, and T. Lu, "Deep neural networks for pattern recognition," arXiv:1809.09645 (2018).
15. J. Zou, M. Huss, A. Abid, P. Mohammadi, A. Torkamani, and A. Telenti, "A primer on deep learning in genomics," Nat. Genet. **51**, 12–18 (2019).
16. M. W. Libbrecht and W. S. Noble, "Machine learning in genetics and genomics," Nat. Rev. Genet. **16**, 321–332 (2015).
17. D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. Van Den Driessche, T. Graepel, and D. Hassabis, "Mastering the game of Go without human knowledge," Nature **550**, 354–359 (2017).
18. D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, and D. Hassabis, "A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play," Science **362**, 1140–1144 (2018).
19. S. Singh, A. Okun, and A. Jackson, "Artificial intelligence: learning to play Go from scratch," Nature **550**, 336–337 (2017).
20. W. Ma, F. Cheng, and Y. Liu, "Deep-learning-enabled on-demand design of chiral metamaterials," ACS Nano **12**, 6326–6334 (2018).
21. I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via Deep Learning," Light Sci. Appl. **7**, 60 (2018).
22. T. F. De Lima, H. T. Peng, A. N. Tait, M. A. Nahmias, H. B. Miller, B. J. Shastri, and P. R. Prucnal, "Machine learning with neuromorphic photonics," J. Lightwave Technol. **37**, 1515–1534 (2019).
23. Z. S. Ballard, D. Shir, A. Bhardwaj, S. Bazargan, S. Sathianathan, and A. Ozcan, "Computational sensing using low-cost and mobile plasmonic readers designed by machine learning," ACS Nano **11**, 2266–2274 (2017).
24. N. Dordević, J. S. Beckwith, M. Yarema, O. Yarema, A. Rosspeintner, N. Yazdani, J. Leuthold, E. Vauthey, and V. Wood, "Machine learning for analysis of time-resolved luminescence data," ACS Photon. **5**, 4888–4895 (2018).
25. L. Pilozzi, F. A. Farrelly, G. Marcucci, and C. Conti, "Machine learning inverse problem for topological photonics," Commun. Phys. **1**, 57 (2018).
26. K. Yao, R. Unni, and Y. Zheng, "Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale," Nanophotonics **8**, 339–366 (2019).
27. Z. Liu, D. Zhu, S. P. Rodrigues, K. T. Lee, and W. Cai, "Generative model for the inverse design of metasurfaces," Nano Lett. **18**, 6570–6576 (2018).
28. I. Sajedian, J. Kim, and J. Rho, "Finding the optical properties of plasmonic structures by image processing using a combination of convolutional neural networks and recurrent neural networks," Microsyst. Nanoeng. **5**, 27 (2019).
29. Z. Liu, D. Zhu, K. T. Lee, A. S. Kim, L. Raju, and W. Cai, "Compounding meta-atoms into metamolecules with hybrid artificial intelligence techniques," Adv. Mater. **32**, 1–7 (2020).
30. W. Ma, F. Cheng, Y. Xu, Q. Wen, and Y. Liu, "Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy," Adv. Mater. **31**, 1901111 (2019).
31. P. R. Wiecha and O. L. Muskens, "Deep learning meets nanophotonics: a generalized accurate predictor for near fields and far fields of arbitrary 3D nanostructures," Nano Lett. **20**, 329–338 (2019).
32. H. Peng, S. Thomson, and N. A. Smith, "Deep multitask learning for semantic dependency parsing," arXiv:1704.06855 (2017).
33. W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: deep belief networks with multitask learning," IEEE Trans. Intell. Transp. Syst. **15**, 2191–2201 (2014).
34. X. Gibert, V. M. Patel, and R. Chellappa, "Deep multitask learning for railway track inspection," IEEE Trans. Intell. Transp. Syst. **18**, 153–164 (2017).
35. D. C. Luvizon, D. Picard, and H. Tabia, "2D/3D pose estimation and action recognition using multitask deep learning," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2018), pp. 5137–5146.
36. A. Maurer, M. Pontil, and B. Romera-Paredes, "The benefit of multitask representation learning," J. Mach. Learn. Res. **17**, 2853–2884 (2016).
37. H. B. Lee, E. Yang, and S. J. Hwang, "Deep asymmetric multi-task feature learning," arXiv:1708.00260 (2018).
38. D. Dong, H. Wu, W. He, D. Yu, and H. Wang, "Multi-task learning for multiple language translation," in *53rd Annual Meeting of the Association for Computational Linguistics and 7th International Joint Conference on Natural Language Processing* (2015), Vol. **1**, pp. 1723–1732.
39. S. Liu, S. J. Pan, and Q. Ho, "Distributed multi-task relationship learning," in *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining Part F1296* (2017), pp. 937–946.
40. P. B. Johnson and R. W. Christy, "Optical constants of the noble metals," Phys. Rev. B **6**, 4370–4379 (1972).
41. P. Yu, L. V. Besteiro, Y. Huang, J. Wu, L. Fu, H. H. Tan, C. Jagadish, G. P. Wiederrecht, A. O. Govorov, and Z. Wang, "Broadband metamaterial absorbers," Adv. Opt. Mater. **7**, 1800995 (2018).
42. P. Yu, J. Wu, E. Ashalley, A. Govorov, and Z. Wang, "Dual-band absorber for multispectral plasmon-enhanced infrared photodetection," J. Phys. D **49**, 365101 (2016).
43. P. Yu, L. V. Besteiro, J. Wu, Y. Huang, Y. Wang, A. O. Govorov, and Z. Wang, "Metamaterial perfect absorber with unabated size-independent absorption," Opt. Express **26**, 20471–20480 (2018).
44. W. Li and J. Valentine, "Metamaterial perfect absorber based hot electron photodetection," Nano Lett. **14**, 3510–3514 (2014).

45. H. Zhu, F. Yi, and E. Cubukcu, "Plasmonic metamaterial absorber for broadband manipulation of mechanical resonances," Nat. Photonics **10**, 709–714 (2016).
46. K. A. Bachman, J. J. Peltzer, P. D. Flammer, T. E. Furtak, R. T. Collins, and R. E. Hollingsworth, "Spiral plasmonic nanoantennas as circular polarization transmission filters," Opt. Express **20**, 1308–1319 (2012).
47. B. Frank, X. Yin, M. Schäferling, J. Zhao, S. M. Hein, P. V. Braun, and H. Giessen, "Large-area 3D chiral plasmonic structures," ACS Nano **7**, 6321–6329 (2013).
48. J. M. Linnanto and J. E. I. Korppi-Tommola, "Investigation on chlorosomal antenna geometries: tube, lamella and spiral-type self-aggregates," Photosynth. Res. **96**, 227–245 (2008).
49. T. Kan, A. Isozaki, N. Kanda, N. Nemoto, K. Konishi, M. Kuwata-Gonokami, K. Matsumoto, and I. Shimoyama, "Spiral metamaterial for tunable circular dichroism," in *IEEE International Conference on Micro Electro Mechanical Systems (MEMS)* (2013), pp. 701–704
50. G. Li, S. Zhang, and T. Zentgraf, "Nonlinear photonic metasurfaces," Nat. Rev. Mater. **2**, 17010 (2017).
51. T. Cao, L. Zhang, R. E. Simpson, C. Wei, and M. J. Cryan, "Strongly tunable circular dichroism in gammadion chiral phase-change metamaterials," Opt. Express **21**, 27841–27851 (2013).
52. C. Zhang, Z. Q. Li, X. Yang, Z. Chen, and Z. Wang, "Controlling third harmonic generation with gammadion-shaped chiral metamaterials," AIP Adv. **6**, 125014 (2016).
53. S. Wu, P. P. Qu, J. Q. Liu, D. D. Lei, K. Y. Zhang, S. T. Zhao, and Y. Y. Zhu, "Giant circular dichroism and its reversal in solid and inverse plasmonic gammadion-shaped structures," Opt. Express **24**, 27763–27770 (2016).
54. Y. Z. Cheng, M. L. Huang, H. R. Chen, Y. J. Zhou, X. S. Mao, and R. Z. Gong, "Influence of the geometry of a gammadion stereo-structure chiral metamaterial on optical properties," J. Mod. Opt. **64**, 1487–1494 (2017).
55. Y. Wang, J. Deng, G. Wang, T. Fu, Y. Qu, and Z. Zhang, "Plasmonic chirality of L-shaped nanostructure composed of two slices with different thickness," Opt. Express **24**, 2307–2317 (2016).
56. Y. Yang and T. M. Hospedales, "Deep multi-task representation learning: a tensor factorisation approach," arXiv:1605.06391 (2017).
57. M. Rei, "Semi-supervised multitask learning for sequence labeling," in *55th Annual Meeting of the Association for Computational Linguist (ACL)* (2017), pp. 2121–2130.
58. I. Misra, A. Shrivastava, A. Gupta, and M. Hebert, "Cross-stitch networks for multi-task learning," arXiv:1604.03539 (2016).
59. B. Neyshabur, S. Bhojanapalli, D. McAllester, and N. Srebro, "Exploring generalization in deep learning," arXiv:1706.08947 (2017).
60. A. Y. Ng, "Feature selection, L_1 vs. L_2 regularization, and rotational invariance," in *21st International Conference on Machine Learning* (2004), pp. 615–622.
61. R. Cipolla, Y. Gal, and A. Kendall, "Multi-task learning using uncertainty to weigh losses for scene geometry and semantics," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2018), pp. 7482–7491.
62. D. P. Kingma and J. L. Ba, "Adam: a method for stochastic optimization," in *3rd International Conference on Learning Representations—Conference Track Proceedings* (2015), pp. 1–15.
63. J. Yao, M. Xu, Z. Yan, W. Wu, and C. Yang, "Supramolecular photochirogenesis with cyclodextrins," Chin. J. Org. Chem. **34**, 26–35 (2014).
64. T. Li, H. G. Park, H. S. Lee, and S. H. Choi, "Circular dichroism study of chiral biomolecules conjugated with silver nanoparticles," Nanotechnology **15**, S660–S663 (2004).
65. J. X. Bao, N. Liu, H. W. Tian, Q. Wang, T. J. Cui, W. X. Jiang, S. Zhang, and T. Cao, "Chirality enhancement using Fabry–Pérot-like cavity," Research **2020**, 7873581 (2020).
66. L. Mao, K. Liu, S. Zhang, and T. Cao, "Extrinsically 2D-chiral metamirror in near-infrared region," ACS Photon. **7**, 375–383 (2020).
67. A. O. Govorov, "Plasmon-induced circular dichroism of a chiral molecule in the vicinity of metal nanocrystals. Application to various geometries," J. Phys. Chem. C **115**, 7914–7923 (2011).
68. Y. Tang and A. E. Cohen, "Optical chirality and its interaction with matter," Phys. Rev. Lett. **104**, 163901 (2010).
69. C. Gilroy, S. Hashiyada, K. Endo, A. S. Karimullah, L. D. Barron, H. Okamoto, Y. Togawa, and M. Kadodwala, "Roles of superchirality and interference in chiral plasmonic biodetection," J. Phys. Chem. C **123**, 15195–15203 (2019).
70. Y. Zhao, A. N. Askarpour, L. Sun, J. Shi, X. Li, and A. Alù, "Chirality detection of enantiomers using twisted optical metamaterials," Nat. Commun. **8**, 6 (2017).
71. T. J. Davis and E. Hendry, "Superchiral electromagnetic fields created by surface plasmons in nonchiral metallic nanostructures," Phys. Rev. B **87**, 085405 (2013).