Differentiable Weightless Neural Networks for Differential Sensing

Andrew Jacob
andrewjacob@utexas.edu
The University of Texas at Austin
Austin, Texas, USA

Richard Huang
richardhuang@utexas.edu
The University of Texas at Austin
Austin, Texas, USA

Johnny Shen
shenj@utexas.edu
The University of Texas at Austin
Austin, Texas, USA

Shashank Nag
shashanknag@utexas.edu
The University of Texas at Austin
Austin, Texas, USA

Matthew Dixon
mhd634@austin.utexas.edu
The University of Texas at Austin
Austin, Texas, USA

Praveen Pasupathy
praveen@engr.utexas.edu
The University of Texas at Austin
Austin, Texas, USA

Diana Zamora-Olivares
diana_z.o@utexas.edu
The University of Texas at Austin
Austin, Texas, USA

Eric V Anslyn
anslyn@austin.utexas.edu
The University of Texas at Austin
Austin, Texas, USA

Lizy K. John
ljohn@ece.utexas.edu
The University of Texas at Austin
Austin, Texas, USA

ABSTRACT

In analytical chemistry, differential sensing is an increasingly popular technique for analyzing and classifying complex mixtures of analytes. Inspired by human taste and smell, differential sensing uses arrays of cross-reactive sensors to produce data "fingerprints" from sample interactions. While individual receptors lack high selectivity, the ensemble response can be analyzed using chemometric and machine learning methods to characterize samples. In this work, we propose Differentiable Weightless Neural Networks (DWNs) as a size and energy-efficient model for analyzing differential sensing datasets, with the future goal of integrating DWNs into co-designed sensor systems that support edge inference of chemical targets. DWNs are compact neural networks that use lookup tables instead of computationally heavy arithmetic operations, making them ideal for memory limited systems and sensor-driven workloads.

We present a proof-of-concept demonstration of DWNs for both analyte identification and concentration estimation using data from two peer-reviewed differential sensing studies. The first study involves classification of Pinot Noir wines by vineyard and region; the second analyzes characteristics of common mitogen-activated protein kinases. Across both studies, DWNs achieve competitive predictive performance relative to baseline methods, while exhibiting a reduction in model parameter size ranging from approximately 3.6x to 1668x compared to their Random Forest counterparts, illustrating their suitability for highly constrained environments.

These results highlight the potential of DWNs as a viable tool for differential sensing data analysis in various chemical sensing applications, and point toward deployable, low-power systems in which analysis models are integrated directly into custom hardware for real-time inference at the edge.

KEYWORDS

differential sensing, chemometrics, compact machine learning, weightless neural networks

1 INTRODUCTION

Traditionally, chemists designing chemical sensors have strived for lock-and-key specificity by searching for receptors which bind one analyte with highly competitive selectivity. Differential sensing takes a fundamentally different approach, using an ensemble of broadly responsive sensors whose collective, pattern-rich signals encode information about a sample [6, 10]. In the same way that olfactory and gustatory systems rely on combinatorial receptor activation to interpret complex stimuli, differential sensing takes advantage of cross-reactive sensor arrays to form data fingerprints that can be decoded computationally.

At the heart of differential sensing is an array of cross-reactive molecular sensors with overlapping affinities and interaction modes, so multiple sensors respond to the same and/or multiple chemical features with distinct sensitivities and signal characteristics [6]. Rather than relying on a single, highly specific receptor, the array comprises various sensor types such as dyes, supramolecular hosts, polymers, nanoparticles, peptides, or conductive materials, with each sensor broadly responding to multiple chemical features [5, 6, 9]. When exposed to a sample, the array produces a multisensor response profile, and because no single sensor is required to uniquely identify the analyte, the system is both adaptable and resilient, having the ability to discriminate complex mixtures and classes of compounds by their emergent signatures. Differential sensing arrays can be read through several complementary output channels, enabling deployment across a range of chemistry environments and instruments. Examples include colorimetric signals, fluorometric outputs, and electrical/electrochemical measurements from chemiresistive, nanomaterial, or electrode-based sensors. Figure 1 provides a visual representation of the differential sensing analysis, illustrating how one or more analytes interact with distinct differential receptors, resulting in a multicolor output.

The multi-dimensional output of cross-reactive arrays lends itself naturally to data-driven analysis. Chemometric methods, such as principal component analysis, linear discriminant analysis, and partial least squares, help reveal key features and relationships

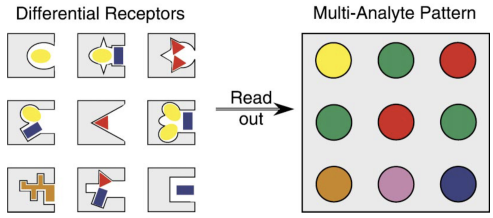

Figure 1: Depiction of Differential Sensing [6]

in the signals [9, 10]. Machine learning approaches have become more prevalent with differential sensing data analysis, and either build on these outputs or operate directly on the array readouts to deliver robust classification and regression, capturing nonlinearity and sensor–sensor interactions [5, 9]. These tools can be used to transform raw cross-reactive signal sensor response data into actionable information. For example, by identifying sample classes, quantifying analyte concentrations, or characterizing subtle chemical phenotypes. [9, 10].

Traditionally, these data analyses are performed on powerful computers or cloud servers, which can limit the deployment of sensing systems in resource constrained environments, and introduce latency, privacy, and connectivity concerns. As demand grows for portable, real-time chemical sensing systems, for example, in environmental monitoring, food authentication, or point-of-care diagnostics, there is a critical need for machine learning models that can be directly integrated into sensor hardware and operate efficiently at the edge.

In this paper, we propose the use of Differentiable Weightless Neural Networks (DWN) [2] as a small, energy-efficient machine learning model that can be integrated in the future directly alongside differential sensor arrays for direct inference at the site where data are collected. Using data from two existing peer-reviewed studies [4, 11], we evaluate the potential of DWNs for the use of differential sensing analysis. In the first study, differential sensing was used to analyze Pinot Noir Wines "derived from the same scion clone" and grown in different locales [4], with the ultimate goal of classifying both the region and the vineyard where the wine was produced. The second study involves the analysis of the presence and concentration of different mitogen-activated protein kinases [11]. For both sets of studies, we evaluate the performance of the DWN against the methods of analysis used in the original studies, as well as competing machine learning models.

The remainder of the paper is organized as follows - Section 2 introduces DWNs. In Section 3 we explain the experimental methodology and our results for both sets of experiments, and Section 4 concludes the paper.

2 DIFFERENTIABLE WEIGHTLESS NEURAL NETWORKS (DWN)

The Differentiable Weightless Neural Network (DWN) is the recent development in the field of Weightless Neural Networks (WNNs) [2]. While Deep Neural Networks (DNNs) mark the standard for many different machine learning applications, DNN performance has come at the cost of expounding model size and complexity. These models would require significant energy to run even simple inference tasks, rendering them unsuitable for applications pertinent to sensor-array integration which are limited in resources[8]. WNNs are a unique class of neural network models that attempt to simplify models by forgoing the use of weighted connections, which are prevalent in DNNs along with their associated multiply and accumulate operations, for the use of lookup tables (LUTs) that act as memory nodes that hold binary data [7]. LUTs are able to efficiently represent nonlinear functions of binary inputs. With this capability, single-layer WNNs are able to learn patterns and interactions that standard deep neural networks would need several non-linear layers to achieve [8].

DWNs address several limitations of historical WNN implementations, which we utilize for our analysis. First, prior implementations have been limited to single-layer architectures due to the lack of gradient-based optimization for learning the discrete bits encoded in the LUTs. To address the non-differentiable nature of LUTs, the DWN introduces the use of a new technique called Extended Finite Difference (EFD), which approximates the derivative of LUT lookup operations. Using EFD, we can use the derivative approximation to efficiently compute gradients for more complex LUTs within multiple layers. This allows DWN to be trained via backpropagation like traditional neural networks [2].

Secondly, previous implementations also required randomized connections between the input and LUTs, and the specific mapping between the input and LUTs could have a significant impact on training results. DWN introduces the concept of Learnable Mapping, in which the connections between the input and LUTs, as well as between different LUT layers, can also be learned using gradient-based optimization. The increased complexity introduced by the Learnable Mapping affects only the training phase, as the connections can be fixed during inference, resulting in no additional computational cost relative to a randomized mapping [2].

2.1 DWN Architecture

We employ the DWN architecture for our differential sensing data analysis, leveraging the unique properties of layers composed entirely of lookup tables (LUTs). As illustrated in Figure 2, the DWN receives a binary input vector, where each bit is mapped through an encoding scheme that generates address patterns. These address patterns are used to index into the LUTs of the initial layer, allowing each LUT to retrieve a binary value from its designated address. The collection of binary outputs from all LUTs in the layer is then assembled into a new binary vector, which serves as the input for the next layer of LUTs.

This process is repeated across multiple layers, with each successive layer further transforming and abstracting the input data. In the final LUT layer, the resulting binary output is partitioned into segments corresponding to the number of output classes. For each class segment, a population count operation is performed, summing the bits within the segment. The resulting counts act as activation scores for each class, and the class with the highest activation is selected as the predicted label.

This architecture enables efficient and highly parallelizable inference, as each LUT operation is independent and can be implemented with minimal computational overhead. By structuring the output

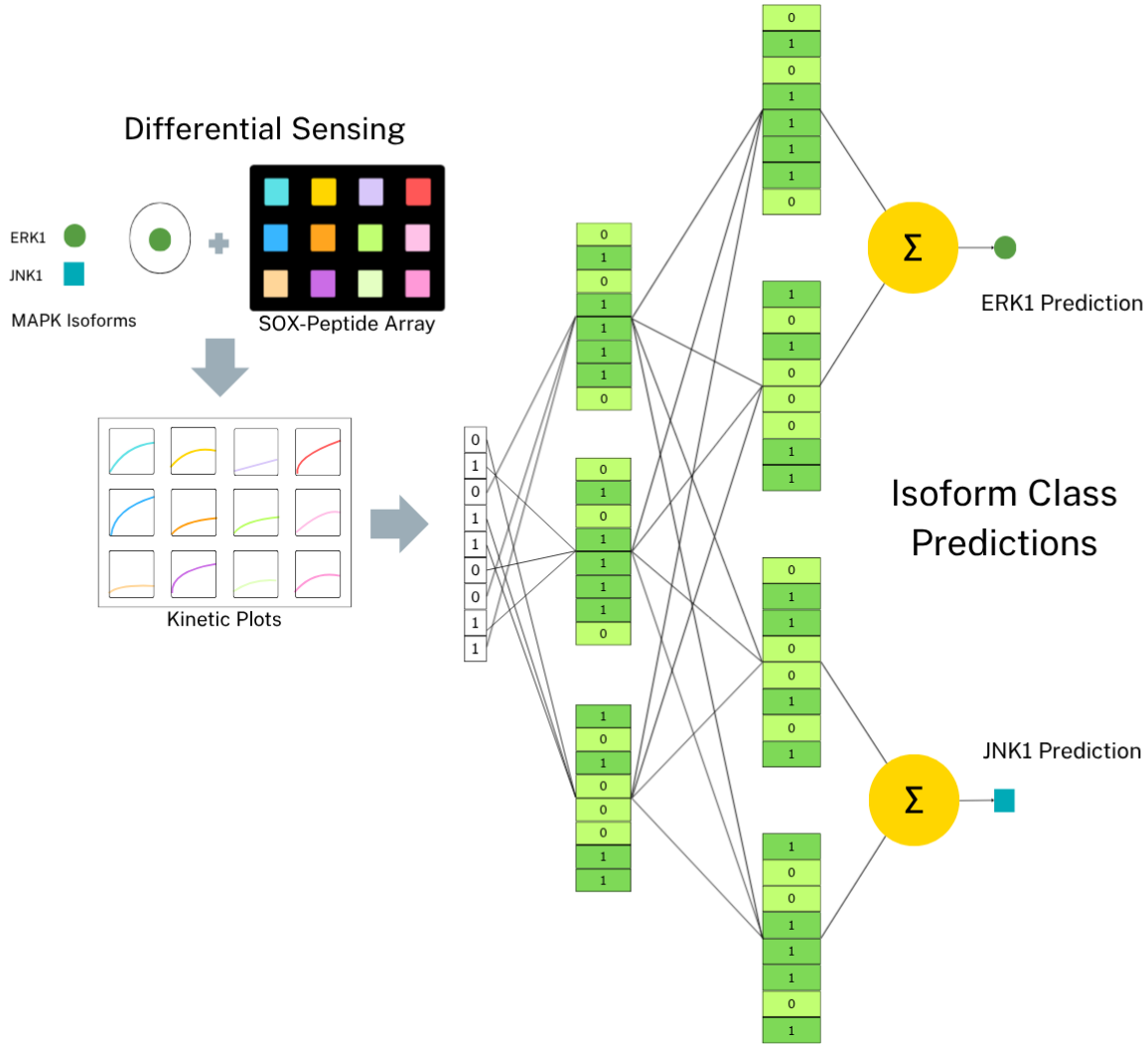

Figure 2: Simple DWN example showcasing kinase isoform classification

in this way, the DWN is able to robustly differentiate between classes using compact, interpretable binary representations, making it particularly well suited for memory and energy-constrained environments where conventional neural networks may be impractical.

2.2 Input Binarization

The DWN and other weightless neural network implementations require binary inputs for operation. However, differential sensing data, as well as data from many general applications, are typically continuous or categorical rather than binary. Therefore, an effective binarization method is necessary to prepare the data for training and inference. One widely used technique for this purpose is thermometer encoding.

In thermometer encoding, each input feature is discretized into levels, and each level is represented by a binary vector where all bits up to the current level are set to 1, and the remaining bits are set to 0. This results in a binary representation that maintains information about the magnitude and ordering of the original data, which can improve the learning capacity of weightless neural networks. [1, 3]

3 EXPERIMENTAL EVALUATION

We evaluated the performance of the DWN on two sets of prior differential sensing studies. The first set of experiments concerns the classification of the vineyard and appellation of similar wines [4]. The second set of experiments categorizes and detects the presence of certain mitogen-activated protein kinases (MAPKs) [11]. From the kinase study, there are two different tasks on which we perform analysis. The first task involves classifying different MAPK isoforms, while the second focuses on quantifying the concentrations of three specific MAPK isoforms within mixtures containing varying proportions of the isoforms.

Table 1: Description of datasets used for DWN evaluation

Dataset	Experiment Description
Wine Classification [4]	15 Pinot Noir wines were produced in 15 different vineyards across 8 different regions on the United States' Pacific Coast. Differential Sensing analysis was performed using 9 peptide sensing ensembles using combinations of two divalent metal ions three colorimetric indicators (PCV, CAS and BPR) and three peptides (WAHEDEFF, FHFPHHF and WEEHEE). Absorbance values were collected in relation to the respective indicators corresponding to their λmax 430 nm, 444 nm, and 560 nm resulting in a total of 27 peptide sensors. Further, 8 analytical replicates were carried out for each wine resulting in a total group of 120 experimental samples. These 15 wines were produced in two vintage years (2015, 2016), with the same experiment carried out across both years, resulting in two datasets. Analysis was performed to classify both the specific vineyard as well as the region that the wine was produced.
Kinase Classification [11]	A kinase is an enzyme and protein kinases act as catalysts in modification of proteins. A cross-reactive array consisting of 14 SOX-peptides was developed to categorize the isoforms of common Mitogen-Activated Protein Kinases (MAPK) across three significant families: ERK (ERK1, ERK2), JNK (JNK1, JNK2, JNK3), and p38 (p38α, p38β, p38γ, p38δ). Nine kinases, each containing a specific isoform listed above, were analyzed using differential sensing with a cross-reactive array. The fluorescence signal produced by each SOX-peptide sensor was monitored for 120 minutes using a fluorescent plate reader. For analysis, the rate constant and the normalized fluorescence signal values at 30, 60, and 90 minutes were used, yielding a total of 56 fluorescence signal measurements. For each kinase isoform, four analytical replicate experiments were performed, resulting in 36 experimental samples.
Kinase Concentration [11]	Ternary MAPK mixtures were produced containing various concentrations of ERK1, JNK3, and p38γ. A cross reactive array of 12 SOX-peptides was produced, with 52 normalized florescence signal measurements captured per SOX-peptide sensor over a period of 120 minutes, resulting in a total of 624 measurements per sample analyzed. 65 MAPK mixtures were produced, with 4 analytical replicate experiments performed per MAPK mixture, for a total of 260 experimental samples. Analysis was performed to determine the specific concentrations of the ERK1, JNK3 and p38γ present in the mixture at nanomolar sensitivity. This experiment was performed twice, with the MAPK mixture analyzed both in vitro and in A549 Cell Lysate solution, resulting in two datasets.

A summary of the three general tasks is referenced in Table 1.

3.1 Wine Classification

In [4], Pinot Noir wine grapes were produced in 15 different vineyards located along the Pacific Coast of the United States. For wine classification, we evaluate the DWN on two datasets from wines produced in 2015 and 2016. We use the identical datasets from prior work [4].

Dataset Inputs. For both datasets, 15 wine samples with 8 analytical replicates for a total of 120 experimental samples were used. The differential analysis peptide sensors produced 27 absorbance measurements that were used as input features for the model.

Wine Classification Tasks. The goal is to classify which one of the 15 vineyards the wine sample was produced, as well as the specific region (American Viticultural Area - AVA) where the vineyard is located.

Model Configurations and Training. Individual models were trained to classify the vineyard and region separately for both datasets, resulting in 4 total DWN models. Each of the DWN models consists of two layers of LUTs and a popcount layer that designates the class selection. A softmax temperature, τ, is applied to the output of the popcount layer. As discussed in section 2.2, a thermometer is used to binarize features of input into a specified number of bits. For training and evaluation of each model, we perform 5-fold cross validation. For each fold, a DWN model was trained for 30 epochs on the remaining folds, decayed by a factor of 0.1 every 14 epochs. The configuration of the 4 training models, the learning rate used during training and the number of thermometer bits used for binarization is located in Table 2.

Results. In the original study, Random Forest was the tool used to perform data analysis. To match the same training and testing environment of the DWN models, we trained our own set of Random Forest models in order to evaluate the performance of the DWN models. A comparison of all our model results for both datasets is located in Table 4. Overall, for the 4 classification tasks, the DWN performed comparably or outperformed the Random Forest classification models, while maintaining significantly smaller model parameter size. The parameters of the DWN models consist of the bits within each of the LUTs located in the LUT layers.

Table 2: DWN Models for Wine Classification

Target	#LUTs Layer 1	#LUTs Layer 2	Thermometer Bits	Tau	Learning Rate	Size (KiB)
2015 Vineyard	2600 (LUT-2)	2500 (LUT-2)	6	6.27	0.064	2.49
2015 Region	2900 (LUT-6)	1400 (LUT-6)	4	10.0	0.01	25.78
2016 Vineyard	900 (LUT-4)	400 (LUT-6)	7	4.0	0.01	4.9
2016 Region	1400 (LUT-2)	600 (LUT-2)	7	6.27	0.005	0.97

3.2 Kinase Classification

The first kinase experiment we evaluated is the classification of different Mitogen-Activated Protein Kinase (MAPK) isoforms, using the identical datasets from prior work [11].

Dataset Inputs. One sample of each of the nine following MAPK isoforms: ERK1, ERK2, JNK1, JNK2, p38α, p38β, p38γ, p38δ, with 4 analytical replicates for a total of 36 experimental samples. Each of the samples had 14 SOX-peptide based sensors, which provided fluorescence signal measurements. For each sample, the rate of fluorescence change and the normalized fluorescence value at 30, 60 and 90 minutes resulted in 56 total measurements which serve as input features to the model.

Kinase Classification Tasks. The goal of the analysis is to determine, based on the 56 fluorescence signals, which of the 9 kinase isoforms is present.

Model Configurations and Training. The DWN architecture for this task is similar to that of the wine classification task. The DWN model consist of two layers of LUTs and a popcount layer. A softmax temperature, τ, is applied to the output of the popcount layer. The DWN model is trained using Leave-One-Out Cross Validation (LOOCV). In this approach, the model is trained 36 separate times, each time using 35 samples for training and the remaining one for testing. The final reported accuracy reflects the number of samples correctly identified out of the 36 total.

The trained DWN model had the following configuration and training parameters:

- **#LUTs Layer 1**: 500 (LUT-2)
- **#LUTs Layer 2**: 200 (LUT-2)
- **Thermometer Bits**: 2
- τ: 8.459
- **Batch Size**: 16
- **Learning Rate**: 0.019, decayed by a factor of 0.1 every 14 epochs.
- **Total Training Epochs**: 30

Results. Using LOOCV, the DWN achieves 100% classification accuracy, which matches the performance of the Linear Discriminant Analysis (LDA) statistical method used to analyze the data in the original study [11]. We performed our own version of LDA analysis, and trained a Random Forest model with near perfect accuracy for size comparison. While the parameter count for all 3 models is viable for portable analysis, the DWN model achieves this level of accuracy with the smallest architecture, requiring only 0.34 KiB of parameters. The accuracy and parameter size of all three models is recorded in Table 4.

3.3 Kinase Concentration

The second kinase experiment [11] evaluated the DWN's ability to quantify the concentrations of three kinase isoforms (ERK1, JNK3 and p38γ) within a mixture containing all three isoforms. Differential sensing analysis was carried out twice, once with the mixtures multiplexed in vitro, and once with the mixtures multiplexed in A549 Cell Lysate solution, resulting in two distinct datasets.

Dataset Inputs. 65 MAPK mixtures contained varying concentrations of ERK1 (up to 1nm), JNK3 (up to 9nm) and p38γ (up to 5nm) were analyzed with 4 replicate experiments for a total of 260 experimental samples. The fluorescence signal produced by 12 SOX-peptide sensors was monitored over a period of 120 minutes. Over the course of 120 minutes, 52 fluorescence signal measurements per peptide were used as input features for the model, for a total of 624 input features.

Kinase Concentration Task. Using the 624 fluorescence signal measurements, the goal is to determine the specific concentration, in nM, of ERK1, JNK3 and p38γ. The model performance is evaluated using the root mean square error (RMSE) of the prediction compared to the actual concentration values.

Model Configurations and Training. For both datasets, three separate DWN models were trained to predict the concentration for each of the individual isoforms, as this method provided lower RMSE over utilizing a single DWN model. The DWN model configuration for each of the 6 models across both datasets consists of two LUT layers, a ternarization layer, and concentration output layer. The final configuration and training parameters used for each DWN model, as well as the number of bits used for thermometer binarization for each model, are referenced in Table 3. The dataset was split into an 80-20 independent train-test split that was used in the original data analysis [11].

Results. The recorded results for all models trained for analysis are located in Table 4. For both datasets, the prior work analysis [11] trained a partial least squares (PLS) model with which we compared the performance of our individual DWN models. For the in vitro dataset, DWN models performed on par or slightly better than the PLS model for ERK1 and JNK3, with root mean square error (RMSE) values of 0.0915 nM (DWN) vs. 0.1023 nM (PLS) for ERK1, and 0.9985 nM (DWN) vs. 1.01 nM (PLS) for JNK3. However, for p38γ, the DWN model exhibited a higher RMSE of 0.637 nM compared to 0.4446 nM for the PLS model.

In the A549 Cell Lysate dataset, DWN models performed on par or better for all three kinase isoforms. The RMSE values were 0.2112 nM (DWN) vs. 0.2253 nM (PLS) for ERK1, 2.096 nM (DWN) vs. 2.471 nM (PLS) for JNK3, and 0.729 nM (DWN) vs. 0.936 nM (PLS) for p38γ.

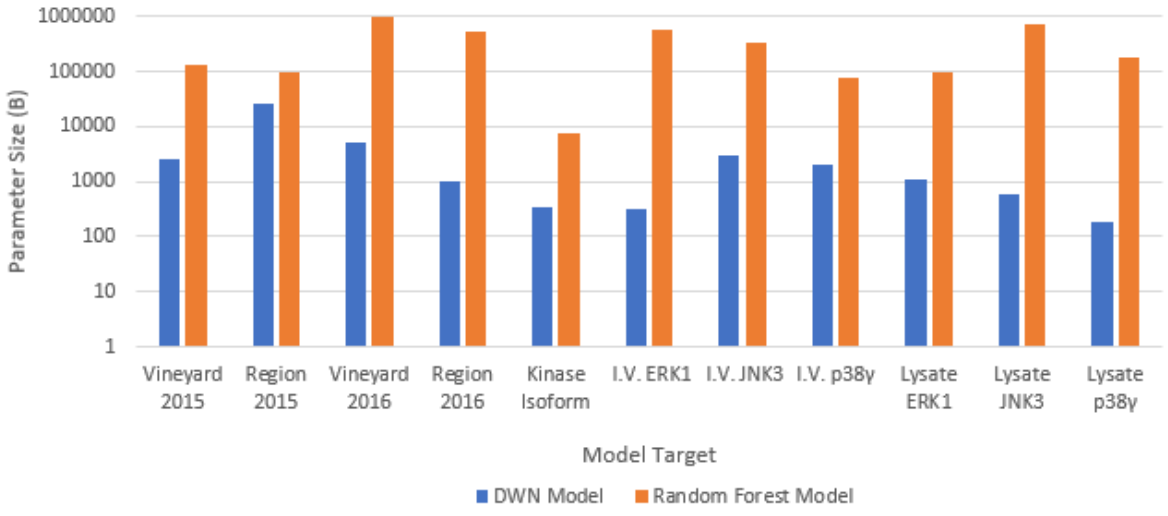

Figure 3: Parameter Size Comparison of DWN model (ours) vs Random Forest Models

To further evaluate the performance and parameter size of the DWN models, we trained separate Random Forest models for each analyte prediction across both datasets. The DWN models consistently outperformed the Random Forest models in terms of predictive accuracy, as measured RMSE in concentration. For the in vitro dataset, the DWN models achieved RMSE values of 0.0915 nM, 0.9985 nM, and 0.637 nM for ERK1, JNK3, and p38γ, respectively, compared to the Random Forest models' RMSEs of 0.1088 nM, 1.194 nM, and 0.8577 nM. In the A549 Cell Lysate dataset, DWN models yielded RMSEs of 0.2112 nM, 2.096 nM, and 0.729 nM, while Random Forest models produced RMSEs of 0.3729 nM, 2.232 nM, and 1.170 nM.

3.4 Model Size Comparison

As shown in Figure 3, analysis of model size across diverse datasets reveals a dramatic advantage for DWN models over Random Forest models. Table 4 details that DWN models range from just 0.18 KiB to 25.78 KiB, while Random Forest models span from 7.48 KiB up to 974.75 KiB. In every case, DWN models are tens to thousands of times smaller than their Random Forest counterparts. For example, when comparing the models predicting p38γ for the lysate dataset, the DWN model is nearly 1,000 times smaller than the corresponding Random Forest model (0.18 KiB vs. 173.88 KiB). This extreme reduction in model footprint is highly relevant for edge AI applications, where memory, storage, and power are limited. Smaller models like the DWN are exceptionally well-suited for deployment on FPGAs and custom hardware platforms, especially when tightly integrated with IoT sensors. In such environments, constraints on memory, computational capacity, and power consumption make the implementation of larger, more complex models both challenging and costly.

4 CONCLUSION AND FUTURE WORK

This work demonstrates that Differentiable Weightless Neural Networks (DWNs) are a robust and practical method for differential sensing analysis, delivering predictive performance comparable to established chemometric and machine learning approaches. Notably in comparison to Random Forest models, DWNs achieve this accuracy with a significantly reduced model parameter size (up to approximately 1668x), emphasizing their efficiency and making them highly suitable for deployment in resource-constrained environments.

Previous research has shown that DWN models can be implemented directly in hardware [2], further enhancing their appeal for edge applications. Building on these findings, future work will focus on the seamless integration of hardware DWN models with

Table 3: DWN Model Configurations for MAPK Concentration Prediction

Model	#LUTs Layer 1	#LUTs Layer 2	Thermometer Bits	Learning Rate
In vitro - ERK1	200 (LUT-2)	300 (LUT-2)	3	0.0184
In vitro - JNK3	1800 (LUT-3)	1000 (LUT-3)	4	0.0151
In vitro - p38γ	300 (LUT-2)	2400 (LUT-2)	4	0.0184
Lysate - ERK1	1200 (LUT-2)	700 (LUT-2)	3	0.018
Lysate - JNK3	600 (LUT-2)	400 (LUT-2)	6	0.0952
Lysate - p38γ	200 (LUT-2)	100 (LUT-2)	6	0.0812

Table 4: Summary of Results

Dataset & Model Target	Model	Accuracy %	RMSE	Parameter Size (KiB)
2015 Wine Vineyard	DWN	98.3	-	2.49
	Random Forest	98.3	-	124.25
2015 Wine Region	DWN	94.8	-	25.78
	Random Forest	94.8	-	93.5
2016 Wine Vineyard	DWN	85	-	4.9
	Random Forest	85.83	-	974.75
2016 Wine Region	DWN	84.17	-	0.97
	Random Forest	79.17	-	530.62
Kinase Isoform	DWN	100	-	0.34
	Linear Discriminant Analysis	100	-	3.76
	Random Forest	97.22	-	7.48
Kinase Multiplex - In Vitro / ERK1	DWN	-	0.0915	0.32
	Partial Least Squares [11]	-	0.1023	-
	Random Forest	-	0.1088	533.81
Kinase Multiplex - In Vitro / JNK3	DWN	-	0.9985	2.98
	Partial Least Squares [11]	-	1.01	-
	Random Forest	-	1.194	326.06
Kinase Multiplex - In Vitro / p38γ	DWN	-	0.637	1.91
	Partial Least Squares [11]	-	0.4446	-
	Random Forest	-	0.8577	73.41
Kinase Multiplex - Lysate / ERK1	DWN	-	0.2112	1.1
	Partial Least Squares [11]	-	0.2253	-
	Random Forest	-	0.3729	93.56
Kinase Multiplex - Lysate / JNK3	DWN	-	2.096	0.59
	Partial Least Squares [11]	-	2.471	-
	Random Forest	-	2.232	709.41
Kinase Multiplex - Lysate / p38γ	DWN	-	0.729	0.18
	Partial Least Squares [11]	-	0.936	-
	Random Forest	-	1.170	173.88

differential sensing sensor arrays. This integration will enable the development of compact, energy-efficient systems capable of real-time chemical analysis and adaptable to a variety of sensing tasks, ultimately advancing the field toward deployable, intelligent chemical sensing platforms.

ACKNOWLEDGMENTS

This research was supported in part by NSF Grants 2326894, 2425655, and NVIDIA Applied Research Accelerator Program Grant. Any opinions, findings, conclusions, or recommendations are those of the authors and not of the funding agencies.

REFERENCES

[1] Alan Bacellar, Zachary Susskind, Luis Villon, Igor Miranda, Leandro Santiago, Diego Dutra, Mauricio Jr, LIZY JOHN, Priscila Lima, and Felipe França. 2022. Distributive Thermometer: A New Unary Encoding for Weightless Neural Networks. 31–36. doi:10.14428/esann/2022.ES2022-94

[2] Alan T. L. Bacellar, Zachary Susskind, Mauricio Breternitz Jr, Eugene John, Lizy Kurian John, Priscila Machado Vieira Lima, and Felipe M.G. França. 2024. Differentiable Weightless Neural Networks. In *Proceedings of the 41st International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 235)*, Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (Eds.). PMLR, 2277–2295. https://proceedings.mlr.press/v235/bacellar24a.html

[3] Hugo Carneiro, Felipe França, and Priscila Lima. 2015. Multilingual part-of-speech tagging with weightless neural networks. *Neural Networks* 66 (03 2015). doi:10.1016/j.neunet.2015.02.012

[4] Alexandra A. Crook, Diana Zamora-Olivares, Fatema Bhinderwala, Jade Woods, Michelle Winkler, Sebastian Rivera, Cassandra E. Shannon, Holden R. Wagner, Deborah L. Zhuang, Jessica E. Lynch, Nathan R. Berryhill, Ron C. Runnebaum, Eric V. Anslyn, and Robert Powers. 2021. Combination of two analytical techniques improves wine classification by Vineyard, Region, and vintage. *Food Chemistry* 354 (2021), 129531. doi:10.1016/j.foodchem.2021.129531

[5] William M. Dawson, Kathryn L. Shelley, Jordan M. Fletcher, G. Davies, F. J. O. Martin, F. J. Wiseman, R. L. Brady, D. Tew, C. W. Wood, and D. N. Woolfson. 2023. Differential sensing with arrays of de novo designed peptide assemblies. *Nature Communications* 14, 1 (2023), 383. doi:10.1038/s41467-023-36024-y

[6] John J. Lavigne and Eric V. Anslyn. 2001. Sensing A Paradigm Shift in the Field of Molecular Recognition: From Selective to Differential Receptors. *Angewandte Chemie International Edition* 40, 17 (2001), 3118–3130. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/1521-3773(20010903)40:17<3118::AID-ANIE3118>3.0.CO;2-Y doi:10.1002/1521-3773(20010903)40:17<3118::AID-ANIE3118>3.0.CO;2-Y

[7] Zachary Susskind, Aman Arora, Igor D.S. Miranda, Luis A.Q. Villon, Rafael F. Katopodis, Leandro S. De Araújo, Diego L.C. Dutra, Priscila M.V. Lima, Felipe M.G. França, Mauricio Breternitz, and Lizy K. John. 2022. Weightless Neural Networks for Efficient Edge Inference. In *31st International Conference on Parallel Architectures and Compilation Techniques (PACT)*. doi:10.1145/3559009.3569680

[8] Zachary Susskind, Aman Arora, Igor D. S. Miranda, Alan T. L. Bacellar, Luis A. Q. Villon, Rafael F. Katopodis, Leandro S. de Araújo, Diego L. C. Dutra, Priscila M. V. Lima, Felipe M. G. França, Mauricio Breternitz Jr., and Lizy K. John. 2023. ULEEN: A Novel Architecture for Ultra-low-energy Edge Neural Networks. *ACM Trans. Archit. Code Optim.* 20, 4, Article 61 (dec 2023), 24 pages. doi:10.1145/3629522

[9] Jia-Hong Tian, Zhe Zheng, Yu-Chen Pan, Yue-Fei Wang, and Dong-Sheng Guo. 2025. Macrocycle-based Differential Sensing: Design Strategies and Applications. *Responsive Materials* 3 (2025). doi:10.1002/rpm.20240036

[10] Alona P. Umali and Eric V. Anslyn. 2010. A general approach to differential sensing using synthetic molecular receptors. *Current Opinion in Chemical Biology* 14, 6 (2010), 685–692. doi:10.1016/j.cbpa.2010.07.022

[11] Lingyu Zeng, Tamer S. Kaoud, Diana Zamora-Olivares, Amanda L. Bohanon, Yiru Li, Jacey R. Pridgen, Yakndara E. Ekpo, Deborah L. Zhuang, Jessica R. Nye, Mitchell Telles, Michelle Winkler, Sebastian Rivera, Federico Marini, Kevin N. Dalby, and Eric V. Anslyn. 2022. Multiplexing the Quantitation of MAP Kinase Activities Using Differential Sensing. *Journal of the American Chemical Society* 144, 9 (2022), 4017–4025. arXiv:https://doi.org/10.1021/jacs.1c12757 doi:10.1021/jacs.1c12757 PMID: 35195411.