

Deep-learning assisted control of Optical Phased Array: a case study

Original

Deep-learning assisted control of Optical Phased Array: a case study / Savio, Daniele; Bardella, Paolo. - In: IEEE ACCESS. - ISSN 2169-3536. - ELETTRONICO. - 10:(2022), pp. 15421-15426. [10.1109/ACCESS.2022.3149115]

Availability:

This version is available at: 11583/2954870 since: 2022-02-08T18:12:32Z

Publisher:

IEEE

Published

DOI:10.1109/ACCESS.2022.3149115

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

Received January 6, 2022, accepted January 25, 2022, date of publication February 4, 2022, date of current version February 11, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3149115

Deep-Learning Assisted Control of Optical Phased Array: A Case Study

DANIELE SAVIO AND PAOLO BARDELLA^{ID}, (Member, IEEE)

Politecnico di Torino, 10129 Turin, Italy

Corresponding author: Paolo Bardella (paolo.bardella@polito.it)

ABSTRACT Modern raster scanning techniques and single pixel applications require a precise control of the field profiles radiated by Optical Phased Arrays, which can be controlled by external electrical signals modifying the optical paths inside these devices. In particular for single pixel imaging solutions, it is generally not straightforward to identify the control signals required to generate the desired far field pattern. We propose therefore a Convolutional Neural Network based approach which allows a user to determine the signals required to accurately reproduce an arbitrary far field profile.

INDEX TERMS Optical phased array, deep-learning, silicon photonics.

I. INTRODUCTION

In the last years, Optical Phased Arrays (OPAs) have gained popularity in a variety of application fields where an integrated, non-mechanical solution is required to precisely control the direction of an optical beam. Similarly to what happens in microwave phased arrays, OPAs can be seen as optical antennas where the output beam is controlled electrically by operating on the optical path of integrated optical components. Thanks to their beam steering capability (both in terms of speed and angle range), together with their reduced weight and dimensions, OPAs find natural application in all the fields where the traditional bulky beam steering systems composed by mechanical servo and optical lens cannot meet the requirements, with applications ranging from inter-satellite communication and sensors to autonomous vehicles [1]–[7].

Unfortunately, it is generally challenging to predict with accuracy the shape of the output beam; analytical expressions can be rarely used in this scenario due to the internal complexity of the OPA. The goal of this work is to show that a Machine Learning (ML) based approach can be effectively used to greatly simplify the control of such devices, regardless of their internal structure.

The paper is organized as follow. First, in Section II a review of the state of the art on OPA solutions and ML design techniques is given. The considered test device is introduced in Section III-A, while details on the used ML engine and

the effectiveness of the proposed approach are presented in Section III-B. Finally, the conclusions are briefly drawn in Section IV.

II. LITERATURE REVIEW

Different technologies have been exploited to realize efficient OPAs, like the micro-electromechanical systems (MEMS) and photonic integrated circuits (PICs). In this paper we will focus on the latter, which present better characteristics in terms of control speed and radiation angle range, together with the possibility of designing a completely integrated system with all the necessary components (e.g. light source, optical amplifiers, phase control region, radiation region) [1].

The most commonly proposed architecture to obtain a 2D control of an emitted beam of light is a 2D array of antennas in which the phases of the emitted signals can be controlled individually to meet the requirements. Nowadays, thanks to the device miniaturization, it is possible to integrate more than a thousand antennas in a small area; for example in [2] a monolithical OPA on SOI technology composed by a 1D array of 1024 independent antennas is reported, being able to create an output beam as narrow as 0.03° together with a steering angle of 45° in the transversal direction. Due to the complexity of controlling each antenna independently, it is of great interest to find ways to lower the redundancy in the control section of antennas without losing steering capabilities. For example in [3] it is demonstrated that an $N \times N$ array of emitters can be controlled by just $2N$ phase shifters instead of N^2 and in [4] the $N \times N$ array of antennas

The associate editor coordinating the review of this manuscript and approving it for publication was Leo Spiekman^{ID}.

is sparsely populated without losing the transmitted beam characteristic, significantly reducing the routing difficulties to control the antennas. In the latter case a beam width of 0.64° is reported, together with a steering capability of 16° in both the directions. A different architecture is proposed in [5], with a serpentine OPA obtained by the serial interconnection of grating waveguides. This particular serpentine configuration allowed the Authors to obtain a steering range of 35.8° in the longitudinal direction (with respect to the waveguide branches), whereas a 5.5° steering angle is reported in the lateral direction; the spot size was $0.11^\circ \times 0.2^\circ$.

While the above mentioned devices take advantage of the well established silicon-on-insulator (SOI) technology, many alternatives have been proposed exploiting different platforms. For example, a silicon nitride OPA able to operate from the near-infrared up to 500 nm with a steering angle up to 65° is proposed in [6].

In this context a paradigm shift is represented by the device proposed in [7]. As for the previously mentioned devices, a phase control region (control that can be implemented with different technological approaches according to the speed requirements) allows to obtain a set of independently controlled signals which are not fed into an array of antennas, but into a single multi-mode fiber (MMF). As a consequence, the obtained output signal is not a single-lobe beam anymore, but a multi-lobe speckle image. This particular output can be of great interest for single-pixel imaging applications, in particular when resorting to compressed sensing instead of the more common raster scan principle [8].

Whichever paradigm or technology is used to develop the desired OPA device, a crucial block in the device operation is the phase control region. In particular this feature can be implemented resorting to many different well established techniques, for example electrical and thermal control. While the former mechanism can result in better speed and tunability and as a consequence it is crucial e.g. in the communication field, the latter can be more easily implemented both in terms of costs and routing design.

Regardless of the considered OPA structure, the goal of this paper is to present how a Deep Learning (DL) approach can be successfully introduced in order to precisely control the properties of the output beam, in terms of position of the radiation peak (for beam steering applications) or radiation pattern (for beam shaping and single pixel solutions).

Indeed, in the recent years, DL in particular and Machine Learning (ML) in general have significantly affected the way photonic devices are designed and controlled, as shown by the large amount of papers and reviews that can be found in the literature ([9]–[14]). The development of efficient algorithms in the DL field has enabled the possibility to define the devices' topology meeting the desired requirements. Inverse design algorithms are widely exploited to optimize the component characteristics and are able to find non-intuitive and irregular solutions that can outperform the analytically or

empirically designed topologies. Moreover these approaches are also widely used for the characterization of the device's behavior since the simulations of photonic components based on the common Beam Propagation Method (BPM) or Finite Difference Time Domain (FDTD) schemes are generally extremely time consuming: on the contrary a properly trained machine can reasonably approximate the analyzed behavior in fractions of a second.

A basic architecture that is widely used in the characterization of these devices is the Multi-Layer Perceptron [9]. This particular type of Neural Network (NN) is composed by a cascade of Fully Connected Layers (FCLs) and it is able to fit any type of continuous function with a finite number of layers and neurons per layer, making it the perfect approximator for characterization applications. MLPs can also be applied in design problems being aware of providing a second MLP in charge of studying the direct problem (mapping the topology to the output response) in order to obtain a convergence in the design of the custom topology. This architecture is called multi-directional MLP and it is widely used in the definition of geometries in the photonic and plasmonic field. An exemplary case is represented by [12] and [13], where the goal is the generation of a machine able to design a compact 1-to-2 MMI splitter in SOI technology with a given splitting ratio. In order to minimize back-reflection of the incident field, the morphology of the splitter was characterized by the presence or absence of SiO_2 holes in the silicon according to a boolean 20×20 pattern. By training a NN with a proper amount of randomly generated patterns and the corresponding splitting ratios obtained through FDTD simulations, it is possible to generate in negligible time the morphology of the splitter meeting the requirements.

III. CASE STUDY

A. DEVICE DESCRIPTION

The studied OPA was part of a larger integrated sensor designed for biomedical applications. In the considered optical circuit, the light beam coming from a laser source is divided into a fixed number of branches: the relative phases of the beams in each branch are individually controlled through thermal tuning and the fields are recombined in the final radiating region. From now on we will focus on this latter region (see Fig. 1a), discussing the chosen topology and its characterization. The combiner here discussed has to be thought as a proof of concept of the feasibility of the characterization algorithm, and has to be optimized and re-cast to a target technology to match the desired specification.

The specific biomedical target sets a constraint on the frequency range to the near and mid-infrared as in these regions many typical footprints of molecules interesting in the biomedical field can be found [15]–[17]. We chose a SOI technology with a 500 nm overlayer (instead of the more standard 220 nm), which is a common solution in the sensor field [18]. This solution is in accordance with the operating

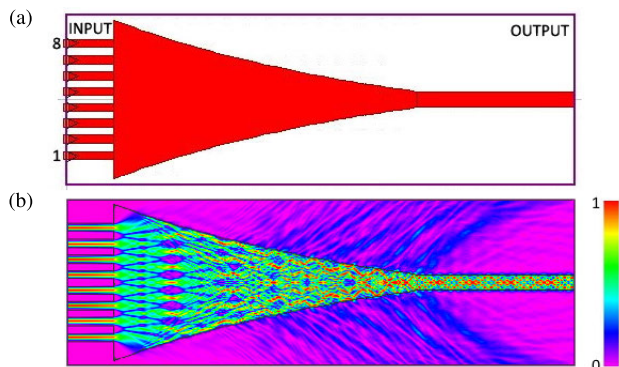


FIGURE 1. (a) Top view of the considered device. Red region: silicon, white region: silicon oxide. The size of the computational box, shown by the purple external rectangle, is $52 \mu\text{m} \times 160 \mu\text{m}$. (b) Example of RSoft BeamPROP simulation result, obtained when no optical phase shift occurs in the input branches. The normalized modulus of the calculated field is shown in linear scale.

frequency which is fixed to $3.5 \mu\text{m}$, close to the upper limit for SOI technology [15]–[17]; as a consequence the width of the waveguides is fixed to $2.5 \mu\text{m}$ to ensure single mode behaviour. The radiation collected from the optical source was divided in eight branches: this number is sufficiently high to exhibit the device capability and sufficiently low to maintain under control the complexity of the successive steps. The device can be easily extended to have more branches and provide a 2D control thanks to additional structures such as arrays or stacks of antennas.

In Fig. 1(a) the beams entering the computational domain in the 8 leftmost branches are then re-combined in an MMI segment whose width is tapered from 20 times to twice the input waveguides width over a length of $100 \mu\text{m}$, and finally they focused into a multi-modal output waveguide.

We used the BeamPROPTM tool of the RSoft CAD EnvironmentTM to perform 3D beam propagation simulations of this device. In each branch k the fundamental mode is applied in input, possibly multiplied by a phase term $e^{j\phi_k}$ introduced by the thermal tuning applied in each branch. An example of computation result, obtained with null input phase difference, is shown in Fig. 1(b).

The far field of the electrical field at the rightmost facet is finally evaluated: some exemplary results are presented in Fig. 2. In Fig. 2(a) the relative phase shift in the branches is zero: in this case the response is a narrow and symmetric single lobe, which can be displaced in the transverse direction by applying a linearly increasing phase shifts in all the branches ($\phi_k = (k - 1)\Delta\phi$, $k \in [1, 8]$), as shown in Fig. 2(b),(c). In these cases the full width at half power of the beam goes in the interval 1° to 7° with a total steering angle larger than 30° . These results prove the steering capability of the system and the possibility to effectively employ it in raster-scan single-pixel imaging systems. Finally, in Fig. 2(d) an arbitrary set of phase variations was applied to the 8 branches, which resulted in a far field characterized by a multi-lobe response. Again, this particular type of output can be well suited for

compressed sensing algorithms where information from the sample is retrieved by illuminating the whole sample with a sequence of pseudo-random multi-lobe spectra instead of scanning a single lobe beam over the sample while fetching information one pixel at the time: a proper set of illuminating spectra together with a computational optimization algorithm can drastically reduce the number of measurement (and, as a consequence, the time) needed to retrieve an image from the tested sample, with respect to a raster scan algorithm [8]. Due to the particular shape of the considered device, the polar representation of the far field profiles in Fig. 2 can not be predicted using analytical expressions; moreover there is no guarantee that a specific far field profile can be generated by a unique combination of the control phases. This remarks justify the introduction of a ML based approach.

B. DEEP LEARNING STUDY

Similarly to what described in [12], [13], we implemented an *inverse* design approach to obtain a set of input waveguides phases ϕ_k able to produce a desired far field pattern. To do that we designed a DNN where the data-set (input) was represented by a collection of far-field images and the outputs to predict were the control phases ϕ_k . Due to the nature of the problem, we opted for a particular type of DNN called Convolutional Neural Network (CNN). This type of network is particularly well suited for processing data represented in high dimensional spaces like the images that we were able to collect from the RSoft simulations, in which high correlation is expected between local points.

We were able to conduct the whole study in the MATLAB environment using the MATLAB Deep Learning Toolbox. The first step was to collect a sufficiently large data-set to train and validate the NN. This was done by automatically generating random set of control phases and invoking RSoft a predefined number of times N_{in} ; the calculated far field intensities were stored in bitmap images with 46 rows

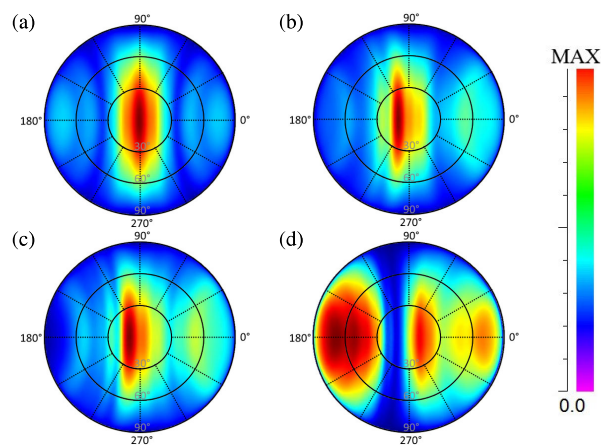


FIGURE 2. Examples of simulated far-field polar profiles for the considered 1D device with 8 input branches. (a) No phase shift in the branches; (b) and (c) linear phase shifts with $\Delta\phi = 2^\circ$ and $\Delta\phi = 3^\circ$ respectively; (d) arbitrary phase shifts.

(corresponding to the radial coordinate, from 0° to 90° with step 2°) and 181 columns (corresponding to the angular coordinate, from 0° to 360° with step 2°). These images were used as input to the NN. The number of RSoft runs N_{in} was fixed to 10000 [12]; the whole data-set was then divided into a training and a validation batch according to a 75%–25% proportion. The training set was used to tune the coefficients of the NN, while the validation set was used to test the validity of the network during the training process.

The chosen CNN was composed by convolutional layers and an output FCL. Normalization and averaging layers were inserted between the convolutional layers to smooth the response and increase the convergence speed. The final architecture of the trained CNN is shown in Fig. 3.

The size of the convolutional layers was decreased from input to output and particular attention was paid to the size of the first layer, since a too small convolutional filter could sweep out the typical feature of the input images, compromising the training convergence. The presence of the output FCL was mandatory to link the results from the convolutional mask to the eight output control phases. The training process of the network was organized in epochs in which the full batch of training images were fed to the NN to match the outputs by tuning the internal coefficient of the machine. The RMSProp algorithm was found to be the best in minimizing the Root Mean Square Error (RMSE) in the outputs prediction. To test the validity of different network architectures we implemented a specific set of training parameters; the results for the tested architectural solutions are reported in Table 1. We set the training duration to 500 epochs, while the initial learn rate (ILR), the learn rate drop period (LRDP) and the learn rate drop factor (LRDF) were set to 0.01, 100 epochs and 0.5, respectively.

The number of convolutional layers was fixed to three as a trade-off between training speed and prediction accuracy:

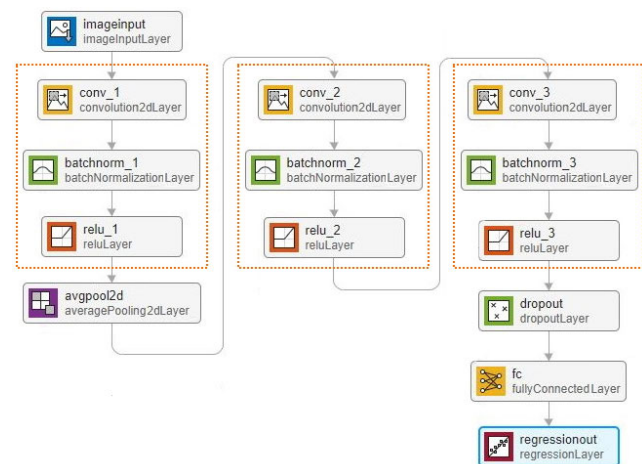


FIGURE 3. Block diagram of the considered CNN architecture. The dashed rectangles indicate the 3 considered layers, whose convolutional 2D filters sizes were finally set to 32, 12, and 4, respectively.

TABLE 1. Summary of the tested CNN architectural solutions. (†) indicates the chosen configuration.

Layers width	Number of masks	Final RMSE	Training time (h)
32-16-8-4	4-8-16-32	0.28	20
40-20-10-5	4-8-16-32	0.30	22
32-8-4	4-8-16	0.37	8
32-12-4 †	4-8-16	0.30	15
16-8-4	4-8-16	0.33	6.5
32-4	10-40	0.30	>24

additional layers did not significantly increase such accuracy at the price of almost doubling the training time and we finally opted for the network with convolutional sizes of 32, 12, and 4, respectively, due to the better compromise between results quality and training time. The effect of a variation of the training parameters (ILR, LRDP, LRDF) was also investigated for this specific configuration, but the values previously listed proved to represent an excellent trade-off between the final RMSE value and the simulation time. In particular with fixed ILR a larger LRDF leads to worse RMSE and worse prediction results while a smaller LRDF leads to longer simulation time with no improvement in the final RMSE. Moreover we observed that starting the training process with a larger ILR and different LRDF did not result in an improvement of the final RMSE.

For the chosen NN architecture, in addition to the RMSE indicator, which is a compact representation of the accuracy of the NN, we visually tested the validity of the predictions. To do that we asked the NN to predict the control phases for some far field images taken from the validation batch: this group of images was never presented to the NN for the training process *per se* but only in the validation step, and as a consequence the machine was never trained to reproduce these particular results. The predicted control phases were then used in RSoft BeamProp simulations to calculate the corresponding far field profiles. In Fig. 4, we report in row (a) the original far field profiles calculated by RSoft using random sets of control phases ϕ_k and in row (b) the far field profiles calculated using the predicted phases χ_k returned by the trained NN. The values of ϕ_k and χ_k used for each case are shown in Fig. 4, row (d); for each considered case, the average ($\langle \varepsilon_k \rangle$), the maximum ($\max(\varepsilon_k)$) and the standard deviation ($\text{std}(\varepsilon_k)$) of the errors $\varepsilon_k = |\phi_k - \chi_k|$ are finally reported in row (d).

These results are really satisfactory, and show a good accuracy in the prediction of the control phases, which leads to a good accuracy in the far fields profiles. Indeed, high accuracy is achieved both in the cases where the field presents a single-lobe behavior and where a more complex speckle response is present.

It has to be pointed out that in a real application scenario the user cannot be sure *a priori* that a certain profile can be effectively generated using the considered OPA layout, the

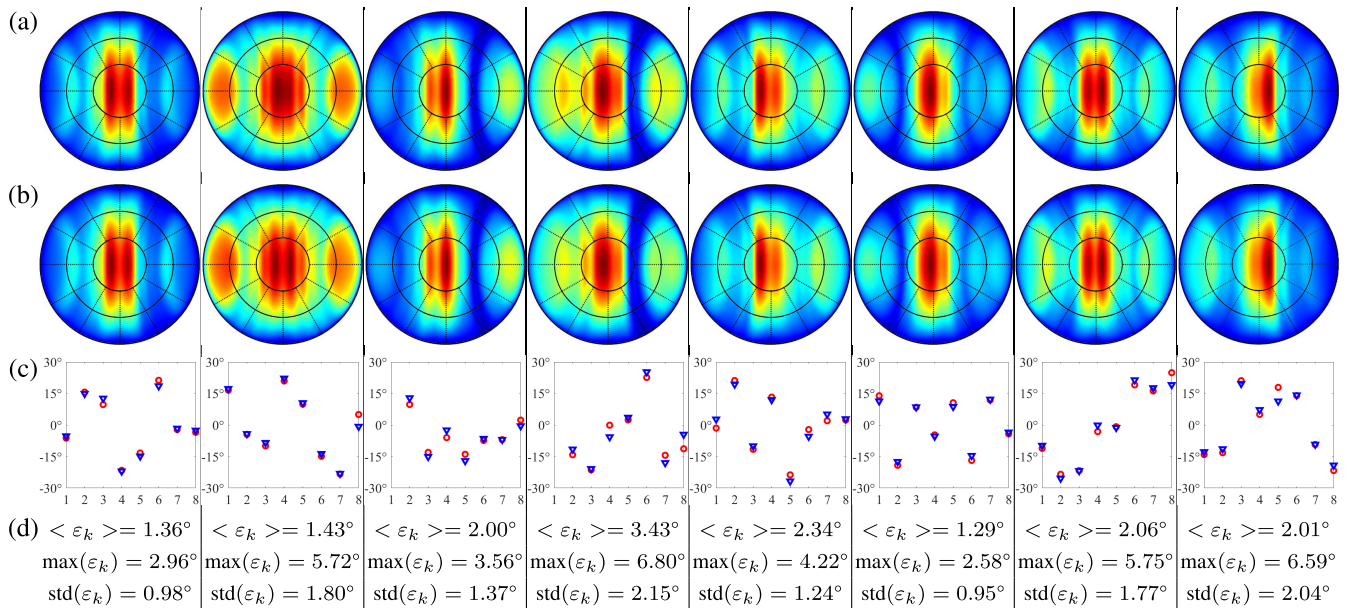


FIGURE 4. Row (a): far fields profiles associated to randomly generated ϕ_k parameters; row (b): far fields profiles calculated with the phases χ_k predicted by the NN once it is presented with the far field distributions on row (a). Simulations performed with the RSoft BeamPROP tool; same color scale as in Fig. 2. row (c): corresponding ϕ_k and χ_k control phases shown with red circular markers and blue triangular markers, respectively, for the 8 control branches. row (d): summary of the prediction errors.

ensemble of allowed far field configurations depending on the device geometry and on the number of input branches. If a profile lying out of the device capabilities is requested, the trained agent returns in any case a set of control phases which can lead to a significantly different far field. It is thus indispensable to associate to the proposed inverse neural network a direct one to generate the far field profile based on the predicted control phases, to finally compare it with the original requirements without the expansive computational cost of the BPM simulation.

IV. CONCLUSION

We presented the application of a Deep Neural network to successfully predict the far field profile in output to an OPA device with great accuracy by generating the proper control signals to match the specifications.

The proposed approach can be easily extended to more complex devices using a more advanced antenna system (for example 3D antenna arrays), to control with more degrees of freedom the output pattern.

ACKNOWLEDGMENT

The software tool BeamPROP has been provided by Synopsys within the OSG University Program—site ID 31272.

REFERENCES

- [1] J. He, T. Dong, and Y. Xu, "Review of photonic integrated optical phased arrays for space optical communication," *IEEE Access*, vol. 8, pp. 188284–188298, 2020.
- [2] S. Chung, H. Abediasl, and H. Hashemi, "A monolithically integrated large-scale optical phased array in silicon-on-insulator CMOS," *IEEE J. Solid-State Circuits*, vol. 53, no. 1, pp. 275–296, Jan. 2018.
- [3] F. Ashtiani and F. Aflatouni, "Optical beam steering using an $N \times N$ optical phased array with $2N$ phase shifters," *Opt. Exp.*, vol. 27, no. 19, pp. 27183–27190, Sep. 2019.
- [4] R. Fatemi, A. Khachaturian, and A. Hajimiri, "A low power PWM optical phased array transmitter with 16° Field-of-View and 0.8° beamwidth," in *Proc. IEEE Radio Freq. Integr. Circuits Symp. (RFIC)*, Jun. 2018, pp. 28–31.
- [5] N. Dostart, B. Zhang, A. Khilo, M. Brand, K. A. Qubaisi, D. Onural, D. Feldkhun, H. K. Wagner, and A. C. Miloš Popovi, "Serpentine optical phased arrays for scalable integrated photonic lidar beam steering," *Optica*, vol. 7, no. 6, pp. 726–733, Jun. 2020.
- [6] H. Wang, Z. Chen, C. Sun, S. Deng, X. Tang, L. Zhang, R. Jiang, W. Shi, Z. Chen, Z. Li, and A. Zhang, "Broadband silicon nitride nanophotonic phased arrays for wide-angle beam steering," *Opt. Lett.*, vol. 46, no. 2, pp. 286–289, Jan. 2021.
- [7] T. Fukui, Y. Kohno, R. Tang, Y. Nakano, and T. Tanemura, "Single-pixel imaging using multimode fiber and silicon photonic phased array," *J. Lightw. Technol.*, vol. 39, no. 3, pp. 839–844, Feb. 1, 2021.
- [8] M. P. Edgar, G. M. Gibson, and M. J. Padgett, "Principles and prospects for single-pixel imaging," *Nature Photon.*, vol. 13, pp. 13–20, Jan. 2019.
- [9] W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nature Photon.*, vol. 15, no. 2, pp. 77–90, Feb. 2021.
- [10] P. R. Wiecha, A. Arbouet, C. Girard, and O. L. Muskens, "Deep learning in nano-photonics: Inverse design and beyond," *Photon. Res.*, vol. 9, no. 5, p. B182, May 2021.
- [11] I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via deep learning," *Light, Sci. Appl.*, vol. 7, no. 1, p. 60, Sep. 2018.
- [12] M. H. Tahersima, K. Kojima, T. Koike-Akino, D. Jha, B. Wang, C. Lin, and K. Parsons, "Deep neural network inverse design of integrated photonic power splitters," *Sci. Rep.*, vol. 9, no. 1, p. 1368, Feb. 2019.
- [13] K. Kojima, M. H. Tahersima, T. Koike-Akino, D. K. Jha, Y. Tang, Y. Wang, and K. Parsons, "Deep neural networks for inverse design of nanophotonic devices," *J. Lightw. Technol.*, vol. 39, no. 4, pp. 1010–1019, Feb. 15, 2021.
- [14] R. Trivedi, L. Su, J. Lu, M. F. Schubert, and J. Vuckovic, "Data-driven acceleration of photonic simulations," *Sci. Rep.*, vol. 9, no. 1, p. 19728, Dec. 2019.

- [15] D. Popa and F. Udrea, "Towards integrated mid-infrared gas sensors," *Sensors*, vol. 19, no. 9, p. 2076, May 2019.
- [16] R. Soref, "Mid-infrared photonics in silicon and germanium," *Nature Photon.*, vol. 4, pp. 495–497, Aug. 2010.
- [17] R. A. Soref, S. J. Emelett, and W. R. Buchwald, "Silicon waveguided components for the long-wave infrared region," *J. Opt. A, Pure Appl. Opt.*, vol. 8, no. 10, pp. 840–848, Oct. 2006.
- [18] D. Thomson, A. Zilkie, J. E. Bowers, T. Komljenovic, G. T. Reed, L. Vivien, D. Marris-Morini, E. Cassan, L. Viot, J.-M. Fédéli, J.-M. Hartmann, J. H. Schmid, D.-X. Xu, F. Boeuf, P. O'Brien, G. Z. Mashanovich, and M. Nedeljkovic, "Roadmap on silicon photonics," *J. Opt.*, vol. 18, no. 7, Jun. 2016, Art. no. 073003.



PAOLO BARDELLA (Member, IEEE) received the Laurea degree in electronic engineering and the Ph.D. degree in electronic and telecommunication engineering from the Politecnico di Torino, Torino, Italy, in 2002 and 2006, respectively. He is currently an Associate Professor with the Politecnico di Torino. His research interests include the simulation of quantum dot-based lasers operating in mode-locking and self-pulsating regime, the study and optimization of semiconductor waveguide for lasers and SLDs, and the design of silicon photonics-based passive components.

• • •



DANIELE SAVIO received the bachelor's degree in physics engineering and the master's degree in electronics of micro and nano systems from the Politecnico of Turin, Italy, in 2018 and 2021, respectively. He is currently an Early Stage Researcher with the Politecnico of Turin. His research interests include silicon photonics and its application in the biomedical field.