

Improving Object Classification Accuracy from Electromagnetic Data Using Attention Mechanisms

Ergun Simsek* and Harish R. Manyam†

* Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD, USA

simsek@umbc.edu † Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD, USA
hmanyam1@umbc.edu

Abstract—Object classification using electromagnetic waves is crucial in various applications, including remote sensing, security screening, and biomedical imaging. However, accurately classifying arbitrarily oriented objects from electromagnetic scattering data remains a significant challenge. In this work, we propose an attention-based machine-learning framework designed to improve the robustness and accuracy of electromagnetic object classification. Our model leverages an attention mechanism to focus on the most informative scattering features dynamically, enabling enhanced feature extraction and improved generalization across different object orientations. We demonstrate the effectiveness of attention-based models in enhancing object classification robustness using a numerical dataset, showing that the proposed method outperforms conventional machine learning models regarding classification accuracy.

Index Terms—Machine learning, electromagnetic waves, classification, neural networks.

I. INTRODUCTION

Electromagnetic inversion and electromagnetic classification are two distinct yet closely related problems in the analysis of wave-matter interactions. Electromagnetic inversion refers to the process of reconstructing the material properties or structural characteristics of an object or medium from measured electromagnetic field data [1]–[6]. This typically involves solving an ill-posed inverse problem to determine spatially resolved parameters such as permittivity, conductivity, or permeability distributions within a given domain. Due to the complexity and nonlinearity of electromagnetic wave propagation, traditional inversion methods often require sophisticated regularization techniques and extensive computational resources to obtain accurate reconstructions.

In contrast, electromagnetic classification aims to identify and categorize objects based on their electromagnetic response without necessarily reconstructing their detailed physical properties [7]–[10]. Rather than solving for continuous material distributions, classification seeks to assign discrete labels to objects by analyzing their scattered field characteristics. This task is particularly relevant in applications where the primary interest is recognizing object types rather than reconstructing their fine-grained internal structures. Given the high dimensionality and intricate nature of electromagnetic scattering data, machine learning (ML) techniques have emerged as powerful tools both for inversion [6], [11]–[14] and classification [7]–[10], leveraging data-driven approaches to extract

meaningful patterns from measured signals. Previous studies have demonstrated that ML models such as support vector machines (SVM), gradient boosting (XGB), and neural networks (NN) can achieve high classification accuracy—up to 90% [7]—for objects with a fixed alignment. However, their performance often deteriorates when objects exhibit arbitrary orientations (e.g., in [7], the accuracy decreases to 57% for arbitrarily aligned objects), limiting their effectiveness in practical deployment. This challenge underscores the need for more advanced classification frameworks capable of handling variations in object positioning and orientation, thereby enhancing the reliability of electromagnetic-based object recognition across diverse real-world scenarios.

To address this limitation, we propose an enhanced classification pipeline incorporating an attention mechanism. By integrating attention-based models, our approach selectively emphasizes the most relevant features of the input data, mitigating the adverse effects of arbitrary orientations on classification accuracy. To demonstrate the effectiveness of attention-based models in enhancing object classification robustness, we first create a dataset based on the MNIST dataset, where we transform the grayscale pixel values into relative electrical permittivity values to form scatterers and calculate the electromagnetic waves scattered from these objects using a two-dimensional (2D) electromagnetic solver developed based on the hybrid spectral integral - finite element (SI-FE) method [15]. Then, we train various machine learning models with this dataset to classify the objects. When we compare these models' classification accuracy and efficiency, we observe that the neural networks with an attention mechanism outperform others, achieving an 87% classification accuracy solely from the data without projecting the input data into a latent space.

The remainder of this paper is organized as follows. Section II provides a brief description of the hybrid SI-FE method. Section III describes the creation of our dataset. Section IV outlines the implementation of our attention-based neural network model, explaining the architecture and key components of the attention mechanism. Section V presents the numerical results, comparing the performance of our proposed model with various conventional machine learning techniques. Finally, Section VI concludes the paper with a summary of our findings and discusses potential future research directions.

II. SPECTRAL INTEGRAL-FINITE ELEMENT METHOD

Consider a general 2D inhomogeneous object or a collection of objects in an unbounded medium, as illustrated in Figure 1. To calculate the electromagnetic waves scattered from these objects with the finite-element method (FEM), a radiation boundary condition (RBC) is necessary to truncate the computational domain, allowing FEM to be applied within the interior region (Region I). By imposing an appropriate RBC on the boundary $\partial\Gamma$, the electromagnetic field in the homogeneous exterior region (Region II) can be determined once the solution for Region I, including its boundary, is obtained. In Region I, the material properties vary spatially, characterized by a relative permeability $\mu_r(x, y)$ and a relative permittivity $\varepsilon_r(x, y)$. In contrast, Region II consists of a homogeneous medium with constant relative magnetic permeability $\mu_{r,b}$ and relative permittivity $\varepsilon_{r,b}$, encompassing Region I.

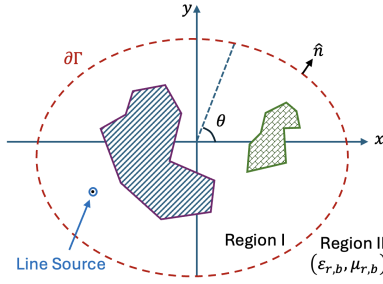


Fig. 1. Schematic of a 2D electromagnetic scattering problem. The artificial boundary ($\partial\Gamma$), shown with a red dashed curve, divides the computational domain into Regions I and II. The two shapes with patterned fills represent the scatterers. The blue circle on the lower left is a line current source creating transverse magnetic waves.

For transverse magnetic (TM) waves, the total electric field $E(x, y)$ in Region I (bounded by $\partial\Gamma$) is determined by solving the following scalar wave equation:

$$\nabla \cdot (\mu_r^{-1}(x, y) \nabla E) + k_0^2 \varepsilon_r(x, y) E = S_e(x, y), \quad (1)$$

where k_0 is the wavenumber in free space, and $S_e(x, y)$ represents the source inside the boundary. To discretize Eq. (1), we multiply the equation by a testing function $W_m(x, y)$, where m is the index of the testing function, and integrate over Region I. Applying vector identities and Gauss's theorem, we derive the weak-form equation:

$$\begin{aligned} \int_{\Gamma} \int_{\Gamma} \left[k_0^2 \varepsilon_r W_m E - \frac{1}{\mu_r} \nabla W_m \cdot \nabla E \right] dx dy \\ + \oint_{\partial\Gamma} W_m \frac{1}{\mu_r} \frac{\partial E}{\partial n} ds = \int_{\Gamma} \int_{\Gamma} W_m S_e dx dy, \end{aligned} \quad (2)$$

where $\hat{n}$ is the outward unit normal vector on the boundary $\partial\Gamma$. Note that in Eq. (2), ε_r , μ_r , W_m , and S_e are all functions of x and y .

To solve Eq. (2) using the FEM scheme, Region I is discretized into triangular elements, as later illustrated in Fig. 2. A linear pyramid basis function $P_n(x, y)$ is used to expand the electric field $E(x, y)$ in the interior region, while a triangular basis function $T_n(x, y)$ is employed to expand the boundary value E_n on $\partial\Gamma$. The nodal points of the boundary

basis function are collocated with those of the pyramid basis function on the boundary. We use pyramid functions for testing as well and obtain the final set of linear equations [15]. Since there are more unknowns than the number of equations, we need additional conditions to obtain a unique solution of the system. In this work, the spectral integral method (SIM) serves as an RBC to provide these additional conditions by relating the fields inside Region I to those on the outer boundary $\partial\Gamma$. To achieve this, we utilize Green's theorem to derive a surface integral equation that expresses the field outside $\partial\Gamma$ in terms of the field values on the boundary.

We define the boundary integral equation using Green's function $G(k_b, R)$ and its normal derivative $\partial G(k_b, R)/\partial n'$ [16]. We express the fields along $\partial\Gamma$ in terms of a smooth boundary parameterized by θ . By leveraging Fourier series expansions, we approximate the field $E(\theta')$ and its normal derivative $\partial E(\theta')/\partial n'$ using a finite number of Fourier coefficients. This transforms the integral equation into a system of algebraic equations in Fourier space. However, since Green's function is singular at $\theta = \theta'$, we apply a singularity subtraction technique to improve numerical convergence [16].

Once we compute the Fourier coefficients of the unknown field and its normal derivative, we obtain the final equation in matrix form, allowing us to enforce the RBC at discrete boundary points. By utilizing fast Fourier transform (FFT) and spectral interpolation, we efficiently compute these coefficients, which brings fast convergence and high spectral accuracy. Note that the SIM RBC effectively suppresses spurious solutions (such as fictitious resonant frequencies) commonly encountered in integral equation solvers. This makes our solver a robust and efficient approach for coupling FEM with an accurate radiation boundary condition.

III. DATASET CREATION

The Modified National Institute of Standards and Technology (MNIST) dataset, a widely recognized resource in machine learning and computer vision, provides a collection of handwritten digits, as detailed in [17]. This dataset is frequently employed for training and evaluating algorithms, particularly those designed for image classification and character recognition. Each image within MNIST consists of a 28×28 -pixel square grayscale representation of a handwritten digit, ranging from 0 to 9, with each image accompanied by a corresponding label indicating the digit's identity.

In this study, we utilize 6,000 digital images from the MNIST dataset to construct a scatterer database. Specifically, the grayscale intensity of each pixel, denoted as $d_{u,v}$, where u and v represent the row and column indices ranging from 1 to 28, is converted into relative electrical permittivity values. These pixel values, scaled between 0 and 255, where 0 signifies white and 255 signifies black, are transformed into permittivity values ranging from $\epsilon_r^{\min}$ to $\epsilon_r^{\max}$ using the following linear relationship:

$$\epsilon_r(u, v) = \epsilon_r^{\min} + (\epsilon_r^{\max} - \epsilon_r^{\min}) \frac{d_{u,v}}{255} \quad (3)$$

where u and v are the row and column numbers. We set $\epsilon_r^{\min} = 1$ and $\epsilon_r^{\max} = 4$.

The computational domain, illustrated in Fig. 2, features 12 transmitter and 12 receiver antennae, shown with green and red circles, respectively. These antennae are uniformly positioned at radial distances of $\rho = 1.5\lambda$ and $\rho = 2\lambda$, respectively. The permittivity values within the triangular elements of the scatterer are determined through a 2D interpolation of the MNIST images, allowing for an approximate representation of the digit's shape within the simulation environment.

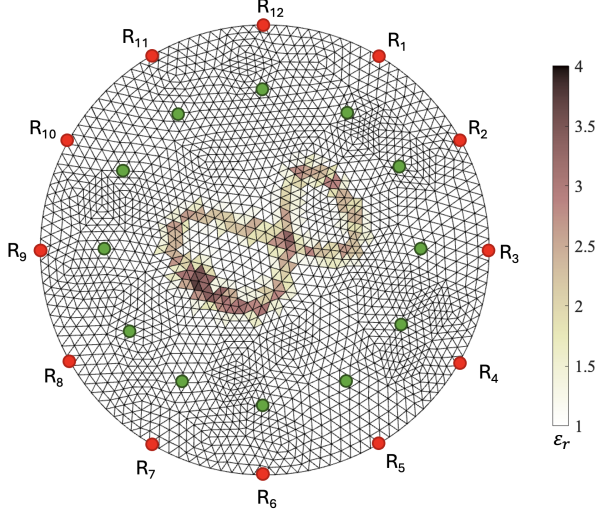


Fig. 2. Illustration of the computational domain. Twelve receiver antennae, shown with red circles, are placed uniformly on the circular boundary with a radius of 2λ . Twelve transmitter antennae are placed uniformly at $\rho = 1.5\lambda$, shown with green circles. The permittivity values of the triangles where the scatterer resides are determined via a 2D interpolation from the MNIST images using Eq. (3).

To compute the scattered electric fields at the receiver antennae, our Spectral Integral/Finite Element (SI-FE) solver is employed. We conduct twelve simulations, each activating one transmitter antenna at a time. This process yields 144 electric field values, representing the fields measured by the 12 receivers for each of the 12 transmitter configurations. Hence, the sizes of the input and output of our ML models will be 144 and 9, as explained in the next section.

IV. NEURAL NETWORK IMPLEMENTATION

Attention mechanisms have proven to be efficient tools for constructing accurate deep models when the features are given in sequential forms [18]. This is why we decided to use a series of neural networks enhanced with an attention mechanism, as illustrated in Fig. 3, to achieve object classification based on scattered electromagnetic wave data as follows.

The process begins with the collection of data from twelve distinct sources, R_1 through R_{12} , each providing complementary information about the target and its alignment. The input data from these sources is fed into a neural network, denoted as $F_b(\cdot)$, which transforms the raw inputs into a set of hidden feature representations $\{h_1, h_2, \dots, h_n\}$. These feature activations capture complex relationships and patterns within the data, reflecting various aspects of the digit shapes, strokes, and their alignment. However, not all features contribute equally to the classification decision, prompting the

need for an adaptive weighting mechanism that can highlight the most informative representations. A scoring function $F_s(\cdot)$ evaluates the relevance of each hidden feature h_i , assigning a scalar score that indicates its importance. These scores are subsequently normalized through a softmax function, converting them into probability weights. The attention mechanism thus enables the model to focus on the most salient features while diminishing the influence of less relevant ones. The weighted features are then aggregated to form a context vector that encapsulates the most discriminative information necessary for accurate digit classification. The resulting context vector is passed through a final classification layer, which outputs a probability distribution over the possible digit classes $\{0, 1, 2, 3, 4, 5, 6 \text{ or } 9, 7, 8\}$. The digit corresponding to the highest probability is selected as the predicted class. Note that due to arbitrary alignment, we cannot differentiate 6s and 9s, so we group them together. By incorporating the attention mechanism, the network dynamically adjusts its focus to the most informative features for each input instance.

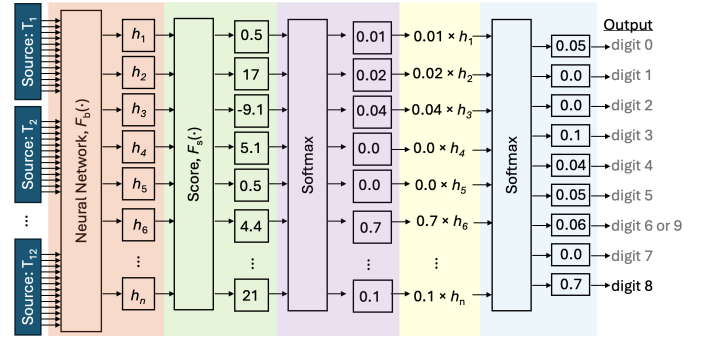


Fig. 3. An illustration of the neural network implemented. In the beginning, each member of the input is equally important. A neural network is applied to each item independently, creating a new sequence representation h_i . A score function is applied to every input, giving each item a value indicating its importance. The softmax function is applied to all the scores, creating a probability distribution. The weighted features are then aggregated to form a context vector that encapsulates the most discriminative information necessary for accurate digit classification.

V. NUMERICAL RESULTS

In our study, the backbone network consists of 4 hidden layers with 256, 256, 512, and 512 neurons. The input of each layer passes a batch normalization before reaching the fully-connected linear layer. The output of each hidden layer passes a standard ReLU (rectified linear unit) activation function. The model's learning process is guided by the Adam optimizer [19], with a learning rate set at 0.001. The categorical cross-entropy loss [20] is chosen as the optimization objective, and categorical accuracy is monitored as a metric to gauge performance. All the code is executed on Google Colaboratory using T4 GPU accelerators.

We allocate 50% of the dataset for training and the other 50% for testing. Figures 4 (a) and (b) show the training and validation loss and accuracy over epochs. The slight divergence between training and validation curves suggests minimal overfitting, indicating the neural network's architecture is slightly larger than ideal. The convergence of these

metrics confirms that the training used an adequate number of epochs. Furthermore, the training stopped at epoch 106, despite being set to 200, due to the “early stopping” feature implemented in the model. In Fig. 4 (b), we also observe that the NN architecture with the attention mechanism achieves an 87% classification accuracy.

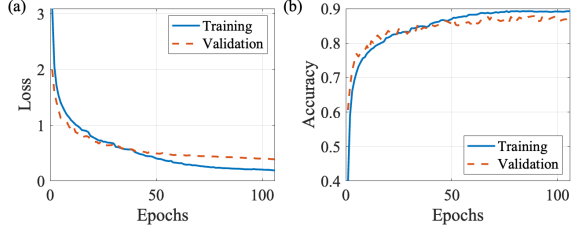


Fig. 4. (a) Loss and (b) accuracy vs. epoch number of the NN implementation.

In addition to neural network (NN) implementation, we also employ other machine-learning models, namely k-nearest neighbor (kNN), random forest (RF), Gaussian Naive Bayes (GNB), support vector machine (SVM), and gradient boosting (XGB), to assess whether they exhibit similar levels of learning from the data or not. For the first four models (kNN, RF, GNB, and SVM), we utilize scikit-learn [21], which is a free software machine learning library for the Python programming language. For the XGB implementation, we use another freely available library [22], which is developed based on XGBoost [23]. Hyperparameter tuning is done using the Grid Search Cross-Validation (GridSearchCV) module [24] that is available in the scikit-learn [21] library.

Table I presents the training time, accuracy, precision, recall, and F1-score for six different classification methods applied to a dataset split evenly between training and testing. Among these methods, the neural network implementation achieved the highest accuracy of 87%, outperforming other classifiers in all evaluated metrics. This superior performance can be attributed to the NN’s ability to capture complex, non-linear patterns and interactions within the data through multiple layers of representation learning. It is worth noting that the classification accuracy obtained with the NN is higher than the previously reported value [8] of 85.3%.

TABLE I
THE TIME SPENT DURING TRAINING, ACCURACY, PRECISION, RECALL, AND F1-SCORE OF THE KNN, RF, GNB, SVM, XGB, AND NN WITH ATTENTION MECHANISM IMPLEMENTATIONS.

Method	Time (s)	Accuracy	Precision	Recall	F1 score
kNN	0.1	0.26	0.27	0.26	0.26
RF	124	0.47	0.46	0.46	0.46
GNB	0.2	0.41	0.41	0.40	0.40
SVM	84	0.82	0.82	0.82	0.82
XGB	48	0.81	0.81	0.81	0.81
NN	112	0.87	0.87	0.87	0.87

In contrast, the kNN, RF, and GNB methods demonstrated comparatively poor performance, with accuracies of 0.26, 0.47, and 0.41, respectively. The low accuracy of kNN can be explained by its reliance on local similarity measures, which

are susceptible to high-dimensional data and noisy features. Without sufficient feature scaling or dimensionality reduction, kNN struggles to differentiate between classes effectively. Similarly, the GNB classifier assumes feature independence and a Gaussian distribution of the data, assumptions that are often violated in real-world datasets. These violations lead to suboptimal decision boundaries and, consequently, lower classification accuracy. The random forest method, while generally robust, may have underperformed due to its low effectiveness when feature interactions are complex and non-linear patterns are prevalent, scenarios where neural networks excel. The SVM and XGB methods achieved relatively high accuracies of 0.82 and 0.81, respectively, indicating their effectiveness in handling non-linear decision boundaries and feature interactions. However, the NN outperformed both, likely due to its deeper architecture and ability to learn hierarchical representations, enabling better generalization across varied samples. Regarding training time, kNN and GNB exhibited the fastest training durations (0.1 and 0.2 seconds, respectively) due to their simplistic modeling approaches. In contrast, methods like RF, SVM, XGB, and NN required significantly longer training times, reflecting the computational complexity involved in optimizing their models. The neural network’s training time of 112 seconds, while substantial, is justified by the notable improvement in classification performance. Overall, the results highlight a trade-off between training time and predictive accuracy. While simpler models offer quick training, they often fall short in capturing complex data structures. Conversely, more sophisticated methods like neural networks demand longer training times but yield superior classification outcomes, mainly when the data exhibits intricate patterns and non-linearities.

VI. CONCLUSION

In this work, we introduced an attention-based neural network model to enhance the classification of objects using electromagnetic wave scattering data. Our approach addresses the limitations of traditional machine learning models, which struggle with arbitrary object orientations. By incorporating an attention mechanism, our model dynamically identifies and emphasizes the most relevant features, leading to improved classification accuracy. Our numerical results demonstrate that the proposed model outperforms conventional methods, achieving a significant increase in robustness and accuracy. Our work highlights the potential of attention-based architectures for electromagnetic object classification, paving the way for more reliable and adaptable recognition systems in challenging real-world scenarios. Future work may explore the extension of this approach to three-dimensional object classification and the integration of additional domain-specific priors to further enhance model performance.

REFERENCES

- [1] D. L. Colton, R. Kress, and R. Kress, *Inverse acoustic and electromagnetic scattering theory*. Springer, 1998, vol. 93.
- [2] O. M. Bucci and T. Isernia, “Electromagnetic inverse scattering: Retrieval information and measurement strategies,” *Radio Science*, vol. 32, no. 6, pp. 2123–2137, 1997.

- [3] Y. M. Wang and W. C. Chew, "An iterative solution of the two-dimensional electromagnetic inverse scattering problem," *Int. J. Imaging Syst. Technol.*, vol. 1, no. 1, pp. 100–108, 1989.
- [4] A. Abubakar, P. Van den Berg, and S. Y. Semenov, "Two- and three-dimensional algorithms for microwave imaging and inverse scattering," *J. Electromagn. Waves Appl.*, vol. 17, no. 2, pp. 209–231, 2003.
- [5] B. Wei, E. Simsek, C. Yu, and Q. H. Liu, "Three-dimensional electromagnetic nonlinear inversion in layered media by a hybrid diagonal tensor approximation: Stabilized biconjugate gradient fast fourier transform method," *Waves in Random and Complex Media*, vol. 17, no. 2, pp. 129–147, 2007.
- [6] E. Simsek, "Determining optical constants of 2D materials with neural networks from multi-angle reflectometry data," *Machine Learning: Science and Technology*, vol. 1, no. 1, p. 01LT01, 2020.
- [7] E. Simsek and H. R. Manyam, "Classification with electromagnetic waves," *IET Microwaves, Antennas & Propagation*, vol. 18, no. 12, pp. 898–910, 2025/02/24 2024.
- [8] A. Momeni and R. Fleury, "Electromagnetic wave-based extreme deep learning with nonlinear time-Floquet entanglement," *Nature Communications*, vol. 13, no. 1, p. 2651, 2022.
- [9] D. Covarrubias-Martínez, H. Lobato-Morales, J. M. Ramírez-Cortés, and G. A. Álvarez-Botero, "Classification of plastic materials using machine-learning algorithms and microwave resonant sensor," *Journal of Electromagnetic Waves and Applications*, vol. 36, no. 12, pp. 1760–1775, 2022.
- [10] T. M. Ting, N. S. Ahmad, and P. Goh, "Material classification via embedded rf antenna array and machine learning for intelligent mobile robots," *Alexandria Engineering Journal*, vol. 106, pp. 60–70, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1110016824006926>
- [11] H. M. Yao, L. Jiang, and W. E. I. Sha, "Enhanced deep learning approach based on the deep convolutional encoder-decoder architecture for electromagnetic inverse scattering problems," *IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 7, pp. 1211–1215, 2020.
- [12] Y. Sanghvi, Y. Kalepu, and U. K. Khankhoje, "Embedding deep learning in inverse scattering problems," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 46–56, 2020.
- [13] K. Xu, L. Wu, X. Ye, and X. Chen, "Deep learning-based inversion methods for solving inverse scattering problems with phaseless data," *IEEE Trans. Antennas Propag.*, vol. 68, no. 11, pp. 7457–7470, 2020.
- [14] L.-Y. Xiao, J. Li, F. Han, W. Shao, and Q. H. Liu, "Dual-module NMM-IEM machine learning for fast electromagnetic inversion of inhomogeneous scatterers with high contrasts and large electrical dimensions," *IEEE Trans. Antennas Propag.*, vol. 68, no. 8, pp. 6245–6255, 2020.
- [15] E. Simsek, J. Liu, and Q. Liu, "A spectral integral method and hybrid SIM/FEM for layered media," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 11, pp. 3878–3884, 2006.
- [16] E. Simsek, J. Liu, and Q. H. Liu, "A spectral integral method (SIM) for layered media," *IEEE Trans. Antennas Propag.*, vol. 54, no. 6, pp. 1742–1749, 2006.
- [17] Y. LeCun and C. Cortes, "MNIST handwritten digit database," 2010.
- [18] M. Soroush, E. Simsek, G. Moille, K. Srinivasan, and C. R. Menyuk, "Predicting broadband resonator-waveguide coupling for microresonator frequency combs through fully connected and recurrent neural networks and attention mechanism," *ACS Photonics*, vol. 10, no. 6, pp. 1795–1805, 06 2023.
- [19] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014.
- [20] G. Tsoumakas and I. Katakis, "Multi-label classification: An overview," *International Journal of Data Warehousing and Mining (IJDWM)*, vol. 3, no. 3, pp. 1–13, 2007.
- [21] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [22] xgboost developers, "XGBoost, an optimized distributed gradient boosting library," <https://xgboost.readthedocs.io/>, 2022.
- [23] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: ACM, 2016, pp. 785–794.
- [24] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of machine learning research*, vol. 13, no. 2, 2012.