Generative adversarial neural networks model of photonic crystal fiber based Surface Plasmon Resonance Sensor

Aimen Zelaci, Ahmet Yasli, Cem Kalyoncu, and Huseyin Ademgil, *Senior Member, IEEE*

Abstract—Photonic crystal fibers (PCF) for specific applications are designed and optimized by both industry experts and researchers. However, the potential number of design combinations and flexible propagation features are offering a wide range of application areas, resulting in an exponential increase in the search space. This issue combined with the speed of the commonly used Full Vectorial Finite Element Method (FV-FEM) simulators causes the task to take a significant amount of time. As stated in the previous works, artificial neural networks (ANN) can be employed to predict the result of numerical simulations much faster. However, there are two issues with the methods proposed previously: the required number of samples for training and the generality of these methods. In this paper, we propose the use of generative adversarial networks (GAN) to augment the real dataset to train an ANN model. The experimental analysis suggested that the proposed combination not only accurately predicts the confinement loss even with limited amounts of data but also GAN can be used to improve existing methods in the literature. Moreover, the proposed system can also predict the confinement loss over a range of analytes and wavelengths in a completely new geometric configuration and generic enough not to require any additional tuning when used in a new dataset. This property is demonstrated by experimenting on an existing PCF dataset, which the proposed system surpassed the accuracy of the method proposed for it.

Index Terms—Artificial neural network (ANN), fiber optic sensor, generative adversarial networks (GAN), machine learning, photonic crystal fiber (PCF), surface plasmon resonance (SPR).

I. INTRODUCTION

Due to their unique propagation features, Photonic Crystal Fibers (PCFs) have been integrated into a number of optical based devices and applications. Design flexibilities and precise light guidance abilities have made them more promising in comparison to conventional step index fibers in various optical applications including sensing [1].

In the last decade, PCF based structures are extensively employed as sensing devices. Possibility of filling the cladding holes with gas and liquid analytes makes them favorable in

A. Z. Author is with the European University of Lefke, Faculty of Engineering, Electrical and Electronics Engineering Department, Lefke, TR-10, Turkey (e-mail: aimen.zelaci@gmail.com).
A. Y. Author is with the European University of Lefke, Faculty of Engineering, Electrical and Electronics Engineering Department, Lefke, TR-10, Turkey (e-mail: ayasli@eul.edu.tr).
C. K. Author is with the European University of Lefke, Faculty of Engineering, Computer Engineering Department, Lefke, TR-10, Turkey (e-mail: ckalyoncu@eul.edu.tr).
H. A. Author is with the European University of Lefke, Faculty of Engineering, Computer Engineering Department, Lefke, TR-10, Turkey (e-mail: hademgil@eul.edu.tr).

both chemical and biological sensing applications [1]–[3]. Moreover, due to adaptable light guidance with possibilities of controlling the confinement loss and the chromatic dispersion at any operating wavelength are distinguishing features of PCF based structures in the literature. These charming propagation features can be achieved by employing, circular, cylindrical and even square and rectangular shaped air holes in the cladding and in the core region of the structures. Furthermore, the cladding pattern arrangements such as hexagonal, octagonal, decagonal, square and circular [1], [4] can be engaged to achieve the desired features for preferred applications. The hole-to-hole spacing (pitch) and the air hole sizes are critical parameters of tuning the aforementioned propagation characteristics. On the other hand, surface plasmon resonance (SPR) is a unique technique that occurs with light and free electrons interaction at metallic boundaries. This interaction results in the excitement of surface plasmon polariton (SPP). Relative permittivity variations of the analytes will change the characteristics of the SPP wave, where sensing is triggered. Kretschmann Raether prism geometry is the fundamental construction principle of SPR based sensors [5]. This geometry practices Snell's Total Internal Reflection rule, where still a very small amount of light refracts from the metallic layer into the evanescent field. The refracted photons are coupling with electron charge oscillations (plasmon) at the surface of the metallic layer, which are known as surface plasmon waves (SPW). The maximum energy transfer from core mode to SPWs occurs with the effective index equalization of these corresponding modes. This phenomenon is known as resonance phase matching, where losses are reaching the maximum levels. The possible changes in the analyte index will directly affect the resonance matching wavelengths. These possible variations are the main parameters for sensitivity calculations. The deep research on SPR based sensors have shown that Kretschmann Raether's prism geometry fails to satisfy the remote sensing and miniaturization requirements of critical detecting applications. One of the most common solutions for overcoming these limitations is replacing the prism with the core of an optical fiber. Considering the recent technological developments, PCF is offering a tremendous amount of design flexibility with the possibility of complete control of light guidance. Hence, SPR technique can be integrated into this unique microstructured fibers for developing a comprehensive multi-purpose sensor [2], [6]. Following the exclusive study by Hassani and Skorobogatiy [7], a number of structures have reported the merits of PCF based SPR sensors. In the past,

structures that contain features such as multi-channel single analyte and multi-channel multi-analyte have been investigated [2], [6], [8], [9]. Also, the effects of various metallic layers such as Au, Ag [8], Ag- Graphene [2], Ta2TO5 [10], ITO [3], TiO2 [2] with various design parameters have also been investigated in detail. Several design parameters such as plasmonic material, the thickness of the layers, air hole size and positions are decisive on resonance matching wavelength of the SPR based structures. Optimizing and refinement of these parameters is an extremely time consuming process with considerable computational time [11]–[13]. In the past, genetic algorithms [12], [14] method was used to optimize the structural design parameters to achieve better propagation features of photonic devices [2]. Nowadays, in order to optimize such structures, artificial neural networks (ANN), and other machine learning (ML) algorithms are being applied [11], [13], [15]. Although ANNs and ML have been applied to PCFs [11], [13] and SPR structures [16], [17] independently, to the best of our knowledge, there seems to have been no published papers that investigated the PCF based SPR structures with this well tested methods.

Machine Learning (ML) techniques are becoming a common tool in many fields and surpassing human performance in many tasks, such as automatic speech recognition, image recognition, natural language processing, drug discovery and toxicology. Additionally, ANN can approximate any function proven by the Universality theorem [18]. This fact propelled researchers to widen the applications of ANN even further, including the study of nano-photonic structures [19], optimization of photonic crystal nano-cavities [20], and more recently, computing optical properties of a PCFs [13] and SPR structures [16], [17].

One of the most difficult challenges that deep learning models are facing is that they benefit from large amounts of data to train, which is often costly and/or difficult to acquire. An alternative solution is to artificially expand the original training dataset by the means of generative networks. In this regard, Generative Adversarial Networks (GAN) introduced by Goodfellow et al [21], has proven to be successful in data generation [22]–[24].

In this work, we focused on estimating the confinement loss behavior of multichannel PCF based SPR sensors by employing the ANNs. Specifically, we have constructed our system on a PCF based SPR model. However, in the experiments section, we have demonstrated that the designed system is generic enough to be applied to multiple similar PCF based designs. The most important contribution of this research is the use of the GAN phase, where the available data is expanded to be used in the training phase.

To our knowledge, Ref. [15] was the first work proposed ANN technique on PCF structures. The results demonstrated the accuracy of the network on a specific PCF structure. On the other hand, the ANN method was employed for predicting the chromatic dispersion of conventional hexagonal lattice PCF structure [25]. In Ref. [26],authors employed a neural network for the refinement and design of Erbium doped PCF. Presented results have shown that conventional methods consume large amounts of time compared to the runtime of the

neural network. Perhaps the most rigorous experiments about neural network use in PCF design is presented in [13].In this work, a feed-forward multi-layer perceptron neural network is used to calculate multiple propagation features of a PCF design. It can be seen that this proposed network can achieve both acceptable accuracy and superior speed compared to the existing numerical methods. On the other hand, genetic algorithms and ANN methods have also been tested on SPR sensor structures [16], [17]. These methods also deliver robust performances.

The accuracy of existing image classifiers of the neural networks can be improved by using GAN [24]. Frid-Adar, et. al. [23] have applied the same method to improve the accuracy of liver lesion classification. It is found that balancing the training data with GAN has a higher improvement in accuracy compared to classical data generation strategies. The same technique was applied to augment and anonymize the training data of the brain MRI [27].

This paper is organized as follows. The details of GAN procedure to generate additional training samples for ANN as well as the proposed neural network architecture is presented in Section II. The design parameters and details of PCF based SPR structure is described in section III. Detailed analysis of the numerical results is discussed in section IV. Finally, concluding remarks are presented in Section V.

II. Network Architecture

In this section, details of the proposed method are discussed. Training of the system contains two phases: a GAN phase and a regression training phase. At the beginning of the training, a GAN is trained to generate additional data by using training samples. These generated samples are filtered to ensure they are within the numeric range of the application. Original training samples and generated samples are joined to train a fully-connected feedforward multi-layer perceptron (MLP) neural network. This satisfies the regression need for estimating the confinement loss. At the end of the training, the ANN will be used to decide the confinement loss of a previously unseen set of input parameters. The illustration of the proposed architecture is presented in Figure 1. The details of the proposed GAN and ANN architecture is discussed in the following sub-sections.

A. Generative adversarial network design

In our study, GAN is employed to augment the number of samples that are used in training [21]. A GAN is composed of two neural networks: generator and discriminator. The aim of the generator is to transform fully randomized data into form that follows the distribution of the original dataset. The discriminator assesses the performance of the generator and provides feedback for training. Instead of directly training the generator, it is trained through this feedback. This paradigm avoids overfitting the data.

The GAN can be trained by diverse metrics [21], [28]]. In this work, Wasserstein distance metric (WGAN) is employed [29]. In this system, the discriminator is named as critic and it measures the distance between the generated and the real
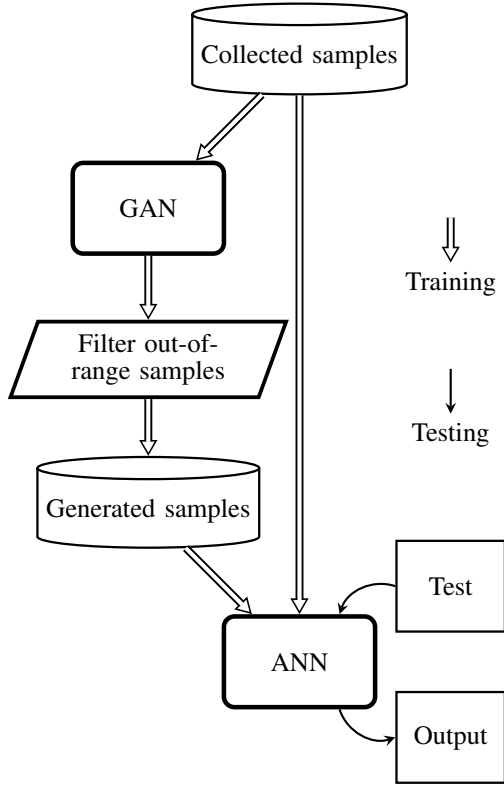
Figure 1: System architecture

data. The distinguishing feature of the WGAN compared to other methods is the ability to determine stopping criteria. In a regular GAN system, training is stopped when the generated data is deemed viable by an observer. GAN is frequently used in image, video or audio generation, where human is used effectively in the loop. In our case, it is not viable for a human to judge the generated data, therefore, automating this procedure to remove the human in-the-loop can lead to over or under-fitting, which may degrade the performance. On the other hand, WGAN uses an adaptive stopping criteria that can easily overcome the aforementioned disadvantages. Additionally, Gradient penalty is incorporated to improve the WGAN as described in [30]. This proposed enhanced WGAN system converges in a stable manner without requiring further fine tuning the hyper-parameters of the system.

In our work, both the critic and the generator are fully connected feed forward MLP models. The detailed parameters of the networks are given in Table I. The output of the system would be input and output pairs that are used to train ANN part of the system. In our proposed model, we have 7 input parameters and a single output parameter, making a total of 8 that will be used in the GAN phase. In Section IV, various PCF structures are tested with different inputs, resulting in a different number of system parameters. The generator is supplied with the same number of parameters. These parameters are generated using Gaussian noise and this input is called a latent variable.

Previous works [31], [32], demonstrated that it is possible for random data augmentation to weaken the performance of the model. Hence, it is important to sample the generated

data in a way to prevent performance degradation [33]. In our proposed system, we have included a filtering step for the generated data. This step uses a simple condition to discard the values that fall out of the desired range of both the independent variables $(n_{analyte}, \lambda, \Lambda, d1, d2, d3, n_c)$, and the confinement loss. Once the training phase is completed, the generator and the filter are used to augment the number of training samples that are available for ANN to train on.

B. Artificial neural network design

The ANN model employed in this work is a fully connected feed-forward MLP network consisting of an input layer, an output layer, and 5 hidden layers. The ANN is designed and trained to estimate the confinement loss in a regression configuration. We have applied log_{10} to confinement loss calculation in order to keep its numeric stability across a wide range [13]. Each hidden layer contains 50 neurons and uses Rectified Linear Unit (ReLU) activation function. Table I summarizes the details of this model. The number of hidden layers and neurons are empirically determined to be able to estimate the loss curve while keeping the complexity of the network to a minimum. With these changes, the proposed network is able to predict more complex functions, however, it will also require more training samples.

In our model, we have used mini-batch training with 48 samples as the batch size. In contrast to full-batch training, mini-batch training allows more frequent gradient calculations, improving the stability and reliability of the training [36], [37]. The weakness of using mini-batches is added computational complexity during training. However, it has no adverse effect on the testing phase.

We have adopted Adam [34]as the optimizer and mean squared error (MSE) as the loss/cost function. In addition, Batch Normalization algorithm [35] is employed to accelerate the training, reduce the effects of initial randomized state and mitigate the problem of internal covariate shift. Also, the ANN is trained using the training samples from the original dataset and the augmented samples generated by the GAN phase.

III. PHOTONIC CRYSTAL FIBER DESIGN

In order to build the desired dataset, the numerical analysis of the proposed sensor structure is carried out with the commercial software package Comsol Multiphysics. The Full Vectorial Finite Element Method (FV-FEM) with the perfectly matched layers (PMLs) [38] are employed.

The cross-section of the proposed structure is illustrated in Figure 3. As can be seen from this figure, the sensor consists of an analyte channel (sensing medium), plasmonic layer (Au) and various sized cladding air holes. The cladding region is formed by nineteen various sized air holes, sited on a silica background with a hexagonal lattice arrangement. The air hole diameters are set to; $d_1 = 0.45\mu m$, $d_2 = 0.75\mu m$, $d_3 = 0.35\mu m$, and $d_c = 0.15$. The distance between air holes (Λ) is fixed to $2\mu m$. Initially, trapping the light (fundamental mode) into the core region is vital for mode quality. On the other hand, some amount of light needs to be leaked towards the plasmonic layer, where the sensitivity measurement occurs. This is the

Table I: Details of ANN models

| Parameter	Generator	Critic	Regressor
Hidden layers	5	5	5
Neurons in hidden layers	128	$2^{layer\_index} \times 32$	50
Batch size	32	32	48
Activation function	ReLU	Leaky ReLU	ReLU
Optimizer	Adam [34]	Adam	Adam
Input normalization	None	None	None
Layer normalization	None	Batch [35]	Batch
Loss function	$-Critic[gen]$	$Critic[gen] - Critic[col] + Penalty_{grad}$	MSE

gen: generated, col: collected, grad: gradient
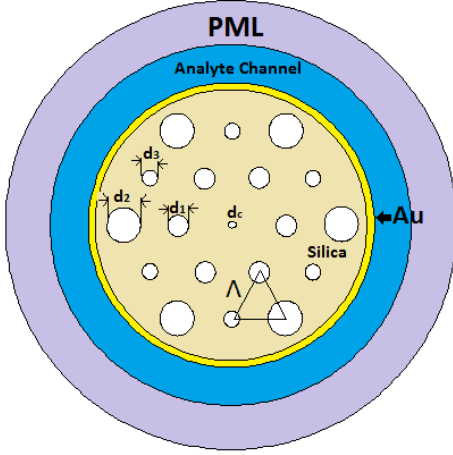
Figure 2: The cross-section of proposed PCF based SPR sensor.

key phenomenon for resonance phase matching occurrence. In this respect, the various sized cladding holes are employed for controlling the light guidance, where each set of air holes are affecting different propagation characteristics of the sensor. The Gold (Au) material is used as a plasmonic layer and set to be $40\mu m$. The Johnson and Christy Data [39] is used for the permittivity values. The Sellmeier Equation (Eq.(1)) is used for calculating the refractive index of Silica material [40].

$$n^2(\lambda) = 1 + \frac{B_1(\lambda)^2}{(\lambda)^2 - C_1} + \frac{B_2(\lambda)^2}{(\lambda)^2 - C_2} + \frac{B_3(\lambda)^2}{(\lambda)^2 - C_3} \quad (1)$$

where λ is the operating wavelength, n is the Refractive Index (RI), $B_{1,2,3}$ and $C_{1,2,3}$ are the Sellmeier constants [40].

As mentioned in the introductory part, during the resonance phase matching condition, the confinement losses are reaching the highest rate. Therefore, the precise calculation of the confinement losses is critical. Equation 2 is used in this study for that purpose [6].

$$\frac{40\pi}{ln(10)\lambda} Im(n_{eff}) \times 10^4 [dB/cm] \quad (2)$$

where the imaginary part of the effective RI shown with $Im(n_{eff})$.

A labeled dataset of 432 samples was provided by FVFEM numerical analysis. Our dataset contains a number of independent variables, such as the operating wavelength λ, analyte index $n_{analyte}$, hole to hole distance Λ, the air holes sizes (d_1, d_2 and d_3), and the real part of the refractive index (n_c).

The labels are specified as the confinement loss of the sensor structure. The set consists of nine different configurations of the geometric properties (Λ, d_1, d_2, d_3), for each configuration the confinement loss was calculated for three different analytes (Water (n=1.33), Ethanol (n=1.35) and several commercial hydrocarbon mixtures (n=1.36)).

The optimizing and refinement of these above parameters are extremely time-consuming and requires considerable computational time. In order to optimize the structural design parameters and achieve the desired propagation features, we have fixed the geometry of the sensor structure (hexagonal arrangement with circular air holes) and vary the operating wavelength λ, index of refraction $n_{analyte}$, hole to hole spacing Λ, and the size of the cladding holes (d_1, d_2 and d_3).

IV. EXPERIMENTS

A. Experimental setup

In order to prove the effectiveness of the proposed setup, we have performed multiple experiments over two datasets. We have built the first dataset that is used in the experiments. Details of this dataset is explained in Section III and it is referred to as SPR dataset. We have performed 9-fold testing on this dataset, due to having nine configurations. Each fold tests a single configuration using the other configurations as training data. Slicing the data allows us to guide the network to predict the loss on a configuration that has not been used in training. Please note that we have applied log_{10} to confinement loss for scaling.

Also, the mean square error (MSE) metric is used for the comparisons. All algorithms are implemented using TensorFlow with Keras library in Python and experiments are performed on a device with Intel i7-5600 CPU and 8GB of RAM. TensorFlow library is set to use only CPU for the experiments.

Additionally, we have used the dataset that has been used in [13]. This dataset contains 1117 samples. There are 5 inputs: refractive index of the core (n_c), wavelength (λ), hole diameter (d), pitch (Λ), and number of rings (N_r); and outputs are effective index (n_{eff}), effective mode area (A_{eff}), dispersion (D), and the confinement loss (α_c). We have only used confinement loss as the output. This dataset is referred to as PCF. As a final note, the range and domain of these datasets are different; therefore, all numeric comparisons should be drawn within the same dataset.

In the following subsections, we have performed experiments to show the performance of the base ANN system

by its own comparing with the state-of-the-art method in the literature [13], improvement gained by employing GAN phase, and finally the computational advantages of using ANN over the simulator. 5000 training epochs are used for [13] as suggested in the original paper. All numeric comparison results are averaged over multiple executions: 9-fold testing for SPR dataset and 10-fold testing for PCF. The number of testing samples used in SPR dataset experiments is 48, PCF experiments are 112. Non-averaged data used in figures, such as configuration estimation, are selected from non-extreme cases for the compared methods.

B. Performance of ANN

In this set of experiments, we have analyzed the performance of the proposed ANN design. The first experiment involves training loss over the SPR dataset. In this experiment, we have trained the ANN model for up to 3000 epochs and reported the training loss in Figure 3. According to this experiment, training our ANN architecture more than 2500 epochs does not yield any improvement. Therefore, the training epoch limit is set to 2500 in all subsequent experiments. However, this number depends on the dataset and additional training may improve the results. Additionally, the benefit of augmented samples from the GAN phase is evident even at the training as they reduce the fluctuations in the training loss and validation MSE.

We have performed experiments comparing the proposed ANN architecture with the state-of-the-art method proposed in [13]. Results of this experiment are demonstrated in Figure 5. In the SPR dataset experiment, ANN models are predicting a new configuration and the number of samples available for training is lower. Therefore, they both have higher error rates compared to the PCF dataset. According to Figure 5c, $n_{analyte} = 1.35$ case, the ANN method proposed in [13] is not a good predictor of the shape of the curve caused by the change in the wavelength, even though it fits this curve much better than the proposed method, showing that the proposed method requires more samples for training. Regardless of this, both models have similar MSE due to the shift in the curve fit. Since the network is never exposed to this particular geometric configuration shift in the curve fit occurs.

C. Performance boost of GAN

The first set of experiments performed on the GAN phase is the number of the augmented samples. In these experiments, we have augmented the proposed ANN system with a varying number of samples. According to these experiments we have reached the minimum MSE at 1000 samples. Adding further samples slightly reduces the performance while increasing the training time. The results of this experiment is demonstrated in Figure 6 showing that the increasing number of augmented samples increases the performance to a point where it will start hurting the performance (> 3000 samples). The reason behind this augmented samples are not perfect. They deviate from the actual data. However, up to a point, the performance impact of having a low number of training samples out-weighs the imperfect data generated by GAN. The number of augmented
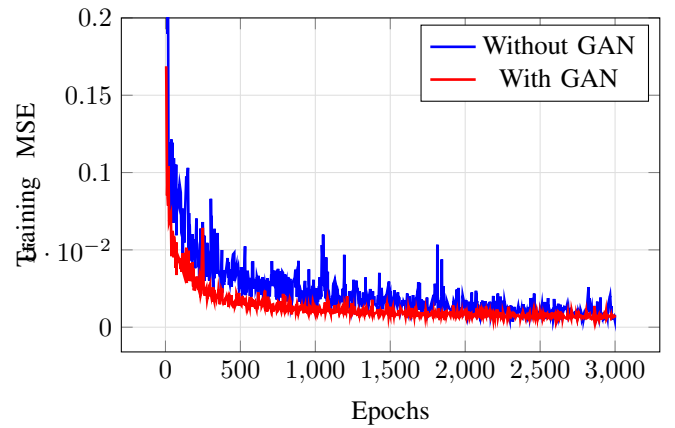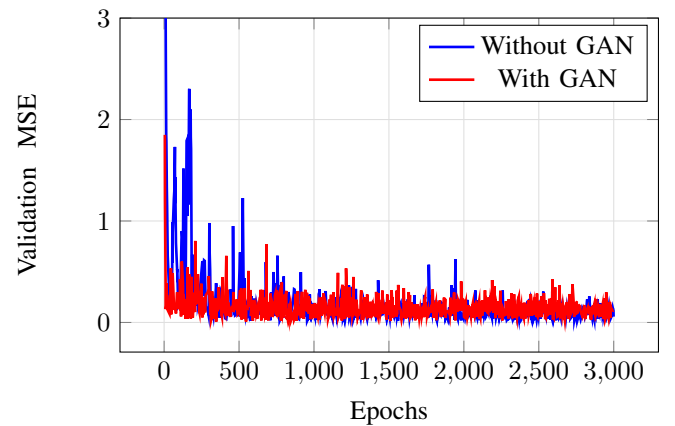

Figure 3: Training MSE of the ANN model

Figure 4: Validation MSE of the ANN model

samples depends on the training samples that are available and can vary depending on the dataset used.

We have performed experiments to showcase the improvement gained by augmenting training samples using the GAN phase. In Figure 7, the results of these experiments are shown. An additional note to take away from this experiment is that the dataset with the limited number of samples has the highest improvement, which is demonstrated in a clearer fashion in Table II. This difference is due to the fact that the number of training samples in this dataset is insufficient to train the regressor network causing reduced performance.

In Figure 7c, prediction of a new configuration with and without GAN phase is compared. In the case with $n_{analyte} = 1.33$, the network without the GAN phase has failed to predict the curve correctly. Even though the GAN boosted network has predicted the curve lower than it is, the peak is in the correct wavelength, which could help PCF designers to pinpoint highest loss.

We have also applied the GAN phase to the network architecture proposed in [13]. In this experiment, the GAN phase reduced the performance slightly. However, as we have discussed earlier, full batch training strategy is the cause of this due to increased number of samples. In fact, when we reduce the batch size down to 900 samples, the GAN phase
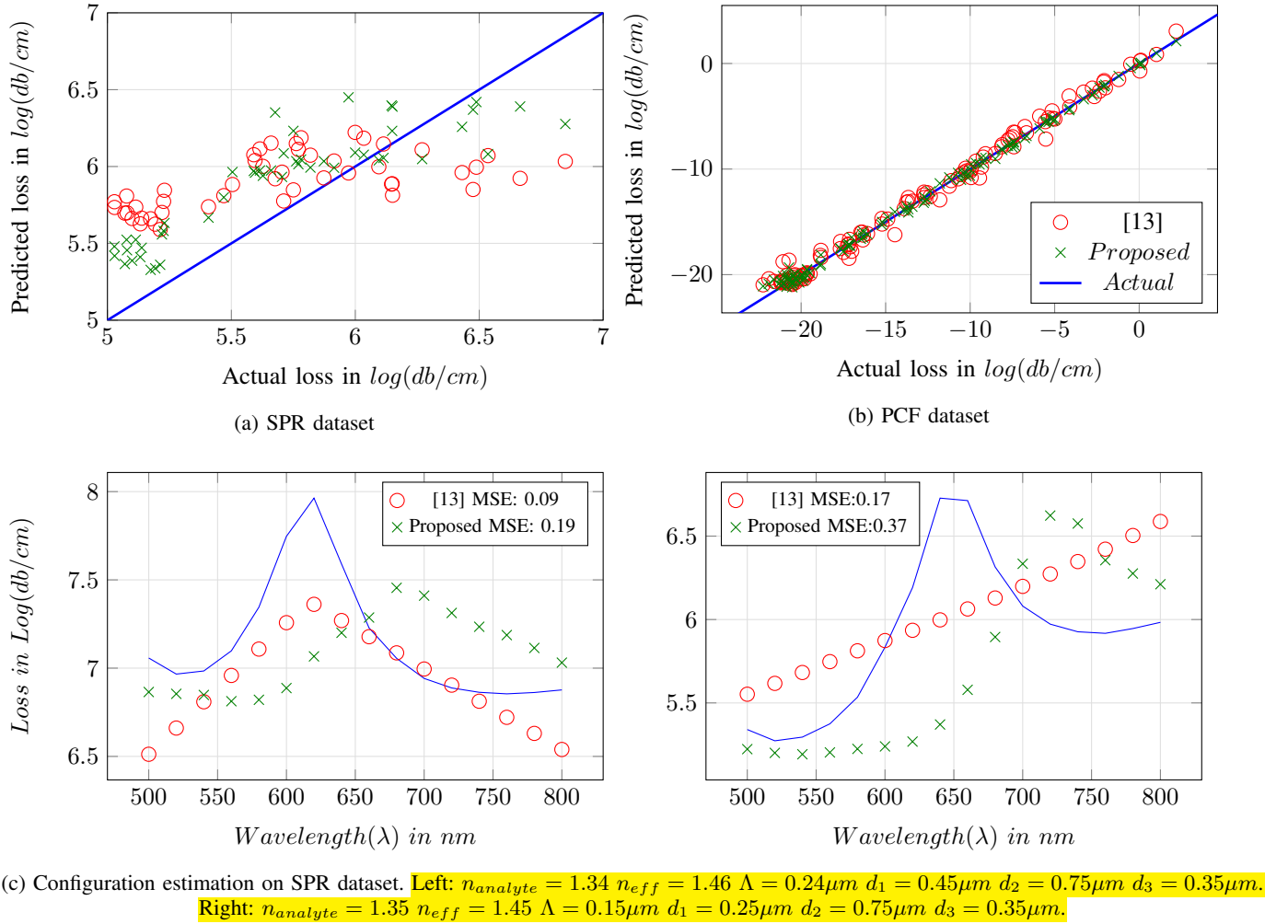
(a) SPR dataset

(b) PCF dataset

(c) Configuration estimation on SPR dataset. Left: $n_{analyte} = 1.34$ $n_{eff} = 1.46$ $\Lambda = 0.24\mu m$ $d_1 = 0.45\mu m$ $d_2 = 0.75\mu m$ $d_3 = 0.35\mu m$. Right: $n_{analyte} = 1.35$ $n_{eff} = 1.45$ $\Lambda = 0.15\mu m$ $d_1 = 0.25\mu m$ $d_2 = 0.75\mu m$ $d_3 = 0.35\mu m$.

Figure 5: Testing results without GAN augmentation phase
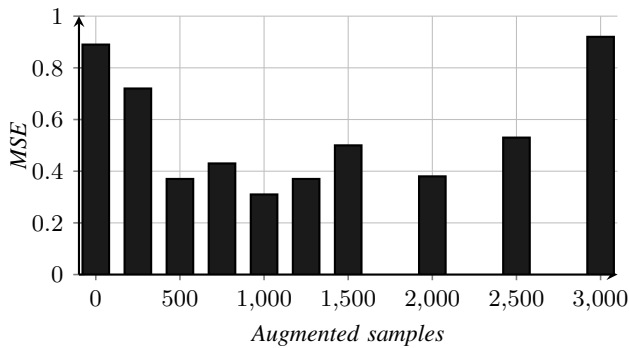
Figure 6: Effect of increasing augmented samples on MSE.

Table II: Detailed performance analysis in MSE

Dataset	SPR			PCF		
Method	Proposed		[13]	Proposed		[13]
GAN Phase	Yes	No	No	Yes	No	No
Average	**0.314**	0.894	0.987	**0.134**	0.156	0.167
Best	**0.006**	0.042	0.047	0.095	**0.093**	0.106
Worst	**1.102**	6.792	5.955	**0.181**	0.265	0.416
Std. dev.	**0.437**	2.213	1.940	**0.031**	0.046	0.090

the network (PCF), using the GAN phase reduces the worst case error by 32%. This added stability can also be observed in Figure 3, where the training loss has less fluctuations when the data augmentation is employed.

D. Computational performance

In this section, we have determined both testing and training time of the proposed system in comparison to the simulation approach. In Table III, the results of these experiments are published. Training time is total training time while testing is per sample. These tests are performed over the SPR dataset. Please note that the training time is considered offline time; i.e., it is the time that will only be used once, before the system is deployed. Once trained, the ANN can work alone without

starts to improve the score of this architecture.

The final set of experiments shows the stability that the GAN phase adds. In Table II, we have listed minimum, maximum, average MSE along with the standard deviation between folds in the testing. It is clear that the GAN phase not only improves the average case but significantly improves the worst case while reducing the deviation, which is important in a live setting. Even in a dataset with enough examples to train

(a) SPR dataset

(b) PCF dataset

(c) Configuration estimation on SPR dataset. Left: $n_{analyte} = 1.34$ $n_{eff} = 1.46$ $\Lambda = 0.24\mu m$ $d_1 = 0.45\mu m$ $d_2 = 0.75\mu m$ $d_3 = 0.35\mu m$. Right: $n_{analyte} = 1.33$ $n_{eff} = 1.45$ $\Lambda = 0.15\mu m$ $d_1 = 0.25\mu m$ $d_2 = 0.55\mu m$ $d_3 = 0.15\mu m$.
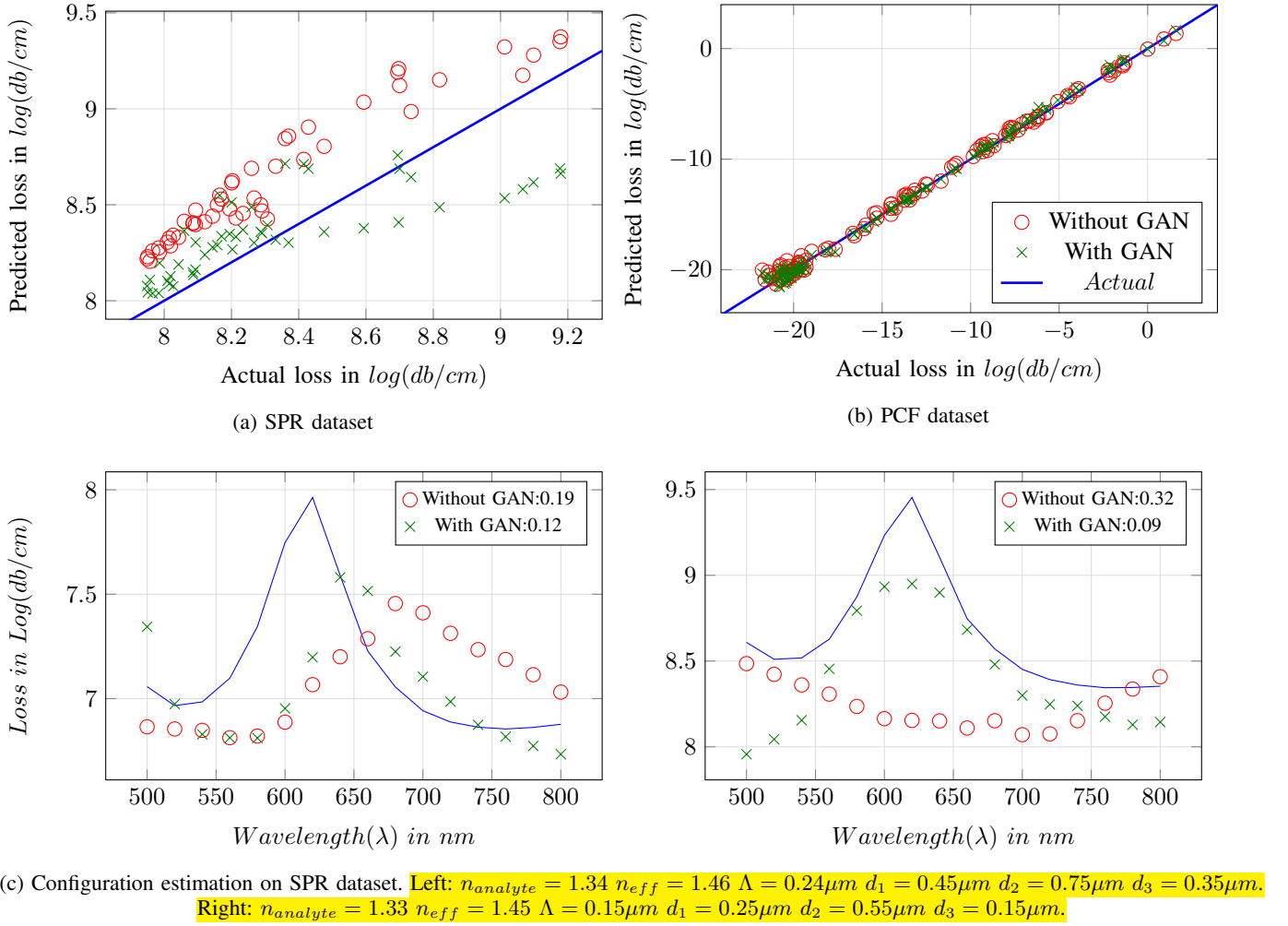
Figure 7: GAN Phase experimental results

the need for GAN or training again. With this information, it is evident that the ANN-based methods are far faster than the simulation approach by several orders of magnitude. Even when training time is included, total time spent to calculate the confinement loss will be faster on a GAN boosted ANN compared to an FV-FEM simulator after only 5 samples.

The training time difference between ANN systems is caused by the increased depth in the proposed architecture as well as the mini-batch training strategy. Training time is dominated by the GAN phase; however, this phase does not add any overhead to the testing time. The testing time difference between two networks is caused by the depth of the neural network (5 hidden layers instead of 3).

We have also experimented with TensorFlow using GPU via an additional system that contains NVidia GTX 1660Ti GPU. However, due to the relatively small size of the network, using GPU slightly reduced the computational performance instead of increasing it.

V. CONCLUSION

In this paper, we have demonstrated that the use of generative adversarial networks can be beneficial to neural networks that are designed to compute the propagation characteristics

Table III: Execution time of algorithms

Method	Training (s)	Testing (ms)
Proposed ANN	42	3.6
Proposed ANN with GAN	449	3.6
GAN phase only	370	-
[13]	20	2.7
FV-FEM	-	106958

of PCF devices. Additionally, we have also proposed and demonstrated a neural network that is both more generic and more accurate to estimate the confinement loss compared to the state-of-the-art network. The GAN phase also reduces the required samples to fully train the ANN, reducing the work necessary to employ the proposed system. The proposed combination is shown to be general enough that even though it is designed for the SPR dataset built for this research article. The proposed combination is shown to be general enough that even though it is designed and optimized for the SPR dataset; in the experiments, it surpassed the accuracy of the method proposed in [13] on the dataset that the method is optimized on.

Another important aspect of this work is the experiments performed on a dataset that is designed to test the ability

of artificial neural networks to estimate the properties of a new configuration. For this purpose, we have built a database based on Surface Plasmon Resonance with varying geometric configurations. During the testing, a new configuration that is not used for the training is employed. The proposed ANN, augmented with the GAN phase is shown to estimate the new configuration to an acceptable error rate.

The most vital advantage of using machine learning methods instead of numerical simulation is the speed factor. In the experiments, it is shown that the machine learning methods work few orders of magnitude faster than the simulation method.

The code for this work is available at: https://github.com/Aimen-Zelaci/SPRPCF_ANN/.

REFERENCES

[1] Y. Wang, M. Yang, D. Wang, and C. Liao, "Selectively infiltrated photonic crystal fiber with ultrahigh temperature sensitivity," *IEEE Photonics Technology Letters*, vol. 23, no. 20, pp. 1520–1522, 2011.

[2] A. A. Rifat, G. A. Mahdiraji, Y. M. Sua, R. Ahmed, Y. Shee, and F. M. Adikan, "Highly sensitive multi-core flat fiber surface plasmon resonance refractive index sensor," *Optics express*, vol. 24, no. 3, pp. 2485–2495, 2016.

[3] J. N. Dash and R. Jha, "Spr biosensor based on polymer pcf coated with conducting metal oxide," *IEEE Photonics Technology Letters*, vol. 26, no. 6, pp. 595–598, 2014.

[4] H. Ademgil, "Highly sensitive octagonal photonic crystal fiber based sensor," *Optik*, vol. 125, no. 20, pp. 6274–6278, 2014.

[5] E. Kretschmann and H. Raether, "Radiative decay of non radiative surface plasmons excited by light," *Zeitschrift für Naturforschung A*, vol. 23, no. 12, pp. 2135–2136, 1968.

[6] A. Yasli and H. Ademgil, "Effect of plasmonic materials on photonic crystal fiber based surface plasmon resonance sensors," *Modern Physics Letters B*, vol. 33, no. 13, p. 1950157, 2019.

[7] A. Hassani and M. Skorobogatiy, "Design of the microstructured optical fiber-based surface plasmon resonance sensors with enhanced microfluidics," *Optics Express*, vol. 14, no. 24, pp. 11 616–11 621, 2006.

[8] A. Yasli, H. Ademgil, S. Haxha, and A. Aggoun, "Multi-channel photonic crystal fiber based surface plasmon resonance sensor for multi-analyte sensing," *IEEE Photonics Journal*, 2019.

[9] S. I. Azzam, M. F. O. Hameed, R. E. A. Shehata, A. Heikal, and S. S. Obayya, "Multichannel photonic crystal fiber surface plasmon resonance based sensor," *Optical and Quantum Electronics*, vol. 48, no. 2, p. 142, 2016.

[10] R. Otupiri, E. K. Akowuah, and S. Haxha, "Multi-channel spr biosensor based on pcf for multi-analyte sensing applications," *Optics express*, vol. 23, no. 12, pp. 15 716–15 727, 2015.

[11] G. Fornarelli, L. Mescia, F. Prudenzano, M. De Sario, and F. Vacca, "A neural network model of erbium-doped photonic crystal fibre amplifiers," *Optics & Laser Technology*, vol. 41, no. 5, pp. 580–585, 2009.

[12] I. Abdelaziz, F. AbdelMalek, S. Haxha, H. Ademgil, and H. Bouchriha, "Photonic crystal fiber with an ultrahigh birefringence and flattened dispersion by using genetic algorithms," *Journal of Lightwave Technology*, vol. 31, no. 2, pp. 343–348, 2012.

[13] S. Chugh, A. Gulistan, S. Ghosh, and B. Rahman, "Machine learning approach for computing optical properties of a photonic crystal fiber," *Optics Express*, vol. 27, no. 25, pp. 36 414–36 425, 2019.

[14] P. I. Borel, A. Harpøth, L. H. Frandsen, M. Kristensen, P. Shi, J. S. Jensen, and O. Sigmund, "Topology optimization and fabrication of photonic crystal structures," *Optics express*, vol. 12, no. 9, pp. 1996–2001, 2004.

[15] M. Hameed, S. Obayya, K. Al-Begain, A. Nasr, and M. A. El Maaty, "Accurate radial basis function based neural network approach for analysis of photonic crystal fibers," *Optical and quantum electronics*, vol. 40, no. 11-12, p. 891, 2008.

[16] P.-H. Fu, S.-C. Lo, P.-C. Tsai, K.-L. Lee, and P.-K. Wei, "Optimization for gold nanostructure-based surface plasmon biosensors using a microgenetic algorithm," *ACS Photonics*, vol. 5, no. 6, pp. 2320–2327, 2018.

[17] P. D. McAtee, S. T. Bukkapatnam, and A. Lakhtakia, "Artificial neural network to estimate the refractive index of a liquid infiltrating a chiral sculptured thin film," *Journal of Nanophotonics*, vol. 13, no. 4, p. 046006, 2019.

[18] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks*, vol. 4, no. 2, pp. 251 – 257, 1991. [Online]. Available: http://www.sciencedirect.com/science/article/pii/089360809190009T

[19] Y. Kiarashinejad, M. Zandehshahvar, S. Abdollahramezani, O. Hemmatyar, R. Pourabolghasem, and A. Adibi, "Knowledge discovery in nanophotonics using geometric deep learning," *Advanced Intelligent Systems*, vol. 2, no. 2, p. 1900132, 2020.

[20] T. Asano and S. Noda, "Optimization of photonic crystal nanocavities based on deep learning," *Optics express*, vol. 26, no. 25, pp. 32 704–32 717, 2018.

[21] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[22] T. Schlegl, P. Seeböck, S. M. Waldstein, U. Schmidt-Erfurth, and G. Langs, "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery," in *International conference on information processing in medical imaging*. Springer, 2017, pp. 146–157.

[23] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "Synthetic data augmentation using gan for improved liver lesion classification," in *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*. IEEE, 2018, pp. 289–293.

[24] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *arXiv preprint arXiv:1712.04621*, 2017.

[25] V. Rodríguez-Esquerre, J. Isídio-Lima, and A. Dourado-Sisnando, "Efficient neural network modeling of photonic crystal fiber chromatic dispersion," in *Latin America Optics and Photonics Conference*. Optical Society of America, 2010, p. PDPTuK1.

[26] L. Mescia, G. Fornarelli, D. Magarielli, F. Prudenzano, M. De Sario, and F. Vacca, "Refinement and design of rare earth doped photonic crystal fibre amplifier using an ann approach," *Optics & Laser Technology*, vol. 43, no. 7, pp. 1096–1103, 2011.

[27] H.-C. Shin, N. A. Tenenholtz, J. K. Rogers, C. G. Schwarz, M. L. Senjem, J. L. Gunter, K. P. Andriole, and M. Michalski, "Medical image synthesis for data augmentation and anonymization using generative adversarial networks," in *Simulation and Synthesis in Medical Imaging*, A. Gooya, O. Goksel, I. Oguz, and N. Burgos, Eds. Cham: Springer International Publishing, 2018, pp. 1–11.

[28] M. Lucic, K. Kurach, M. Michalski, S. Gelly, and O. Bousquet, "Are gans created equal? a large-scale study," in *Advances in neural information processing systems*, 2018, pp. 700–709.

[29] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein gan," *arXiv preprint arXiv:1701.07875*, 2017.

[30] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of wasserstein gans," in *Advances in neural information processing systems*, 2017, pp. 5767–5777.

[31] S. Ravuri and O. Vinyals, "Seeing is not necessarily believing: Limitations of biggans for data augmentation," 2019.

[32] K. Shmelkov, C. Schmid, and K. Alahari, "How good is my gan?" in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 213–229.

[33] B. Bhattarai, S. Baek, R. Bodur, and T.-K. Kim, "Sampling strategies for gan synthetic data," *arXiv preprint arXiv:1909.04689*, 2019.

[34] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[35] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[36] D. Masters and C. Luschi, "Revisiting small batch training for deep neural networks," *arXiv preprint arXiv:1804.07612*, 2018.

[37] N. S. Keskar, D. Mudigere, J. Nocedal, M. Smelyanskiy, and P. T. P. Tang, "On large-batch training for deep learning: Generalization gap and sharp minima," *arXiv preprint arXiv:1609.04836*, 2016.

[38] M. Koshiba, "Full-vector analysis of photonic crystal fibers using the finite element method," *IEICE transactions on electronics*, vol. 85, no. 4, pp. 881–888, 2002.

[39] P. B. Johnson and R.-W. Christy, "Optical constants of the noble metals," *Physical review B*, vol. 6, no. 12, p. 4370, 1972.

[40] A. S. B. Anders Bjarklev, Jes Broeng, *Photonic Crystal Fibres*. Kluwer academic publisher, 2003.