Identifying Optimal Photonic Crystal Sensor Designs with Machine Learning

Eric Y. Zhu[1,2,*], **Li Qian**[1], **Ofer Levi**[1,2,*]

[1] *Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, Ontario M5S 3G4, Canada*
[2] *Institute of Biomaterials and Biomedical Engineering, University of Toronto, 164 College Street, Toronto, Ontario M5S 3G9, Canada*
[*] *Corresponding authors: eric.zhu@utoronto.ca, ofer.levi@utoronto.ca*

Abstract: We train a neural network to predict the optical properties (center wavelength λ_0, linewidth, sensitivity S) of photonic crystal slab structures. We are able to faithfully model the results to within 1% for λ_0 and S.

OCIS codes: (130.0130) Integrated optics; (350.4238) Nanophotonics and photonic crystals

While deep learning has been associated primarily with classification, it can also be used for regression, or more colloquially, curve fitting. Recent work has shown that one can model the properties of photonic crystal fibers with neural networks [1]. Since photonic crystal structures can be complex, and require solving mode equations frequently, there is an opportunity to create and train a deep learning model that can represent the resonance properties of a photonic crystal slab (PCS) sensor, which are often nonlinearly dependent on the device geometry and material. Deep learning can also provide insight into creating devices with higher biosensitivity.

Applications of photonic crystal slab (PCS) sensors include index-based biosensing [2] and, more recently, ultrasound sensing [3]. The archetypical structure of the PCS is shown in Fig. 1a and 1b. A square array of nanoholes is etched into a 'core' layer of stoichiometric silicon nitride (Si_3N_4) with thickness t, lattice constant a, and nanohole radius r. The Si_3N_4 layer is sandwiched between a layer of thermal silica (SiO_2) and a top cladding layer n_{clad} that can change depending on the application of the sensor. The entire device is fabricated on a silicon substrate.

When used as a biosensor, the top cladding layer will likely be an aqueous (water-based) solution, with $n_{\text{clad}} \sim 1.3-1.4$. When used as an ultrasound sensor, the top cladding material can be chosen for its high photoelastic constant as well as its ability to 'drag' out more of the electric field from the core region of the structure [4]; in this case, the top cladding index can vary from $n_{\text{clad}} \sim 1.3-1.6$. In either application, the PCS sensor is interrogated from the back with laser light (Fig. 1c).

The sensing mechanism for the sensor relies on the ability to resolve and measure an optical resonance of linewidth Γ (measured here in units of wavelength, Fig. 1d). When the index of the top layer n_{clad} changes, the center wavelength λ_0 of the resonance shifts in response; this shift is quantified by $S \equiv \frac{d\lambda_0}{dn_{\text{clad}}}$, the index sensitivity, and is itself a function of n_{clad}. The ultimate sensitivity, or detection limit, of the PCS can be given by a simple unitless figure-of-merit (FOM): $S \times \Gamma^{-1}$. It is then our goal to find a particular structure of the PCS to maximize this FOM.

The optical spectra of our PCS devices can be faithfully simulated using rigorous coupled wave analysis (RCWA). We use the open-source software S^4 [5] to perform RCWA. Given a particular geometry and material composition of the PCS, an optical spectrum of the transverse-electric (TE) optical resonance of the PCS can be generated. A Fano lineshape is fit to the resonance, and three of its properties – the spectral position λ_0, linewidth Γ, and index sensitivity S – are extracted.

ger

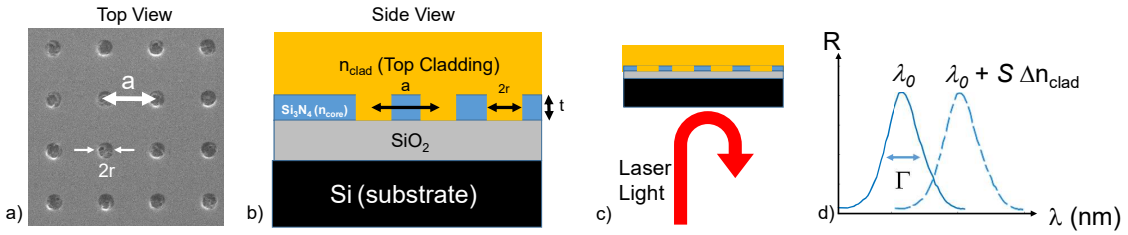

Fig. 1. a) Electron micrograph of a fabricated PCS sensor. The lattice constant (a) and nanohole diameter ($2r$) are labeled. b) Side view of sensor. The PCS is covered by a top cladding whose index n_{clad} is tracked by the position of the optical resonance of the PCS. c) Interrogation of the PCS is performed by laser light from the backside. d) The optical resonance of the PCS, with linewidth Γ, shifts by $\Delta\lambda = S \times \Delta n_{\text{clad}}$ n response to a change Δn_{clad}. The properties of the PCS are chosen so that the optical resonances lie in the 1.5 μm telecom band.

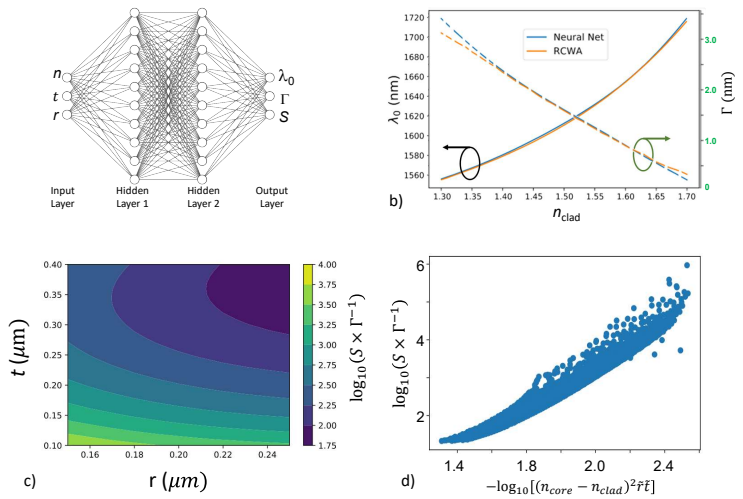

Fig. 2. a) Fully-connected neural network (NN) structure used to model the properties of the PCS, which contains 160 weights. b) The NN's prediction vs RCWA simulation results for peak wavelength (solid lines) and linewidth (dashed lines) vs n_{clad} for a particular PCS geometry ($r = 0.2\mu$m, $t = 0.25\mu$m). We observe good agreement. c) Using the NN model, and setting $n_{\mathrm{clad}} = 1.5$, we find that the FOM $S \times \Gamma^{-1}$ is maximized when t and r are minimized. d) Performing principal component analysis (PCA) on the raw RCWA data reveals a monotonic dependency of the FOM on the independent variables; $\tilde{r}$ ($\tilde{t}$) represent the nanohole radius (Si_3N_4 layer thickness) normalized by a, the lattice constant.

- the Si_3N_4 layer thickness t is varied from 0.15 to 0.25 μm, and
- the nanohole radius r is varied from 0.1 to 0.4 μm.

For each entry, the properties of the resonance (λ_0, Γ, and S) are extracted from the simulated optical spectrum and recorded. To simplify calculations and without loss of generality, the lattice constant is set to $a = 1.0$ μm. This is because, all other input parameters being the same, the properties of the resonance simply scale linearly with the value of a.

The dataset is then used to train a fully-connected neural network (NN) (Fig. 2a). The input layer takes the (normalized) PCS parameters (n_{clad}, t, r). This is followed by two fully-connected 10-node hidden layers, and a 3-node layer which returns (normalized and non-dimensional scaled equivalents of) λ_0, Γ, and S at the outputs. The `scikit-learn` Python library [6] is used to build and train the neural network. The hyper-parameters of the model – such as the number of hidden layers, the type of nonlinear activation function used, and cost minimization algorithm – were varied to find the minimal fitting error. The two-layer structure (Fig. 2a) was found to give better results than a single- or three-hidden layer NN, and the $\tanh$ activation function was found to work best. The limited-memory BFGS optimization algorithm was chosen because it worked faster, returning results within 1 second using a modern CPU, and gave better results. We observe in Fig. 2b that the NN-based model agrees well with the RCWA that it was trained on for a particular PCS geometry ($r = 0.2\mu$m, $t = 0.25\mu$m).

We can use this model and look at PCS geometries (different r and t) for which the FOM $S \times \Gamma^{-1}$ is maximized. Figure 2c shows such a plot when $n_{\mathrm{clad}} = 1.5$, and indicates that the optimal geometry involves minimizing both t and r. By performing principal component analysis (PCA) on the raw RCWA signal, we arrive at a similar result (Fig. 2d): $S \times \Gamma^{-1} \propto (rt)^{-4}$.

In summary, we have demonstrated the ability to model the optical properties of a PCS sensor by using a fully-connected neural net. The nonlinear dependence of the sensor on its geometry can be distilled down from 4000 data points to 160 weights in the NN. This method also helps us to optimize the design of the PCS according to the FOM. Future work will involve extending this approach to PCS designs that have more degrees of design freedom, and generating non-intuitive designs for sensors that may be more expensive computationally when done conventionally.

References

1. S. Chugh, A. Gulistan, S. Ghosh, and B. Rahman, Optics Express **27**, 36414–36425 (2019).
2. M. El Beheiry, V. Liu, S. Fan, and O. Levi, Optics express **18**, 22702–22714 (2010).
3. E. Y. Zhu, C. Rewcastle, R. Gad, L. Qian, and O. Levi, Optics letters **44**, 2609–2612 (2019).
4. E. Y. Zhu, M. C. Charles, C. Rewcastle, et al. in "CLEO: Applications and Technology," (Optical Society of America, 2020), pp. ATh4K–2.
5. V. Liu and S. Fan, Computer Physics Communications **183**, 2233–2244 (2012).
6. F. Pedregosa, G. Varoquaux, A. Gramfort, et. al., Journal of Machine Learning Research **12**, 2825–2830 (2011).